The gene-culture coevolution of human decision-making

Aysha Bellamy

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology

Royal Holloway, University of London

2022

Abstract

This thesis aimed to investigate both (i) the flexibility of social learning and (ii) the complexity of decision-making from a gene-culture coevolutionary perspective. I investigated three key areas of human decision-making: asocial skills, social norms, and cooperation by utilising a game against nature, a coordination game and a Prisoner's Dilemma respectively. To address the first aim, chapter 3 investigated the social learning of asocial skills and social norms, and chapter 4 investigated the social learning of cooperation, in an empirical setting. The participants were more flexible than previously found. They adjusted their frequency-dependent social learning strategies to a third-order complexity, including: (i) the frequency of the choices made by the group from whom they learned; (ii) whether this group were identified as learning in a similar or different environment to the participant and (iii) the reliability of this similarity signal. There was an upper limit to this flexibility, as all participants found it easier to master asocial skills, social norms, and cooperative behaviours from groups of reliably similar others. To address the second research aim, chapters 5-6 used agent-based models to investigate whether the cognitive and motivational processes underlying decision-making were likely to be fully modular, partly modular, or domain-general. Fully modular psychology was necessary to acquire skills and norms in 2 distinct domains, though modular cognition was more important to skill acquisition. Domain-general agents were instead needed to uphold the costly levels of cooperation seen across human societies. Any agent may uphold suboptimal behaviour simply via drift. Together, these findings have implications for our understanding of how maladaptive behaviour is maintained via cultural evolutionary processes. Maladaptive behaviour may be upheld via both a trade-off in social learning flexibility and a trade-off in the complexity of decision-making, which in turn is impacted by drift.

Declaration of Authorship for Co-authored Work

If you are presenting partly co-authored work, please indicate below your individual contribution to the thesis.

Name of candidate: ....Aysha Bellamy...........................

Thesis title: ....The gene-culture coevolution of human decision-making ....

I confirm that the thesis that I am presenting has been co-authored with:

........Prof Ryan McKay, Prof Charles Efferson and Prof Sonja Vogt ....

Within this partly co-authored work, I declare that the following contributions are entirely my own work:

(Here you should indicate, in précis style, the datasets that you gathered, interpreted and discussed; methods that you developed; complete first drafts that you wrote; content that is entirely your own work; etc. It is often appropriate to organise this statement by chapter)

Chapter 1: Introduction
This is entirely my own work

Chapter 2: Methods
This is entirely my own work

Chapter 3: What is the extent of a frequency-dependent social learning strategy space?
I formulated this idea with the assistance of Charles Efferson. Charles wrote the code for the Z-Tree programme used in the testing sessions. I wrote the R script used to analyse the data gathered. I wrote the manuscript with feedback from Charles, Ryan McKay, and Sonja Vogt.

Chapter 4: When will you learn to help? Social learners can respond flexibly to a range of social information when learning to cooperate
I formulated this idea with the assistance of Charles Efferson. Charles wrote the code for the Z-Tree programme used in the testing sessions. I wrote the R script used to analyse the data gathered. I wrote the manuscript with feedback from Ryan McKay.

Chapter 5: Jack of all trades: Modular cognition and motivation may underlie our ability to master skills over distinct domains
I formulated this idea with the assistance of Charles Efferson. I wrote the code for the agent-based model, and the R scripts for analysis, with assistance from Charles. I wrote the manuscript with feedback from Charles Efferson.

Chapter 6: It takes two: Modular psychology allows us to coordinate on various social norms including maladaptive ones
I formulated this idea with the assistance of Charles Efferson. I wrote the code for the agent-based model, and the R scripts for analysis, with assistance from Charles. I wrote the manuscript with feedback from Charles Efferson.

Chapter 7: I only want to help: Disentangling the influence of cognition and motivation on the emergence of cooperation
I formulated this idea with the assistance of Charles Efferson. I wrote the code for the agent-based model, and the R scripts for analysis, with assistance from Charles. I wrote the manuscript with feedback from Charles Efferson.

Chapter 8: Conclusion
This is entirely my own work.

Signed: ABellamy ....ABellamy........ Date:.......31/01/2022........
(Candidate)

Signed: Ryan McKay Ryan McKay Date: 31/01/2022
(Supervisor)

This form should be signed by the candidate and the candidate's supervisor and returned to the Doctoral School when the electronic copy of the thesis is submitted Student Administration, Royal Holloway, Egham, Surrey TW20 0EX with the copies of the thesis.

Acknowledgements

No (wo)man is an island and I have had the pleasure of working in such a supportive atmosphere throughout my PhD process, that the list of people I wish to thank may end up being another book itself! To save some time (and also trees) I am having to stick to the main players, but please know that if I have had the pleasure of working with you over the past five years but your name has not made the list below, I still value your support and friendship more than two pages can convey.

First and foremost, thanks to my wonderful supervisors Charles and Ryan. Your feedback over the past few years has not only improved my writing and public speaking immensely but has made me a more confident person than I was at the start of this journey. Charles, thank you so much for the support from Switzerland. Although I was initially worried about having a long-distance supervisor, it has paid off in so many ways. A massive thank you for introducing me to all the brilliant scientists working in the PACE lab, including the wonderful Sonja. Thanks to Ryan for stepping up to primary, and for always being available to chat over ideas (including papers that did not even make the final cut of the thesis as unfortunately I'm already pushing the maximum word count that my conscience will permit me to submit!)

I would also like to thank all of the Psychology department who have supported me throughout this journey: thank you so much to Dawn and Saloni for always lending a supportive ear, thanks to the Psychology PhD department who I can honestly say have been the most fabulous cohort to work with, and thanks to the lovely teaching staff that have made me feel so welcome as I transitioned to a new role here. Especially, a massive thanks to Bea, Louisa, Farah, and Becky for being the best Doctor friends available and for always being ready to pick me up with a glass of wine (or indeed, bottle).

A massive thanks to all the students who I have taught on the MSc and PS2010 cohort this year for your continued support and understanding as I ploughed through with the final chapters of my doctorate.

Moving beyond the department, I would love to thank my best friend Jo for all the support that she has given me over the past 13 years of our life and for always knowing what to say to pick me up when I feel like I've hit a wall (to be honest, it usually a Disney quote).

Many thanks to my wonderful friends who have supported me ever since university take 1: Jess, Ben, Paige, Sammi, Tom and Nikki. If your support can get me through three degrees, then let's be honest- there's nothing you can't do.

Of course, I have to thank my wonderful family for their continued support throughout not just the past five or so years of study, but throughout twenty-seven years of study because- let's be honest- I have never stopped. Many thanks to my wonderful supportive parents, Robert and Angy, and to my brothers, Callum and Shane, who never fail to help me see the funny side of life. A massive thank you to my sister-in-law Chloe (whose notes from undergrad I still have somewhere!) and to my niece, Emily (who knows? By the time I've submitted, she may be able to read this!).

Even though she is no longer here with us (and knew nothing about my PhD when she was), I'd love to thank my best friend from childhood, Ebony, who taught me that there isn't a problem in the world that can't be solved by hugging your fluffy friend.

I would also love to thank my grandparents for their continued support and for always supplying me with treats whenever I go to see them. Thank you Violet and Margaret. Finally, I wish to dedicate this book to my Grandad, William, from whom I inherited my love of reading (and fried foods).

Table of contents

Table of tables

Table of figures

0. Thus, the payoff to the agent for matching her skillset to the environment is driven entirely by the size of the f_{Bzero} and f_{Bone} parameters.	
Figure 3. The agent outcomes on runs where the agent made decisions in two similar skill domains. The x axis gives the possible outcomes. Both_opt represents cases where the agent chose the optimal behaviour in both domains, S0A indicates correct answers in state 0 of domain A, S0B in state 0 of domain B, S1A in state 1 of domain A and S1B in state 1 of domain B. Results for the fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents are graphed separately. The graphs display the agents' response to i) runs where the priors and fitness tied to matching one's skillsets to the environment favour state 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4) and, ii) runs where the most common environmental state to match on is not necessarily the state which gives the highest payoff (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.	**23**
Figure 4. The agent outcomes on runs where the agents made decisions in two distinct skill domains. The x axis gives the possible outcomes. Both_opt represents cases where the agent chose the optimal behaviour in both domains, S0A indicates correct answers in state 0 of domain A, S0B in state 0 of domain B, S1A in state 1 of domain A and S1B in state 1 of domain B. Results for the fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents are graphed separately. The graphs display the agents response to i) runs where the priors and fitness tied to matching one's skillset to the environment favoured state 0 in domain A but state 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25) and ii) runs where the most common environment was not necessarily the environment which gave the highest payoff, over two distinct domains (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.	**28**
Figure 5 and Figure 6. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) partly modular agents with modular motivation but domain-general cognition; (iii) partly modular agents with modular cognition but domain-general motivation and (iv) fully domain-general agents when deciding in two similar skill domains. Figure 5 gives the results for runs where both the priors and the fitness tied to matching one's skillsets to the environment favour behaviour 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). Figure 6 gives the results for runs where the priors favour state 1 but the fitness tied to matching one's skillsets to the	**35-50**

Figure 4. Bar charts showing the agent outcomes on runs where the agents try to coordinate in two distinct domains. Note that the agent outcomes on the x axis are dependent on the focal agent's behaviour. Therefore, miscoord_Opt represents cases of miscoordination where the focal agent chose what would have been the Pareto optimum and miscoord_Subopt represents cases of miscoordination where the focal agent chose what would have been the suboptimal option. Coord_subopt represents both agents coordinating on the suboptimal social norm, and Coord_opt represents coordinating on the Pareto optimum. These bar charts give the behavioural strategies for fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents in response to i) distinct domains in which behaviour 0 is favoured in domain A and 1 in B; $p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$ and ii) coordination over two distinct domains where the most common environmental state is not the most fit to coordinate on; $p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$. Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.	**30**
Figure 5 and Figure 6. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) agents with modular motivation only; (iii) agents with modular cognition only and (iv) fully domain-general agents when deciding to coordinate in two similar domains. Figure 5 gives the results for runs where both the priors and the fitness tied to suboptimal coordination favour coordination on behaviour 0 ($p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), while Figure 6 gives the results for runs where the priors favour state 1 but the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 ($p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).	**36-51**
Figure 7 and Figure 8. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) agents with modular motivation only; (iii) agents with modular cognition only and (iv) fully domain-general agents when deciding to coordinate in two distinct domains. Figure 7 gives the results for runs where both the priors and the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 of domain A but behaviour 1 in state 1 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$). Figure 8 gives the results for runs where the priors favour state 0 in domain A and state 1 in domain B, but the fitness tied to suboptimal coordination favour coordination on behaviour 1 in state 1 of domain A and behaviour 0 in state 0 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).	**54-69**

bootstrapped confidence interval clustered on social learners, to reflect the multiple observations gathered per learner. The regions shaded in grey depict where the social learners' data would fall if they used a conformist strategy, while the dashed lines give points of reference for proportions of learners choosing % at 0, 0.5, and 1. The colour denotes the games being played, with game against nature (asocial skills) in red, coordination game (social norms) in blue and Prisoner's Dilemma (cooperation) in green.	

Chapter 1:

Introduction

1.1: Gene-culture coevolution

The dual-inheritence theory posits that genetic evolution and cultural evolution both drive human behaviour (Richerson & Boyd, 1978; 2008; 2020; Richerson et al., 2010). Just as genetic predispositions can shape our culture, so too can our cultural systems create selection pressures that act on our genes in a process called gene-culture coevolution (Feldman & Laland, 1996; Gintis, 2011). In a similar way to how natural selection acts to increase genetic information which codes for certain advantages, cultural selection acts to increase social information which upholds advantageous behaviour (Henrich, 2015; Mesoudi, 2017; Mesoudi et al., 2004). Here, social information refers to any information which can be learned from and shared with others (Mesoudi et al., 2006a).

This definition of cultural selection implies that any animal species capable of sharing social information will experience 'culture' (Riebel et al., 2015; Whitehead et al., 2019). However, humans are still unique in the dual-inheritience framework as only we have the capacity for 'cumulative culture' (Boyd et al., 2011; Henrich & McElreath, 2007; Legare, 2019; McGuigan et al., 2017; Muthukrishna et al., 2018). This means that wisdom is acquired and transmitted across many generations and is modified and extended in relevant ways to keep up with the changing demands of the environment over time (Chudek & Henrich, 2011; Mesoudi, 2011a; Miu & Morgan, 2020). Cumulative culture is believed to be underlaid by a 'ratchet effect' (Tennie et al., 2009). Once an invention has been made, it is effectively 'locked into place' by sharing it with others. Due to amendments over time, the item in question may grow more complex, though it will very rarely experience a downwards trend in complexity (Mesoudi & Thornton, 2018). In this way, social information becomes increasingly complex as it is shared (Boyd et al., 2011; Legare, 2017).

To address why only humans have such extended cultural repertoires– including globalised travel, the smartphone and the Internet– some research builds agent-based models. Models can investigate the trajectory of certain behaviours in a theoretically evolving population. Models of cultural evolution often generate a positive feedback loop. As soon as we became capable of sharing complex social information, then larger brains would have been selected for to process and store said social information. As soon as larger brains were selected for, then more cultural information could be produced. Put another way, as soon as cumulative culture emerges then a runaway selection pressure is likely to select for more advanced cultures (Rendell et al., 2011; Markov & Markov, 2020). Interestingly, there is likely to be an 'upper limit' of brain size that can evolve this way (Markov & Markov, 2020). Once this is reached, then our cultural systems of storing information become more important (e.g., museums and libraries as public record-keeping) (Lotem et al., 2017; Mesoudi & Thornton, 2018).

Evolutionary theories need to address why human culture is so advanced. Traditional evolutionary theories often see culture as being 'evoked' by genes (Gangestad et al., 2006; Nettle, 2009; Tooby & Cosmides, 1992). Alternatively, specific behaviours may arise due to an interaction between genes and one's local environment, though Evolutionary Psychologists reject culture as a causal mechanism (Confer et al., 2010). Dual-inheritence theory is unique as it postulates a degree of gene-culture coevolution; that is, it places cultural evolution on equal footing to genetic evolution. Genetics may influence the kind of environments that we seek out and how we shape our environments. Likewise, changing the environment and engaging in culture can create new niches, or roles, to fulfil which in turn acts upon genetic selection in a cyclical process (Laland et al., 2010).

To illustrate how culture may affect genetic selection with an example, consider the lactase enzyme. Adults from Western European, West African, and Middle Eastern descent have genes coding for enzymes which can digest lactose in dairy (Gerbault et al., 2011; Itan et al., 2009). Ancestral humans from these three regions began to farm cattle for dairy approximately 10,000 years ago. Dairy provided a rich source of nutrients for those capable of digesting it. The cultural act of farming dairy cattle thus created a selection pressure for individuals to be able to digest lactose, which in turn affected their genetic evolution. Interestingly, mutations occur at three distinct regions of promoter genes at these sites (Western Europe, West Africa, and the Middle East), which differentially impact whether the genes for lactose tolerance stay switched on (Tishkoff et al., 2007). Those from other cultures – who did not engage in dairy farming until more recently – do not have the same genes to 'switch on' lactose tolerance. This is an example of gene-culture coevolution as it highlights how a cultural process (farming dairy) resulted in a selection pressure for a unique gene (in this case, for lactose tolerance).

Interestingly, there is evidence to suggest that cattle breeds from North and Western Europe have genes to produce richer milk (Beja-Pereira et al., 2003). The way that we bred cattle for our farming needs thus produced a coevolutionary feedback mechanism between human and cattle evolution at these sites. Artificial breeding is an example of cultural evolution as it changes a species via human involvement, though it often affects the evolution of a breed and human society in a myriad of unexpected ways (Mesoudi et al., 2004).

Lactose tolerance and artificial breeding are just two examples of how culture may influence genetic evolution. Other examples include the prevalence of enzymes to digest carbohydrates as a result of agriculture (Hale, 2017; Hancock et al., 2010), the prevalence of enzymes to digest alcohol after its discovery (Thomasson et al., 1991) and our preference to obey certain social norms, or rules (Gintis, 2003; Gintis, 2004).

To understand cultural evolution in more detail, gene-culture coevolutionary researchers use the metaphor of genetic evolution. This example is of course illustrative and there are likely to be key differences between genetic and cultural evolution (Billiard & Alvergne, 2018). I have already highlighted natural selection as a force which allows for beneficial genes to replicate, and cultural selection as a force which allows for beneficial social information to replicate. Cultural selection – just like natural selection – is Darwinian (Mesoudi et al., 2004). Just as the principles of competition, variation, and inheritance drive natural selection, so too do they drive cultural selection (Laland & Brown, 2011; Mesoudi, 2017; Mesoudi et al., 2004).

In regard to Darwinian natural selection, competition describes the fact that all organisms produce more offspring than can realistically be expected to survive in the environment. Variation describes the differences in phenotypes between all individual organisms. Sometimes, this variation may help the organism to outcompete others for access to resources or mates. Finally, the concept of inheritance means that beneficial variants are more likely to be passed from parent to offspring than less beneficial variants. Offspring with said beneficial variants would in turn be more likely to survive long enough to have their own offspring and so the process continues.

To illustrate this process with a classic example, there is competition for access to vegetation amongst the herbivores of the African savannah. There would have been variation amongst the ancestors of modern-day giraffes. Some giraffes had longer necks than others. These longer necks allowed the giraffes to eat leaves off of trees, a food source with less competition than ground level shrubs. As long-necked giraffes had a less competitive food source, then they were more likely to survive long enough to reproduce than short-necked giraffes. Due to inheritance, the long-necked giraffes typically had offspring who also had longer necks. This process occurred over thousands of years to produce the typical long-necked giraffes that are recognisable today. Put another way, natural selection favoured giraffes with longer necks as this allowed them to access a less competitive food source. Thus, the giraffe's long necks can be considered an evolved adaptation to their savannah surroundings.

To illustrate how these processes act upon social information, consider social media as an example. There is clearly more information being shared on social media sites than any one person could attend to. This is the concept of competition. Of course, the information being shared also displays variation. Some content may be emotional, humorous, or serious. Due to these qualities, some social information may be more attention-grabbing than others (Brady et al., 2020). Then, the process of inheritance occurs as some social information is more likely to be passed on (re-shared) than others. For example, emotional content is more likely to be shared than neutral content (Brady et al., 2020; Stieglitz & Dang-Xuan, 2013). When it comes to inheritance, individuals are not limited to just sharing the opinions of our parents. Indeed, social information can be inherited from elders besides one's parents (oblique transmission; Laland, 1993; McElreath & Strimling, 2008; Mullon et al., 2020; Muthukrishna et al., 2018) and from one's same-age peers (horizontal transmission; Gintis et al., 2001; Markov & Markov, 2020; Molleman et al., 2019a).

As there is such an excess of social information, then individuals need a way of filtering this. Broadly speaking, this can be achieved by content-based selection or context-based selection (Acerbi & Tehrani, 2018). Content-based selection occurs when there is a trait that is inherent in the social information itself that makes it more likely to be transmitted by individuals and attended to by further individuals. For example, we have a preference to attend to content that is emotional (Eriksson & Coultas, 2014; Stubbersfield et al., 2017), relevant to our survival (Boyd, 2009 [part 3]) and easy to digest (Bartlett, 1920).

Context-based selection instead occurs when we have a preference for *who* shared the social information, rather than the quality of the information itself. For example, it may be best to learn a specific skill from someone who is already qualified in that area (Henrich et al., 2015; Henrich & Gil-White, 2001). If I wish to become a better hunter, then it makes sense to copy the actions of the individual who makes the most kills. Other times, we copy something because the individual who shared it is prestigious. Prestigious individuals tend to be very successful in one or more cultural domains, and so are afforded a position of social influence (Henrich et al., 2015; Price & van Vugt, 2014). Perhaps a recent example of this would be the England football team. They are successful due to their prowess at playing a sport, but also use their prestigious position off-pitch to campaign for certain social issues, such as an end to racism and online abuse.

Beyond *who* shares social information, context-based selection can also account for *how many* people share social information (Richerson & Henrich, 2012). Frequency-dependent social learning strategies involve basing our behaviour on the number of others in the group who also display this behaviour (Efferson et al., 2016; McElreath et al., 2008; Mesoudi & Lycett, 2009). For example, we may copy others in a linear based fashion and become more likely to employ a behaviour as more individuals adopt this (Morgan & Laland, 2012).

Perhaps we instead copy the behaviour that only a minority of others endorse (Efferson et al., 2008a). This may occur when the minority is more confident than the majority. For example, Evans et al. (2018) conducted a study where adult demonstrators would enact irrelevant behaviours when teaching children how to open a puzzle box (a box which can be opened with a complex series of levers, sliders, and buttons to yield rewards). Importantly, the children would only copy these irrelevant actions if all adults demonstrated them. When a confident minority could open the puzzle box without resorting to irrelevant behaviour, then the children would instead copy the minority. Moreover, certain individuals can exert a strong minority influence. Those who are more extraverted and hold positions of social influence can promote the spread of an initially unpopular opinion (Muthukrishna & Schaller, 2020).

More broadly, there are times when it would be useful to distinguish between a minority and a majority. To illustrate with an example, picture a cooperative group who is known to share its abundant wealth equally among its members. Now picture that this group are identifiable as all its members wear red clothing. It may be relatively easy for outsiders to acquire a piece of red clothing and infiltrate this group. These outsiders have no intention of sharing. Instead, they wear this clothing to fake a signal of their generous intentions while intending to take resources from others. The majority of the group are generous, but the minority– the infiltrators– are selfish. It would make sense for a learner to be able to distinguish between the cooperative majority, and this minority of selfish infiltrators. For example, perhaps the infiltrators are identifiable as their clothing is an off-shade of red. Once the learner can distinguish between the minority and the majority, she must then decide whether cooperating like the majority, or selfishly extracting resources like the minority, is more appealing (Burton-Chellew et al., 2017). I will return to this example throughout the empirical studies of my thesis (see Chapters 3 and 4).

To recap, majority-based and minority-based social learning strategies are important types of frequency-dependent social learning strategies. Perhaps the most important type of frequency-dependent social learning strategy in the gene-culture coevolutionary literature is conformity. Conformity is defined as the *disproportionate* tendency to copy the majority (Efferson et al., 2016; Morgan & Laland, 2012; Morgan et al., 2019). To illustrate, imagine a social group where 75% of the group members wear their hair in a braid. A conformist learner will wear her hair in a braid with a probability greater than 0.75.

Previous research into the flexibility of conformity may have been hampered by the preciseness of its definition. In this thesis, I use conformity to mean a *disproportionate* tendency to copy the majority *of the group* (as per Boyd & Richerson, 1985 [chapter 7]; Efferson et al., 2016). The definition of conformity in previous work has also included copying the most frequent instance of a behaviour rather than the majority of a group per se (Morgan et al., 2019; Uchiyama et al., 2021). I focus on the majority of the group rather than the majority of instances, as the studies in this thesis focus on how individuals use conformity and other social learning strategies to learn from *groups* who present different social information.

While all social learning strategies influence cultural evolution, conformity has received the most attention in the literature (Boyd & Richerson, 1985 [chapter 7], Efferson et al., 2016; Lachlan et al., 2018). Conformity is thought to be uniquely beneficial as it can uphold social information even when the individual is uncertain about what she is learning (Mesoudi, 2018). Conformity is also ‘blind’. That is to say, conformity is about *how many people* already display a behaviour, rather than focusing on the payoffs of the behaviour (Molleman et al., 2013a). Thus, conformity may have a unique role in upholding cooperation, where individuals accept a personal cost to benefit another (Henrich & Boyd, 2001); and when upholding maladaptive behaviour (Boyd & Richerson, 2007), such as foot-binding (Gavrilets, 2020). This thesis seeks to investigate how conformity may uphold certain behaviours in conditions of uncertainty, perhaps including costly behaviours, in more detail. First, I discuss conformity as a social learning strategy in more depth in Section 1.2.

1.2: Conformity and cultural evolution

Conformity is important in driving cultural evolution. First, conformity can help to explain the cultural diversities between groups in a way that genetic evolution alone cannot (Bell et al., 2009). As there was likely to be a high rate of migration (Henrich & Boyd, 1998) and intergroup contact (Boyd & Richerson, 1985 [chapter 7]; Efferson et al., 2008b) between social groups throughout the ancestral past, then any genetic differences between social groups would have been mitigated by migrating to a new group before having children (Henrich & Muthukrishna, 2021). Instead, conforming to the actions of one’s local group could allow in-group members to become more similar. While the in-group members become more homogeneous via conformity, the differences between diverse social groups increase over time (Henrich & Boyd, 1998).

Second, conformity allows individuals to copy the majority behaviour with a high fidelity (McElreath et al., 2018; Muthukrishna et al., 2017). This high fidelity would allow even recent migrants and young children to uphold the seemingly complex and arbitrary social rules that guide behaviour within a group, even when they are uncertain about the rules that they should uphold (Legare, 2019; Molleman et al., 2013a). When uncertain, copying what most others do with a high fidelity would enable one to fit in (Kendal et al., 2018; Wood et al., 2013). Conformity could therefore stabilise a wide range of social behaviours (Molleman et al., 2013b).

Conformity could even uphold arbitrary or maladaptive behaviours that are costly for individuals to maintain (Boyd & Richerson, 2007). This is because cultural evolution can be blind when conformist strategies are bluntly applied (Mesoudi, 2008b; pg 251). That is, maladaptive norms will proliferate via conformity whenever – for whatever reason – the majority of the group have converged on a behaviour that is costly (Avarguès-Weber et al., 2018; Mesoudi, 2008a).

To illustrate how conformity may uphold a maladaptive behaviour, consider cases where a sudden temporal or spatial shift means that the behaviour displayed by the majority of the group suddenly becomes suboptimal. In these cases, conformity could uphold an outdated strategy (Deffner et al., 2020; Feldman et al., 1996; Wakano & Akoi, 2006). For example, imagine a group who signal their group membership via dance. If a sudden shift in the environment results in a loss of a major food source, then the group may find themselves living on a calorific knife-edge. In this case, dancing would be costly as it wastes calories. However, a group which consists of pure conformists would continue to dance even though this is costly, as they uphold the behaviour that was already the majority.

Conformity may be uniquely important in upholding costly behaviour as it removes a focus on the payoffs to one's behaviour (Burton-Chellew et al., 2015; Molleman et al., 2013a). Other harmful behaviour which may be upheld via conformity include Female Genital Cutting (FGC; Efferson et al., 2020a), witchcraft beliefs (Tanaka et al., 2009) and mob behaviour in crowds (Raafat et al., 2009). These behaviours would be difficult to explain if individuals acted to maximise their own (genetic) fitness. However, they can be accounted for by conformity to a maladaptive social norm (Kendal et al., 2018). Moreover, individuals may knowingly conform to a norm that they know is wrong. In Asch's (1955) infamous line experiment, participants gave clearly incorrect visual judgements if the majority of confederates had given their incorrect answers first.

As the above paragraphs highlight, conformity can affect group behaviour in a myriad of both positive and negative ways. If researchers are to understand how conformity works as a mechanism, it is important that we understand how people decide to conform (Kendal et al., 2018). By understanding how a social learning strategy works, we can speculate on the likely function of the strategy and its likely evolutionary origins (Kendal et al., 2018). This is important, as conformity has a pronounced effect on the behaviours that are upheld at a group level.

Previous research that has investigated the flexibility of conformity tends to take one of three approaches. First, some researchers investigate when individuals will conform versus when they will rely on their own trial-and-error (Boyd & Richerson, 1985 [chapter 7]; Mesoudi, 2011b; Mesoudi et al., 2015; Miu & Morgan, 2020; Reader, 2003). Second, some researchers investigate when individuals will use conformity as opposed to a different social learning strategy, such as copying the individuals with the highest payoff on an experimental task (McElreath et al., 2008; Molleman & Gächter, 2018). Third, some researchers focus on the individual factors that affect one's conformity preferences (Efferson et al., 2008a; Muthukrishna et al., 2016).

This thesis takes a different route. Conditional on frequency-dependent social information, I ask just how flexible an individual's decision to conform will be in light of multiple social signals. Specifically, I investigate whether the participants' decision to conform can remain flexible to (i) the frequency of choices made by a group; (ii) a signal indicating whether these group members make decisions in a similar or different environment to the participant and (iii) the reliability of this similarity signal.

It is important to investigate the flexibility of conformity, as there is a divide in the literature over how exactly this social learning strategy should be classified (Heyes, 2016; Lamba, 2014). Some researchers consider conformity and other social learning strategies to operate like heuristics, or rules-of-thumb (Gintis, 2003; Henrich & Boyd, 1998; Molleman et al., 2014). A rule-of-thumb would be beneficial in helping us to process a wealth of social information but is also likely to lead to trade-offs (Henrich & Boyd, 2001; McKay & Efferson, 2010). For example, 'copying the majority' would often be a successful rule-of-thumb as the majority are likely to agree on the way of doing something for a reason (Barrett, 2015 [chapter 9]). However, copying any majority can lead to upholding incorrect (Asch, 1955) or harmful behaviour (Boyd & Richerson, 2007). For example, copying the majority after a sudden environmental shift can uphold outdated behaviour (Deffner et al., 2020).

Conformity has traditionally been modelled in small social groups which are assumed to be like the groups that we lived in throughout the ancestral past. For example, in Boyd and Richerson's (1985, Chapter 7) seminal model, they only allowed the agent to conform to three other demonstrators at a time. While this was the simplest scenario for testing their hypotheses, it is worth pointing out that we are exposed to the opinions and behaviours of thousands of people in this increasingly globalised age (Price, 2008). Perhaps conformity is upholding maladaptive behaviour as it is being used to process social information from much larger groups now, with much more varied sources of social information to filter through. For example, suicide is thankfully rare. However, information cascades surrounding celebrity suicides in the media can lead to a rise in copycat suicides (Mesoudi, 2009). In a similar fashion, certain websites can promote terrorist ideology (Jin et al., 2013) or eating disorders (Coletto et al., 2017) by making these harmful behaviours seem common in an isolated network of individuals.

To recap, conformity as a rule of thumb to 'follow the majority' would sometimes be useful. For example, it allows the individual to uphold the same behaviour as the majority, who typically uphold optimal behaviours (Barrett, 2015 [chapter 9]). Conformity can also allow the learner to uphold relevant behaviour even when uncertain (Mesoudi, 2018). However, in rare cases when the majority have already settled on a suboptimal behaviour then conformity could uphold costly behaviours. This would be the case when there is a sudden change in the environment (Deffner et al., 2020) or during certain information cascades online (Jin et al., 2013).

To demonstrate how rule-of-thumb thinking may apply to social learning strategies in more depth, consider the spillover effect. The spillover effect occurs when individuals start copying successful others in domains beyond that which the other person was originally successful in. In Brand et al’s. (2020) study, participants were recruited to play an online general knowledge quiz with distinct sub-categories (e.g., geography, art, sports, history etc.) The participants could see how well the other quiz players were doing. Participants copied whoever had the highest performance in the first sub-category of the quiz throughout, even though those who performed well in the first category did not necessarily perform well in the following categories. To illustrate, it would perhaps be unreasonable to expect an individual who is skilled in the arts to know just as much about sports and yet the participants acted like this was the case.

Marks et al. (2019) show a spillover effect of a different nature. In their study, participants had to identify difficult geometric patterns. They would copy other participants who had a similar political ideation to themselves, regardless of the other’s ability at the geometric task. This displays rule-of-thumb thinking. Typically, those who endorse similar opinions to oneself should be copied but not in cases when there is no relation between the original task and the test task (i.e., politics and identifying geometric patterns bear no relation to each other). Despite this, the individuals still rigidly applied their rule-of-thumb strategy to a domain for which it was never intended.

At the other end of the spectrum, some researchers argue that social learning strategies remain flexible to the myriad of situations that we may find ourselves in (Burdett et al., 2018; Efferson et al., 2016; Miu et al., 2020). For example, some individuals may only conform to the majority when they are uncertain of the task at hand (Kendal et al., 2018; Toelch et al., 2014).

The view that social learning strategies operate like rules-of-thumb, and the view that social learning strategies are used flexibly, sound like contrasting ideas. However, the two opposing views can be combined if one considers social learning as an individual preference. Some individuals will follow a simple rule to ‘always copy the majority’, while other individuals consider when exactly they should copy the majority over other social learning strategies across various situations (Efferson et al., 2008a; Muthukrishna et al., 2016).

The same individual could even switch between following a rule-of-thumb and showing a flexible use of social learning strategies across different tasks. Heyes (2016) applies the dual-systems approach to social learner cognition as a way of highlighting how individuals can engage in both heuristic-based and flexible social learning. In dual-system approaches to human cognition, System 1 describes a ‘fast’ thought process and System 2 a ‘slow’ thought process (Evans & Stanovich, 2013; Kahneman, 2011). System 1 thinking is automatic and often occurs without much conscious thought. This could uphold social learning strategies as heuristics (Heyes, 2016). For example, when the individual always copies the majority, this is likely to reflect a blunt application of a conformity heuristic. System 2 cognition is instead slower and more deliberated. This can help the learner to achieve complex decision-making in regard to when it is best to conform or to use another social learning strategy (Heyes, 2018). While both heuristic and flexible social learning strategies are possible, it is still important to clarify whether the decision to conform is closer to one end of the spectrum or the other. Understanding the flexibility of conformity will help us to understand how individuals react to a wealth of social information (Efferson et al., 2016).

The flexibility of one's decision to conform is also likely to be influenced by one's cultural orientation. Hofstede (1980) classified countries as being individualistic or collectivist. Individualistic countries are those that focus on the individual's autonomy, where independence of thought is valued. For example, the US and the UK. Collectivist countries instead focus on the group or family unit, and value harmony and compliance. For example, China and Japan. One's individualism or collectivism orientation can impact one's social learning preferences. Collectivist individuals may conform more (Bond & Smith, 1996). This would allow collectivist individuals to uphold the social norms of their group with a higher fidelity and thus ensure group harmony.

Moreover, there can be shifts in one's social learning styles over time. This is because we are the only species who learn socially how to socially learn. In Mesoudi et al's. (2015) task, participants could design their own virtual arrowheads with access to social information. British participants (whose culture is individualistic) preferred to rely on their own trial-and-error. Chinese participants (whose culture is collectivist) preferred to use social information. Chinese participants who had been living in the UK for a while preferred trial-and-error learning, similar to individualistic participants. This suggests that these Chinese participants shifted their social learning strategies to match the individualistic country in which they now lived.

The relationship between cultural orientation and social learning is not always a clear one, however. Molleman & Gächter (2018) recruited British and Chinese participants to play a series of economic games through both trial-and-error and social learning. Chinese participants used social information more than British participants did but preferred a payoff-based social learning strategy. Payoff based social learning strategies involve disproportionately copying the highest earner of the group. British participants may have used social information less, but when they did use social information, they were more likely to conform to the majority. This shows a nuanced relationship between cultural orientation and social learning preferences. Further research investigating social learning across more diverse samples than the WEIRD participants typically collected may illuminate these findings (Western, Educated, Industrialised, Rich and Democratic; Henrich, 2020; Henrich et al., 2010; Muthukrishna et al., 2020).

Another factor that will influence the flexibility with which one decides to conform is the behaviour that one learns. We make thousands of decisions across our life span. These decisions range from the small (such as deciding which toothpaste to buy) through to the large (such as deciding whether to move to a new country for a job opportunity). Classifying the complexity of all human behaviour is beyond the scope of this thesis (and perhaps any thesis!). Throughout this thesis, I therefore focus on the flexibility of human decision-making in three key areas:

1. Asocial skills, where the outcomes of one's behaviour affect one's own fitness only (Legare & Nielsen, 2015). For example, learning how to use a tool. It does not matter too much how one uses the tool, as long as the end result is beneficial. For example, chimpanzees (Whiten, 2019) and corvids (Rutz et al., 2018) use a range of tools to extract food from their environment.
2. Social norms, where the success of one's behaviour depends on how well one coordinates with others (Legare, 2017). The decision of when it is acceptable to wear formal wear is a good example of a social norm.

3. Cooperation, or the mutual exchange of resources over time (Henrich et al., 2001).

When it comes to learning asocial skills, one's own trial-and-error is of course important (Mesoudi, 2008a; Mesoudi, 2011b). In fact, a combination of individual and social learning strategies is perhaps best placed to account for skill-learning (Aoki et al., 2012; McElreath, 2004; Romero-Mujalli et al., 2017). To illustrate why social learning is also important, consider the example of a hunter who wishes to improve her ability. It would make sense for her to copy the hunter who makes the most kills (Henrich et al., 2015).

Let's walk through an in-depth example of when conformity would be useful in upholding an asocial skill. Consider an individual who has recently joined a new social group. This group lives in quite a different environment to her previous group. There is a wealth of new and strange looking plants in this location, and she does not know which are safe to eat (Henrich, 2015). This is an example of a domain where trial-and-error is inappropriate. Eating one poisonous plant is enough to potentially kill. She could identify the most successful or prestigious gatherer and copy what this gatherer does. However, this approach requires time and energy. First of all, she must calculate which gatherer she feels is the most successful or prestigious. Once identified, she will then have to do something to gain access to this individual, like gain her favour. In fact, it may be easier for this individual to observe the plants that all the other gatherers pick and then pick only those plants which the majority favour. After all, the majority of individuals would not knowingly collect plants which could make themselves or others sick.

Social norms are rules that govern social behaviour and allow us to coordinate on a collective action (Bowles & Gintis, 2004; Wen et al., 2019). These norms are likely to encompass group rituals, which are important to signal our belonging to a group

(Wen et al., 2020). As these rituals are important for group bonding, then they may be copied with extra fidelity (Henrich & Muthukrishna, 2021). This means that conformity would be expected to increase when learning a social norm.

Legare and colleagues show that this is the case using an ingenious ‘necklace designing’ paradigm. When children were informed that making a necklace is a social norm, then they copied a demonstrator’s actions with high fidelity. When the children were instead told that necklace building is a skill ‘up to the individual’, then they showed more creativity when building necklaces (Clegg & Legare, 2016; Legare et al., 2015). Again, this is impacted by culture. Ni-Vanuatu children (collectivist) conformed in both the social norm and skill condition, though were more conforming in the former. American children (individualistic) conformed to a social norm but were more likely to follow their own initiative and avoid conformity in the skill condition (Wen et al., 2020). It is worth acknowledging that these studies typically use just one adult demonstrator. In real social groups, we often have to learn social norms from a multitude of others.

Finally, it is important to consider the role of conformity in cooperation. Cooperation describes the mutual exchange of resources over time and may be difficult to explain from the individual’s perspective. If an agent should act to maximise her own resources, then why would a behaviour ever be selected for that involves donating resources to others? (Colman, 2006). Of course, it is important to acknowledge that cooperation can be beneficial at the level of the social group (Ihara, 2011). More cooperative groups can bring down better game (Price, 2006); build better shelter (Chudek & Henrich, 2011) and can band together to out-compete others in times of warfare (Ihara, 2011).

It is prudent to note that conformity cannot kickstart cooperation in an otherwise uncooperative group. This is because conformity can only increase the prevalence of the behaviour that is already the majority. If the majority of the group are uncooperative, then conformity will only ever uphold uncooperative behaviour (Molleman et al., 2013b). There are other ways that cooperation can initially become prevalent, including being kickstarted by a prestigious leader (Henrich et al., 2015).

Once cooperation has already emerged in the group, then it can be maintained with conformity (Henrich & Boyd, 2001; Szolnoki & Perc, 2015). Conformity is particularly important as it removes the focus on the payoffs of one's behaviour (Burton-Chellew et al., 2015; Molleman et al., 2013a). Defection or free-riding (refusing to cooperate with others) is typically the most beneficial strategy at the individual level (Holt & Roth, 2004). This is because free-riders accept the benefits of cooperation without ever engaging in a costly donation of resources themselves (Burton-Chellew et al., 2017). This could explain why individuals using a payoff-based social learning strategy are more likely to freeride than those that conform (Burton-Chellew et al., 2017; Molleman et al., 2013a).

Conformity is apparent in most participants' cooperative strategies. Approximately 60% of participants are 'conditional cooperators', meaning that they will only donate resources provided that the majority of others in the group do so first (Li et al., 2021; Molleman et al., 2019b; Price & Johnson, 2011). It is important to note that not everybody in the group has to be a conditional cooperator, and the urge to conform does not necessarily have to be strong. This is because most social groups punish free-riding. This is an action designed to impose a cost onto the free-rider so that her actions do not pay off. If all the individuals in the group collectively punished the free-rider, then the individual may feel unfairly targeted and seek retaliation (Molleman et al., 2019b; van den Berg et al., 2012). To avoid over punishing defectors, it may only be necessary for individuals to have a weak conformity preference to uphold cooperation and norms regarding punishment (Henrich & Boyd, 2001).

In order to reduce the temptation to freeride, perhaps our cooperative preferences 'hitchhiked' onto certain genes (Gintis, 2003; Tomasello & Gonzalez-Corbua, 2017). These adaptations may form a 'norm psychology' effect which makes humans uniquely cooperative and coordinative (Gintis, 2003; 2004). Norm psychology would create a positive feedback loop. When more social norms emerge, then there would be a selection pressure for psychology which can uphold coordination. Once these psychological preferences have emerged, then of course societies will create more norms to follow which further reinforces the need for norm psychology (Markov & Markov, 2020).

Essentially, this means that the desire to follow norms is so strongly internalised that one feels rewarded for following rules, and guilty when breaking rules, even when others are not watching (Gintis, 2003; 2004). Of course, the desire to 'cheat' on certain social rules may also be reduced by Big God religions which suggest that we are always being watched and may incur supernatural punishments for any misdeeds (Gray & Watts, 2017; Lenfesty & Morgan, 2019). This may explain why individuals engage in such extreme levels of cooperation (Gintis et al., 2008). Our desire to follow rules – and our guilt when we do not – could have made humans an extremely cooperative species.

This thesis will further investigate the flexibility of an individual's decision to conform across the three key areas of decision-making. I design some economic games to mirror the learning of asocial skills, social norms, and cooperation in-lab (see Methods chapter for more details). Moreover, I recruit participants for our first study in

India to address concerns over testing WEIRD samples. Specifically, I test whether an individual's decision to conform remains flexible to up to three pieces of social information simultaneously. Note that previous research tends to only test one or two social signals simultaneously (Efferson et al., 2016) and so this is a novel test of social learner flexibility when conforming. Understanding whether people can learn differently from multiple sources of social information would likely have impact for the trajectory of cultural evolution (see Chapters 3 and 4).

Whilst these studies highlight the flexibility of social learner *behaviour,* they do not necessarily highlight how flexible social learner *cognition* is. Section 1.3 turns to the flexibility of human cognition across the three key areas of decision-making that I focus on throughout this thesis.

1.3 The flexibility of cognition in decision-making

The research presented in sections 1.1 and 1.2 displays that human behaviour can be flexible due to the influence of social learning and cultural evolution. It is equally important to understand the flexibility of human decision-making at a cognitive level. Flexible behaviour does not necessarily arise from a flexible cognitive system. Indeed, seemingly complex behaviour can arise from simplistic cognitive systems which are not particularly specialised (Amodio et al., 2019; Mikhalevich et al, 2017). This is known as the 'inverse problem' (Deffner et al., 2020; Tarantola, 2006). That is, a range of behaviour may be consistent with a wide range of underlying psychologies. Clever Hans demonstrates this critique nicely. Clever Hans was a horse which could do sums by tapping out the answers with his hoof. This behaviour could have been systematic of a complex cognitive system capable of doing abstract math. However, it was more likely that the horse understood when to stop tapping the ground based on the reaction of his audience. The horse's ability to answer sums was not an indication of mathematical ability but instead could have arisen via simple positive reinforcement (Lapuschkin et al., 2019; Pfungst, 1911).

There has also been a tendency to 'blackbox' cognition in the social learning literature (Heyes, 2016; Kendal et al., 2018). Previous research tends to focus on the complexity of human behaviour in social learning studies (Efferson et al., 2008a; McGuigan et al., 2017; Molleman & Gächter, 2018) or anthropological data (Henrich, 2015; Mackie et al., 2015; Smith, 2014), while not applying as precise a focus on the cognitive systems that may support such behaviour.

Any study of human cognition thus faces the 'inverse problem' (in that we cannot infer flexible cognition from flexible behaviour alone) and 'blackboxing' (in that the role of cognition may be under researched due to issues in making tangible inferences about this process). One way around both these problems is agent-based modelling. Agent-based modelling forces the researcher to make explicit assumptions about human behaviour and its underlying psychology (Bolhuis et al., 2011; Laland, 1993; Laland et al., 2007; 2009; Mesoudi et al., 2004). Once an exploratory framework is agreed upon, the finer details of the model can be built. Eventually, the modeller will leave one or two psychological components to evolve endogenously in a group of agents. These agents are then left to make decisions across thousands of generations. Computer simulations thus allow us to see the influence of evolution on the behaviour of a theoretical population of agents (Muthukrishna & Henrich, 2019). This thesis uses agent-based modelling to investigate the complexity of human psychology when making decisions across three key areas of interest: asocial skills (Chapter 5), social norms (Chapter 6) and cooperation (Chapter 7).

Evolutionary researchers have a challenging time coming up with a cognitive theory which could plausibly account for the myriad of behaviours that humans display. What cognitive system could explain both our extreme levels of cooperation and charity, coupled with our unique propensity for wars? What cognitive system could explain how we are capable of upholding complex skills such as reading, writing and numeracy (Tamariz & Kirby, 2016); and yet can explain why we partake in maladaptive behaviour? For example, individuals smoke and over-eat processed, fatty foods despite knowing that it is bad for their health. We spend lots of money on lottery tickets, despite knowing that our chances of winning are fantastically low.

The first thing to acknowledge is that we make a wealth of decisions and most of these are relatively novel in terms of our ancestral past (Barrett & Kurzban, 2006; Whitehead et al., 2019). We cannot plausibly have a genetic bias telling us who to vote for in the next election, or which television show to watch next, as these decisions occurred so recently in human history. Moreover, we often have to make decisions to deal with unexpected scenarios that we have not planned for. For example, figuring out what to do next when unexpectedly involved in a minor accident. As we have to make a range of novel and complex decisions, then Bolhuis et al. (2011) argue that human cognition is likely to be domain-general. A domain-general system is a central processor that must remain flexible to a wide range of environmental inputs in order to achieve a complex range of behavioural outputs (Kan et al., 2013; Vergauwe et al., 2010). Domain-general systems have the capacity to respond to novel stimuli, and to unexpected events (Bolhuis et al., 2011).

The view that human cognition is domain-general directly contrasts with the dominant view in Evolutionary Psychology (EP). Evolutionary Psychology is an alternative viewpoint to dual-inheritence theory, which also applies evolutionary logic to the study of the human mind and behaviour (Tooby & Cosmides, 2005). Note the capitalised EP is used to denote this field as a paradigm of research in itself, which makes some unique assumptions about human behaviour (Buss, 1995).

First of all, Evolutionary Psychologists focus on cognitive psychology (Cosmides & Tooby, 1994a; Tooby & Cosmides, 2005). The importance of cognitive processing is up-weighed and thought to drive other aspects of psychology, such as motivation, across relevant domains (Tooby et al., 2006). Evolutionary Psychologists see the mind as being 'massively modular' (Cosmides & Tooby, 1994(b); Sperber, 2001). This means that the human mind is comprised of a series of modules, or pockets of cognitive processing that are designed to have functional specialization (Pietraszewski & Wertz, 2021). Put simply, these cognitive processors are designed to work on one specific input in order to achieve a certain output. Each module is designed to deal with a recurrent issue that humans would have faced throughout the ancestral past. Putative examples of modules include: colour processing (a module which works on the input of foveal cell activation and produces the perception of colour; Schalk et al., 2017); face detection (a module which works on the input of face-like stimuli and produces recognition for faces; Rosenthal et al., 2017) and jealousy (a module which works on input to suggest that a mate could potentially cheat on us, and biases us towards behaviours designed to decrease the likelihood of that occurring; Buss, [1995] – though see Harris [2004] for a critique).

The debate over modularity has ranged from the philosophical (Fodor, 1983; Frankenhuis & Ploeger, 2007; Gould & Lewontin, 1979; Pietraszewski & Wertz, 2021) to the neuropsychological grounds (Spunt & Adolphs, 2017; Stokes & Bergeron, 2015). Fodor (1983) originally proposed a list of criteria that cognitive processing should display to be considered modular. Fodor then argued that only peripheral processing, such as vision, is likely to be modular while the rest of human cognition is overseen by a domain-general or central processor. Here, peripheral processes describe those modules that can work relatively automatically. We do not have to think about perceiving colour to do it. Central processes are those that govern conscious decision-making, and these are likely to be domain-general.

It may therefore seem surprising that Evolutionary Psychologists postulate that the brain is comprised entirely of modules. In fact, this is due to a relaxing of Fodor's criteria (Pietraszewski & Wertz, 2021). The only criterion that cognition need display to be considered modular in EP is functional specialization (or the ability to work on only a specific type of input). Therefore, any attempts to discredit modularity in a Fodorian sense (for example, by showing that brain areas are not informationally encapsulated, or separated in processing; Spunt & Adolphs, [2017]) are not relevant to the discussion of modularity in EP.

To highlight how modules may drive behaviour, consider the putative cheater-detection module. This module works on the input of a social contract *("If I do prerequisite X, then I may take reward Y")* and produces the output of identifying cheaters who violate these contracts (though see Sperber & Girotto [2002] for a rebuttal). In a series of experiments, Cosmides, Tooby and colleagues demonstrated that this cheater-detection module is uniquely activated by social contracts (Cosmides et al., 2010); is activated more strongly when there is an intention to cheat (Cosmides et al., 2010; Price, 2006), and boosted participant performance on otherwise difficult 'if-then' logic tasks (van Lier et al., 2013). This module may also direct memory, with participants more likely to recall the faces of cheaters than of criminals (Delton et al., 2012). Beyond cheater-detection, we may have modules underlying a range of skills that we used recurrently throughout the ancestral past (Buss, 1995; Sperber & Mercier, 2018).

Another key assumption of EP is that we evolved to the Environment of Evolutionary Adaptedness (EEA; Bennett, 2018). This was likely to have been a time period during the Pleistocene some 10,000 years ago as this was the last period of human history where the environment was thought to be stable enough to induce selection pressures that could code for genetic adaptations (Cosmides & Tooby, 1994a). As the environment has changed rapidly since then, we may display behaviour that is maladaptive today but would have been beneficial during this time in the ancestral past (Tooby & Cosmides, 1990). To put it simply, “our modern skulls house a Stone Age mind” (Cosmides & Tooby, 2007; p. 10).

This ‘mismatch’ between the EEA and the modern-day environment is how Evolutionary Psychologists explain maladaptive behaviour, as opposed to the focus on social norms and conformity in gene-culture coevolutionary research (Chudek et al., 2013). As an example of a ‘mismatch’, consider our preference for high-fat, high-sugar foods. This preference would have been adaptive for our hunter-gatherer ancestors who had to survive on a calorific knife-edge but is maladaptive today, as access to these foods has increased and we lead increasingly sedentary lifestyles (Power & Schulkin, 2013; van den Bos & de Ridder, 2006).

A ‘mismatch’ between the current environment and our ancestral past is how Evolutionary Psychologists account for the type of cooperation that is often observed in experiments (Price, 2008). Participants typically make substantial donations to anonymous strangers who they are unlikely to meet again in these experiments (Fehr et al., 2002; Fehr & Fischbacher, 2005; Sally, 1995; Vogt et al., 2015). This behaviour is called one-shot cooperation. As well as potentially being influenced by a social desirability effect (Burton-Chellew et al., 2016; Price, 2008), this costly level of cooperation may be symptomatic of the conditions of the EEA. We lived in small hunter-gatherer troupes where each individual was known to each other, and cooperative interactions could be tracked (Price, [2008]; Price & Johnson, [2011], though see Boyd & Richerson [2006] and Gintis et al. [2008] for a rebuttal). Moreover, individuals within the same tribe were likely to be slightly genetically related to each other (Curtin et al., [2020]; though see Bell et al. [2009] for an alternate view). These conditions would have created a strong selection pressure to cooperate. Thus, we cooperate now because the environment has changed too quickly for us to adapt to it. Our preference for cooperation is outdated in large societies where most interactions are one-shot or anonymous (Price, 2008).

Error Management Theory (EMT) is an evolutionary theory that seeks to explain one-shot cooperation. EMT predicts that we must lean on the side of caution when making decisions under uncertainty (Nesse, 2005). We must aim to make the least costly of two errors whenever we are uncertain of the payoffs of our behaviour (Haselton et al., 2015).

To illustrate with cooperation: whenever an individual met a new person throughout the ancestral past there would have been some uncertainty over whether she would meet this other person again. The agent may have made one of two mistakes here. She could falsely assume that she would meet this other individual again, and so offer a one-off donation of her resources. This is one-shot cooperation. Alternatively, she may falsely assume that she will never meet this other person again and free-ride. This behaviour would anger the other individual if she had intended on starting a long-lasting mutually beneficial relationship with the focal agent. Put simply, the agent's brash behaviour could have scuppered the chance for a long-lasting cooperative relationship. As the latter error is more costly than the former, then individuals evolve to lean towards one-shot cooperation as the least costly of two errors (Delton et al., 2011; Delton et al., 2013; Krasnow & Delton, 2016; though see Zimmerman & Efferson [2017] for a rebuttal).

Thus far, EP has applied modularity to explain our skills in performing certain tasks (Dhum et al., 2017; Sperber & Mercier, 2018), and in driving a bias to cooperate (Delton et al., 2011; Price, 2008). However, EP may face more of a challenge in accounting for the wealth of social norms that different cultural groups display. Indeed, there is likely to be more variation in the norms of different social groups than can be explained by genes alone (Bell et al., 2009; Henrich & McElreath, 2007; Richerson & Boyd, 2008).

To account for social norms and culture without evoking gene-culture coevolution, Sperber and colleagues have proposed that culture can emerge from higher-order modules. The human mind consists of numerous peripheral modules, which take in a discrete set of environmental inputs. The output of these lower-order modules then forms the inputs to second-order modules and so on. They argue that a hierarchal modular system in increasing order of complexity could give the appearance of complex and flexible behaviour (Sperber, 1994; Sperber & Hirschfeld, 1999; 2004).

Moreover, Sperber and colleagues account for complex and novel human behaviours by distinguishing between a module's *proper* domain and its *actual* domain. The proper domain describes the set of inputs that a module has actually evolved to process. The actual domain describes the inputs that are similar enough to the module's proper domain to 'trigger' the module. Thus, the module ends up working on a broader range of environmental inputs than it was originally designed for. For example, consider the human pitch module (Sperber, 1994). This was likely to have evolved to process early human speech (its proper domain). Before we had the propensity for languages, we were likely to communicate meaning via the pitch associated to utterances. Birdsong also uses pitch to convey meaning. Whilst the pitch module had evolved to process the pitch in early human speech (its proper domain), it can still process birdsong (its actual domain).

Complex language and writing systems may have reduced the need for pitch-based expressions in human language and yet the module persists. Of course, this could be a vestigial module that persists as it is more costly to remove once already emerged. This is similar to a vestigial organ like the appendix which is no longer necessary now that humans do not eat tough vegetation (Downes, 2015; Smith & Wright, 2018). However, one of the many reasons as to why the pitch module still persists may be due to the invention of music. Thus, music is said to be the *cultural* domain of the pitch module (Claidière et al., 2018; Sperber, 1994). The cultural domain describes any man-made or recent stimuli that may trigger module activation which would not have been present when the module originally emerged in the ancestral past. Music is likely to be advantageous in order to continue this pitch module. Music can soothe infants and is considered to have healing properties (Mehr et al., 2018; 2019).

Of course, a full debate on the plausibility of massive modularity in EP and the domain-general cognitive systems proposed in gene-culture coevolution is beyond the scope of the current thesis. The agent-based models presented in this thesis do not solve this debate. Instead, they should be viewed as an initial comparison of domain-general and modular decision-mking systems, which will hopefully pave the way for more detailed future comparisons. For this initial investigation, a domain-general agent must manage decision-making flexibly over multiple domains whilst modular agents have systems that can specialise per each decision-mking domain.

Note that domain-general decision-making is often thought to be driven by associative learning (Bolhuis et al., 2011; Macintosh, 1974; Reader et al., 2011). For simplicity, this thesis will not replicate these models but instead, the simplest comparison is to view a domain-general system as driving the decision to act- or not act- over various domins simultaneously whilst modular systems specilise. This should be seen as the first necessary step towards a full theoretical comparison of domain-general and modular decision-making systems (see Section 2.4 for further details).

These models are thus novel in their aim to directly compare the evolutionary trajectory of these two types of agents. Another novel aspect of these models is that I also consider partly modular agents. That is, I consider agents with a mixture of domain-general and modular processing systems rather than only committing to extreme ends of the spectrum.

These agent-based models are exploratory and as of such I have no formal hypotheses to test (Muthukrishna & Schaller, 2020). I compare the evolutionary trajectory of decision-making across the three key areas of interest when cognition is domain-general versus modular. These models will focus on the psychology that emerged to support decision-making in asocial skills (chapter 5), social norms (chapter 6) and cooperation (chapter 7).

1.4: Summary of the current thesis

This thesis seeks to further the field of gene-culture coevolution. This work can broadly be thought of as having two research aims:

1. To investigate the flexibility of the decision to conform. As highlighted in section 1.2, investigating the flexibility of one's decision to conform is essential to understand conformity as a mechanism and how this influences cultural evolution (Kendal et al.,

2018). I investigate the flexibility of conformity when learning asocial skills, social norms (Chapter 3) and cooperation (Chapter 4). These studies are novel as they investigate how the social learner responds to up to three pieces of social information simultaneously. This is the highest test of social learning flexibility that I am aware of, as previous studies focus on one or two pieces of social information (Efferson et al., 2016).

2. To investigate the flexibility of human cognition when making complex decisions. Specifically, I investigate whether cognition should be considered domain-general or modular when the agent learns asocial skills (Chapter 5), social norms (Chapter 6) and cooperative behaviour (Chapter 7). These models are novel as they are the first to compare the flexibility in agents' behaviour as based on whether they have domain-general or modular cognition.

To summarise, Chapters 3 and 4 focus on my first research aim regarding the flexibility of social learning behaviour whilst Chapters 5-7 focus on my second research aim regarding the flexibility of human cognition. To tie these social learning studies and agent-based models together, remember that I aim to capture the complexity of human decision-making across three key areas: asocial skills, social norms, and cooperation. Finally, Chapter 8 will summarise the findings across these papers and place them within the broader context of dual-inheritence and EP frameworks. I will also make some suggestions for future work and clarify some limitations of the current thesis. To begin, Chapter 2 focuses on how I chose to measure asocial skills, social norms, and cooperation with a series of economic games. I then outline how these games were applied to both the social learning studies and the agent-based models that I conducted. I also highlight some further details and evaluations of these techniques.

Chapter 2:
Methods

2.1 Methods Overview

This thesis takes a gene-culture coevolutionary approach to investigate human decision-making, placing it broadly within the dual-inheritence framework. The chapters are linked by investigating three key areas of decision-making: asocial skills, social norms, and cooperation. This work has two key aims. First, I investigate the flexibility with which individuals choose to conform or use other frequency-dependent social learning strategies when learning from groups with access to multiple sources of social information. I investigate this with two empirical laboratory-based studies. Second, I aim to understand the complexity of human psychology that is likely to underlie our decision-making. I investigate this with three agent-based models. Although these are distinct methods (Kendal et al., 2018; Laland, 1993), the thesis is linked by a series of three economic games that are employed throughout. I outline the three economic games of interest that will be used throughout this thesis in section 2.2.

To begin, I briefly outline why I choose to use a mixture of both social learning studies and agent-based modelling. Put simply, each technique can compensate for the weaknesses of the other (Fagiolo et al., 2007; Miu & Morgan, 2020; Morgan et al., 2019). While agent-based models are theoretical, empirical studies can allow causal inferences to be drawn from the data (Rendell et al., 2011; Deffner et al., 2020; Pearl, 2015).

Models can also address some limitations of empirical work. First and foremost, studying the evolutionary trajectory of human behaviour in an actual population is both unfeasible and unethical. Agent-based modelling provides a workaround, as computer packages are capable of running many simulations of a theoretically evolving populations of agents (Acerbi et al., 2020; Foramitti, 2021; Mason et al., 2021). Moreover, there is uncertainty over whether participants understand complex empirical studies (Burton-Chellew et al., 2016) or whether the participants merely display the behaviour that they believe the experimenter wants to see (Delton et al., 2011). Modelling removes such confounds. Moreover, researcher bias is minimised as modelling forces the researcher to be explicit about her hypotheses, and how she expects certain variables to interact (Laland, 1993). Models can generate complex data by considering how multiple variables would have likely coevolved throughout our ancestral past rather than focusing on a specific snapshot of human behaviour as is necessary in empirical work (Mesoudi et al., 2006b; Muthukrishna & Schaller, 2020).

Finally, the use of social learning studies and agent-based models together gives this work a broader interdisciplinary appeal to multiple areas of evolutionary sciences (Efferson et al., 2020b; Mesoudi et al., 2006b). Indeed, the use of both these methods is paramount to the gene-culture coevolutionary literature in which I place this work (Efferson et al., 2020b).

Specifically, the social learning studies utilised in chapters 3 and 4 aim to address my first research aim, regarding the flexibility with which one conforms– or uses another social learning strategy. These studies test whether participants can adjust their social learning strategies to incorporate up to three pieces of social information that are simultaneously presented about the social group from whom they learn. These studies are novel as they are the first to consider three pieces of social information when investigating how participants socially learn asocial skills, social norms, and cooperative behaviours. These studies will have a meaningful academic output, as understanding the flexibility of conformity and other social learning strategies could help expand our knowledge of their function (Kendal et al., 2018). Section 2.3 outlines the design of these empirical laboratory-based studies.

The agent-based models outlined in chapters 5-7 address my second research aim regarding the flexibility of agent psychology when one comes to master the three decision-making areas of interest. Each chapter uses a different economic game to inform its design. These models will have a meaningful academic output in addressing the ongoing debate regarding whether psychology is domain-general or modular (Fodor, 2001; Frankenhuis & Ploeger, 2007; Pietraszewski & Wertz, 2021; Spunt & Adolphs, 2017; Stephen, 2014; Stokes & Bergeron, 2015) though these models cannot completely solve the modularity debate. Instead, they should be seen as illustrative of the psychological mechanisms that could underlie our ability to master a complex range of skills, coordinate on many varied social norms, and (perhaps) come to cooperate. Section 2.4 addresses the benefits of modelling, plus the variables that I manipulate and measure, in each design.

2.2 The economic games

Economic games have been used to investigate social learning in the gene-culture coevolutionary literature (Efferson et al., 2008b; 2016; Mesoudi, 2016b; Molleman & Gächter, 2018) and are often considered the gold standard of investigating complex social interactions (Thielmann et al., 2021), which justifies their use here. While concerns have been raised about their external validity (Pisor et al., 2020), these economic games are beneficial as they simplify complex interactions down to simple choices that are directly observable. This is preferable to the researcher trying to infer participant preferences based on complex and uncontrolled interactions in a naturalistic setting (Haselhuhn & Mellers, 2005; Pisor et al., 2020). Moreover, economic games can reduce the influence of social desirability on results as it is easy to anonymise interactions (Pisor et al., 2020). This is particularly helpful when studying cooperation, as some participants may only cooperate in order to improve their reputation if they are identifiable (Burton-Chellew et al., 2016; Price, 2008).

To measure asocial skills, I use an economic game which I call a game against nature (see figure 1). I call it a ‘game against nature’ as the payoff of the focal individual’s choices is not influenced by her partner’s behaviour, or by the behaviour of the rest of the group, in any way. Thus, the game is asocial. Strictly speaking, economic games in game theory have an element of interaction. While my participants do interact in pairs (see figure 1), as a single participant’s behaviour does not affect the group or vice-versa, then this would be a decision problem as opposed to an economic game in the strictest sense (Hansen & Samuelson, 1988). However, I label these interactions as ‘games’ throughout to give a consistent terminology to the methods used throughout this thesis. Previous research also uses asocial tasks in a similar manner to a ‘game’ (Molleman & Gächter, 2018), supporting the use of this terminology here. In the game against nature, one option is worth more points for the individual to choose than the other option. The social learning study presented in chapter 3, and the agent-based model presented in chapter 5, use the game against nature task.

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 150** Partner's expected points: 150	**Your expected points: 150** Partner's expected points: 120
You choose @	**Your expected points: 120** Partner's expected points: 150	**Your expected points: 120** Partner's expected points: 120

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 120** Partner's expected points: 120	**Your expected points: 120** Partner's expected points: 150
You choose @	**Your expected points: 150** Partner's expected points: 120	**Your expected points: 150** Partner's expected points: 150

Figure 1. An example of the game against nature task, taken from the study reported in Chapter 3. Notice how the payoff of the focal participant's behaviour does not depend on what her partner chooses or vice-versa. In Game Left, % is optimal and @ is suboptimal. In Game Right, @ is optimal and % is suboptimal.

To investigate social norms, I use a coordination game (see figure 2). A coordination game is an economic game where the participant receives a higher payoff if she chooses the same behaviour as a partner with whom she has been paired (Bernard et al., 2020). Her partner is anonymous, and she cannot communicate directly with them, which increases the difficulty of this task (Efferson et al., 2008b). This pressure to coordinate on a certain behaviour when multiple equilibria is possible may drive participants to agree on a certain social norm regarding which option to choose. This makes the game comparable to how social norms drive behaviour in social groups (Efferson et al., 2008b; McElreath et al., 2003; Molleman & Gächter, 2018).

In the coordination game, one behaviour is worth more points to coordinate on than the other. Coordinating on the behaviour associated with the optimal mutual payoff is referred to as coordinating on the Pareto dominant equilibrium, or Pareto optimal equilibrium (Kets et al., 2021). Coordinating on the other behaviour is referred to as suboptimal coordination (Bernard et al., 2020). A failure to choose the same behaviour as one's partner is called miscoordination (Kets et al., 2021). Miscoordination always gives the lowest payoff, thus creating a pressure to coordinate with the other individual. The social learning study presented in chapter 3, and the agent-based model presented in chapter 6, use the coordination game.

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 325** Partner's expected points: 325	**Your expected points: 100** Partner's expected points: 100
You choose @	**Your expected points: 100** Partner's expected points: 100	**Your expected points: 250** Partner's expected points: 250

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 250** Partner's expected points: 250	**Your expected points: 100** Partner's expected points: 100
You choose @	**Your expected points: 100** Partner's expected points: 100	**Your expected points: 325** Partner's expected points: 325

Figure 2. An example of a coordination game, in this case taken from the study reported in Chapter 3. The participant receives the lowest payoff for failing to coordinate with her partner. In game left, % is the Pareto dominant equilibrium and coordinating on @ would be suboptimal coordination. In game right, @ is the Pareto dominant equilibrium and coordinating on % would be suboptimal coordination.

To measure cooperation, I use a Prisoner’s Dilemma (see figure 3; Axelrod, [1980]; Holt and Roth, [2004]; Sally, [1995]). In this game, the participant chooses to cooperate or defect with an anonymous partner. The logic is this. If she defects when her partner cooperates, then she receives a higher payoff than if she had cooperated. If she defects when her partner defects, then she still does better than if she had cooperated. Defecting when her partner cooperates gives the highest payoff, called the free rider payoff. Cooperating when her partner defects gives the lowest payoff, called the sucker payoff. As defection is the best strategy at the individual level, then the game is thought to have only one Nash equilibrium: mutual defection (Holt & Roth, 2004). However, mutual cooperation gives a higher payoff than mutual defection and thus the game represents a dilemma (Molleman & Gächter, 2018; Molleman et al., 2014; Sally, 1995). The social learning study presented in chapter 4 and the agent-based model presented in chapter 7 use the Prisoner’s Dilemma.

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 145** Partner’s expected points: 145	**Your expected points: 280** Partner’s expected points: 100
You choose @	**Your expected points: 100** Partner’s expected points: 280	**Your expected points: 235** Partner’s expected points: 235

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 235** Partner’s expected points: 235	**Your expected points: 100** Partner’s expected points: 280
You choose @	**Your expected points: 280** Partner’s expected points: 100	**Your expected points: 145** Partner’s expected points: 145

Figure 3. An example of a Prisoner’s Dilemma, in this case taken from the study

reported in Chapter 4. The game tables (left or right) depict which option would be the Nash equilibrium, if both players chose this (i.e., the equivalent to choosing to defect). In game left, % is the Nash equilibrium if both players choose this and @ is cooperation. In game right, @ is the Nash equilibrium if both players choose this and % is cooperation.

As can be seen in Figures 1-3, the agent must choose between two arbitrary symbols (@ or %) during the economic games used in the studies of Chapters 3 and 4. I used these symbols to avoid certain artefacts that may influence the results (Efferson et al., 2016). For the Prisoner's Dilemma in chapter 4, the participants may merely choose to 'cooperate' if they see this label due to a social desirability effect (Price, 2008). The arbitrary options chosen helped to avoid social desirability.

I did not use ordered labels (such as A and B, or 1 and 2) as this could make coordination in the coordination game easier (Efferson et al., 2016). If a participant is paired with an anonymous partner with whom she cannot speak, and she knows that she must choose the same option as this other person, then it becomes relatively obvious for both the participants to simply choose 1, or A, as these come first. Finally, using labels such as crop names or colours may induce a role of personal preference. Of course, it is possible for agents just to have a preference to choose one of the in-game options (@ or %) which shall be investigated in my analysis of these chapters. Previous research also uses simple symbols (Efferson et al., 2016) and by using these, I hope to minimise confounding effects on my results. Once the arbitrary symbols of @ and % were decided, these were kept consistent across all three games used in Chapters 3 and 4.

2.3 The social learning studies

All the empirical studies involved recruiting a group of participants (*N* ~ 16-30 in each session, with multiple sessions per game type) to play the economic games simultaneously in a computer lab. The participants played one of the three economic games via Z-Tree (Zurich Toolbox for Ready-made Economic Experiments; Fischbacher, 2007). This programme allows group-wide randomisation, meaning that the participants could interact anonymously with others in their group (Ertac & Kotan, 2020).

I test in groups for two reasons. First, it allowed the participants to genuinely interact with another person during the economic games. This avoids deception, as would be the case if the researcher plays the games in a certain way, or the computer plays certain strategies. As well as being unethical, a participant who feels that she is being deceived may behave in a different way than she would do otherwise (Davis & Holt [1993][chapter 1.6]; Wilson, [20146; though see Krasnow et al. [2020] for a rebuttal). Second, investigating these economic games in real time allowed for a meaningful shift in choices to emerge *endogenously* in each social group. This is key, as it allows experiments to investigate how new social groups are formed and how they come to agree on certain skillsets and norms with each other (Efferson et al., 2008b; McElreath et al., 2003).

Six of the participants in each session were assigned to play the role of 'demonstrators' throughout. Demonstrators could form their own preferences regarding which option to pick when playing the economic game in their session. Although the demonstrators did not know which option would be optimal to pick, they received immediate feedback regarding the points that they had earned for their decision in each period, and they played four periods in a row in each block of the game. This allowed for trial-and-error learning (Mesoudi et al., 2015; Molleman & Gächter, 2018). The games in figures 1-3 represent the expected points that a participant could earn from her decision, but these points were also affected by a random disturbance (M=0, SD=20 points). This ensured that the task was difficult enough for the demonstrators trial-and-error learning to be non-trivial. This disturbance also ensured that the sessions ended with a range of demonstrators choosing optimally, rather than having performance fixed at ceiling, which would likely be the case if the random disturbance was not added. Note that I use six demonstrators as this was considered a sizeable enough group over which group-wide preferences for certain choices may have emerged (Efferson et al., 2016).

The remaining participants were all ‘social learners’. The social learners had no opportunity for trial-and-error and could only base their choices on certain social information. It may seem simplistic to separate the participants into two roles for individual and social learning, as real populations often cycle between the two (Miu & Morgan, 2020; Mesoudi, 2008a; Molleman & Gächter, 2018). As these studies were interested in the flexibility of the participants’ decision to conform, then it was vital to ensure that the social learners adjusted their choices to the social information only. Separating the participants into two roles thus allowed me to investigate the flexibility of social learning without the inferential challenges that typically hold when the participants are allowed to learn both socially and via their own trial-and-error (Angrist, 2014; Manski, 2000).

The social information that the social learners could respond to was of the same type across all empirical studies. In chapter 3, half the sessions concerned the game against nature task and the other half concerned the coordination game. In chapter 4, the sessions concerned the Prisoner’s Dilemma game. The social information displayed to the social learners included: (i) the frequency of choices made by the demonstrators; (ii) a signal telling the social learners whether they learned in a similar or different decision-making environment to the demonstrators and (iii) the reliability of this similarity signal.

Frequency is investigated as the simplest test of social information use. A conformist learner, as per the definition that I use in this thesis, should respond to frequency information (Efferson et al., 2016; Lachlan et al., 2018; Morgan et al., 2019). Note that the actual choices made by the demonstrators in each session formed the frequency information. If the social learners saw a signal telling them that four out of six demonstrators chose @, then four of the demonstrators in this session would have genuinely chosen @ in this block. Only the demonstrators' responses in the final period of each block were used, as these responses should represent the demonstrators' choices after trial-and-error (Efferson et al., 2016).

As for the similarity signal, there is a wealth of studies to suggest that individuals prefer to learn from similar others (Efferson et al., 2008b; Richerson et al., 2016) though a lot of empirical studies only investigate the role of one similar or dissimilar demonstrator (Salali et al., 2015; Shutts et al., 2010; Wood et al., 2013). Previous research may also focus on similarity on observable traits, such as age or gender of the participant (Salali et al., 2015; Shutts et al., 2010).

Similarity on these observable traits are important as long as they infer similarity on key unobservable traits (Efferson et al., 2016; Richerson et al., 2016; Smaldino et al., 2018). That is, we like to learn from those who appear similar to ourselves as these individuals are likely to share similar decision-making environments to ourselves, and so are likely to be displaying behaviour which would be optimal for us to copy, too. In my studies, I cut out these observable cues and instead directly tell participants whether or not they share a decision-making environment to the group from whom they learn. This allowed me to objectively investigate how similarity at the level of the social group would influence one's chosen frequency-dependent social learning strategies.

Third, reliability is important to test as signals suggesting a shared group membership have not always been accurate throughout the ancestral past (McElreath et al., 2003; Smaldino et al., 2018). In the study chapters, the reliability signal had three levels: reliably correct, uninformative, or reliably incorrect. Reliably correct signals provided accurate similarity information 9 times in 10 and represented cases where the individual bases their similarity to others on readily observable cues, such as age (Jiménez & Mesoudi, 2019) or gender (Efferson et al., 2016).

The uninformative signal was correct 5 times in 10 and rendered the similarity information at chance level of being correct. At chance probability, the social learner is maximally unsure as to whether the group from whom she learned may be similar or different to herself. She therefore had no guarantee that conformity would lead to higher payoffs than any other social learning strategy. The participants may therefore have chosen options based on arbitrary preferences, or chosen randomly, in response to this uninformative signal.

Finally, a reliably incorrect signal only provided correct similarity information 1 time in 10. To illustrate the reliably incorrect signal with an example, consider the extremely cooperative group who are identifiable as all the members wear red clothing (I also detailed this example in Chapter 1, section 1.1). Outsiders with no intention of cooperation may infiltrate the group to take resources by dressing in a similar fashion. If the clothes that they wear are an off-shade of red, then the colour of the clothing may become a reliably incorrect signal of who could be similar to oneself. It is reliably incorrect, as the off-shade of red is typically a faked signal of group membership. The reliability of similarity signals may become important in cases where an extremely cooperative group may be at risk of free-rider infiltration (Richerson et al., 2016; Sosis et al., 2007), as is the case in my red clothing example.

All the sessions ended with the participants completing an anonymous questionnaire. This asked for demographics such as age and gender, as these may affect the participants' preferred social learning strategies (Mesoudi et al., 2015; Wen et al., 2019). The questionnaire also asked for the participants' main country of residence. Note that chapter 3 recruited in the Foundation for Liberal and Management Education (FLAME) University in Pune, India. This was to generalise beyond the WEIRD samples that are often recruited in evolutionary work (Henrich et al., 2010; Henrich, 2020; Muthukrishna et al., 2020). Many previous studies compared highly-individualistic countries, such as the US and UK, to highly collectivist countries, such as China (Bond & Smith, 1996; Mesoudi et al., 2016; Molleman & Gächter, 2018). I instead chose to recruit in India as this country falls half-way on Hofstede's (1980) individualism-collectivism scale and these middling countries may be under-represented.

However, there was no guarantee that the students who attended the sessions at FLAME university were Indian just as there was no guarantee that those students who attended Royal Holloway, University of London (RHUL) in Chapter 4 were English. It is worth pointing out that the samples in chapters 3 and 4 were not wildly different. As I recruited both studies in universities, then the typical student is likely to be more educated, rich, and democratic than the general population of India or the UK as a whole (Henrich, 2020). The participants were similar enough to suggest that they learned to play these economic games in a similar fashion (Ekuni et al., 2020).

While this is useful for comparability purposes, future studies may wish to investigate social learning processes in a broader sample than university students.

It would have been ideal to recruit both studies in both India and the UK for full comparability. Chapter 3 (testing asocial skills and social norms) recruits in FLAME University in India, while Chapter 4 (testing cooperation) recruits in RHUL, UK. The cultural background of my participants may influence their choices in Chapter 3 compared to Chapter 4, though the Prisoner's Dilemma was deemed a different game structure as there is no optimal option per se and one's outcomes depends entirely on what one's partner chooses. That is, the games are distinct enough to be thought of as separate chapters and thus having separate samples is not too problematic.

To summarise, the aim of these empirical studies is to investigate the flexibility with which the participants conform– or use other social learning strategies– in light of three pieces of social information. Chapter 3 recruits in FLAME University, India and investigates the flexibility of social learning using a game against nature to measure asocial skills and a coordination game to investigate social norms. Chapter 4 recruits in RHUL, UK and investigates the flexibility of social learning using a Prisoner's Dilemma to measure cooperation.

2.4 The agent-based models

Agent-based models are an important tool in gene-culture coevolutionary research as they help avoid the inverse problem (Deffner et al., 2020; Tarantola, 2006). Rather than inferring the complexity of an agent's cognitive processes based on their behaviour alone, agent-based models force the researcher to be explicit in her hypotheses about how psychological mechanisms drive behaviour. Once a framework is agreed, a model is built to reflect the conditions of the ancestral past and then the researcher allows one or two psychological traits to evolve endogenously throughout the model's runtime.

The analysis then visualises the final generation of agents' psychological architecture, plus their behaviour. Thus, models allow the researcher to directly test hypotheses regarding how psychological mechanisms relate to behaviour which avoids the inverse problem. Moreover, by building a representation of psychology in this model, I avoid 'blackboxing' the role of cognition in agent behaviour (Heyes, 2018; Kendal et al., 2018). A final strength to consider is that models allow the researcher to visualise the evolutionary trajectories of agent behaviour on a detailed scale (Laland, 1993).

Perhaps the main criticism of agent-based modelling is that these may represent an 'oversimplification' of complex social phenomena (Mesoudi et al., 2006b). However, other researchers in the gene-culture coevolutionary literature argue that agent-based models are useful *because* they are simple (Laland, 1993; Mesoudi et al., 2006b). By quantifying the relationships between some key variables, then the researcher can clarify the logic of any theoretical arguments.

Moreover, the use of agent-based models may be preferable to purely relying on the researcher's intuition based on what one knows about the ancestral past. These verbal arguments are fallible to just-so storytelling (Ketelaar & Ellis, 2000; Gould & Lewontin, 1979). A model reduces the chance of this as, while the researcher's intuition behind the model may be wrong, the model itself is objective in the data it produces (Efferson et al., 2020b; Mesoudi et al., 2006b).

My agent-based models investigated how agents come to master asocial skills (chapter 5), social norms (chapter 6) and cooperation (chapter 7) respectively. Every agent in each model played two economic games. These were always the same game type. So, the agents in Chapter 5 each played two games against nature; the agents in Chapter 6 each played two coordination games and the agents in Chapter 7 each played two Prisoner's Dilemmas. These games represented two separate domains of decision-making, which I refer to as domain A or domain B throughout.

For asocial skills (Chapter 5), the two game against nature tasks may represent two separate skill domains. For example, the decision of when to hunt with a bow-and-arrow (Tomka, 2013) and when to cook food (Wrangham, 2009). For social norms (Chapter 6), the two coordination games may represent two social norm domains. For example, engaging in a collective dance ritual (Gelfand et al., 2020) or deciding which side of the road to drive on (Hao et al., 2017). For cooperation (Chapter 7), the two Prisoner's Dilemmas can be thought of as the decision to cooperate in two domains. For example, deciding whether to join a collaborative hunting effort or deciding whether to join a communal building project (Chudek & Henrich, 2011). The full purpose of these examples are expanded upon in their respective chapters.

To directly compare both domain-general and modular decision-making, I simply refer to behaviour in each decision-making domain as the decision 'to act' or 'not to act'. Note that in each domain the environment may take one of two states (0 or 1). Put simply, these environmental states controlled which behaviour (0 or 1) would be optimal for the agent to play. To illustrate, consider the use of the bow-and-arrow as domain A in chapter 5. When the state was 0, then it would have been optimal for the agents to not act (i.e, to avoid using the bow-and-arrow). When the state was 1, then the it would be optimal for the agents to act (i.e, use the bow-and-arrow). To illustrate in reference to social norms, consider the social norm regarding collective dance as domain A in chapter 6. When the state was 0, it would have been optimal for the agent and her partner not to act (i.e., to avoid dancing). When the state was 1, then it would have been optimal for both of the agents to act (i.e, to to dance).

Finally, the unique demands of cooperation made for different model specifications. In chapter 7, the state of the environment (0 or 1) instead controlled which behaviour was considered cooperative (0 or 1). To illustrate, consider the decision of whether or not to join a collaborative hunting effort in domain A of chapter 7. When the state was 0, then it was cooperative to avoid over-hunting. For example, when over-hunting would deplete resources and take away from one's neighbours (Safin et al., 2015). Behaviour 0 would then represent the decision not to act (i.e., avoid hunting). When the state was 1, then it would instead have been cooperative to hunt more and share resources with the group (Hill, 2002). Behaviour 1 would then represent the decision to act (i.e, to hunt). Note that behviour 1 always reflects the decision to act, and behviour 1 always reflects the decision to act, in each domain. This level of abstraction allowed the basic assumption of domain-general mechanisms to be worked into the model, in that the agent can make similar, simultaneous decisions over multiple domains.

In order for the agents to display a certain behaviour in these models, I had to code their underling psychology. I took the most simplistic representation of human behaviour as one that involves both a cognitive and motivational system (Delton et al., 2011; Tooby et al., 2006). In order to show a certain behaviour, the agent first had to formulate a belief regarding how likely it was that this behaviour was beneficial in her current environment. If she approved of the behaviour, she then must be motivated to display it (Chiappe & MacDonald, 2005).

I investigated these psychological components as cognition and motivation have historically been modelled as separate processes. To illustrate with an example, consider Delton et al's. (2011) model. They investigated the EMT of one-shot cooperation. That is, agents are assumed to cooperate in one-shot interactions in cases where there is a potential for a long-lasting future relationship with another individual (Krasnow & Delton, 2016). This is because a one-shot donation of resources which is not reciprocated will be less costly than scuppering the chances for a long-lived, mutually beneficial relationship (Haselton et al., 2015).

When investigating the evolution of one-shot cooperation, Delton et al. would fix either cognition or motivation, and leave the other variable to evolve endogenously throughout their model. Cognition could be fixed at an arbitrary baseline, and then the agents' motivation to cooperate on both a one-shot and a repeated interaction was left to evolve. Alternatively, motivation could be fixed so that the agent always wanted to defect on a one-shot interaction or cooperate on a repeated one; but the agent's cognitive thresholds for determining a repeated interaction could evolve.

Delton et al. likely decided to investigate these components separately in an effort to avoid psychological polymorphism. Psychological polymorphism describes the case where, if several variables are allowed to coevolve, then there may be multiple psychologies that evolve endogenously throughout a model's runtime, where a single psychological type can be represented as a point in multidimensional space. Thus, there will be a range of outcomes that can be represented graphically if psychological polymorphism affects my results. Finding this result in my models would suggest that a range of behavioural strategies may become related to multiple psychological spaces, which means that the exact relationship between agent psychology and behaviour becomes hard to disentangle (Laland, 1993).

On the other hand, agent-based models are only useful in as far as they reflect complex human behaviour (Kendal et al., 2018). When early humans came to make decisions throughout the ancestral past, then both their cognitive and motivational systems would have evolved in-tandem. The decision to investigate both processes separately may therefore be an oversimplification. To correct this, I allowed both cognition and motivation to coevolve to understand the complexity with which these systems underlie human behaviour.

Further motivating this design, consider that agent-based models may either be pragmatic or conceptual (Muthukrishna & Schaller, 2020). A pragmatic model is one that exists in its simplest form, in order to explicitly test a theoretical assumption that has already been articulated elsewhere. Delton et al's. (2011) model is an example of a pragmatic model, as it adds weight to their proposed EMT of the evolution of one-shot cooperation. Conceptual models may not make explicit assumption per se, but their ultimate aim is to measure complex and coevolving variables. There has been an increasing interest in these conceptual models recently (Markov & Markov, 2020; Muthukrishna & Schaller, 2020) supporting my current models. My models are conceptual as they are the first to address the theoretical debate of domain-generality versus modularity in a theoretically evolving population (see section 1.3).

Human societies uphold a complex set of skills (Henrich, 2015), social norms (Henrich & Muthukrishna, 2021) and cooperative norms (Chudek et al., 2013). Whether these behaviours are underlined by an impressive generic reasoning ability (Bolhuis et al., 2011; Mesoudi, 2011b; Vergauwe et al., 2021) or by a separate series of modules– each functioning on one problem humans faced throughout the ancestral past (Cosmides & Tooby, 1994b; Goschke & Bolte 2012) – remains an open question. Modelling agent cognition and motivation as separate processes is important to test this debate in more detail, by also investigating cases of partly modular agents between the two extremes of full domain-generality (Bolhuis et al., 2011) and full modularity (Cosmides & Tooby, 1994b). In these models, partly modular agents may have either modular cognition but domain-general motivation, or modular motivation but domain-general cognition.

These parlty modular agents are also important to investigate as they may represent a mid-ground. Some researchers posit that we have a few modules which are then shaped by domain-general processes (Chiappe & Macdonald, 2005; Michael & D'Auslio, 2015). Particularly, Chiappe & Macdonald (2005) argue that we may have modular cognitive biases which are shaped by domain-general motivational systems. My agent-based models are the first to test this theoretical claim by comparing the success of decision-making in partly modular agents to fully modular and domain-general agents.

It is worth acknowledging that the conceptualisation of modularity within Evolutionary Psychology – as well as the use of the term – has attracted much philosophical debate (Fodor, 2001; Frankenhuis & Ploeger, 2007; Spunt & Adolphs, 2017; Stephen, 2014; Stokes & Bergeron, 2015). In a recent review, Pietraszewski and Wertz (2021) argue a distinction between *intentional* systems, which are equivalent to a 'homunculus' in the brain that allows us to flexibly oversee decision-making across a variety of processes; and *functional* modules or mechanisms, which are designed – as per Evolutionary Psychologists – to work on certain inputs to produce certain behavioural outputs.

Whilst this philosophical debate is undoubtedly important, my models were unique in their aims to apply a mathematical notion to the study of such systems. What I call a *domain-general* system (and what other researchers may call an *intentional system* or a little homunculus) is any one system that must account for a range of environmental inputs. What I call a *module* (and what other researchers may call a *functional mechanism* or *domain-specific system*) is any system that can specialise to deal with only one set of environmental input.

It should be noted that the term 'domain-general' in the wider literature is often tied to associative learning (Bolhuis et al., 2011; Macintosh, 1974; Reader et al., 2011). I do not attempt to replicate these models here. As no previous models have compred modular and domain-general psychology before, then I attempt to compare these agents in the most straightforward case. As encapsulated in Pietraszewski and Wertz's (2021) review, evolutionary psychologists often use the term 'domain-general' to describe any decision-making system which much operate on the highest level of abstraction, by making decisions across a range of domains. Throughout my agent-based model, I simply take agents that must make decisions over two domains. As per the concept of functional specialisation (Müller, 2007; Reader, 2006), I take a modular agent has one that has two separate decision-mking systems, which can specilise to the demands of one domain alone. Conversely, domain-general systems have a limited pool of resources to react to various environmental inputs (Örün & Akbulut, 2019; Vergauwe et al., 2021). To capture this theoretical concept, I simply provide a domain-generl system with one system that must work across both domains. Whilt domain A and domain B may seem distinct throughout my models in Chapters 5-7, remember that this comparison is warranted if we see decision-making in each domain as the decision ‘to act’ or ‘not to act’. This level of abstrction makes decision-making achievable for a domain-general agent in my models.

Thus, agents with modular motivation could show a separate desire to play a certain behaviour when they thought that this was optimal, and a separate desire to play a certain behaviour when they thought that this was suboptimal, as they had separate motivational thresholds that could evolve distinctly in domains A and B.

Instead, agents with domain-general thresholds simply had one cognitive threshold to process the likelihood of a certain behaviour being optimal across both domains A and B. Agents with domain-general motivation had a generic drive towards displaying a certain behaviour when they believed that this was optimal, and a generic drive towards displaying a certain behaviour when they believed that this was suboptimal, across both domains A and B.

As I model modularity separately for the cognitive and motivational systems, these models can be thought of as investigating four agent types:

1. Fully modular agents had a separate cognitive threshold for each domain, which allowed them to reach specific conclusions about the likelihood of a behaviour being optimal in each domain. They also had modular motivation, which allowed them to show a distinct desire to play certain behaviours in the separate domains.
2. Partly modular agents with modular cognition only. They had separate thresholds to reach specific conclusions about the likelihood that a behaviour was optimal in each domain, though they had domain-general motivation and so could only show a generic desire towards playing a certain behaviour across both domains.
3. Partly modular agents with modular motivation only. They only had one cognitive threshold to reason about the likelihood of a certain behaviour being optimal across

both domains, though they had modular motivation and so they could specialise their desire to play certain behaviours in each domain.

4. Domain-general agents had one cognitive threshold, so they could only reason about the likelihood of a behaviour being optimal across both domains. They also had domain-general motivation, which meant that they could only show a generic desire towards playing a certain behaviour across both domains.

All models coded 100 agents to play the two economic games of interest (Chapter 5 focuses on asocial skills; Chapter 6 on social norms and Chapter 7 on cooperation). In each run, the agent decided to play a certain behaviour as based on her cognitive or motivational thresholds. The agents then received a certain amount of fitness, as dependent on how her choice aligned to the context of the economic game used in each model (see the methods section of each chapter for specific details).

The agents' fitness was summed across the two domains, and all agents received an exogenous fitness value of 1 on top of their decisions in these domains. This exogenous fitness value reflects the fitness that the agents may gather from interactions outside of the focal behaviour of interest in these three chapters (Molleman & Gächter, 2018) and are also important to ensure that the agent's fitness can never be negative. Negative fitness values are hard to represent and often lead to agent's dying at different rates. While this may reflect a realistic level of entropy in the environment, this is logistically difficult to model and may distract from the main aim of my models, which is to investigate whether domain-general or modular cognition is more likely to underlie complex human behaviour. I therefore considered a Wright-Fisher model, where the offspring replaced the parental generation entirely after inheritance to ensure a consistent death and birth rate (Suchow et al., 2017).

The agent's fitness was transformed into a cumulative proportion of the entire generation's fitness. When offspring agents were made, they drew their cognitive and motivational thresholds from the generation before as dependent on these cumulative fitness values. The parental agents with a higher fitness had a higher chance of having more offspring. Offspring agents inherited these thresholds with a small mutation rate which means that reproduction was proportional to fitness and affected by mutation (Suchow et al., 2017). Cognition and motivation were the variables that I leave to evolve endogenously throughout the model's run time. I ran 20,000 generations over 100 simulations for each combination of parameter values of interest.

Parameter values are those that are varied exogenously by the researcher, to see how these variables affect the evolutionary trajectory of the endogenously evolving variables (Fagiolo et al., 2007). In this case, I varied certain parameters to see how these affected the coevolution of the agents' cognitive and motivational systems. I varied the fitness associated with certain behavioural outcomes, though these are specific to the economic games modelled and thus this will be discussed in-depth in the methods section of the relevant chapters. I also varied the prior probabilities that the environment will be a certain state.

Specifically, I modelled the likelihood that the environmental state was 1 in domain A (p_A) and the likelihood that the environmental state was 1 in domain B (p_B). Note that the likelihood of the environmental state being 0 was thus given by 1-p_A and 1- p_B respectively. The p_A and p_B values were modelled separately for each domain and could take a low (0.1), chance (0.5) or high probability (0.9). Here, I focus on the parameter combinations of interest which I will investigate throughout the modelling chapters of this thesis.

First, I consider cases where the agent made decisions over two similar domains. To achieve this, I set the prior probabilities to be the same across both domains. This represents cases where the agent acquires two similar skills (Chapter 5), coordinates on two similar social norms (Chapter 6) or decides to cooperate in two similar domains (Chapter 7). For the example of asocial skills, if one lives in an environment where the bow-and-arrow is a necessary skillset as it is common to hunt easily startled prey, then this also implies that the same environment would encourage the use of any long-distance weapon, such as a spearthrower (Hughes, 1998). Specifically, I considered runs where state 0 was modelled to be likely across both domains ($p_A = 0.1, p_B = 0.1$) or state 1 was modelled to be likely across both domains ($p_A = 0.9, p_B = 0.9$). It is expected that the agent's cognitive and motivational thresholds evolve on a specific trajectory in both runs, in order to maximise the likelihood of the agent's choosing the behaviour that is likely to match the most common environment across both domains.

Second, I also considered runs over two distinct domains. To achieve this, I set the prior probabilities to take different values across both domains. This would be like trying to master a behaviour across two widely different domains. For example, a social norm regarding dance is likely to be unrelated to a social norm regarding which side of the road to drive on (Henrich & Muthurkrishna, 2021). I arbitrarily considered a run where state 0 was favoured in domain A and state 1 was favoured in domain B ($p_A = 0.1, p_B = 0.9$). Modularity may have been more important here, as the agent has to specialise to two distinct priors.

For every parameter combination of interest, I structured the results section of chapters 5-7 to adhere to the following format. First, I visualised the behavioural outcomes of the final generation of agents with clustered bar charts (section 3.1). This allowed me to visualise the actual decisions made by the final generation of agents. Second, I created binned heatmaps to visualise the psychological architecture of the final generation of agents (section 3.2). Then, I ran a regression to see how the agents' cognitive and motivational thresholds predicted their fitness (section 3.3). These analyses showed me whether cognition and motivation coevolved to meaningfully predict agent behavioural outcomes and fitness. A lack of significant findings here may have suggested some sort of psychological polymorphism where cognition and motivation coevolved to influence behaviour in a myriad of ways which are hard to disentangle (Laland, 1993). Finally, I ran a regression comparing the fitness of the four agent types (section 3.4). This analysis will show whether modular or domain-general psychology was more likely to uphold our decision-making across the two domains in each chapter. As the models provided a large amount of data, I finished each chapter with a summary of three key findings (section 3.5) which then directed the discussion of these chapters.

To summarise, chapters 5-7 use agent-based models to investigate the complexity of the psychology underlying three key areas of decision-making. Specifically, chapter 5 uses the game against nature task to inform a model of how we come to master asocial skills over two domains. Chapter 6 uses the coordination game to inform a model of how we come to coordinate on social norms over two domains. Chapter 7 used the Prisoner's Dilemma to inform a model of how we decide whether or not to cooperate across two domains. These models are novel in comparing how domain-general versus modular psychology come to uphold complex human decision-making across two distinct domains.

2.5. Methods summary

This thesis uses both social learning studies and agent-based models, as both these techniques have complimentarily strengths and are common in the gene-culture coevolutionary literature (Efferson et al., 2020b). I use three economic games (game against nature, coordination game and the Prisoner's Dilemma). These games represent the learning of asocial skills, social norms, and cooperation respectively. The social learning studies in Chapters 3 and 4 apply these economic games to a group-based laboratory experiment. These studies have the aim of investigating the flexibility of the individual's decision to conform in response to three pieces of simultaneous social information. The agent-based models in Chapters 5-7 use these economic games to build a representation of decision-making. These models aim to investigate the likely psychological mechanism (i.e., domain-general, or modular) that underlies complex decision-making.

Chapter 3
What is the extent of a frequency-dependent social learning strategy space?

Aysha Bellamy
Ryan McKay
Sonja Vogt
Charles Efferson

Published in in Evolutionary Human Sciences

Word count: 11,000 excluding references and appendices

Abstract

Traditional models of frequency-dependent social learning strategies only investigate how individuals respond to the frequency of a behaviour in a group, which gives the impression that these strategies only operate on one order of social information. I believe that learners should also adjust their strategies in response to social information about the group from whom they learn. Specifically, I tested 302 participants' ability to master skills and social norms in response to: (i) the proportion of a group who chose certain options; (ii) whether this group were identified as having learned in a similar or different environment to the participant and (iii) the reliability of this similarity information. The similarity information could have been reliably correct, uninformative, or reliably incorrect, to reflect cases when observable cues of similarity did not match similarity on unobservable traits. Participants adjusted their social learning strategies in response to all three pieces of social information, though they performed better when learning from groups who shared a similar decision-making environment, provided this signal was reliably correct. Nonetheless, individuals adjusted their frequency-dependent social learning strategies to three orders of social information and so the extent of the frequency-dependent social learning strategy space has been underestimated.

Key words: social learning, frequency-dependent social learning, conformity, cultural evolution, social norms.

1. Introduction

Cultural evolution describes how social information is shared over time (Mesoudi et al., 2004). Social learning strategies help individuals to filter social information (Efferson et al., 2016; Mesoudi & Lycett, 2009; Molleman et al., 2013). One's choice of social learning strategy may have profound effects on cultural evolutionary outcomes (Henrich & Boyd, 2001; Henrich & Muthukrishna, 2021; Molleman et al., 2014).

The variability in individual choices of social learning strategies has been a topic of great interest in past research (Acerbi & Tehrani, 2018; Brady et al., 2020; Henrich, 2015; Henrich et al., 2015; Lenfesty & Morgan, 2019; Price & van Vugt, 2014). This paper compliments this field, by investigating the variability in individuals' chosen frequency-dependent social learning strategies. Frequency-dependent social learning strategies involve copying a behaviour based on *how many* others in a group display this behaviour. Perhaps the most widely studied frequency-dependent social learning strategy is conformity. There has been a debate over the exact definition of conformity (Efferson et al., 2016; Morgan & Laland, 2012; Morgan et al., 2019). For the sake of clarity, in this paper I consider conformity to mean a *disproportionate tendency* to adopt the same behaviour as the majority of a *group*. For example, if 75% of a social group wear their hair in a braid, then a conformist learner will also wear her hair in a braid with a probability greater than 0.75.

Conformity shapes cultural evolution (Chudek & Henrich, 2011; Henrich & Muthukrishna, 2021). In Boyd and Richerson's (1985) seminal model, conformity led to the homogenising of social groups. In this model, the agent would sample three other group members at random and– based on the size of a parameter *D*– the individuals that had a conformity preference would adapt the same behaviour as the majority of these three agents. The group members became more similar over time. As the agent only copied behaviours based on how frequent they were within the individuals that they sampled, then this model may give the impression that individuals only process frequency-dependent social information when choosing to conform and that some individuals always copy the majority as a rule-of-thumb.

While there is likely to be individual variation in conformity preferences (Muthukrishna et al., 2016), the notion that conformity is a fixed preference does not match recent literature. First, learners can flexibly switch between using a frequency-dependent social learning strategy or following their own intuition (Boyd & Richerson, 1985 [chapter 7]; Mesoudi et al., 2015; Miu & Morgan, 2020; Reader, 2003). For example, Deffner et al. (2020) found that agents were more likely to conform– rather than use their own trial-and-error– when migrating to a new social group, as they must adjust to a new environment. Second, individuals can switch between using frequency-dependent social learning strategies and payoff-based social learning strategies (McElreath et al., 2008; Molleman & Gächter, 2018). Finally, some individuals will only conform when uncertain (Kendal et al., 2018; Toelch et al., 2014).

Individuals are clearly flexible when choosing between frequency-dependent social learning strategies *and other strategies.* This paper takes a different approach. Conditional on frequency-dependent social information, I ask when an individual will use *different* frequency-dependent social learning strategies based on social information about the groups from who she learns. Individuals do not merely conform. They may also follow the majority in a linear fashion (Morgan & Laland, 2012), or copy the minority of a social group (Efferson et al., 2008a; Evans et al., 2018). Individuals differ in their preferences for different frequency-dependent social learning strategies (McElreath et al., 2008) and this is likely to have unique effects on the behaviours preserved at the group level (Berger & Heath, 2007). Rather than responding to one order of social information, as in Boyd and Richerson's (1985) seminal model, it is my intuition that individuals will differ in their preferences to follow a majority– or a minority– based on certain social information about the groups that they are exposed to.

An individual's preferred frequency-dependent social learning strategy may depend upon the decision-making task. In this paper, I focus on asocial skills and social norms (Legare & Nielsen, 2015). Asocial skills refer to any behaviour that maximise one's own payoffs but does not affect others in a group, such as private tool use. I test this with a game against nature (see section 2.2). Social norms are rules that most members of a society should coordinate on (Legare, 2017). I test this with a coordination game (Kets et al., 2021). These tasks create different pressures for my participants. In the game against nature, the participant must choose the option that gives themselves the highest payoff. Following the choices made by the majority of a group makes sense as the majority of a group tend to display similar behaviours for a reason (Barrett, 2015 [chapter 9]). Conformity may be more important during a coordination game as this mirrors the learning of social norms where everyone must coordinate on the same option to receive a high payoff (Legare, 2017; Wen et al., 2019).

As well as investigating the participants' responses to two different decision-making tasks, my main aim is to investigate whether the participants will change their frequency-dependent social learning strategies based on social information about the group from whom they learn. This includes: (i) the frequency of a behaviour among this group; (ii) whether this group are indicated as making decisions in a similar or different environment to the participant and (iii) whether this similarity information is reliable. I expect all participants to adjust their behaviour to frequency-dependent social information (henceforth ‘first-order social information’). Even those who follow a simple rule to ‘always copy the majority’ or ‘minority’ are, by definition, responding to frequency information.

The second-order social information is the signal informing the participants that they make decisions in a similar or different environment to the group from whom they learn. Previous studies have focused on whether groups *appear* similar or different to the participants, on traits such as age (Shutts et al., 2010) or gender (Efferson et al., 2016). Such cues are useful as a short-hand for identifying those who face similar decision-making environments, and have similar optima, to oneself. For example, those of the same gender are likely to master similar skillsets, and uphold similar social norms (Efferson et al., 2016). Focusing on similar external appearances may not capture the full picture, however. Consider a recent migrant to a new social group. She will conform to the behaviour of the group, even if they do not appear similar to herself, as she must learn about the local optima of her new group (Deffner et al., 2020; Henrich, 2015). I simply inform the participants whether they are playing a similar or different game to the group from whom they learn throughout this study, in order to provide direct information about who the participant shares a similar decision-making environment with.

Individuals may conform to a group that shares a similar decision-making environment to themselves (Efferson et al., 2016). Conversely, individuals may be more likely to follow the minority behaviour of a group of different others, in order to signal a different affiliation to this group (Efferson et al., 2016; Smaldino et al., 2017;

Smaldino & Jones, 2021). While some previous research highlighted the different ways in which participants respond to groups labelled as similar versus different, this study is novel as I also test the social learners' response to a third level of social information.

This third order of social information is the reliability information. The signal telling the participant that they play the same or different game to the group from whom they learn can be reliably incorrect, uninformative, or reliably correct. A reliably correct signal would indicate cases where the individual knows whether group members share the same decision-making environment as herself (Deffner et al., 2020) or when the individual bases her decisions of who is similar to herself on readily observable cues, such as age (Jiménez & Mesoudi, 2019) or gender (Efferson et al., 2016).

A reliably incorrect signal renders the signal regarding who shares the same decision-making environment as the participant likely incorrect. For example, picture a cooperative group who are known to share their abundant wealth equally amongst all group members and that this group are identifiable as they all wear red clothing. It would be easy for an outsider to acquire red clothing in order to infiltrate this group. This free-rider wears red clothing as a faked signal of cooperative intentions, in order to exploit resources from the well-meaning cooperative group members.

This example offers a useful context for envisioning the reliability signal. When group markers are easy to fake (e.g., it is relatively easy for someone to acquire red clothing) and when the incentives of faking this marker are high (e.g., the outsider can infiltrate a cooperative group and take some of their resources), then markers are likely to be reliably incorrect. When markers are difficult to fake and the incentives for doing so are low, then any signals of group affiliation are likely to be reliably correct.

Realistically, markers would fill an intermediate case. Cooperative groups typically employ painful or elaborate signals of group membership precisely because these are difficult for outsiders to fake though there would be high incentives for doing so (Smaldino et al., 2018; Sosis et al., 2007). For example, imagine that the cooperative group begin to wear a shade of red that is difficult for any outsiders to produce. Those individuals wearing an off-shade of red are thus likely to be outsiders. The off-shade red clothing becomes a reliably incorrect signal as it suggests who may be *faking* group membership (Stein et al., 2021).

Some signals of group membership may also be covert. Covert markers are useful because they are ambiguous and thus are less likely to draw negative attention from different others (Smaldino, 2019; Smaldino et al., 2018). In this case, signals of similarity or difference would be uninformative. This theoretical work suggests that the reliability of similar or different signals cannot be assumed as markers are not always easily interpreted. Indeed, Toelch et al. (2014) have shown that individuals adjust their preferences to conform or use an individual learning strategy based on the perceived reliability of social information. While this confirms my intuition that the reliability of social information is likely to be important, their study did not consider three orders of social information.

Our participants may adjust to a reliably correct and a reliably incorrect signal as both these signals are informative. However, the participants cannot adjust meaningfully to the uninformative signal, as this renders the similarity information at a chance level of being correct. If the individual is unsure of her similarity to the group from whom she learns, then she has no way to choose a frequency-dependent social learning strategy that would help her extract useful information from this group. Indeed, all strategies should be equivalent in terms of expected payoffs (Efferson et al., 2016). Participants may choose the strategy that they prefer without any expectation that this would affect their pay-offs in response to these uninformative signals.

We test whether participants respond to up to three orders of social information when choosing their frequency-dependent social learning strategies during both a game against nature and a coordination game. If frequency-dependent social learning strategies should only be considered over a first-order strategy space, then it would be expected that some individuals always copy the majority (or minority). This would mirror the use of conformity in Boyd and Richerson's (1985) seminal model. Alternatively, finding that individuals can adjust their frequency-dependent social learning strategies to second or even third-order social information would support my intuition that the extent of a frequency-dependent social learning strategy space may be more complex and flexible than has previously been assumed.

As this is the first study to investigate three orders of social information, it is difficult to predict if and how a learner will use this information. Broadly speaking however, the flexibility of the social learning strategies may match one of the following possibilities:

(i) The participants adjust their strategies to both second and third-order social information, performing equally well on all informationally-equivalent trials.

(ii) Individuals adjust to both second and third-order social information but show asymmetric adjustments. The individuals find some social information easier to respond to than others. For example, they may find reliably correct signals easier to process than reliably incorrect signals, as individuals are more likely to encounter reliable signals of group membership throughout their lives (McElreath et al., 2003; Richerson et al., 2016). Finding (i) or (ii) would suggest that the frequency-dependent social learning strategy space is more flexible than previously assumed.

(iii) There may be a trade-off in the complexity of social information that the participants process. Participants may adjust their strategies to second-order but not third-order social information.

(iv) The second and third-order social information are too complex to process. The social learners respond to first-order social information only. Finding (iv) would support the original representation of frequency-dependent social learning strategies over a simple social learning space comprising of only one order of social information, as per Boyd and Richerson's (1985) seminal model.

2. Methods

2.1. Participants

302 participants were recruited at the Centre for Experimental Social Sciences (CESS) partner lab in the Foundation for Liberal and Management Education (FLAME) University in Pune, India (mean age = 20.73, SD = 2.54, males = 88; see appendix 1 for rationale and protocol). An ex-ante power calculation supported this sample size (see pre-registration at OSF | prereg_10Dec2018_bellamyEtAl.docx).

This sample consisted of university students. University students have an adequate ability to learn to play the economic games (Ekuni et al., 2020), though it is prudent to remember that university students may be more rich, democratic, and educated than the rest of the population of India (Henrich, 2020). Moreover, there was no guarantee that the students who attended these sessions were Indian. A small number of participants (*N*=2 in the game against nature) were international students (see supplementary material 1). I therefore settled on English as the main language as it is (assumedly) a common language amongst international university students (Andrade, 2009).

These 302 participants were divided to play one of two roles. Sixty-six participants became demonstrators, who learned to play the tasks via trial-and-error. The remaining 236 participants became social learners, who could only base their decisions on social information. Separating the participants into these roles may seem arbitrary as real populations shift between individual and social learning strategies (Miu et al., 2020). However, this design crucially allowed the experimenter to draw clean causal inferences regarding how social information affected behaviour without the typical inferential challenges (Angrist, 2014; Manski, 2000).

2.2. The games

One-hundred-and-fifty participants took part in what I call a game against nature, which measured asocial skills. Although the participants played in pairs, the focal participant's choice could not affect her partner's pay-offs or vice-versa and so the game was asocial. The optimal option to select in each game version was the one associated with the higher pay-off of 150 points (see Figure 1).

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 150** Partner's expected points: 150	**Your expected points: 150** Partner's expected points: 120
You choose @	**Your expected points: 120** Partner's expected points: 150	**Your expected points: 120** Partner's expected points: 120

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 120** Partner's expected points: 120	**Your expected points: 120** Partner's expected points: 150
You choose @	**Your expected points: 150** Partner's expected points: 120	**Your expected points: 150** Partner's expected points: 150

Figure 1. The payoff matrix shown to participants for the game against nature. Text in **bold** represents the expected payoffs from the focal participant's choices.

One-hundred-and-fifty-two participants took part in the coordination game measuring social norms (see Figure 2). The participants earned more points if they chose the same option as an anonymous partner with whom they had been paired but could not communicate. Conditional on coordinating, one option had a higher pay-off. This game structure gave many possible equilibria (coordinating on the optimal option, coordinating on the sub-optimal option, or a mixed strategy equilibrium which, if adopted, may lead to miscoordination some of the time; Kets et al., 2021). Through repeated periods of play, it was possible for the participant to develop a norm regarding the strategy that she expected others to play. Once such a norm emerges, coordination becomes straightforward and so the structure of the coordination game allowed for norms to emerge.

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 325** Partner's expected points: 325	**Your expected points: 100** Partner's expected points: 100
You choose @	**Your expected points: 100** Partner's expected points: 100	**Your expected points: 250** Partner's expected points: 250

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 250** Partner's expected points: 250	**Your expected points: 100** Partner's expected points: 100
You choose @	**Your expected points: 100** Partner's expected points: 100	**Your expected points: 325** Partner's expected points: 325

Figure 2. The payoff matrix shown to participants for the coordination game. Text in **bold** represents the expected payoffs from the focal participant's choices.

During both economic games, the participants chose between two arbitrary options: @ or %. To allow the option with the highest payoff to change between blocks, I divided both the game against nature and the coordination game into 'Game Left' (where % was optimal) or 'Game Right' (where @ was optimal). I avoided ordered labels to prevent the social learners from converging on an option before any endogenous norms had formed during the coordination game. For example, labelling the options 'A or B' would have allowed participants to follow a rule of 'just choose A'. Arbitrary symbols were used across both games to prevent this.

The points shown in Figures 1 and 2 reflect the expected points that the participants made from their decision. All choices were influenced by a random disturbance drawn from a random normal distribution, with M=0 and SD=20 points. This random disturbance was similar to Molleman and Gächter's (2018) design and was independently drawn for each period and independently drawn for each participant. This disturbance meant that the participants could earn more or less points than expected and, occasionally, choosing the 'suboptimal' option would give the participant more points than choosing the 'optimal' option (McElreath et al., 2005). This disturbance ensured that the learning problem was not trivial.

2.3. Materials

The participants read an instructional booklet specific to either the game against nature or the coordination game, depending on which was being played in their session (see appendix 2). The participants completed the game via Z-Tree version 3.5 (Fischbacher, 2007) at individual PC terminals to ensure anonymity (see appendix 3). The participants answered an end-game survey, with questions regarding demographics and social learning preferences (see supplementary material S1).

2.4. Procedure and design

Participants were tested in groups of 20 - 30. The numbers were kept even, as the games were played in pairs. The participants answered questions to confirm their understanding of the game (see appendix 2). At the start of each session, the computer randomly chose six participants to play as demonstrators (labelled Type A in game). The remaining participants played as social learners (labelled Type B in game).

Each session lasted for 22 blocks of 4 periods. The computer began each block by randomly assigning the participants to play in pairs (see Figures 1 and 2). The pairings were constrained so that the demonstrators only played with other demonstrators and the social learners only played with other social learners. The pairings could change between blocks but not within a block. Finally, the computer randomly selected whether the demonstrators played Game Left or Game Right (and thus decided whether % or @ was optimal for the demonstrators). The optimal option could change between blocks but not within blocks.

The demonstrators played first in each block. They did not know which game version they were playing and thus did not know which option was optimal. However, they could use the immediate feedback from their choices to learn which option was likely to be optimal. The random disturbance in points was crucial for the demonstrators. If the points were always as displayed in Figures 1 and 2, then all demonstrators would answer optimally by the final period of a block. The disturbance in points made the task sufficiently difficult to ensure that some blocks would end with some demonstrators choosing optimally and some sub-optimally.

The social learners only made one choice in the final period of a block. They did not see any feedback from their decisions and could only base their choices on social information. This included: (i) the frequency of demonstrators who chose % and @ in the final period of a block (first-order social information); (ii) a signal informing the learner that they played either the same game version or a different game version to these demonstrators (i.e., the similarity information, or second-order social information) and (iii) whether this similarity information was reliably incorrect, uninformative, or reliably correct (i.e., the reliability information, or third-order social information). Figure 3 depicts a typical round for the social learners.

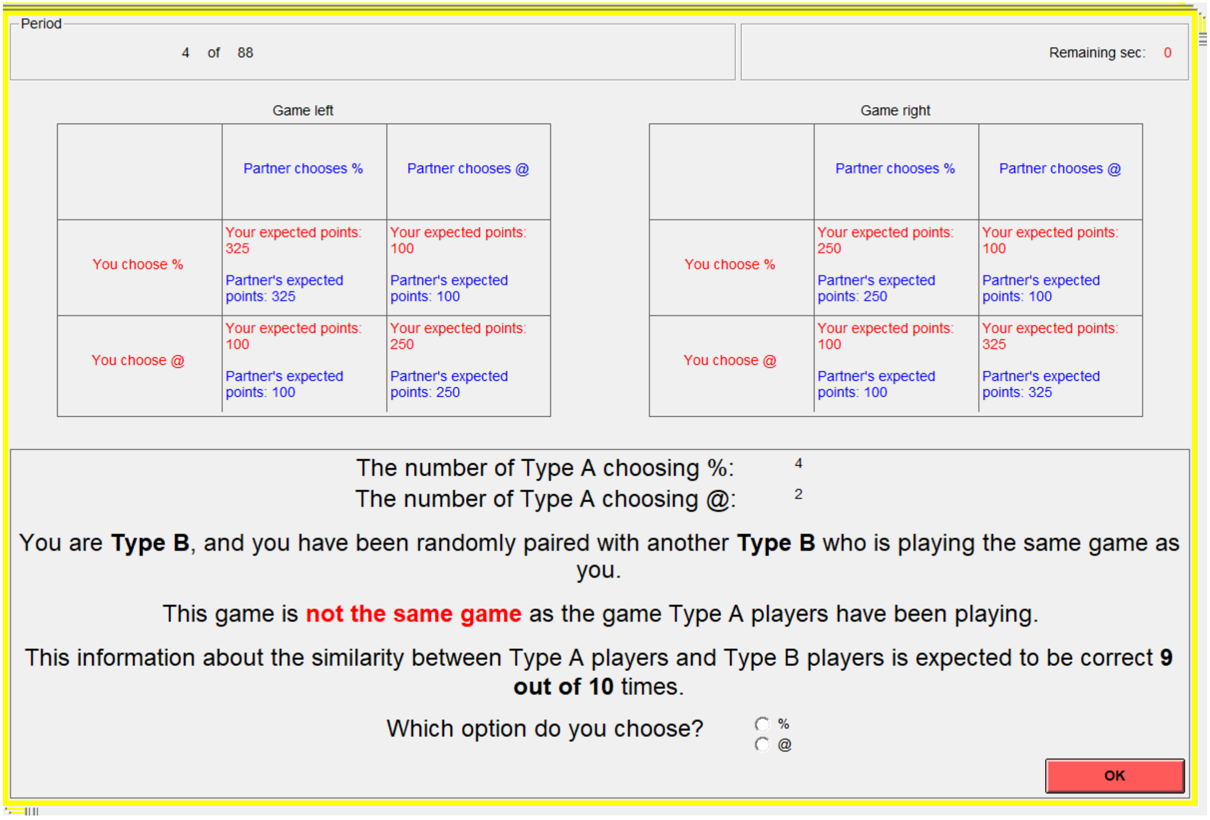


Figure 3. A screenshot of a typical round for the social learners. Note that Type A participants were demonstrators and Type B participants were social learners. I avoided the term 'demonstrator' or 'social learner' in case it led the participants respond to the task in certain ways. The top half of the screen reminds the participants of the expected pay-offs from Game Left and Game Right (for the coordination game in this case). The bottom half of the screen contains frequency-dependent information (i.e., the number of demonstrators who chose @ or %), similarity information (i.e., whether the demonstrators were identified as playing the same or different game version to the social learners) and the reliability information in bold.

This study manipulated both the similarity information (2 levels: same game version or different) and the reliability information (3 levels: reliably incorrect, uninformative, or reliably correct) in a 2X3 within-subjects design. The similarity information conveyed whether the social learners played the same game version as the demonstrators (e.g., both participants played Game Left), or a different game version (e.g., the demonstrators played Game Left, but the social learners played Game

Right). The reliability of the similarity signals was conveyed as probabilities. The similarity signal was correct 1 in 10 times for a reliably incorrect signal, 5 in 10 times for an uninformative signal and 9 in 10 times for a reliably correct signal.

The computer started each block by randomly assigning a third of social learners to play a 'reliably incorrect' round, a third to play an 'uninformative' round and a third to play a 'reliably correct' round. The computer then randomly assigned all social learners to play either Game Left or Game Right. After randomly assigning the demonstrator game, the computer tracked whether the demonstrators and social learners were playing the same game. The similarity signal which was shown to the social learners was based on the probabilities that were implemented as the reliability signal. To illustrate with an example, imagine that the social learners were playing the same game as the demonstrators, and they were in the reliably correct condition (which was correct 9 times in 10). They would see a similarity signal telling them that they were playing the 'same game version as the demonstrators' with a probability of 0.9, or the signal would tell them that they were playing a 'different game version to the demonstrators' with a probability of 0.1. This ensured that the similarity and reliability information were meaningful.

Finally, the participants completed the survey and were paid separately based on the points that they had earned. The participants earned ₹14 per point during the game against nature, or ₹23 per point during the coordination game. Participants were also paid a ₹100 show-up fee. On average the participants made ₹958.53 (SD = ₹24.34) during the game against nature and ₹827.23 (SD = ₹86.81) during a coordination game. Based on the current exchange rate, this meant that the participants earned the equivalent of £9.60 (SD = 24p) on average for the game against nature and £8.95 (SD = 56p) on average for the coordination game. This study was self-certified via the Royal Holloway University of London's School of Psychology ethical criteria (see appendix 4). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

2.5. Analysis

There were three steps to analysis. First, I confirmed that the demonstrators provided varied– but on the whole, accurate– social information (see section 3.1). Second, I investigated whether the social learners adjusted their behaviour based on the three pieces of social information. This analysis showed how *flexible* the social learners were but could not show how *efficient* they were at extracting optimal social information from blocks with informationally-meaningful signals. Therefore, my final analysis focused on whether the social learners chose their social learner optimum as based on the social information (see appendix 5 for scripts).

Our second analysis in Section 3.2 aimed to investigate social learner flexibility, by analysing whether the social learners adjusted their frequency-dependent social learning strategies to each level of the social information. To investigate this, I ran a logistic regression modelling whether the social learners chose %. The predictors in this regression were: (i) the centered proportion of demonstrators who chose %, (ii) dummies representing each combination of the similarity and reliability information and (iii) interactions between these dummies and the centered proportion of demonstrators who chose %. These dummies were: reliably incorrect - different signals; uninformative - similar signals; uninformative - different signals; reliably correct - similar signals and reliably correct - different signals. The omitted category was reliably incorrect - similar signals. I investigated similarity and reliability information as combined dummies to avoid modelling three-way interactions, which can be difficult to interpret in large analyses.

The social learners may have adjusted to the social information but still experienced trade-offs, as they may have found some information easier to respond to than others. To investigate this, the third analysis in section 3.3 ran a logistic regression to model whether the social learners chose their social learner optimum. The predictors of this regression included: (i) the centered proportion of demonstrators who chose their demonstrator optimum, (ii) the dummies representing each combination of the similarity and reliability information (as above) and (iii) interactions between these dummies and the centered proportion of demonstrators who chose the demonstrator optimum.

Finally, certain demographics such as age (Wen et al., 2019) and gender (Mesoudi et a*l.,* 2015) may influence conformity preferences. All regression analyses were repeated with demographics included as control predictors, which can be seen in appendices 6 and 8. Due to a crash during the coordination game, one group of participants did not complete the full session and so did not provide demographics. The models that contain control variables for the coordination game thus only include the data for these participants (see appendices 6 and 8). Removing the crashed data gives N=278 participants overall. All the coordination game models reported in the results section use full datasets, though all analyses are repeated with the crashed sessions removed for transparency purposes in the Supplementary Materials S2.

Note that these methods were pre-registered, including an a-priori sampling plan and power analysis. However, I did not pre-register a specific plan for the analysis as I wished to remain explorative in terms of the social information, and the other participant variables, that I considered. This exploratory approach was justified by the novelty of testing social information up to third-order complexity.

2.6. Predicted social learning strategies

No studies have investigated social learning flexibility to a third-order complexity and so I did not make specific predictions of how flexible the social learners would be. However, it is possible to predict the frequency-dependent strategies that the social learners may have used if they wished to extract optimal information from the informationally-equivalent signals. As the uninformative signals rendered the similarity information at chance level of being correct, then informationally-equivalent trials refer to reliably correct and reliably incorrect signals of similarity and difference only.

Table 1 depicts the social learning strategies that allow the learner to extract optimal information from the informationally-equivalent signals, given that the demonstrators were usually accurate at choosing the demonstrator optimum. If the social learners were likely to be playing the same game version as the demonstrators, then the social learners could copy the majority choice made by the demonstrators as the demonstrator optimum and the social learner optimum was likely to match. Both a reliably correct signal of similarity and a reliably incorrect signal of difference implied that the social learners and the demonstrators had the same optimum, and so the social learners could have followed the majority under both signals. Conversely, a reliably correct signal of difference and a reliably incorrect signal of similarity both implied that the social learners played a different game version to the demonstrators and so had to choose the different option to extract optimal pay-offs. Therefore, social learners may have followed the minority under both reliably correct signals of difference and reliably incorrect signals of similarity.

Table 1. The strategies that would allow the social learners to extract optimal pay-offs when seeing one of the four informationally meaningful trials.

		Third-order social information	
		Reliably incorrect	**Reliably correct**
Second-order social information	**Similar**	Follow the minority	Follow the majority
	Different	Follow the majority	Follow the minority

<u>3. Results</u>

3.1: Did the demonstrators provide varied– yet accurate– social information?

The demonstrators' choices formed the frequency-dependent social information shown to the social learners and so the game did not involve deception. This analysis clarifies that the demonstrators learned to choose the demonstrator optimum by the final period of a block, similar to how individuals master skills and social norms via their own trial-and-error. If the demonstrators show only a trivial preference, then there would be no way for the social learners to be able to answer optimally during these tasks (Efferson et al., 2016).

Table 2 confirms that the demonstrators were more likely to choose the optimal option on the final period of each block than any other period ('finalPeriodDummy': game against nature estimate = 0.378, $p<0.001$; coordination game estimate = 0.325, $p = 0.0004$). This shows that the demonstrators learned via their own trial-and-error across both games. I also include the dummy labelled, '% as optimal' to understand whether the demonstrators simply preferred to choose one arbitrary option. A significant finding here would indicate an arbitrary preference on the part of the demonstrators to choose the % symbol. No such bias was found during the game against nature ('% as optimal': estimate = 0.178, $p = 0.54$) though the demonstrators seemingly had a preference to choose % in the coordination game ('% as optimal': estimate = 1.146, $p<0.001$). Perhaps the demonstrators simply chose % during the coordination game to increase their chances of coordinating with their partner, regardless of the payoffs of coordination.

To investigate this arbitrary preference further, I calculated how many demonstrators would answer optimally by the final period within a block (see appendix 6). Approximately two thirds of demonstrators would answer optimally by the final period of the average block (68.64% for the game against nature and 68.11% for the coordination game). I conducted one-sample t-tests to confirm that these percentages were significantly greater than 50%, as half the demonstrators could have chosen optimally by chance. These confirmed that the demonstrators played the games better than chance (game against nature: ($t(21)=-2619.9$, $p<0.001$); coordination game: $t(21)=-2223.70$, $p<0.001$). This means that the social learners would be able to use the social information from the demonstrators to help them to answer optimally on approximately two-thirds of the blocks, provided that they saw an informationally meaningful (i.e., a reliably correct or a reliably incorrect) signal. Thus, the demonstrators provided varied– but on the whole, accurate– social information.

Table 2. A logistic regression modelling whether the demonstrators chose their demonstrator optimum. Includes a dummy for the final period of the blocks and % as optimal as predictors. Model 1 displays the data for the game against nature (asocial skills) and Model 2 displays the data for the coordination game (social norms). Robust standard errors in parentheses are clustered on the demonstrator. I also include the lower and upper limit 95% confidence interval for each estimate below this standard error.

Parameter	**Estimate (Game against nature)**	**Estimate (Coordination game)**
Intercept	0.321 *	-0.130
	(0.161)	(0.154)
	95% CI [0.01, 0.64]	95% CI [-0.43, 0.17]
Final Period dummy	0.378 ***	0.325 ***
	(0.091)	(0.092)
	95% CI [0.20, 0.56]	95% CI [0.15, 0.51]
% as optimal	0.178	1.146 ***
	(0.290)	(0.260)
	95% CI [-0.39, 0.75]	95% CI [0.64, 1.66]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

3.2. Did the social learners flexibly adjust their strategies to each level of the social information presented?

To visualise whether the social learners' strategies match those predicted in Table 1, I plot the proportion of social learners who chose % as a function of the number of demonstrators who chose %. I plot this for each combination of the similarity and reliability information in Figure 4. This allows me to compare the frequency-dependent social learning strategies that the social learners used across each level of the similarity and reliability information.

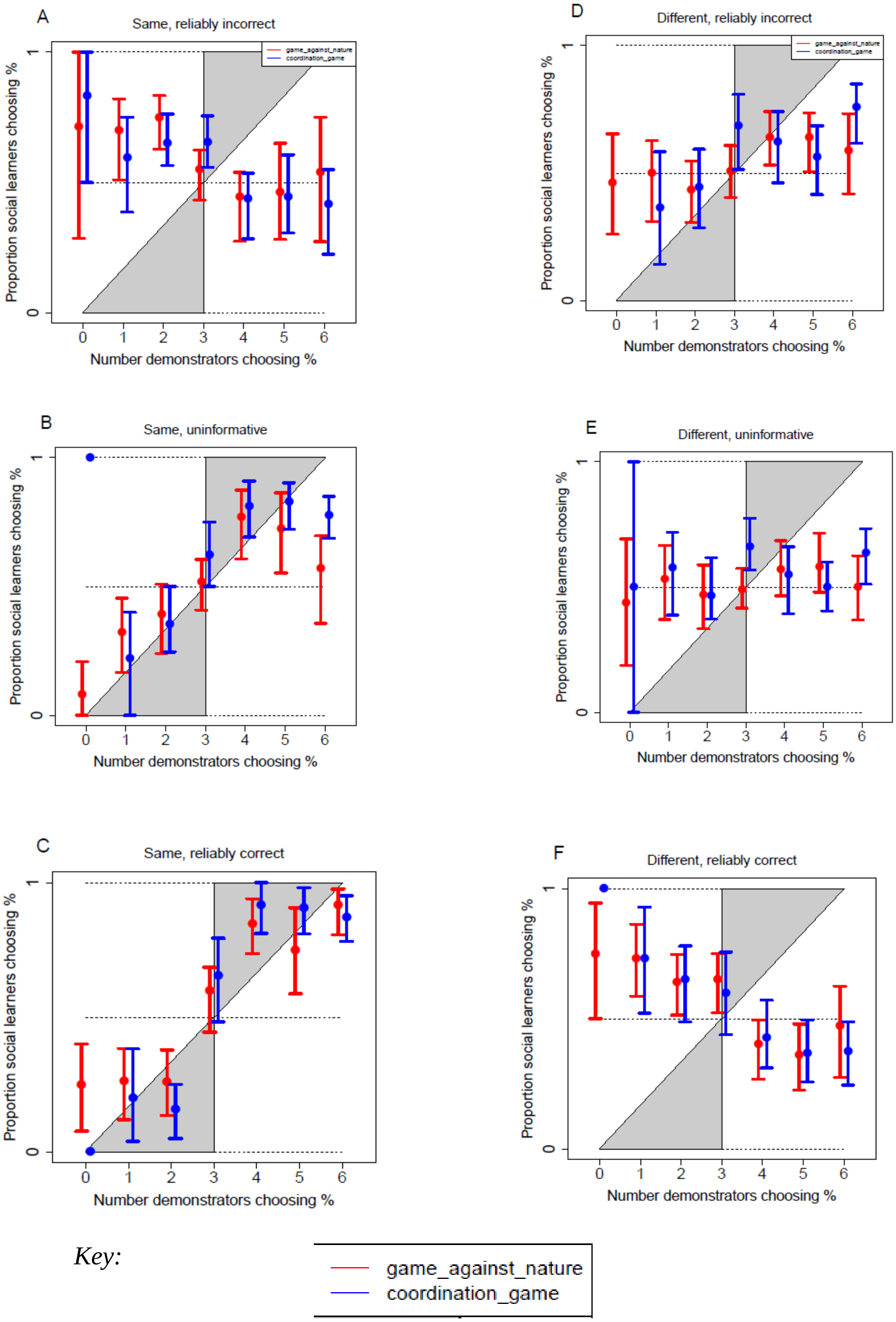


Figure 4. The proportion of social learners who chose % based on the number of demonstrators who chose %. The different panels show the social learners' responses

to frequency-dependent social information by each level of the second- and third-order social information, for both the game against nature (learning skills, in red) and the coordination game (learning social norms, in blue). The error bars give the 95% bootstrapped confidence interval clustered on social learners, to reflect the multiple observations gathered per learner. The regions shaded in grey depict where the social learners' data would fall if they used a conformist strategy, whilst the dashed lines give points of reference for proportions of learners choosing % at 0, 0.5, and 1.

For clarity going forward, I refer to groups of demonstrators who share a similar decision-making environment to the social learners as cases where the social learners learned from 'similar' others. Cases where the demonstrators made decisions in a different decision-making environment to the social learners are cases where the social learners learned from 'different' others. The results depicted in Figure 4 show that the social learners typically used frequency-dependent social learning strategies, as they changed their preference for choosing % based on whether the minority or the majority of demonstrators had chosen %. The response to frequency information was less pronounced under reliably incorrect signals from different others (figure 4D) however, and the response to uninformative signals from different others (figure 4E) was random and may not respond to frequency information.

To understand how the social learners responded to second-order (i.e., similarity) information, one need only compare the figures in the same row of Figure 4. Take reliably correct signals in the bottom row, for example. On average, the social learners copied the majority of reliably similar others (Figure 4C) though instead followed the minority of reliably different others (Figure 4F) for both games. These strategies matched those given in Table 1, meaning that the social learners responded to frequency-dependent social information based on whether the group from whom they learned were identified as sharing a similar or different decision-making environment to themselves.

To understand how the social learners responded to third-order information (i.e., the reliability of the similarity signals), one need only compare the figures within the same column of Figure 4. Take the social learners' responses to similar others in the left-hand column, for example. On average, the social learners copied the majority of similar others with an uninformative (Figure 4B) or a reliably correct signal (Figure 4C). The learners instead followed the minority choice among similar demonstrators if they saw a reliably incorrect signal (Figure 4A). These strategies matched those in Table 1, showing that the social learners changed their response to frequency-dependent information as based on their perceived similarity of their decision-making environments to a group, and the reliability of this signal. The logistic regression in Table 3 further investigates how the social learners adjusted their strategies to each level of the social information.

Table 3. The logistic regressions modelling whether social learners chose %. Predictors included (i) the centred number of demonstrators who chose % on their final period, (ii) each combination of the similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, and (iii) the interactions between each of these dummies and the centered proportion of demonstrators who chose %. I centered the proportion of demonstrators choosing % so any block where 3/6 demonstrators chose % became the omitted category of the regression, and thus were reflected in the intercept. The robust standard errors given in parentheses were clustered on the social learner to reflect the multiple observations gathered per learner. See appendix 7 for the regressions with control predictors. As the only significant control predictor was an increased likelihood to choose % as the blocks progressed during the game against nature only, then the models reported below just

focus on the social information of interest. I also give the 95% confidence interval lower and upper bounds for each estimate.

Parameter	Estimate (game against nature)	Estimate (coordination game)
Intercept	0.337 ** (0.12) 95% CI [0.10, 0.57]	0.308 * (0.120) 95% CI [0.07, 0.54]
Centred proportion of demonstrators choosing %	-1.493 * (0.581) 95% CI [-2.63, --0.36]	-1.400 ** (0.492) 95% CI [-2.36, -0.44]
Reliably incorrect- different dummy [signal indicates different and is correct with 0.1 probability]	-0.177 (0.145) 95% CI [-0.46, -0.11]	-0.112 (0.175) 95% CI[-0.45, 0.23]
Uninformative - same dummy [signal indicates same and is correct with 0.5 probability]	-0.251 . (0.143) 95% CI [-0.53, -0.03]	-0.007 (0.154) 95% CI [-0.31, 0.29]
Uninformative - different dummy [signal indicates different and is correct with 0.5 probability]	-0.274 . (0.158) 95% CI [-0.58, -0.03]	-0.065 (0.154) 95% CI [-0.37, 0.24]
Reliably correct - same dummy [signal indicates same and is correct with 0.9 probability]	0.006 (0.152) 95% CI [-0.29, 0.30]	-0.089 (0.197) 95% CI [-0.47, -0.30]
Reliably correct - different dummy [signal indicates different and is correct with 0.9 probability]	-0.034 (0.137) 95% CI [-0.30, 0.14]	-0.025 (0.168) 95% CI [-0.35, 0.30]
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	2.371 ** (0.750) 95% CI [0.90, 3.84]	2.911 *** (0.751) 95% CI [1.44, 4.38]
Centred proportion of demonstrators choosing % X uninformative-same dummy	4.161 *** (0.802) 95% CI [2.59, 5.73]	4.242 *** (0.752) 95% CI [2.77, 5.71]

Centred proportion of demonstrators choosing % X uninformative-different dummy	1.785 * (0.695) 95% CI [0.43, 3.14]	1.630 * (0.677) 95% CI [0.31, 2.95]
Centred proportion of demonstrators choosing % X reliably correct-same dummy	5.714 *** (0.992) 95% CI [3.77, 7.65]	6.700 *** (0.989) 95% CI [4.77, 8.63]
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-0.438 (0.640) 95% CI [-1.69, -0.81]	-0.576 (0.642) 95% CI [-1.83, 0.68]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· (trend: p=0.05 – 0.10 significance)

The regression in Table 3 shows remarkable consistency between the game against nature and the coordination game. The social learners became less likely to choose % as the centered proportion of demonstrators choosing % increased across both games (game against nature: effect = -1.49, z = -2.57, p=0.01; coordination game: effect = -1.40, z = -2.85, p =0.004). This reflects the estimate for the omitted category of reliably incorrect-similar signals. That is, the learners followed the minority choice made by the demonstrators on blocks with a reliably incorrect signal of similarity. This makes sense, as demonstrators who were unlikely to be similar to oneself were in fact likely to be different. The social learners also followed the minority around reliably different others (Figure 4F), suggesting they understood that these signals were informationally equivalent.

The significant effects in Table 3 show that the social learners became more likely to choose % as more demonstrators did under the following blocks: reliably incorrect-different; uninformative- similar; uninformative- different and reliably correct-similar. The social learners clearly followed the majority on blocks with uninformative and reliably correct signals from similar others (Figure 4B and 4C). There was a slight trend to follow the majority when responding to reliably incorrect signals from different others (Figure 4D). A reliably correct signal of similarity and a reliably incorrect signal of difference both implied that the demonstrators had played the same game version as the social learners. Learners followed the majority on both these blocks, which shows a complex processing of third-order information.

However, the social learners' responded at chance for an uninformative signal from different others (see Figure 4E). Note that the logistic regression merely confirms that this response is significantly distinct to the omitted category of reliably incorrect signals from similar others. The social learners response to reliably incorrect signals from similar others was meaningful (figure 4A) though their response to the uninformative signal from different others was not (figure 4E).

We run linear combinations to test whether the social learners employed a significantly distinct strategy in response to each level of the similarity and reliability information. These linear combinations focused on cases when either the majority of demonstrators chose % ($\geq$4) or none did. These highlighted the differences between flexible social learning strategies, and those that follow simple learning rules. Flexible learners would adjust to the similarity and reliability information, whilst learners who followed a rule to 'always copy the majority' would choose % whenever the majority of demonstrators had. These linear combinations revealed that the social learners employed a significantly distinct strategy to each level of social information. The only exception is that the social learners did not adjust to reliably incorrect signals versus uninformative signals from different others for the game against nature only (see appendix 8).

To summarise, the social learners adjusted somewhat to each level of the social information. Section 3.3 focuses on whether these adjustments are complete and symmetric to all informationally-equivalent blocks.

3.3. Did the social learners choose their social learner optimum?

Social learners who adjusted symmetrically to the social information available would extract their social learner optimum on any block with an informationally-meaningful signal. This excluded blocks with uninformative signals, as these rendered the similarity information at chance level of being correct. The regression in table 4 analyses whether the social learners adjusted symmetrically to all informationally-equivalent blocks when playing both a game against nature and a coordination game.

Table 4. Logistic regressions modelling whether the social learners chose the social learner optimum. Predictors included: (i) the centered proportion of demonstrators who chose the demonstrator optimum, (ii) dummies for each combination of similarity and reliability information, minus the omitted category of reliably incorrect-similar signals, and (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose the demonstrator optimum. Robust standard error clustered on social learner. See appendix 9 for the regressions with control predictors, though the only significant control predictor was that the social learners were more likely to answer optimally on blocks where the % symbol was optimal, suggesting an arbitrary preference to choose this symbol across both games. I also give the 95% confidence interval lower and upper bounds for each estimate.

Parameter	Estimate (game against nature, all signals)	Estimate (coordination game, all signals, full data)
Intercept	0.210 (0.148) 95% CI [-0.08, 0.50]	0.047 (0.109) 95% CI [-0.17, 0.26]
Centred proportion of demonstrators choosing the demonstrator optimum	0.166 (0.642) 95% CI [-1.09, 1.42]	0.613 (0.462) 95% CI [-0.29, 1.52]
Reliably incorrect-different dummy [indicates different and is correct with 0.1 probability]	-0.136 (0.184) 95% CI [-0.50, 0.22]	0.184 (0.165) 95% CI [-0.14, 0.51]
Uninformative-same dummy [indicates same and is correct with 0.5 probability]	-0.193 (0.201) 95% CI [-0.59, 0.20]	-0.026 (0.159) 95% CI [-0.34, 0.29]
Uninformative-different dummy [indicates different and is correct with 0.5 probability]	-0.131 (0.193) 95% CI [-0.51, -0.25]	-0.136 (0.160) 95% CI [-0.45, 0.18]
Reliably correct-same dummy [indicates same and is correct with 0.9 probability]	-0.278 (0.222) 95% CI [-0.71, -0.16]	0.254 (0.204) 95% CI [-0.14, 0.65]
Reliably correct-different dummy [indicates different and is correct with 0.9 probability]	-0.048 (0.192) 95% CI [-0.42, -0.33]	-0.017 (0.168) 95% CI [-0.35, 0.31]
Centred proportion of demonstrators choosing optimum X reliably incorrect-different dummy	0.374 (0.815) 95% CI [-1.22, 1.97]	0.786 (0.622) 95% CI [-0.43, 2.00]
Centred proportion of demonstrators choosing optimum X uninformative-same dummy	-0.787 (0.819) 95% CI [-2.39, 0.81]	-1.197 * (0.595) 95% CI [-2.36, -0.03]

Centred proportion of demonstrators choosing optimum X uninformative-different dummy	-0.901 (0.805) 95% CI [-2.48, 0.67]	-0.023 (0.616) 95% CI [-1.23, 1.18]
Centred proportion of demonstrators choosing optimum X reliably correct-same dummy	2.53 ** (0.911) 95% CI [0.75, 4.31]	1.460 . (0.770) 95% CI [-0.05, 2.96]
Centred proportion of demonstrators choosing optimum X reliably correct-different dummy	0.724 (0.785) 95% CI [-0.81, 2.26]	0.573 (0.605) 95% CI [-0.61, 1.76]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

Beginning with the game against nature, the social learners were significantly more likely to choose their social learner optimum as more demonstrators chose the demonstrator optimum for a reliably correct- similar dummy (effect = 2.53, z = 2.78, p=0.005). The social learners conformed to the majority of similar others with a reliably correct signal (Figure 4C). This strategy allowed them to choose the social learner optimum.

We further confirmed this preference to learn from reliably similar others with linear combinations (see appendix 10). The social learners were more likely to respond optimally when viewing reliably correct signals rather than uninformative signals of similarity, though this is expected as the social learners could only respond optimally to the uninformative signals by chance. Interestingly, the social learners were more likely to answer optimally when viewing a reliably correct signal of similarity than they were when viewing a reliably incorrect signal of similarity, or a reliably correct signal of difference. This suggests a bias in processing information, as these signals were informationally equivalent. This implies a trade-off when the social learners processed complex social information. The social learners were more likely to master an asocial skill when learning from similar others with reliably correct signals.

For the coordination game, the only significant effect was that the social learners were less likely to coordinate on the optimal option when more demonstrators coordinated on their demonstrator optimum for blocks with uninformative-similar signals (effect = -1.20, z=-2.01, p=0.04). The uninformative condition provided a baseline trial, in which the social learners could just choose the social learning strategy that they preferred without any expectation that this could affect their pay-offs. Despite this, the social learners seemingly employed a strategy of ‘copy similar others even if the blocks give uninformative information’ (Figure 4B). This perhaps helped the social learners to coordinate though they were more likely to coordinate on the sub-optimal option.

We then used linear combinations to investigate if there were any differences in the social learners’ ability to answer optimally to each level of social information during the coordination game (see appendix 11). The social learners were more likely to respond optimally to reliably correct signals from similar others than reliably correct signals from different others. The social learners were also more likely to respond optimally to reliably correct signals from similar others than to uninformative or reliably incorrect signals from similar others. The bias to learn from similar others with reliably correct signals was not as pronounced during the coordination game as it was during the game against nature, however.

A final caveat to consider is that the Z-tree code worked with the probabilities given in the economic games. This refers to the reliability of the similarity signal: 0.9

for reliably correct signals; 0.5 for uninformative signals and 0.1 for reliably incorrect signals. Working with these probabilities meant that the blocks sometimes displayed incorrect similarity information. For example, a reliably incorrect signal informing the participant that she played the same game version as the demonstrators would correctly depict that the social learner was playing the same game version as the demonstrators with a probability of 0.1. However, it was more likely that she played the different game version to the demonstrators with a probability of 0.9. In this case, the signal suggesting that the social learners and the demonstrators played the 'same' game type would be incorrect. This caveat could have affected the social learners' ability to choose their social learner optimum.

We analysed how the social learners responded to the similarity signals that happened to be correct compared to those signals that happened to be incorrect for both the game against nature (see appendix 12) and the coordination game (see appendix 13). This analysis broadly revealed that the social learners responded to a reliably incorrect signal as if it always provided wrong similarity information, and that they responded to a reliably correct signal as if it always provided correct similarity information. This shows a complex understanding of the informationally-equivalent reliability signals on the part of the social learner. The social learners also seemed to treat an uninformative signal of similarity as if it were always correct during a coordination game, perhaps as this strategy made coordination easier.

To summarise, the social learners may have occasionally coordinated on the sub-optimal option during the coordination game, but they were usually flexible at adjusting to social information when acquiring a social norm. The bias to learn from reliably similar others was perhaps less pronounced on the coordination game as it was in the game against nature. Otherwise, there was a remarkable consistency between how the learners acquired both an asocial skill and a social norm.

3.4. Summary of the social learning strategies

The social learners showed a complex adjustment when choosing their frequency-dependent social learning strategies. They chose these strategies based on (i) frequency information, (ii) signals cuing whether the group played a similar or different game to the participant and (iii) the reliability of similarity signals. However, there was a trade-off in the extent to which learners can respond to this information. The social learners were more likely to respond optimally when learning from similar others with reliably correct signals.

4. Discussion

This study investigated how flexible social learners were when choosing between different frequency-dependent social learning strategies. I found that individuals adjusted to first-order social information (frequency-dependent information), second-order social information (signals indicating whether the learner shared a similar or different decision-making environment to the group from whom she learns) and third-order social information (the reliability of similarity signals). This suggests that the boundaries of a frequency-dependent social learning strategy space are more complex than simply responding to first-order information. These results are consistent with a growing body of literature suggesting that social learning strategies are employed flexibly across different scenarios (Deffner et al., 2020; Kendal et al., 2018; Rendell et al., 2011).

The social learners responded differently to frequency-dependent information based on whether the group from whom they learned were similar or different to themselves. Groups of (reliably) similar others were more likely to be copied. Perhaps a high rate of intergroup contact (Boyd & Richerson, 1985 [chapter 7]; Deffner et al., 2020; Efferson et al., 2008b) may have created enough exposure to different others for social learners to be able to account for this second-order social information and avoid copying (reliably) different others.

This study characterised the similarity information as a signal informing the participants that they learned from a group who made decisions in a ‘similar’ or ‘different’ environment to themselves. Previous literature focuses on whether the participants *look* similar to others on observable traits (House et al., 2013; Jiménez & Mesoudi, 2019; Molleman et al., 2019a; Salali et al., 2015; Shutts et al., 2010). Individuals cue their cultural group memberships via ethnic markers (Efferson et al., 2008b; 2016). It makes sense to learn from those with similar markers as they are likely to share similar norms and skillsets as oneself (Richerson et al., 2016; Fischer, 2009; Wood et al., 2013). By directly telling the participants that they were similar or different to the group, I remove the need for the participant to calculate their perceived similarity to others on observable traits, which may be difficult to test explicitly.

Future work should explore whether the preference to learn from (reliably) similar others in the current study is heightened for those that *look* similar to ourselves, perhaps due to ethnocentrism (Hales & Edmonds, 2019) or out-group prejudices (Efferson et al., 2008b; Konrad & Morath, 2012; Vogt et al., 2013). To achieve this, a design similar to Efferson et al’s. (2008b) can be used, where the participants can represent themselves with on-screen avatars. Sometimes, these avatars are uninformative; and sometimes, they become linked to the game being played in such a way that they become a reliably correct or reliably incorrect signal of who is playing the same game as the participant. This could highlight how in-group preferences occur in a group where ethnic markers and social norms emerge endogenously, and can highlight how reliably incorrect (i.e., easy to fake) markers may affect the rate of outgroup prejudices.

The social learners responded to third-order complexity, though they were not infinitely flexible in their adjustments. The learners were more likely to master an asocial skill and a social norm when learning from reliably similar others. These results seemingly support prediction (ii) in the introduction: learners' chosen frequency-dependent social learning strategies were based on three orders of social information, but they adjusted to these three orders asymmetrically.

The social learners followed the minority of similar others on trials with reliably incorrect signals. This suggests that observable cues of group affiliation have not always reliably signalled who shares similar or different behaviours to oneself. To illustrate with an example, some individuals may fake signals of group membership in order to gain access to another group for some benefit (Sosis et al., 2007). Additionally, some individuals signal their cultural identities subtly or hide them entirely (Smaldino et al., 2018). If such scenarios have been common, then this may have created a need for individuals to process the reliability of any cues suggesting similarity.

While the social learners adjusted to the reliability signal given alongside similar others, they did not adjust to the reliability signal given alongside different others. While it may make sense for outsiders to pretend to be similar to others to exploit some kind of perceived benefit of belonging to another social group (Sosis et al., 2007), it is difficult to envision a case where an individual would pretend to be different to someone else. After all, failing to coordinate one's actions to the rest of the group may have devastating social consequences (Chudek & Henrich, 2011; Molleman et al., 2019b).

There would not necessarily be a realistic analogue to the reliably incorrect signal of difference in my study. This is not a weakness of the design as such, but an interesting test of cognition more generally. If human cognition is best viewed as an infinitely flexible domain-general processing system (Bolhuis et al., 2011; Shenhav & Greene, 2010) then I would expect the social learners to be able to respond to all levels of the reliability signal, even if unusual. The fact that the social learners did not show a strong response to reliably incorrect signals from different others instead suggests that social learner cognition is only as flexible as far as there is likely to be a realistic counterpart to these in-game signals.

The origins of this asymmetric adjustment may be genetic or cultural in nature. For example, a genetic bias to conform around reliably similar others may have emerged as we were more likely to encounter and learn from our reliably similar in-group members throughout the ancestral past (Henrich & Boyd, 1998; Mercier & Morin, 2019; Molleman et al., 2014). This may have led to the bias to follow the majority around (reliably) similar others only in the current study. Alternatively, individuals can socially learn how to learn from others (Kendal et al., 2018; Heyes, 2016; Mesoudi et al., 2016). The social learners may respond to both reliably correct and reliably incorrect signals from similar others as they have encountered both cases enough times in their own lives to have learned a response to these signals. Perhaps reliably incorrect signals of difference are rarely encountered and so the social learner has no framework when responding to this signal.

There has been a tendency in gene-culture coevolutionary research to assume that biased and rigid behaviour is evidence of a genetic bias while flexible behaviour is socially learned or influenced by cultural input (Barrett, 2015 [chapter 9]). I made no such assumption that a genetic propensity to conform must stay rigid across one's entire life, or that culturally learned social rules are always flexible– indeed, they may result in social learning behaviour that appears biased. Future research should investigate the likely origins of this asymmetric adjustment favouring reliably similar others, but my study cannot highlight the genetic and/or cultural origins of this adjustment. Instead, the novel contribution of this study was that social learners were found to adjust their frequency-dependent social learning strategies to a third order social information space.

Regardless of the origin, this study found that social learners made more money when learning from similar others with reliably correct signals than they did when learning from different others with reliably correct or incorrect signals. There is an outstanding question as to whether all participants experienced this asymmetric trade-off, or whether some participants adjusted to third-order social information completely while other participants merely followed a rule to 'always copy the majority (or minority)'. This would have cancelled out to an asymmetric adjustment at the aggregate level. To investigate this caveat, I build scatterplots and heatmaps to highlight the likelihood of each participant to follow the majority on periods with 'similar' signals and to follow the majority on periods with 'different' signals. I create these for both reliably correct and reliably incorrect signals, for both the game against nature and the coordination game (see appendices 14-15). This further analysis highlighted that most individuals varied their strategies based on the reliability of the signal being shown. Thus, most of my participants did adjust their frequency-dependent social learning strategies to a third-order complexity.

However, very few participants did so optimally. For reliably correct signals, the expected optimal strategy would be if the participants always followed the majority of similar others, but never followed the majority of different others. I found that some participants did this (see appendix 15). I also found that some participants did the exact opposite of this, and instead chose to follow the majority of different others and to never follow the majority of similar others. On top of this, I also found some strategies that fell between the two extremes. These individuals who partially adjusted to the signals were not necessarily the same participants who reported that they ignored the social information in the end-survey (see supplementary materials). This asymmetric adjustment was thus a meaningful– if suboptimal– strategy.

This suboptimal strategy is not just a caveat to my study. McElreath et al. (2008) and Efferson et al. (2008a) find that a subset of participants will avoid conforming even when it would be suboptimal to do so, while Goeree and Yariv (2015) find that a subset of participants will conform even when conformity is suboptimal. Some of my participants routinely committed to a suboptimal social learning strategy despite this trade-off in processing social information having real payoff consequences as the participants' pay was based on points. Future research should investigate if this preference exists outside of laboratory studies, as people who are less responsive to frequency-dependent social information and do not use it optimally are likely to have profound effects on the cultural evolutionary outcomes of information spread in their group.

The fact that some social learners could adjust to third-order social information completely and others showed an asymmetric adjustment may be evidence of different cognitive strategies in the participants. A dual system approach to social learner cognition would enable the individual to both follow simple biases and to show flexible adjustments to social information (Heyes, 2016). System 1 processing is fast and effortless, but likely to be driven by rule-of-thumbs. This may be consistent with how conformity has sometimes been modelled (Boyd & Richerson, 1985 [chapter 7]; Henrich, 2004; Henrich & Boyd, 2001) and may explain the conformity around reliably similar others in the current study. System 2 processing is instead slow and effortful. This may be consistent with a growing body of literature finding that social learning strategies flexibly guide who, what and when to copy (Efferson et al., 2016; Kendal et al., 2018; Rendell et al., 2011; Wood et al., 2013). System 2 processing may underlie the complex adjustments to both second and third-order social information in response to all informationally-equivalent signals that some of my participants achieved. As System 2 processing is difficult to engage, then not all of the participants could adjust to third-order social information completely. Indeed, a failure to update beyond simplistic System 1 biases in complex learning environments may explain why individually maladaptive behaviour, such as mob behaviour, can be upheld at the group level (Kendal et al., 2018; Mesoudi, 2009; Boyd & Richerson, 2007).

Thus far, the asymmetric adjustment found in this study has been discussed in relation to social learner cognition though this could apply to human cognition more broadly. For example, I argued that the lack of response to reliably incorrect signals from different others may be due to the fact that the social learners were unlikely to have encountered others who pretended to be different to themselves. Equally, the participants may struggle with this framing as it is a double-negative (Cutmore et al., 2015; Johnson-Laird & Tridgell, 1972). Participants should copy (reliably) similar others to answer optimally, though they must choose the other option to groups of (reliably) different others to answer optimally. For a reliably incorrect signal of difference, the social learner must flip the signal again to understand that someone who is unlikely to be playing a different game to oneself is in fact likely to be playing the same game. Thus, the social learners should copy the majority under any reliably incorrect signals of difference. Social learning may be just one task required of a cognitive system which cannot process infinite information (Heyes, 2016; Krafft et al., 2021; Mesoudi, 2011).

For the sake of this paper, I treat reliably correct signals and reliably incorrect signals as informationally equivalent. Strictly speaking, this would only be true on tasks with two options. Minority-based social learning strategies are less efficient on a task with multiple equilibria, as they can only rule out one inefficient behaviour at a time. To illustrate with an example, imagine a hunter who does not know whether to use a net, a spear, or a club to hunt the local game. She observes a group of fishermen using the net. This rules out the net– clearly, that was designed for fishing. However, she is still none the wiser as to whether the spear or the club can be used to hunt. As mastering skillsets or coordinating on social norms is likely to involve multiple equilibria (Mesoudi et al., 2015) then this paper was perhaps limited in scope. As my study had the broader aim of investigating frequency-dependent social learning strategies in response to three orders of social information, I kept the design simple so that any departures from perfect judgement could be easily visualised.

This study's results support a complex and flexible social learning strategy space for frequency-dependent social learning strategies which can be based on up to three orders of social information. The use of frequency-dependent social learning strategies to a third-order complexity should be further confirmed with agent-based models. Modelling is important to check theoretical assumptions (Muthukrishna & Henrich, 2019). Further, empirical studies should investigate whether a similar flexibility exists for other types of social learning strategies. For example, whether an individual's preference to copy prestigious others (Chudek et al., 2012), or popular or socially-dominant others (Flynn & Whiten, 2012), is also influenced by similarity and reliability information.

In summary, the social learners flexibly adjusted their frequency-dependent social learning strategies to (i) the number of others who had chosen certain behaviours in a group, (ii) whether the members of this group were identified as learning in a similar or different environment to oneself and (iii) the reliability of this similarity information. The social learners showed asymmetric adjustments as they were more likely to master a skill or social norm from reliably similar others. Taken together, these results suggest that frequency-based social learning strategies are more complex than simply accounting for the frequency of a behaviour amongst a group. The boundaries of a frequency-dependent social learning space should be extended, to consider that social learners flexibly base their chosen strategies on second and third-order social information. The social learners responses to third-order social information were not fully flexible in this study, however. There is likely to be an extent to the breadth of social information that one can process when choosing between frequency-dependent social learning strategies.

Data availability

The data collected for this study, alongside full materials, appendices and supporting materials, can be found at OSF | What is a social-learning strategy, anyway?

5. References

Acerbi, A., & Tehrani, J. J. (2018). Did Einstein Really Say that? Testing Content Versus Context in the Cultural Selection of Quotations, *Journal of Cognition and Culture*, *18*(3-4), 293-311. doi: https://doi.org/10.1163/15685373-12340032.

Andrade, M. S. (2009). The effects of English language proficiency on adjustment to university life. *International Multilingual Research Journal*, *3*(1), 16-34. DOI: 10.1080/19313150802668249.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98-108. https://doi.org/10.1016/j.labeco.2014.05.008.

Barrett, H.C. (2015). Culture. In H.C.Barrett (Ed.) *The shape of thought: How mental adpatations evolve,* (p 224-243). New York: Oxford University Press.

Berger, J., & Heath, C. (2007). Where consumers diverge from others: Identity signaling and product domains. *Journal of consumer research*, *34*(2), 121-134. https://doi.org/10.1086/519142.

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biology*, *9*(7), e1001109. https://doi.org/10.1371/journal.pbio.1001109.

Boyd, R., & Richerson, P. J. (1985). Frequency-dependent Bias and the Evolution of Cooperation. In R. Boyd & P.J. Richerson (Eds.) *Culture and the Evolutionary Process,* (p 204-241). Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (2007). Cultural adaptation and maladaptation: Of Kayaks and Commissars. In S.W. Gangestad & J.A. Simpson (Eds.) *The evolution of the mind: Fundamental questions and controversies* (p327-332). New York: Guilford Press.

Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General, 149*(4), 746–756. https://doi.org/10.1037/xge0000673.

Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: Bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, *33*(1), 46-56. https://doi.org/10.1016/j.evolhumbehav.2011.05.005.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*(5), 218-226. https://doi.org/10.1016/j.tics.2011.03.003.

Cutmore, T. R.H., Halford, G. S., Wang, Y., Ramm, B. J., Spokes, T., & Shum, D. H.K. (2015). Neural correlates of deductive reasoning: An ERP study with the Wason Selection Task. *International Journal of Psychophysiology*, *98*(3), 381-388. https://doi.org/10.1016/j.ijpsycho.2015.07.004.

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734.

Efferson, C., Lalive, R., Cacault, M. P., & Kistler, D. (2016). The evolution of facultative conformity based on similarity. *PloS one*, *11*(12), e0168551. https://doi.org/10.1371/journal.pone.0168551.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008a). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior, 29*(1), 56-64. https://doi.org/10.1016/j.evolhumbehav.2007.08.003.

Efferson, C., Lalive, R., & Fehr, E. (2008b). The coevolution of cultural groups and ingroup favoritism. *Science, 321*(5897), 1844-1849. DOI: 10.1126/science.1155805.

Ekuni, R., de Souza, B. M. N., Agarwal, P. K., & Pompeia, S. (2020). A conceptual replication of survey research on study strategies in a diverse, non-WEIRD student population. *Scholarship of Teaching and Learning in Psychology.* Advance online publication. https://doi.org/10.1037/stl0000191.

Evans, C. L., Laland, K. N., Carpenter, M., & Kendal, R. L. (2018). Selective copying of the majority suggests children are broadly “optimal-” rather than “over-” imitators. *Developmental science, 21*(5), e12637. https://doi.org/10.1111/desc.12637.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10,* 171–178. https://doi.org/10.1007/s10683-006-9159-4.

Fischer, I. (2009). Friend or foe: Subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General, 138*(3), 341–350. https://doi.org/10.1037/a0016073.

Flynn, E., & Whiten, A. (2012). Experimental “microcultures” in young children: Identifying biographic, cognitive, and social predictors of information transmission. *Child Development, 83*(3), 911-925. DOI: 10.1111/j.1467-8624.2012.01747.x.

Goeree, J. K., & Yariv, L. (2015). Conformity in the lab. *Journal of the Economic Science Association*, *1*(1), 15-28. https://doi.org/10.1007/s40881-015-0001-7.

Hales, D., & Edmonds, B. (2019). Intragenerational cultural evolution and ethnocentrism. *Journal of Conflict Resolution*, *63*(5), 1283-1309. DOI: 10.1177/0022002718780481

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behaviour and Organisation*, *53*(1), 3– 35. doi:10.1016/S0167-2681(03)00094-5.

Henrich, J. (2015). *The secret of our success*. USA: Princeton University Press.

Henrich, J. (2020). *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous.* New York: Farrar, Straus and Giroux.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and human behavior*, *19*(4), 215-241. https://doi.org/10.1016/S1090-5138(98)00018-X.

Henrich, J, & Boyd R. (2001). Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, *208*(1), 79– 89. https://doi.org/10.1006/jtbi.2000.2202.

Henrich, J., Chudek, M., & Boyd, R. (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1683), 20150013. https://doi.org/10.1098/rstb.2015.0013.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. doi:10.1017/S0140525X0999152X.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes, C. (2016). Who knows? Metacognitive social learning strategies. *Trends in cognitive sciences*, *20*(3), 204-213. https://doi.org/10.1016/j.tics.2015.12.007.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*(4), 15-41. https://doi.org/10.1080/00208825.1980.11656300.

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B.S., McElreath, R. & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, *110*(36), 14586-14591. https://doi.org/10.1073/pnas.1221217110.

Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(20), 1-12. https://doi.org/10.1057/s41599-019-0228-7.

Johnson-Laird, P. N., & Tridgell, J. M. (1972). When negation is easier than affirmation. *Quarterly Journal of Experimental Psychology*, *24*(1), 87-91. https://doi.org/10.1080/14640747208400271.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Kets, W., Kager, W., & Sandroni, A. (2021). The value of a coordination game. *Available at SSRN 3796188*. http://dx.doi.org/10.2139/ssrn.3796188.

Konrad, K. A., & Morath, F. (2012). Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology*, *306*, 61-67. https://doi.org/10.1016/j.jtbi.2012.04.013.

Krafft, P. M., Shmueli, E., Griffiths, T. L., Tenenbaum, J. B., & Pentland, A. (2021). Bayesian collective learning emerges from heuristic social learning. *Cognition, 212,* 104469. https://doi.org/10.1016/j.cognition.2020.104469.

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences*, *114*(30), 7877-7883. https://doi.org/10.1073/pnas.1620743114.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in cognitive sciences*, *19*(11), 688-699. https://doi.org/10.1016/j.tics.2015.08.005.

Lenfesty, H., & Morgan, T. J.H. (2019). By reverence, not fear: prestige, religion, and autonomic regulation in the evolution of cooperation. *Frontiers in Psychology*, *10*, 2750. https://doi.org/10.3389/fpsyg.2019.02750.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives*, *14*(3), 115-136. DOI: 10.1257/jep.14.3.115.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1509), 3515-3528. https://doi.org/10.1098/rstb.2008.0131.

McElreath, R., Boyd, R., & Richerson, P. (2003). Shared norms and the evolution of ethnic markers. *Current anthropology*, *44*(1), 122-130. https://doi.org/10.1086/345689.

McElreath, R., Lubell, M., Richerson, P. J., Waring, T. M., Baum, W., Edsten, E., Efferson, C. & Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior*, *26*(6), 483-508. https://doi.org/10.1016/j.evolhumbehav.2005.04.003.

Mercier, H., & Morin, O. (2019). Majority rules: how good are we at aggregating convergent opinions?. *Evolutionary Human Sciences*, *1*, e6. doi:10.1017/ehs.2019.6.

Mesoudi, A. (2009). How cultural evolutionary theory can inform social psychology and vice versa. *Psychological Review, 116*(4), 929–952. https://doi.org/10.1037/a0017062.

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334-342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001.

Mesoudi, A., Chang, L., Dall, S. R.X., & Thornton, A. (2016). The evolution of individual and cultural variation in social learning. *Trends in ecology & evolution*, *31*(3), 215-225. https://doi.org/10.1016/j.tree.2015.12.012.

Mesoudi, A., Cheng, L., Murray, K., & Lu, H. J. (2015). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences, 282*(1798), 20142209. https://doi.org/10.1098/rspb.2014.2209.

Mesoudi, A., & Lycett, S. J. (2009). Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior, 30*(1), 41-48. https://doi.org/10.1016/j.evolhumbehav.2008.07.005.

Mesoudi, A., Whiten, A., & Laland, K. N. (2004). Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the perspective of The Origin of Species. *Evolution, 58*(1), 1-11. https://doi.org/10.1111/j.0014-3820.2004.tb01568.x.

Miu, E., Gulley, N., Laland, K. N., & Rendell, L. (2020). Flexible learning, rather than inveterate innovation or copying, drives cumulative knowledge gain. *Science Advances, 6*(23), eaaz0286. DOI: 10.1126/sciadv.aaz0286.

Miu, E., & Morgan, T. J.H. (2020). Cultural adaptation is maximised when intelligent individuals rarely think for themselves. *Evolutionary Human Sciences, 2,* e43. doi:10.1017/ehs.2020.42.

Molleman, L., & Gächter, S. (2018). Societal background influences social learning in cooperative decision making. *Evolution and Human Behavior, 39*(5), 547-555. https://doi.org/10.1016/j.evolhumbehav.2018.05.007.

Molleman L, Kanngiesser P, van den Bos W (2019a) Social information use in adolescents: The impact of adults, peers and household composition. *PLoS ONE 14*(11): e0225498. https://doi.org/10.1371/journal.pone.0225498

Molleman, L., Kölle, F., Starmer, C., & Gächter, S. (2019b). People prefer coordinated punishment in cooperative interactions. *Nature human behaviour, 3*(11), 1145-1153. https://doi.org/10.1038/s41562-019-0707-2.

Molleman L, Pen I, Weissing FJ (2013). Effects of Conformism on the Cultural Evolution of Social Behaviour. *PLoS ONE 8*(7): e68153. https://doi.org/10.1371/journal.pone.0068153.

Molleman, L., Van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications, 5*(3570), 1-9. https://doi.org/10.1038/ncomms4570.

Morgan, T. J.H., Acerbi, A., & Van Leeuwen, E. J. (2019). Copy-the-majority of instances or individuals? Two approaches to the majority and their consequences for conformist decision-making. *PloS one*, *14*(1), e0210748. https://doi.org/10.1371/journal.pone.0210748.

Morgan, T. J. H., & Laland, K. N. (2012). The biological bases of conformity. *Frontiers in neuroscience*, *6,* 87. DOI: 10.3389/fnins.2012.00087.

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour, 3,* 221-229. https://doi.org/10.1038/s41562-018-0522-1.

Muthukrishna, M., Morgan, T. J.H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior, 37*(1), 10-20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004.

Price, M. E., & Van Vugt, M. (2014). The evolution of leader–follower reciprocity: the theory of service-for-prestige. *Frontiers in human neuroscience, 8,* 363. https://doi.org/10.3389/fnhum.2014.00363.

Reader, S. (2003). Innovation and social learning: individual variation and brain evolution. *Animal biology*, *53*(2), 147-158. DOI: https://doi.org/10.1163/157075603769700340.

Rendell, L., Fogarty, L., Hoppitt, W. J.E., Morgan, T. J.H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in cognitive sciences*, *15*(2), 68-76. https://doi.org/10.1016/j.tics.2010.12.002.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E.K., Naar, N., Newson, L., Ross, C., Smaldino, P.E., Waring, T.M. & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39,* e30. doi:10.1017/S0140525X1400106X.

Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior*, *36*(2), 86-94. https://doi.org/10.1016/j.evolhumbehav.2014.09.004.

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*(4), 667-677. https://doi.org/10.1016/j.neuron.2010.07.020.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental science*, *13*(4), 599-610. DOI: 10.1111/j.1467-7687.2009.00913.x.

Smaldino, P. E. (2019). Social identity and cooperation in cultural evolution. *Behavioural processes*, *161*, 108-116. https://doi.org/10.1016/j.beproc.2017.11.015.

Smaldino, P. E., Flamson, T. J., & McElreath, R. (2018). The evolution of covert signaling. *Scientific reports*, *8*(4905), 1-10. DOI: https://doi.org/10.1038/s41598-018-22926-1.

Smaldino, P. E., Janssen, M. A., Hillis, V., & Bednar, J. (2017). Adoption as a social marker: Innovation diffusion with outgroup aversion. *The Journal of Mathematical Sociology*, *41*(1), 26-45. DOI: 10.1080/0022250X.2016.1250083

Smaldino, P., & Jones, J. (2021). Coupled dynamics of behaviour and disease contagion among antagonistic groups. *Evolutionary Human Sciences, 3*, E28. doi:10.1017/ehs.2021.22.

Sosis, R., Kress, H. C., & Boster, J. S. (2007). Scars for war: Evaluating alternative signaling explanations for cross-cultural variance in ritual costs. *Evolution and Human Behavior, 28*(4), 234-247. https://doi.org/10.1016/j.evolhumbehav.2007.02.007.

Stein, D. H., Hobson, N. M., & Schroeder, J. (2021). A sacred commitment: How rituals promote group survival. *Current Opinion in Psychology*, *40*, 114-120. https://doi.org/10.1016/j.copsyc.2020.09.005.

Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2014). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences, 281*(1776), 20132864. https://doi.org/10.1098/rspb.2013.2864.

Vogt, S., Efferson, C., & Fehr, E. (2013). Can we see inside? Predicting strategic behavior given limited information. *Evolution and Human Behavior, 34*(4), 258-264. https://doi.org/10.1016/j.evolhumbehav.2013.03.003.

Wen, N. J., Clegg, J. M., & Legare, C. H. (2019). Smart conformists: Children and adolescents associate conformity with intelligence across cultures. *Child development, 90*(3), 746-758. DOI: 10.1111/cdev.12935.

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, *33*(4), 341-356. https://doi.org/10.1016/j.dr.2013.08.002.

6.Appendix

Note all extra references used in the appendices are present in the reference list for Chapter 3.

Appendix 1A: Rationale for testing in India.

There is a bias to recruit WEIRD participants which current research trends must move away from (Henrich et al., 2010). Previous researchers have begun investigating the differences in social learning style preferences between British and Chinese participants (Molleman & Gachter, 2018). British participants are highly individualistic whilst Chinese participants are from highly collectivist cultures. This means that social learning preferences have been tested at polarised ends of Hofstede's (1980) individualism scale. This data may not represent countries who are less differentiated and more central on this scale, and so the current study decided to recruit in India to address this gap. India ranks 21st on Hofstede's individualism scale with a score of 48 (see https://clearlycultural.com/geert-hofstede-cultural-dimensions/individualism/ for classifications).

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. doi:10.1017/S0140525X0999152X.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*(4), 15-41. https://doi.org/10.1080/00208825.1980.11656300.

Appendix 1B: Protocol given to researchers in CESS lab to ensure standardised testing across research labs. Note that the italicised text represents additional instructions that were used for the coordination game only.

<u>Instructions for experimenters (experimenter use only)</u>

1.Accept only an even number of participants. If odd, randomly select one person who showed up to send home.

2.Run the welcome treatment (.ztt file) provided on enough PCs for the invited number of participants. Turn off the PC screens before participants arrive. The welcome treatment displays a screen saying "Welcome, press OK", but participants will not be able to see it initially because their screens are off.

3.Provide an instructions booklet and pen at every client PC depending on the game type played.

4.As participants enter, direct them to their seat and tell them to start reading their instructions booklet. Tell them that when they have finished reading, they should fill in the multiple-choice questions and consent form at the back of the booklet if they wish to take part. They should then raise their hand to inform the researcher when they have finished.

Please ensure that all participants are over the age of eighteen years old, capable of giving fully-informed consent and can speak English.

5.When participants raise their hand, use the answer sheet provided overleaf to check their understanding on the multiple-choice questions. If they have **not** answered all questions correctly, then ask them to try again referring back to the instructions booklet. If they have answered correctly, turn on their PC screen. At this point, they will see that they need to press "OK." With the clients table open in z-Tree, this OK press will be visible to the experimenter. In this way, the experimenter can keep track of who has answered all comprehension questions correctly.

6. A protocol regarding verbal instructions for those with poorer English abilities is provided overleaf. These participants will still have to answer the multiple-choice questions to show understanding however, and if they cannot understand these then they should be removed from the session with their show-up fee.

7. When the experimenter has verified that everyone has answered all comprehension questions correct, which again can be seen in the clients table, read the "Summary for participants" aloud in front of the entire session of participants.

8. Run the appropriate *coordGameWithinSubjectsIndependentSignals*ztt* or *bestChoiceWithinSubjectsIndependentSignals*ztt* file, depending on the game type played. Encourage participants to enter their unique ID that they input on the first screen on to the top of their consent form. After this, do not talk to participants unless they have a question to ask. Use the instructions, debrief, or this protocol to help you answer any questions that a participant may have. If there are any questions that you cannot answer, please tell the participant to contact me using the email address that I have provided both in their instructions and debrief.

9. Once all participants have finished the game, call participants up to pay them privately according to your standard procedure. Remember to take their consent form. Please provide the debrief form to the participants with every payment.

Protocol for verbal instructions

Read the below aloud in cases where the participant is slow to read English. They will still need to look at the screen-shots and answer pre-game questions. If they still do not understand after the verbal summary, then they may have to be removed. Where possible try to avoid testing participants who may be of poor English ability. It is preferable that participants read rather than relying on this verbal summary. If they have further questions, then use the protocol/instructions booklet to answer them. Do not provide them with information that is not in the booklet, or help them answer questions, as this may give them an unfair advantage. Note that text in *italics* need only be read for the coordination game specifically.

"You are about to play a game on the computer which involves choosing between two options (@ or %). The order of these options on-screen may change throughout the game, so pay attention when choosing. You will play the game for 88 periods, which are divided into 22 blocks of 4 periods each. Your exact number of turns depends on your participant type. The computer starts by randomly choosing six of you to be Type A Participants. The others are Type B Participants.

Type A Participants make a decision in every period and thus make 88 choices. Type B Participants decide in the last period of every block, making 22 choices overall. Don't worry- Type B Participants will have the chance to earn equal points (and therefore, money) to Type A Participants.

The game is played in pairs. Your pair will be the same type as you (meaning that As play with As, and Bs play with Bs). You will be assigned a new partner in each new block of the game. Your points earned will also depend on what your partner chooses (see the points table in the booklet).

The computer will randomly decide which game Type A Participants play, and which game Type B Participants play. The two games are Game Left and Game Right. *Regardless of game-type, you will earn more points if you match your choice to your partner's (e.g. both pick %).* Which option (% or @) it is worth the most points to match on depends on which game you are playing. Type A and Type B Participants may be playing the same game, or different. You do not know which game you are playing.

Type A Participants choose every period and immediately see the points that they have earned based on their choice. Note that a random shock will be applied to these points, so you may earn more or less points than the expected points shown in the table on-screen. This shock represents the effect of things beyond your control that affect your decisions in every-day life.

Type B Participants only choose in the final period of each block. Before choosing, they will see (i) how many Type A Participants have chosen @ or % on their final (fourth) turn, (ii) a signal indicating whether they are playing the same game (Left or Right) as Type A Participants, and (iii) the probability that this signal is correct. If you are a Type B participant, you and your partner will always see the same information. The information may change, however, from one block to the next, so please pay attention to the on-screen information.

Type B Participants will not see the points that they have earned based on their choices, until they see their total points earned at the end of the game. Each of Type B's choices are played four times, with four separate random shocks, so that they earn similar points to Type A Participants.

After playing the game, you will complete a short survey about how you played the game. Fill in the multiple-choice answers by clicking the option that most applies to you with your mouse, and fill in text answers by clicking inside the purple box and typing your answer.

We ask that you remain seated at the end of the study and wait until you hear your ID number being called to accept your pay and debrief before you leave.

If you have any questions, or would like to leave for whatever reason, then please ask now. Remember not to talk to your fellow participants during the game. If you wish to leave during the game, or have any questions, then raise your hand to let a researcher know.

If you are happy to take part, then please now fill in your consent form and answer the multiple-choice questions in the back of your booklet. You must answer all questions correctly to start playing".

Predicted questions and how to answer (for experimenter use only)

•What is the experiment actually about?

This study will help us understand how people use information to make decisions, and the diversity of this process, particularly in regards to decisions made in groups.

•What will the findings be used for?

The findings will be used as part of a PhD project being ran at Royal Holloway University of London. The lead researcher's details are in your instructions booklet and on the debrief given to you at the end. This study may be published but we have taken care to ensure that your decisions are anonymous and will not be traced back to you individually.

•I'm unsure what a Type A Participant does

You will see some tables denoting game left or game right. The computer has chosen one of these games to play, though you do not know which. *You are playing the same game as your partner.* You should choose between @ or %, and hit OK when you're happy with your choice. You can then see the points that you have earned *based on both your decisions, and your partner's decisions,* after everyone has answered. One option will be worth more points than the other option, *provided that you and your partner choose the same option.* You have four turns at choosing in every block. Read the instructions again for further information.

- **I'm unsure what a Type B Participant does**

Type B Participant have one chance to match their choice of % or @ with their partner's in every block. You are playing Game left or Game right, and this may be the same game as Type A participant's or different. You will see some information on-screen to help you make your choice. This includes the number of Type A Participants choosing @ or %, and whether you are playing the same or different game to these participants. You will also be told how likely it is that this information about playing the same or different game is to be correct. Remember that the partner you are paired to play the game with always sees the same information as you. The on-screen information and your assigned partner can change between blocks however, so you should play close attention each time you make a decision. Read the instructions again for more information.

- **How do I answer the survey?**

Try to answer these questions as honestly as you can. Answer the multiple-choice questions by using your mouse to click the white button next to your desired answer. Please select the answer that you feel most applies to you. The final field requires text answers. Click inside the purple text box and manually type your responses. When you are done hit OK and then raise your hand to let the researcher know.

•You may have to remind people to type their age and years lived in their city as numbers instead of text (i.e. “22” instead of “twenty-two”).

•If the participant sees an error message in a different language: this means that they have not clicked the option or typed their answer properly and must do so to proceed. Tell them to hit OK and answer again.

•If participants see an extra question screen at the end repeating the age questions, then the individual has entered their years lived in a country as more than their age by mistake in the first survey. Tell them that they are being asked to rectify any small mistakes in a separate screen at the end.

•Any other questions, ask them to email me after the experiment using the address that I included on both the instructions and the debrief.

Appendix 2: The instructions given to the participants including pre-game questions to check understanding.

Appendix 2A: Instructions for the game against nature.

Welcome! You are invited to participate in a study for approximately **1 ½ hours.** You can earn points during this study, which will be converted to money at the following rate:

14 points = ₹1

You will also be paid a show-up fee of **₹100** on top of the money you earn. The choices participants make during the study will be anonymous. This means you will not be able to identify the specific participants in the room who make certain choices, and none of the participants will be able to trace your choices back to you.

Please do not communicate with the other participants. If you have questions, or need to withdraw, then please raise your hand and tell the researcher.

Please read this instruction sheet **carefully**. You will then answer some questions to check that you have understood the study. We will not be able to proceed until everyone answers **all** questions correctly. You will also respond to a brief **survey** after the main study.

The study:

To begin, the computer will randomly choose six of you to be Type A Participants. Others will be Type B Participants. As explained later, your type will determine how often you make choices and the information you have when you do so. The study lasts for 88 periods. We will divide these 88 periods into 22 blocks of 4 periods each. Type A Participants will choose every period, which means they will make 88 choices. Type B Participants will only make a choice in the final period of each block, which means that Type B Participants will make 22 choices. Don't worry.

Though Type A and Type B Participants do not make the same number of choices, they will have exactly the same opportunity to earn points. We will explain this in detail later. The upper left-hand corner of your screen will have a counter that displays the current period you are in.

The games:

At the beginning of each block of 4 periods, the computer will randomly pair you with another participant of the same type to play a game. A Type A Participant will always be paired with another Type A, and a Type B Participant will always be paired with another Type B Participant. Every time you play, both you and your partner must choose between one of two options, either option "%" or option "@".

Specifically, there are two games, which we call "Game Left" and "Game Right". At the beginning of each block, the computer will randomly pick which game Type A Participants play and which game Type B Participants play. The computer decides this completely randomly, giving four possible combinations, which are all equally likely to occur (each with a 1 in 4 probability). The four possibilities are:

(i) both types play Game Left,

(ii) Type A Participants play Game Left, Type B Participants play Game Right,

(iii) Type A Participants play Game Right, Type B Participants play Game Left,

(iv) both types play Game Right.

Note that we do not tell you if you are playing Game Left or Game Right. The following tables show you how your points will depend on the choices you make, for each of the games that you might play.

Game left

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 150 Partner's expected points: 150	Your expected points: 150 Partner's expected points: 120
You choose @	Your expected points: 120 Partner's expected points: 150	Your expected points: 120 Partner's expected points: 120

Game right

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 120 Partner's expected points: 120	Your expected points: 120 Partner's expected points: 150
You choose @	Your expected points: 150 Partner's expected points: 120	Your expected points: 150 Partner's expected points: 150

As you can see, one option is worth more points and will therefore result in you earning more money, on average, if you pick this option. The option (% or @) that is worth the most points is different depending on whether you are playing Game Left or Game Right.

You can also see that your payoff will **not** depend on what your partner chooses in any way. Nor will your partner's payoff depend on what you choose.

Lastly, points will also be affected by forces outside of your control, as in real life. The tables above shows the expected points you will earn, but random shocks will be applied to these values. These random shocks can lead you to earn **more OR less** points than the expected values shown in the tables.

For example, assume you are playing Game Left. It is possible that you could earn more by choosing @ than you could by choosing % for a single choice. It is more likely, however, that % will earn more than @, and so when choosing repeatedly % is highly likely to produce the most points.

Similarly, assume you are playing Game Right. It is possible that you could earn more by choosing % than by choosing @ for a single choice. It is more likely, however, that @ will earn more than %, and so when choosing repeatedly @ is highly likely to produce the most points.

In summary, the points that you earn will depend on three things.

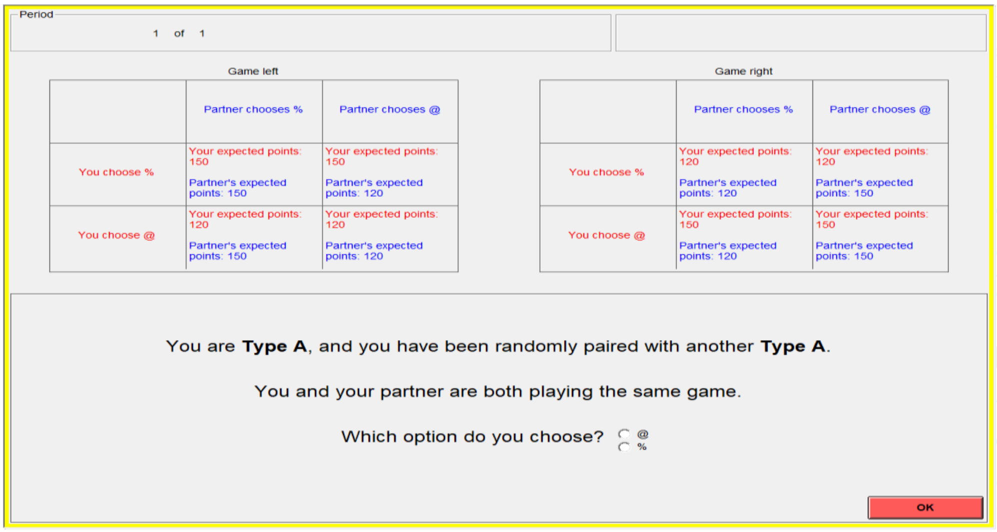
1. **The game being played:** Game Left or Game Right.
2. **The option you choose:** Option % or Option @.
3. **The random shock:** Random shocks are added to an expected payoff. Random shocks are independent of each other. This means that sometimes they will lead to more points than expected, and sometimes they lead to less points than expected.

IMPORTANT: The computer chooses the games being played at the beginning of each block. The games being played can change from one block to the next, and you will be paired with a different partner of the same type as you in each new block.

Type A vs. Type B Participants

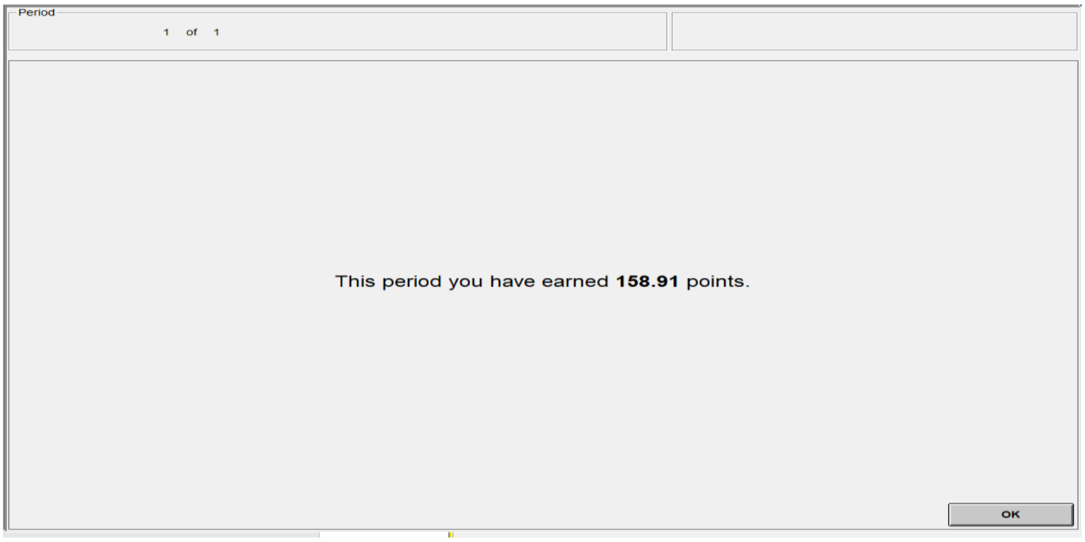
As explained above, before the study begins, the computer will randomly select six participants to be Type A participants. Others will be Type B participants. Your type will not change.

Type A Participants -> Type A Participants choose every period, and each Type A Participant immediately sees the points he or she earns after making a choice. **IMPORTANT:** The computer chooses the games being played at the beginning of each block of four periods. Note that the game being played does **NOT** change across the four periods within a block. Here is an example of a choice screen for a Type A Participant:

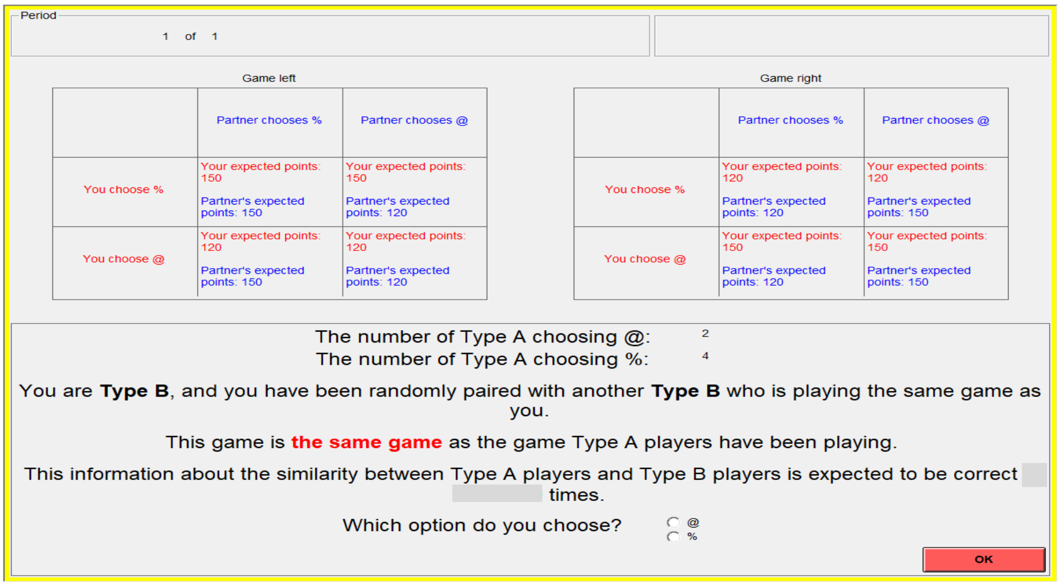


In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This is true for everyone. In any given period some people will see % listed first, while others will see @ listed first. This means that you should pay close attention when choosing.

After making a choice and receiving a payoff, each Type A participant will immediately learn how many points he or she earned. Here is an example of a feedback screen that a Type A Participant might see:



Type B Participants -> Type B Participants do NOT choose every period. Instead, they only choose in the final period of every block. In earlier periods of a block, Type B Participants simply wait. Here is an example choice screen that a Type B Participant might see in the final period of a block:



In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This is true for everyone. In any given period some people will see % listed first, while others will see @ listed first. This means that you should pay close attention when choosing.

Importantly, when Type B Participants make a choice in the final period of a block, they will see the following information.

- The number of Type A Participants who chose option % and the number who chose option @ in the final (fourth) period of the block.
- A SIGNAL indicating if Type B Participants are playing the SAME game as Type A Participants or a DIFFERENT game (Game Left or Game Right). **IMPORTANTLY, you and your partner will always see the same signal.** As explained above, the computer begins each block by randomly choosing which game all Type A Participants play. It then randomly chooses the game that all Type B Participants play separately. This means that Type B Participants **may or may not** be playing the same game as Type A Participants.
- The probability that the above signal is correct. The signal indicating if Type A Participants play the same game as Type B Participants is not always correct. This signal will only be correct with a certain probability. We do not tell you what this probability is here, but you will see it on the screen in **bold** every time you make a choice. The above screen shot provides an example in which we have blurred out the probability that the signal is correct. **IMPORTANTLY, the probability the signal is correct may change from one block to the next, so please pay attention every time you make a choice. The probability will ALWAYS be the same for you and your partner.**

The following table summarizes the relationship between games and signals:

	Type A Participants play Game Left	**Type A Participants play Game Right**
Type B Participants play Game Left	Correct signal: SAME Incorrect Signal: DIFFERENT	Correct signal: DIFFERENT Incorrect Signal: SAME
Type B Participants play Game Right	Correct Signal: DIFFERENT Incorrect Signal: SAME	Correct Signal: SAME Incorrect Signal: DIFFERENT

After making a choice, the Type B Participants will receive four separate pay-offs based on their choice and the game they are playing. Four separate random shock values will be added to these pay-offs. This means that, even though Type B Participants make fewer choices than Type A Participants, they have the exact same number of opportunities to earn points.

Type B Participants will not see the points they earn after choosing. Instead, they will only see the total number of points earned across all blocks at the very end of the study.

***

Final instructions

Once you have played the last block of the game, you will complete a short survey. Then, please wait until the researcher calls your seat number to receive your payment. Your earnings will not be told to any other participant.

Now please sign the consent form and answer the 10 pre-game questions. Raise your hand to alert the researcher when you are finished. Everyone must answer **ALL** pre-game questions correctly before we can begin.

Please keep these instructions to refer back to during the study.

If you have any further questions, feel free to contact the lead PhD researcher (Aysha Bellamy) at: **pejt007@live.rhul.ac.uk**

Comprehension questions

Please answer the following multiple-choice questions, by circling your chosen answer. Everyone must answer all 10 questions correctly before we can begin. You may use your instructions to help you:

***Q1*: Which of the following statements is true, in regards to the number of choices that each participant (Type A and B) makes?**

a) Type A participants make fewer choices than Type B participants.

b) Type A participants make 4X as many choices as Type B participants, but both types of participant have the same opportunities to earn points.

c) Type A Participants make 4X as many choices as Type B Participants, and thus earn 4X as much.

***Q2*: Which of the following statements is true, in regards to the points that you can earn if you were playing Game Left?**

a) You would expect to receive 150 points for choosing the % option.

b) You would expect to receive 120 points for choosing the % option.

***Q3*: Which of the following statements is true, in regards to the partner you are assigned to play the game with?**

a) I get paired with a new partner who is the same type as me in every block.

b) I will play with the same partner throughout the whole study.

c) I get paired with a new partner who is a different type from me in every block.

Q4: **Which of the following statements is true, in regards to the things that may affect the points that you can earn?**

a) My points earned depend only on which game I am playing, and which option I choose.

b) My expected points depend on **both** the game I am playing and my partner's decisions, though the points that I can earn will also be affected by a random shock.

c) My expected points depend on which game I am playing, and which option I choose, though the points that I can earn will also be affected by a random shock.

Q5: **Which of the following statements is true, in regards to whether you are playing the same game (Left or Right) as other participants?**

a) I always play the same game as my partner, though Type A and Type B Participants may play the same game or different games.

b) I play a different game to my partner, though Type A and Type B Participants play the same game.

c) All participants play the same game.

Q6**: Which of the following statements is true, in regards to the feedback that Type A Participants receive?**

a) Type A participants see no information

b) Type A participants see the points made by other Type A participants, but not their own points.

c) Type A participants see their own points, but not the points of other Type A participants.

Q7: **If you are a Type B Participant, and the signal tells you that you're playing a different game from Type A Participants, is this information necessarily correct?**

a) This information is always correct.

b) This information will sometimes be correct, with a certain probability, and sometimes incorrect, with the remaining probability.

c) This information is never correct.

Q8: **If you are a Type B Participant, when will you see the probability that the signal (telling you that you are playing the same game as Type A's, or a different game) is correct?**

a) At the very end of the game.

b) After I make each choice.

c) It will be with the information on-screen before I make my choice.

Q9: **If you are a Type B Participant, you and your partner will always see the same signal indicating whether you're playing the same game as Type A Participants.**

a) True

b) False

Q10: **If you are a Type B Participant, the probability that this signal (see Q9) is correct will always be the same for you and your partner.**

a) True

b) False

Please now raise your hand and alert the researcher, who will check your answers to the multiple-choice comprehension questions.

<u>Summary for participants (Experimenter will read aloud just after everyone has answered all comprehension questions correctly)</u>

- Welcome to the main portion of the study! Today's session will consist of 88 periods divided into 22 blocks of 4 periods each. When we begin, the computer will randomly assign you to play as a Type A or Type B Participant.

- At the beginning of each block, you will be assigned to play with a partner of the same type as you for all periods in the block.

- For each choice, you will choose between two options (% or @). One option is expected to result in more points, though the option expected to be worth more depends on whether you are playing Game Left or Right. You do not know which of these games you are playing. The game can change between blocks, but not within blocks. You and your partner will always be playing the same game.

- Type A Participants choose every period and see the points earned immediately after every choice.

- Type B Participants only choose in the final period of each block. Before choosing, they will see (i) how many Type A Participants have chosen @ or %, (ii) a signal indicating whether they are playing the same game (Left or Right) as Type A Participants, and (iii) the probability that this signal is correct. If you are a Type B participant, you and your partner will always see the same information. The information may change, however, from one block to the next, so please pay attention to the on-screen information.

- If you have any questions, please ask them now. Remember not to talk to your fellow participants during the game.

✂ ____

Consent form continued

Client ID number: ________________

Thank you very much for reading the instructions sheet. If you have any questions, then feel free to raise your hand and ask the researcher.

If you are happy to take part in this study, then please sign below:

I have now read the instructions sheet and understood the study. I can confirm that I would still like to take part in this study.

Name in block letters:

____

Sign here: ___________________________________ Date: _____________________

<u>Comprehension questions answer sheet (for experimenter use only)</u>

This sheet contains only the correct options. Use this to check the participants understanding. If they have failed a certain question, direct them towards the relevant section of the instructions booklet.

***Q1*: b**

***Q2*: a**

***Q3*: a**

***Q4:* c**

***Q5:* a**

***Q6*: c**

***Q7:* b**

***Q8:* c**

***Q9:* a**

***Q10:* a**

Appendix 2B: Instructions for the coordination game.

Welcome! You are invited to play a game for approximately **1 ½ hours.** You can earn points during this study, which will be converted to money at the following rate:

25 points = ₹1

You will be paid a show-up fee of **₹100** on top of the money you earn. The choices that participants make during the study will be anonymous. This means you will not be able to identify the specific participants in the room who make certain choices, and none of the participants will be able to trace your choices back to you.

Please do not communicate with the other participants. If you have questions, or need to withdraw, then please raise your hand and tell the researcher.

Please read this instruction sheet **carefully**. You will then answer some questions to check that you have understood the study. We will not be able to proceed until everyone answers **all** questions correctly. You will also respond to a brief **survey** after the main study.

The study:

To begin, the computer will randomly choose six of you to be Type A Participants. Others will be Type B Participants. As explained later, your type will determine how often you make choices and the information you have when you do so. The study lasts for 88 periods. We will divide these 88 periods into 22 blocks of 4 periods each. Type A Participants will choose every period, which means they will make 88 choices. Type B Participants will only make a choice in the final period of each block, which means that Type B Participants will make 22 choices. Don't worry. Though Type A and Type B Participants do not make the same number of choices, they will have exactly the same opportunity to earn points. We will explain this in detail later. The upper left-hand corner of your screen will have a counter that displays the current period you are in.

The games:

At the beginning of each block of 4 periods, the computer will randomly pair you with another participant of the same type to play a game. A Type A Participant will always be paired with another Type A, and a Type B Participant will always be paired with another Type B Participant. Every time you play, both you and your partner must choose one of two options, either option "%" or option "@".

Specifically, there are two games, which we call "Game Left" and "Game Right". At the beginning of each block, the computer will randomly pick which game Type A Participants play and which game Type B Participants play. The computer decides this completely randomly, giving four possible combinations, which are all equally likely to occur (each with a 1 in 4 probability). The four possibilities are:

(i) both types play Game Left,

(ii) Type A Participants play Game Left, Type B Participants play Game Right,

(iii) Type A Participants play Game Right, Type B Participants play Game Left,

(iv) both types play Game Right.

Note that we do not tell you if you are playing Game Left or Game Right. The following tables show you how your points will depend on the choices made by both you and your partner, for each of the games that you might play:

Game left

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 325 Partner's expected points: 325	Your expected points: 100 Partner's expected points: 100
You choose @	Your expected points: 100 Partner's expected points: 100	Your expected points: 250 Partner's expected points: 250

Game right

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 250 Partner's expected points: 250	Your expected points: 100 Partner's expected points: 100
You choose @	Your expected points: 100 Partner's expected points: 100	Your expected points: 325 Partner's expected points: 325

As you can see, one option is worth more points and will therefore result in you earning more money, on average, if you pick this option. The option (% or @) that is worth the most points is different depending on whether you are playing Game Left or Game Right.

You can also see that your payoff depends on what your **partner** chooses. Likewise, your partner's payoff depends on what you choose.

Lastly, points will also be affected by forces outside of your control, as in real life. The tables above shows the expected points you will earn, but random shocks will be applied to these values. These random shocks can lead you to earn **more OR less** points than the expected values shown in the tables. These random shocks can also lead your **partner** to earn **more OR less** points than those shown in the tables.

For example, assume you are playing Game Left. It is possible that you could earn more by choosing @ than you could by choosing % for a single choice, provided that your partner chooses the same option as you. It is more likely, however, that you and your partner will earn more points for choosing % than @, and so when choosing repeatedly % is highly likely to produce the most points.

Similarly, assume you are playing Game Right. It is possible that you could earn more by choosing % than by choosing @ for a single choice, provided that your partner chooses the same option as you. It is more likely, however, that you and your partner will earn more points for choosing @ than %, and so when choosing repeatedly @ is highly likely to produce the most points.

In summary, the points that you earn will depend on four things.

4. **The game being played:** Game Left or Game Right.
5. **The option you choose:** Option % or Option @.
6. **The option your partner chooses:** Option % or Option @.
7. **The random shock:** Random shocks are added to an expected payoff. Random shocks are independent of each other. This means that sometimes they will lead to more points than expected, and sometimes they lead to less points than expected.

IMPORTANT: The computer chooses the games being played at the beginning of each block. The games being played can change from one block to the next, and you will be paired with a different partner of the same type as you in each new block.

Type A vs. Type B Participants

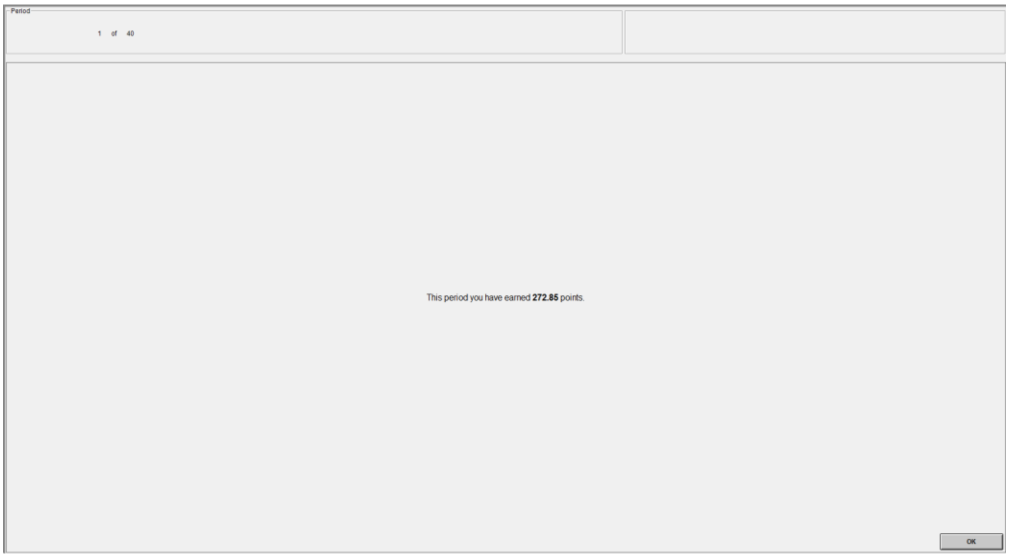
As explained above, before the study begins, the computer will randomly select six participants to be Type A participants. Others will be Type B participants. Your type will not change.

Type A Participants -> Type A Participants choose every period, and each Type A Participant immediately sees the points he or she earns after making a choice.

IMPORTANT: The computer chooses the games being played at the beginning of each block of four periods. Note that the game being played does **NOT** change across the four periods within a block. Here is an example choice screen for a Type A Participant:

In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This is true for everyone. In any given period some people will see % listed first, while others will see @ listed first. This means that you should pay close attention when choosing.

After making a choice and receiving a payoff, each Type A participant will immediately learn how many points he or she earned. Here is an example points screen a Type A Participants might see:



Type B Participants -> Type B Participants do NOT choose every period. Instead, they only choose in the final period of every block. In earlier periods of a block, Type B Participants simply wait. Here is an example of a choice screen a Type B Participant might see in the final period of a block:

In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This is true for everyone. In any given period some people will see % listed first, while others will see @ listed first. This means that you should pay close attention when choosing.

Importantly, when Type B Participants make a choice in the final period of a block, they will see the following information.

- The number of Type A Participants who chose option % and the number who chose option @ in the final (fourth) period of the block.
- A SIGNAL indicating if Type B Participants are playing the SAME game as Type A Participants or a DIFFERENT game (Game Left or Game Right). **IMPORTANTLY, you and your partner will always see the same signal.** As explained above, the computer begins each block by randomly choosing which game all Type A Participants play. It then randomly chooses the game

that all Type B Participants play separately. This means that Type B Participants **may or may** not be playing the same game as Type A Participants.

- The probability that the above signal is correct. The signal indicating if Type A Participants play the same game as Type B Participants is not always correct. We do not tell you what this probability is here, but you will see it on the screen in **bold** every time you make a choice. The screen shot provides an example in which we have blurred out the probability that the signal is correct. **IMPORTANTLY, the probability the signal is correct may change from one block to the next, so please pay attention every time you make a choice. The probability will ALWAYS be the same for you and your partner.**

The following table summarizes the relationship between games and signals:

	Type A Participants play Game Left	**Type A Participants play Game Right**
Type B Participants play Game Left	Correct signal: SAME Incorrect Signal: DIFFERENT	Correct signal: DIFFERENT Incorrect Signal: SAME
Type B Participants play Game Right	Correct Signal: DIFFERENT Incorrect Signal: SAME	Correct Signal: SAME Incorrect Signal: DIFFERENT

After making a choice, the Type B Participants will receive four separate pay-offs based on their choice, their partner's choice and the game they are playing. Four separate random shock values will be added to these pay-offs. This means that, even though Type B Participants make fewer choices than Type A Participants, they have the exact same number of opportunities to earn points.

Type B Participants will not see the points they earn after choosing. Instead, they will only see the total number of points earned across all their blocks at the very end of the study.

***

Final instructions

Once you have played the last block of the game, you will see a short survey you should then complete. This survey asks about how you played the game, and for some background information. You can answer the multiple-choice questions by using the mouse to select your chosen answer, and the remaining questions can be answered by clicking in the purple text box and typing in an answer. The study is finished once you answer this survey, and hit OK. You will then see a screen showing your total points and money earned. Please wait until the researcher calls your seat number to receive your payment. Your earnings will not be told to any other participant.

Now turn the page to sign the consent form and answer the 11 pre-game questions. Raise your hand to alert the researcher when you are finished. Everyone must answer **ALL** pre-game questions correctly before we can begin.

Please keep these instructions to refer back to during the study.

If you have any further questions, feel free to contact the lead PhD researcher (Aysha Bellamy) at: **pejt007@live.rhul.ac.uk**

Comprehension questions

Please answer the following multiple-choice questions, by circling your chosen answer. Everyone must answer all 11 questions correctly before we can begin. You may use your instructions to help you:

***Q1*: Which of the following statements is true, in regards to the number of choices that each participant (Type A and B) makes?**

a) Type A participants make fewer choices than Type B participants.

b) Type A participants make 4X as many choices as Type B participants, but both types of participant have the same opportunities to earn points.

c) Type A Participants make 4X as many choices as Type B Participants, and thus earn 4X as much.

***Q2*: Which of the following statements is true, in regards to the points that you can earn if you were to choose the same option as your partner whilst playing Game Left?**

a) You would expect to receive 325 points for choosing the % option.

b) You would expect to receive 250 points for choosing the % option.

Q3: **Which of the following statements is true, in regards to the number of points that you should earn if you choose a different option to your partner?**

a) You will get 0 points.

b) You should get 100 points, but as earnings in real-life are affected by more than one decision, then a random shock applied to these points may mean that some participants receive more or less points than 100.

c) You should get 250 points, but as earnings in real-life are affected by more than one decision, then a random shock applied to these points may mean that some participants receive more or less points than 250.

Q4**: Which of the following statements is true, in regards to the partner you are assigned to play the game with?**

a) I get paired with a new partner who is the same type as me in every block.

b) I will play with the same partner throughout the whole study.

c) I get paired with a new partner who is a different type from me in every block.

Q5: **Which of the following statements is true, in regards to the things that may affect the points that you can earn?**

a) My points earned depend only on which game I am playing, and which option I choose.

b) My expected points depend on **both** the game I am playing and which option me and my partner choose, though the points that I can earn will also be affected by a random shock.

c) My expected points depend on which game I am playing, and which option I choose, though the points that I can earn will also be affected by a random shock.

***Q6:* Which of the following statements is true, in regards to whether you are playing the same game (Left or Right) as other participants?**

a) I always play the same game as my partner, though Type A and Type B Participants may play the same game or different games.

b) I play a different game to my partner, though Type A and Type B Participants play the same game.

c) All participants play the same game.

***Q7*: Which of the following statements is true, in regards to the feedback that Type A Participants receive?**

a) Type A participants see no information

b) Type A participants see the points made by other Type A participants, but not their own points.

c) Type A participants see their own points, but not the points of other Type A participants.

Q8: **If you are a Type B Participant, and the signal tells you that you're playing a different game from Type A Participants, is this information necessarily correct?**

a) This information is always correct.

b) This information will sometimes be correct, with a certain probability, and sometimes incorrect, with the remaining probability.

c) This information is never correct.

Q9: **If you are a Type B Participant, when will you see the probability that the signal (same game as Type A or different game) is correct?**

a) At the very end of the game.

b) After I make each choice.

c) It will be with the information on-screen before I make my choice.

Q10: **If you are a Type B Participant, you and your partner will always see the same signal indicating whether you're playing the same game as Type A Participants.**

a) True

b) False

Q11: **If you are a Type B Participant, the probability that this signal (see Q10) is correct will always be the same for you and your partner.**

A. True

B. False

Please now raise your hand and alert the researcher, who will check your answers to the multiple-choice comprehension questions.

Summary for participants (Experimenter will read aloud just after everyone has answered all comprehension questions correctly)

- Welcome to the main portion of the study! Today's session will consist of 88 periods divided into 22 blocks of 4 periods each. When we begin, the computer will randomly assign you to play as a Type A or Type B Participant.

- At the beginning of each block, you will be assigned to play with a partner of the same type as you for all periods in the block.

- For each choice, you will choose between two options (% or @). One option is expected to result in more points if both you and your partner choose it at the same time, though the option that is expected to be worth more depends on whether you are playing Game Left or Right. You do not know which of these games you are playing. The game can change between blocks, but not within blocks. You and your partner will always be playing the same game.

- Type A Participants choose every period and see the points earned immediately after every choice.

- Type B Participants only choose in the final period of each block. Before choosing, they will see (i) how many Type A Participants have chosen @ or %, (ii) a signal indicating whether they are playing the same game (Left or Right) as Type A Participants, and (iii) the probability that this signal is correct. If you are a Type B participant, you and your partner will always see the same information. The information may change, however, from one block to the next, so please pay attention to the on-screen information.

- If you have any questions, please ask them now. Remember not to talk to your fellow participants during the game.

✂ ....

Consent form continued

Client ID number: _______________

Thank you very much for reading the instructions sheet. If you would like to clarify anything, then please raise your hand and ask a researcher. If for whatever reason you no longer wish to participate in this study, then please inform a researcher without signing this sheet.

If you are happy to take part in this study, then please sign below:

I have now read the instructions sheet and understood the study. I can confirm that I would still like to take part in this study.

Name in block letters:

____

Sign here: __________________________________ Date: ____________________

***

<u>Comprehension questions answer sheet (for experimenter use only)</u>

This sheet contains only the correct options. Use this to check the participants understanding. If they have failed a certain question, direct them towards the relevant section of the instructions booklet.

Q1: b

Q2: a

Q3: b

Q4: a

Q5: b

Q6: a

Q7: c

Q8: b

Q9: c

Q10: a

Q11: a

Appendix 3: The script used to run the experiment via Z-Tree version 3.5.

Appendix 3A: The link to the script used to run the game against nature, for a total of 30 participants in a session:

https://www.dropbox.com/sh/x5luey0br97kcb1/AABGw-2GQV-3HZHt0294wM_Va/session_2_14120218/session_2_14122018?dl=0&preview=bestChoiceWithinSubjectsIndependentSignals_totalOf30.ztt&subfolder_nav_tracking=1

Appendix S3B: **The link to the script used to run the coordination game, for a total of 30 participants in a session:**

https://www.dropbox.com/sh/x5luey0br97kcb1/AABbKs8hh_PW6krj4EXnePO1a/session_1_13120218/session_1_13122018?dl=0&preview=coorWithinSubjectsIndependentSignals_totalOf30.ztt&subfolder_nav_tracking=1

Appendix 4: Ethical requirements for study.

Appendix 4A: The debrief given to the participants at the end of both games:

ID number ___________

Thank you for taking part in **The gene-culture co-evolution of group identities** study via CESS. The data you have provided will be used in my PhD project. It will help us to understand how people make decisions based on social information. Specifically, we are interested in how your similarity to others during the game, and the reliability of this information (1/5/9 in 10 reliable) affected the way that Type B Participants used social information. We will also use your responses to the survey to further understand when and why people use social information when making decisions.

What happens now?

You have been paid according to the points you earned plus a show-up fee.
CESS will keep a copy of your raw data, and send us an anonymised copy. We will use this data in our analysis, which may be published, though I not identify any one's data specifically. Your unique subject ID codes will ensure that any published data is anonymous (i.e. cannot be traced back to you personally). Your data will be stored securely in accordance with the Data Protection Act 1988.

Thank you!

We are extremely grateful for the time you have given to take part in this study. If you would like any further information about the study, have concerns about your data, or are interested in any of the topics, then please contact me (Aysha Bellamy) with the contact details below. As you leave, **please remember to take this sheet with you.**

Aysha Bellamy *Work email:* pejt007@rhul.live.ac.uk

Appendix 4B: Proof of self-certified ethical clearance from Royal Holloway, University of London.

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Ethics Review Details

You have chosen to self certify your project.	
Name:	Bellamy, Aysha (2017)
Email:	PEJT007@live.rhul.ac.uk
Title of research project or grant:	The gene-culture co-evolution of group identities.
Project type:	Royal Holloway postgraduate research project/grant
Department:	Psychology
Academic supervisor:	Dr. Charles Efferson
Email address of Academic Supervisor:	Charles.Efferson@rhul.ac.uk
Funding Body Category:	No external funder
Funding Body:	
Start date:	01/11/2018
End date:	01/09/2020

Research question summary:

The theory of gene-culture coevolution suggests that cultural evolution is an important influence on human behaviour. For cultural evolution to take place, one of the key assumptions in this field is that we preferentially learn from those of the same culture as ourselves. In experiments, this should translate to participants preferentially learning from a group of similar others (as similar people are more likely to belong to the same social group). Very few previous studies have investigated this assumption. Those studies that did investigate similarity have not taken into account that our ability to calculate our similarity to others may not always be reliable. These experiments will be the first to investigate a more-realistic, graded similarity-signal, to see how this impacts social-learning style preference. Conformity (or the disproportionate trend to adapt the same behaviour as the majority of a group) is thought to be important in homogenising cultural groups (i.e. making people of the same culture more similar), and in allowing costly levels of cooperation to emerge in human societies. Thus, conformity should be more likely in a group of similar others. Experiment 2 will also address whether conformity can allow cooperation to emerge, by investigating how participants learn from similar and different others during a Social Dilemma task structure.

In summary, the main research question behind these two studies is how people use similarity-information to modify their choice of social-learning strategy. Conformity and cooperation will also be investigated, and these studies will run in a lab in India in the aims of expanding our knowledge of social-learning strategies to those from a diverse range of cultures.

Note that instructions given below will be modified only slightly to accommodate the coordination and social dilemma games.

Research method summary:

Experiments run via Centre for Experimental Social Sciences (CESS) labs in Pune, India. Experiment 1 tests a group of participants en-masse in a best-choice or coordination game, and Experiment 2 tests a Social Dilemma game. Regardless of game-type, participants are assigned to play as individual- or social-learners. The aim of all three games is to choose between two options (@ or %). Individual-learners have four turns in a row, with immediate feedback from their decisions (points). Points are converted to pay with a random shock to reflect exogenous factors influencing decision-making. Individual-learners learn to choose the best option. Social-learners will not see the feedback based on their choices until the end of the game. Instead, they see the number of individual-learners who chose @ or % on their final turn. They also see a sentence telling them that they play the same or different game to individual-learners, and that this signal is likely to be correct only 1/5/9 in 10 of the time (representing an unreliable, chance guess and a reliable signal respectively). The same game means that the same option produces the highest points for individual- and social-learners (e.g. both choose @). At the end of the game, a survey asks social-learners what social-learning strategies they chose and why. Every participant gives key demographics, including age, gender, country of residence and length of time living there. This experiment is ran via Z-Tree, and individuals play at separately-screened computers, anonymising data. CESS provides anonymised files for analysis. The 'game' also involves paper-based instructions and multiple-choice questions to check understanding. The coordination game involves trying to sync your choice to another participant, with whom you cannot interact, whilst Social Dilemmas involve choosing between an individual-maximising strategy (defection) in favour of a strategy which is better when both parties choose it (cooperation).

Risks to participants

Does your research involve any of the below?

Children (under the age of 16),

No

Participants with cognitive or physical impairment that may render them unable to give informed consent,

No

Participants who may be vulnerable for personal, emotional, psychological or other reasons,

No

Participants who may become vulnerable as a result of the conduct of the study (e.g. because it raises sensitive issues) or as a result of what is revealed in the study (e.g. criminal behaviour, or behaviour which is culturally or socially questionable),

No

Participants in unequal power relations (e.g. groups that you teach or work with, in which participants may feel coerced or unable to withdraw),

No

Participants who are likely to suffer negative consequences if identified (e.g. professional censure, exposure to stigma or abuse, damage to professional or social standing),

No

Details,

Design and Data

Does your study include any of the following?

Will it be necessary for participants to take part in the study without their knowledge and/or informed consent at the time?,
No

Is there a risk that participants may be or become identifiable?,
No

Is pain or discomfort likely to result from the study?,
No

Could the study induce psychological stress or anxiety, or cause harm or negative consequences beyond the risks encountered in normal life?,
No

Does this research require approval from the NHS?,
No

If so what is the NHS Approval number,

Are drugs, placebos or other substances to be administered to the study participants, or will the study involve invasive, intrusive or potentially harmful procedures of any kind?,
No

Will human tissue including blood, saliva, urine, faeces, sperm or eggs be collected or used in the project?,
No

Will the research involve the use of administrative or secure data that requires permission from the appropriate authorities before use?,
No

Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?,
No

Is there a risk that any of the material, data, or outcomes to be used in this study has been derived from ethically-unsound procedures?,
No

Details,
Financial inducements are based on in-game performance. These cannot be negative (i.e. participants cannot lose pay) and points are set to be high, allowing for all participants to feel they are good at playing the game. Inducements will also have a random shock applied to them, so it will be made clear to participants that their pay earned in no way reflects how 'smart' they are at playing the game, so this use of financial incentives is not deemed to be unethical.

Risks to the Environment / Society

Will the conduct of the research pose risks to the environment, site, society, or artifacts?,
No

Will the research be undertaken on private or government property without permission?,
No

Risks to Researchers/Institution

Does your research present any of the following risks to researchers or to the institution?

Is there a possibility that the researcher could be placed in a vulnerable situation either emotionally or physically (e.g. by being alone with vulnerable, or potentially aggressive participants, by entering an unsafe environment, or by working in countries in which there is unrest)?,
No

Is the topic of the research sensitive or controversial such that the researcher could be ethically or legally compromised (e.g. as a result of disclosures made during the research)?,
No

Will the research involve the investigation or observation of illegal practices, or the participation in illegal practices?,

No

Could any aspects of the research mean that the University has failed in its duty to care for researchers, participants, or the environment / society?,
No

Is there any reputational risk concerning the source of your funding?,
No

Is there any other ethical issue that may arise during the conduct of this study that could bring the institution into disrepute?,

No

Details,

Declaration

By submitting this form, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Certificate produced for user ID, PEJT007

Date:	05/11/2018 10:11
Signed by:	Bellamy, Aysha (2017)
Digital Signature:	Aysha Bellamy
Certificate dated:	11/5/2018 10:41:50 AM
Files uploaded:	Self-Assessment-851-2018-03-12-14-36-PEJT007.pdf ins_india_BC_5Nov2018.docx

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Ethics Self Assessment

Your answers indicate that you do not need ethical approval. If your research includes use of animals as research subjects, you will have been emailed separate guidance which must be followed before you begin your research. Should the circumstances of your research alter in any way please revisit this process to validate your project.

Applicant details

Declaration

By clicking the 'submit form' button, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Project type:	Royal Holloway postgraduate research project/grant
Name:	Bellamy, Aysha (2017)
Email:	PEJT007@live.rhul.ac.uk
Academic supervisor:	Dr. Charles Efferson
Department:	Psychology
Title of research project or grant:	The gene-culture co-evolution of group identities.
Email address of Academic Supervisor:	Charles.Efferson@rhul.ac.uk
Funding Body Category:	No external funder
Funding Body:	

Information about the Research Project

Will the research project involve the use of human participants or human tissue (with or without their knowledge or consent at the time)?, No

Are the results of the research project likely to expose any person or community to physical or psychological harm?, No

Will the research project involve the use of animals as research subjects?, No

Will you have access to personal information that allows you to identify individuals or company confidential information (that is not covered by confidentiality terms within an agreement or by a separate confidentiality agreement)?, No

Does the conduct of the research project present a significant risk to the environment or society?, No

Are there any other ethical issues raised by this research project that in the opinion of the PI require further ethical review?, No

Does the PI believe that the results of this research could reasonably lead to legal action or negative press coverage, for which the PI would require University support?, No

Certificate produced for user ID PEJT007

Certificate dated 3/12/2018 2:36:57 PM

Appendix 5: All analysis scripts used in RStudio.

Note that the following includes the script to run all regressions (*analysis_CAG_GN_together_doubleCheck_30Ap21.R*), the function to create bootstrapped confidence intervals *(estAndBootImitationFunctions_CG_app.R*), and the script to create the graph seen in Figure 5 *(plotBoot_Graphs_BC_CG_Together_line.R*). The *clx*.R* scripts perform the bootstrapped clustering as a function. We also have a script to run the linear combinations (*linearCombo*),* the script to build the scatterplots and heatmaps (*individualVariance.R*),* a code sheet to describe the variables used and the supplementary materials displaying the analysis of self-reported strategies.

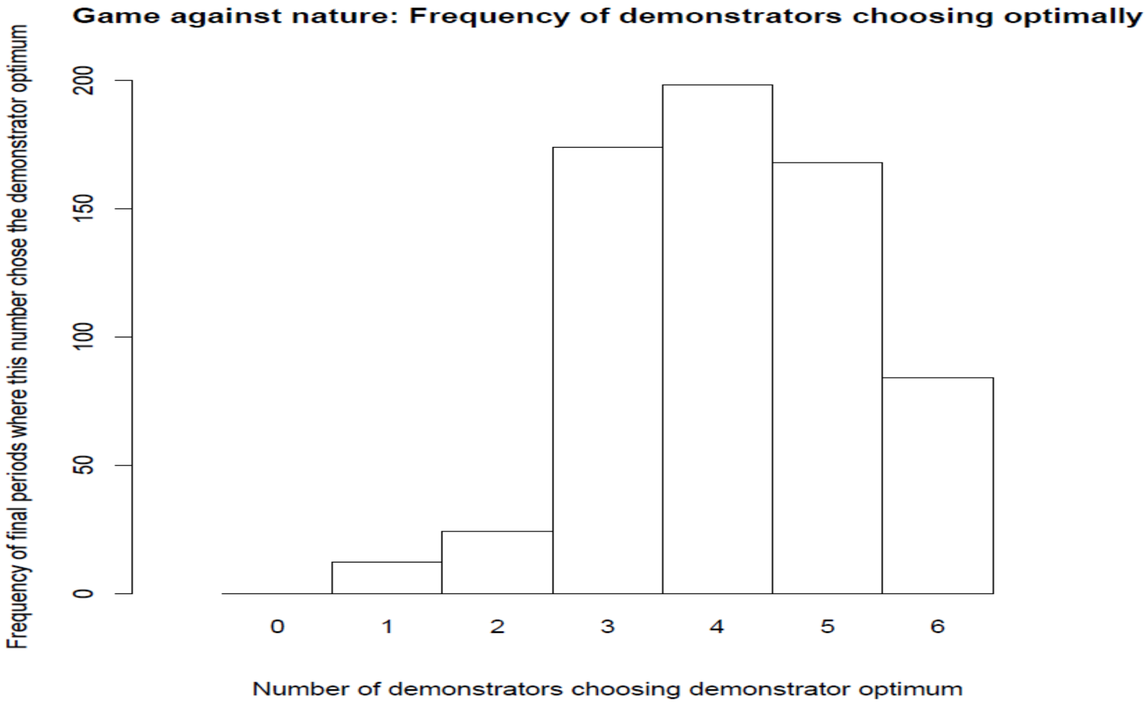
These analysis scripts can be accessed at the following link:

OSF | What is a social-learning strategy, anyway?

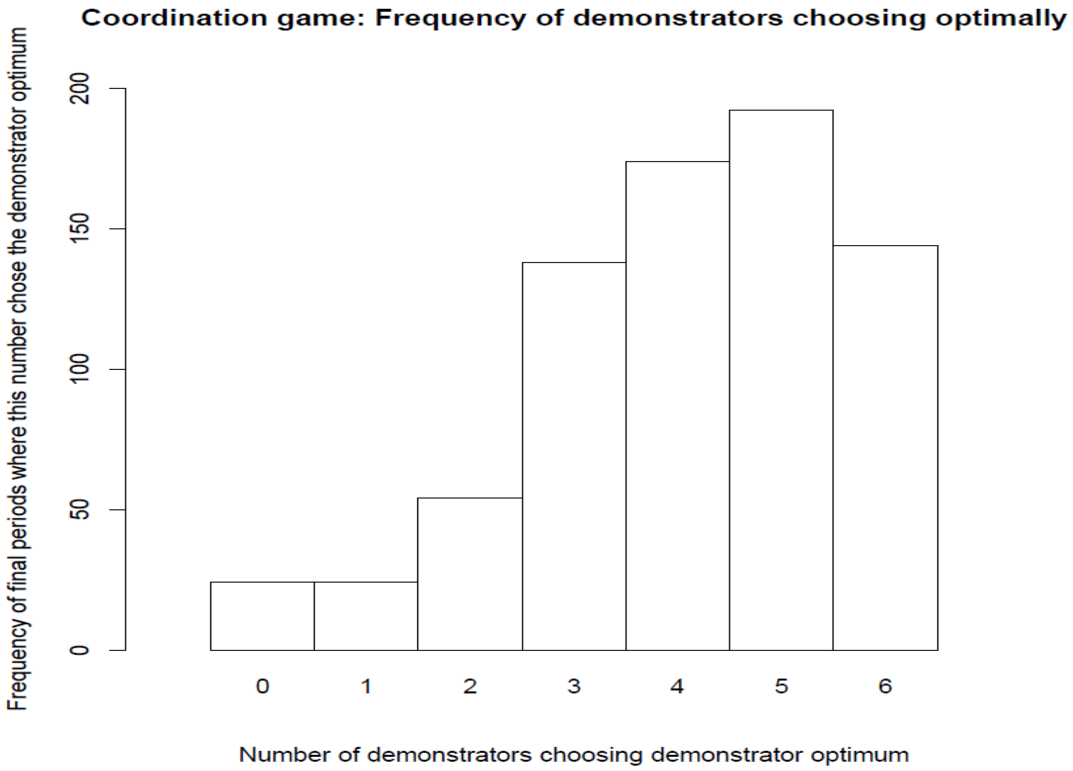
Appendix 6: Histograms displaying individual learner data

These histograms show the frequency of demonstrators who chose the demonstrator optimum distributed across the final periods of all blocks for a) the game against nature and b) the coordination game. The positive skew in both histograms shows that the demonstrator was more likely to answer optimally than not. This confirms that they did provide varied– but on the whole accurate– social information to the social learners.

A



B


Coordination game: Frequency of demonstrators choosing optimally
Frequency of final periods where this number chose the demonstrator optimum
0
50
100
150
200
0
1
2
3
4
5
6
Number of demonstrators choosing demonstrator optimum

Appendix 7: The logistic regression modelling whether the social learners chose % with controls.

Predictors included (i) the centred number of demonstrators who chose % on their final period, (ii) each combination of the similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, (iii) the interactions between each of these dummies and the centered proportion of demonstrators who chose % and (iv) demographic variables and other controls. The robust standard errors given in parentheses were clustered on the social learner to reflect the multiple observations gathered per learner.

Parameter	**Estimate (game against nature, with controls)**	**Estimate (coordination game, with controls)**
Intercept	-0.895 (1.267) 95% CI [-3.37, 1.58]	-0.168 (0.599) 95% CI [-1.34, 1.00]
Centred proportion of demonstrators choosing %	-1.530 ** (0.573) 95% CI [-2.65, --0.41]	-1.685 ** (0.557) 95% CI [-2.77, -0.60]
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	-0.177 (0.145) 95% CI [-0.46, 0.11]	-0.242 (0.208) 95% CI [-0.65, 0.16]
Uninformative-same dummy [signal indicates same and is correct with 0.5 probability]	-0.240 (0.147) 95% CI [-0.53, 0.05]	0.032 (0.191) 95% CI [-0.34, 0.40]
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	-0.274 . (0.160) 95% CI [-0.59, 0.04]	-0.244 (0.176) 95% CI [-0.59, 0.10]

Reliably correct-same dummy [signal indicates same and is correct with 0.9 probability]	0.013 (0.154) 95% CI [-0.29, 0.31]	-0.255 (0.219) 95% CI [-0.68, 0.17]
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.039 (0.140) 95% CI [-0.31, 0.23]	-0.053 (0.193) 95% CI [-0.43, 0.32]
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	2.274 ** (0.750) 95% CI [0.81, 3.74]	3.352 *** (0849) 95% CI [1.69, 5.01]
Centred proportion of demonstrators choosing % X uninformative-same dummy	4.167 *** (0.803) 95% CI [2.60, 5.74]	4.093 *** (0.838) 95% CI [2.46, 5.73]
Centred proportion of demonstrators choosing % X uninformative-different dummy	1.710 * (0.693) 95% CI [0.36, 3.06]	1.883 * (0.745) 95% CI [0.43, 3.43]
Centred proportion of demonstrators choosing % X reliably correct-same dummy	5.692 *** (1.009) 95% CI [3.72, 7.66]	6.997 *** (1.132) 95% CI [4.79, 9.21]
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-0.473 (0.643) 95% CI [-1.73, 0.78]	-0.356 (0.700) 95% CI [-1.72, 1.01]
Percentage as optimal dummy [signal indicates percentage is optimal option]	0.044 (0.089) 95% CI [-0.13, 0.22]	0.020 (0.103) 95% CI [-0.18, 0.22]
Age	0.008 (0.030) 95% CI [-0.05, 0.07]	0.033 (0.028) 95% CI [-0.02, 0.09]
Gender	-0.104 (0.161) 95% CI [-0.42, 0.21]	-0.04 (0.155) 95% CI [-0.34, 0.26]
Time in residence	0.007 (0.011)	-0.011 (0.009)

	95% CI [-0.01, 0.03]	95% CI [-0.03, 0.01]
Block Index	-0.016 * (0.006) 95% CI [-0.03, 0.00]	0.001 (0.007) 95% CI [-0.01, 0.01]
India dummy	1.214 (1.066) 95% CI [-0.87, 3.30]	-

The asterisks denote the level of significance of our p values, with the following key:

**** (p<0.001)*

*** (p<0.01)*

** (p<0.05)*

· *(trend: p=0.05 – 0.10 significance)*

Note that the only significant control predictor was Block Index for the game against nature. That is, the social learners were less likely to choose % in the later blocks of the game against nature. The analysis script in appendix 5 confirmed a non-significant trend for Game Version Left to be less likely to be played in the later periods of the sessions. That is, % was less likely to be optimal as the games progressed and so the social learners' choices may be appropriate to the random effects of the code used.

Appendix 8: Linear combinations

Appendix 8A: The code used to calculate the linear combinations.

The code to calculate the linear combinations *(linearCombo_USE_2021.R)* can be found at the following link:

OSF | What is a social-learning strategy, anyway?

Appendix 8B: The linear combinations produced for social-learner choices for the game against nature.

Note there is a significant difference between the social learners' strategies in response to all levels of social information, with the exception of different-reliably incorrect versus different-uninformative signals. The social learners' strategies to these two signals were not significantly distinct (see Figure 5D and 5E in the main text).

<u>Second-order (similar versus different)</u>

- Reliably incorrect signals (0/6:$F(1,2628)=11.33$, $p=0.0008$;

  6/6: $F(1,2628)= 6.39$, $p=0.01$).

  5/6: $F(1,2628)=4.59$, $p=0.03$)

  4/6: $F(1,2628)=1.33$, $p=0.25$

- Uninformative signals (0/6: $F(1,2628)=9.88$, $p=0.002$;

  6/6: $F(1,2628)=9.63$, $p=0.002$.

  5/6: $F(1,2628)=8.00$, $p=0.005$

  4/6: $F(1,2628)=4.35$, $p=0.04$

- Reliably correct signals (0/6: $F(1,2628) = 40.79$, $p<0.001$;

  6/6: $F(1,2628)=37.00$, $p<0.001$).

5/6: $F(1,2628)=32.07$, $p<0.001$

4/6: $F(1,2628)=19.54$, $p<0.001$

Third-order (comparing reliability)

- Similar, reliably incorrect vs uninformative (0/6: $F(1,2628) = 32.25$, $p<0.001$;

  6/6: $F(1,2628) = 17.25$, $p<0.001$).

  5/6: $F(1,2628)=12.84$, $p=0.003$.

  4/6: $F(1,2628)=4.60$, $p=0.03$

- Similar, reliably incorrect vs correct (0/6: $F(1,2628) = 32.66$, $p<0.001$;

  6/6: $F(1,2628) = 28.38$, $p<0.001$).

  5/6: $F(1,2628)=25.07$, $p<0.001$

  4/6: $F(1,2628) =16.13$, $p<0.001$

- Similar, uninformative vs reliably correct (0/6: $F(1,2628) = 2.02$, $p=0.16$;

  6/6: $F(1,2628)=5.02$, $p=0.025$.

  5/6: $F(1,2628)=5.17$, $p=0.02$

  4/6: $F(1,2628)=5.04$, $p=0.02$

- Different, reliably incorrect vs uninformative (0/6: $F(1,2628) = 0.38$, $p=0.54$;

  6/6: $F(1,2628)=2.15$, $p=0.14$

  5/6: $F(1,2628)=2.16$, $p=0.14$

  4/6: $F(1,2628)=1.66$, $p=0.20$

- Different, reliably incorrect vs correct (0/6: $F(1,2628) = 16.24$, $p<0.001$;

  6/6: $F(1,2628) = 12.72$, $p=0.0004$).

  5/6: $F(1,2628)= 9.86$, $p=0.002$

  4/6: $F(1,2628)=3.72$, $p=0.054$

- Different, uninformative vs reliably correct (0/6: F(1,2628) = 14.42, p=0.0001;

  6/6: F(1,2628) = 7.42, p=0.006)

  5/6: F(1,2628)=4.37, p=0.04

  4/6: F(1,2628)=0.53, p=0.47

Appendix 8C: The linear combinations produced for the social-learner choices for the coordination game.

Note that the social learner shows a significantly distinct strategy for each level of the similarity and reliability information.

<u>Second-order (similar versus different)</u>

- Reliably incorrect signals (0/6: F(1,2372)=11.14, p=0.0009;

  6/6: F(1,2372) = 14.71, p<0.001

  5/6: F(1,2372)=12.14, p=0.0005

  4/6: F(1,2372)=4.66, p=0.03

- Uninformative signals (0/6: F(1,2372)=9.54, p=0.002;

  6/6: F(1,2372) = 16.09, p<0.001

  5/6: F(1,26372)=14.53, p=0.0001

  4/6: F(1,2372)=8.47,p=0.004

- Reliably correct signals (0/6: F(1,2372) =46.85, p<0.001;

  6/6: F(1,2372) = 50.89, p<0.001

  5/6: F(1,2372)=44.78, p<0.001

  4/6:F(1,2372)=25.73, p<0.001

Third-order (comparing reliability)

- Similar, reliably incorrect vs uninformative (0/6: $F(1,2372) = 23.89$, $p<0.001$;

  6/6: $F(1,2372)=31.77$, $p<0.001$

  5/6: $F(1,2372)=28.14$, $p<0.001$

  4/6: $F(1,2372)=15.63$, $p<0.001$

- Similar, reliably incorrect vs correct (0/6: $F(1,2372)=37.95$, $p<0.001$;

  6/6: $F(1,2372) = 41.69$, $p<0.001$

  5/6: $F(1,2372)=35.73$, $p<0.001$

  4/6: $F(1,2372)=18.7$, $p<0.001$

- Similar, uninformative vs reliably correct (0/6: $F(1,2372) =5.85$, $p=0.02$;

  6/6: $F(1,2372)=7.99$, $p=0.005$

  5/6: $F(1,2372)=6.79$, $p=0.009$

  4/6: $F(1,2372)=2.83$, $p=0.09$

- Different, reliably incorrect vs uninformative (0/6: $F(1,2372)=3.02$, $p=0.08$;

  6/6:$F(1,2372)=4.45$, $p=0.04$

  5/6: $F(1,2372)=3.25$, $p=0.07$

  4/6: $F(1,2372)=0.97$,$p=0.33$

- Different, reliably incorrect vs correct (0/6: $F(1,2372) = 14.66$, $p<0.001$;

  6/6:$F(1,2372)=26.77$, $p<0.001$

  5/6: $F(1,2372)=23.16$, $p<0.001$

  4/6: $F(1,2372)=9.14$, $p=0.003$

- Different, uninformative vs reliably correct (0/6: $F(1,2372) =8.90$, $p=0.003$;

  6/6: $F(1,2372)=19.56$, $p<0.001$

5/6: F(1,2372)=14.84, p=0.0001

4/6: F(1,2372)=4.60, p=0.03

Appendix 9: The regressions predicting social-learner optimality with control predictors included.

Appendix 9A: Logistic regression modelling whether the social learner chose the social learner optimum for the game against nature, with controls.

Predictors included: (i) the centered proportion of demonstrators who chose the demonstrator optimum, (ii) dummies for each combination of similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose the demonstrator optimum and (iv) demographic variables and other control predictors. Robust standard error clustered on social learner.

Parameter	Estimate (game against nature, all signals, controls)	Estimate (game against nature, correct signals, controls)	Estimate (game against nature, incorrect signals, controls)
Intercept	-0.250 (0.478) 95% CI [-1.18, 0.68]	-0.408 (0.474) 95% CI [-1.33, 0.52]	-0.232 (0.4442) 95% CI [-1.10, 0.63]
Centred proportion of demonstrators choosing demonstrator optimum	0.270 (0.637) 95% CI [-0.97, 1.51]	-0.428 (0.308) 95% CI [-1.03, 0.17]	1.537 *** (0.323) 95% CI [0.91, 2.17]
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	-0.077 (0.182) 95% CI [-0.43, 0.28]	-0.124 (0.546) 95% CI [-1.19, 0.94]	-0.008 (0.131) 95% CI [-0.26, 0.25]

Uninformative-same dummy [signal indicates same and is correct with 0.5 probability]	-0.200 (0.202) 95% CI [-0.59, 0.19]	-0.177 (0.218) 95% CI [-0.60, 0.25]	-0.109 (0.215) 95% CI [-0.53, 0.31]
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	-0.157 (0.195) 95% CI [-0.54, 0.22]	-0.136 (0.187) 95% CI [-0.50, 0.23]	-0.353 * (0.178) 95% CI [-0.70, -0.01]
Reliably correct-same dummy [signal indicates same and is correct with 0.9 probability]	-0.205 (0.226) 95% CI [-0.65, 0.24]	0.044 (0.170) 95% CI [-0.29, 0.38]	0.980 . (0.575) 95% CI [-2.10, 0.14]
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.055 (0.194) 95% CI [-0.43, 0.32]	0.094 (0.145) 95% CI [-0.19, 0.38]	-0.141 (0.529) 95% CI [-1.17, 0.89]
Centred proportion of demonstrators choosing optimum X reliably incorrect-different dummy	0.079 (0.814) 95% CI [-1.51, 1.67]	-1.348 (2.107) 95% CI [-5.46, 2.77]	-1.040 * (0.526) 95% CI [-2.07, -0.01]
Centred proportion of demonstrators choosing optimum X uninformative-same dummy	-0.698 (0.836) 95% CI [-2.33, 0.93]	3.496 *** (1.027) 95% CI [1.49, 5.50]	-4.255 *** (0.942) 95% CI [-6.09, -2.41]
Centred proportion of demonstrators choosing optimum X uninformative-different dummy	-0.747 (0.806) 95% CI [-2.32, 0.83]	0.076 (0.690) 95% CI [-1.27, 1.42]	-2.385 *** (0.525) 95% CI [-3.41, -1.36]
Centred proportion of demonstrators choosing optimum X reliably correct-same dummy	2.346 * (0.920) 95% CI [0.55, 4.14]	4.016 *** (0.892) 95% CI [2.27, 5.76]	-4.356 (2.835) 95% CI [-9.90, 1.18]
Centred proportion of demonstrators choosing optimum X reliably correct-different dummy	0.766 (0.781) 95% CI [-0.76, 2.29]	1.737 ** (0.617) 95% CI [0.53, 2.94]	-4.518 . (2.589) 95% CI [-9.57, 0.54]
Percentage as optimal dummy [signal indicates percentage is optimal option]	0.398 ** (0.145) 95% CI [0.11, 0.68]	0.458 ** (0.148) 95% CI [0.17, 0.75]	0.390 ** (0.150) 95% CI [0.10, 0.68]
Age	0.007 (0.184)	0.009 (0.019)	0.003 (0.018)

	95% CI [-0.03, 0.04]	95% CI [-0.03, 0.05]	95% CI [-0.03, 0.04]
Gender	0.164 . (0.089)	0.162 . (0.091)	0.174 (0.090)
	95% CI [-0.01, 0.34]	95% CI [-0.02, 0.0]	95% CI [0, 0.35]
Time in residence	-0.011 (0.007)	-0.010 (0.007)	-0.012 (0.007)
	95% CI [-0.03, 0.00]	95% CI [-0.02, 0]	95% CI [-0.03, 0]
Block Index	0.009 (0.006)	0.008 (0.007)	0.008 (0.006)
	95% CI [0.0, 0.02]	95% CI [0, 0.02]	95% CI [-0.01, 0.02]
India dummy	0.065 (0.262)	0.024 (0.257)	0.099 (0.246)
	95% CI [-0.45, 0.58]	95% CI [-0.48, 0.53]	95% CI [-0.38, 0.58]

The asterisks denote the level of significance of our p values, with the following key:

**** (p<0.001)*

*** (p<0.01)*

** (p<0.05)*

· *(trend: p=0.05 – 0.10 significance)*

For the blocks containing just incorrect similarity signals (right-hand column of Table 8A), the social learners' were less likely to coordinate on their social learner optimum in response to uninformative-different blocks. They were also less likely to answer optimally as more demonstrators coordinated on the demonstrator optimum for reliably incorrect-different blocks. This suggests that the learner shows a meaningful adjustment to both reliably incorrect and uninformative signals from different others, but this only pays off when the blocks happen to give incorrect information. Otherwise, the effects match those depicted in table 4 of the main text.

The only significant control predictor is Percentage as optimal dummy. That is, the social learners were significantly more likely to answer optimally when %

happened to be optimal. This suggests a bias to choose % rather than @ throughout the game against nature. This suggests an arbitrary preference for the % symbol.

Appendix 9B: The logistic regression modelling whether the social learners chose the social learner optimum for the coordination game, with controls.

Predictors included: (i) the centered proportion of demonstrators who chose the demonstrator optimum, (ii) dummies for each combination of similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose the demonstrator optimum and (iv) demographic variables and control predictors. Robust standard error clustered on social learner.

Parameter	**Estimate (Coordination game, all signals, controls)**	**Estimate (Coordination game, correct signals, controls)**	**Estimate (coordination game, incorrect signals, controls**
Intercept	-0.633 (0.518) 95% CI [-1.64, 0.38]	-0.439 (0.491) 95% CI [-1.40, 0.52]	-0.328 (0.489) 95% CI [-1.28, 0.63]
Centred proportion of demonstrators choosing demonstrator optimum	1.184 * (0.493) 95% CI [0.22, 2.15]	0.067 (0.232) 95% CI [-0.39, 0.52]	1.879 *** (0.305) 95% CI [1.28, 2.47]
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	0.244 (0.178) 95% CI [-0.10, 0.59]	0.903 (0.663) 95% CI [-0.39, 2.20]	0.203 (0.159) 95% CI [-0.11, 0.51]
Uninformative-same dummy [signal indicates same and is correct with 0.5 probability]	-0.086 (0.184) 95% CI [-0.45, 0.27]	-0.488 * (0.223) 95% CI [-0.92, -0.05]	-0.274 (0.217) 95% CI [-0.70, 0.15]

Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	0.020 (0.190) 95% CI [-0.35, 0.39]	-0.368 (0224) 95% CI [-0.81, 0.07]	0.374 (0.235) 95% CI [-0.08, 0.83]
Reliably correct-same dummy [signal indicates same and is correct with 0.9 probability]	0.678 ** (0.252) 95% CI [0.19, 1.17]	0.216 (0.547) 95% CI [-0.28, 0.71]	0.182 (0.297) 95% CI [-0.40, 0.76]
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.033 (0.189) 95% CI [-0.40, 0.34]	-0.143 (0.148) 95% CI [-0.43, 0.15]	0.928 (1.039) 95% CI [-1.10, 2.96]
Centred proportion of demonstrators choosing optimum X reliably incorrect-different dummy	-0.218 (0.681) 95% CI [-1.55, 1.11]	-2.615 (2.02) 95% CI [-6.56, 1.33]	-0.707 (0.577) 95% CI [-1.83, 0.42]
Centred proportion of demonstrators choosing optimum X uninformative-same dummy	-1.214 * (0.609) 95% CI [-2.40, -0.02]	3.378 *** (0.835) 95% CI [1.75, 5.01]	-4.044 *** (0.812) 95% CI [-5.63, -2.46]
Centred proportion of demonstrators choosing optimum X uninformative-different dummy	-0.876 (0.673) 95% CI [-2.19, 0.44]	0.664 (0.749) 95% CI [-0.80, 2.13]	-2.359 * (0.950) 95% CI [-4.21, -0.50]
Centred proportion of demonstrators choosing optimum X reliably correct-same dummy	-0.209 (0.835) 95% CI [-1.84, 1.42]	4.313 *** (1.143) 95% CI [2.08, 6.54]	-6.955 *** (1.761) 95% CI [-10.39, -3.52]
Centred proportion of demonstrators choosing optimum X reliably correct-different dummy	0.654 (0.635) 95% CI [-0.59, 1.89]	1.934 *** (0.572) 95% CI [0.82, 3.05]	-6.080 * (2.739) 95% CI [-11.43, -0.73]
Percentage as optimal dummy [signal indicates percentage is optimal option]	0.706 *** (0.14) 95% CI [0.41, 1.00]	0.656 *** (0.144) 95% CI [0.37, 0.94]	0.660 *** (0.148) 95% CI [0.37, 0.95]
Age	0.003 (0.235) 95% CI [-0.04, 0.05]	0.00002 (0.023) 95% CI [-0.05, 0.05]	-0.011 (0.023) 95% CI [-0.06, 0.03]

Gender	0.128 (0.108) 95% CI [-0.08, 0.34]	0.129 (0.109) 95% CI [-0.08, 0.34]	0.094 (0.112) 95% CI [-0.13, 0.31]
Time in residence	-0.005 (0.008) 95% CI [-0.02, -0.01]	-0.0018 (0.751) 95% CI [-0.02, 0.01]	0.007 (0.008) 95% CI [-0.01, 0.02]
Block Index	0.011 (0.007) 95% CI [0, 0.02]	0.014 * (0.007) 95% CI [0, 0.03]	0.012 . (0.007) 95% CI [0, 0.03]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

On blocks with incorrect similarity information (right-hand column of Table 8B), the social learners were less likely to answer optimally as more demonstrators did for uninformative-different signals, suggesting that the learners in the full sessions only did show a meaningful adjustment to this signal.

On blocks with correct signals (middle column of Table 8B), the social learners were less likely to coordinate on the social-learner optimum for reliably incorrect-different signals, which suggests that the social learners treat these signals as if they were always incorrect.

On blocks with all correct and incorrect signals collapsed together (left-hand column of Table 8B), the social learners are more likely to coordinate on the social learner optimum as more demonstrators answer optimally (Centered proportion of demonstrators choosing the demonstrator optimum effect). This suggests that the learners respond meaningfully to frequency-dependent social information for the omitted category of reliably incorrect signals from similar others. They are also the most likely to answer optimally to blocks with reliably correct-similar signals. All other significant effects match those displayed in table 4 of the main text, thus confirming similarity between the basic model and full model with controls.

Finally, there is a significant bias to answer optimally if percentage happened to be optimal (Percentage as optimal dummy effect). This suggests a bias whereby the learner followed a rule of 'just choose %', perhaps due to the pressure to coordinate. Secondly, the social learners were more likely to answer optimally as the blocks progressed for blocks where the similarity signals happened to be correct only (Block index effect). This suggests a rate of learning over the blocks as the learners became more used to the coordination game.

Appendix 10: Linear combinations comparing the social-learners' ability to respond optimally across each level of the social information, for the game against nature.

Note that the social learners found it easier to respond optimally to reliably-similar than reliably-different others and found it easier to respond to reliably correct-similar than uninformative-similar or reliably incorrect-similar others. These biases persist for all signals and correct signals.

<u>All signals, game against nature</u>

<u>Second-order social information</u>

Reliably incorrect: 0/6: $F(1,2628)= 0.36$, $p=0.55$

6/6: $F(1,2628)=0.02$, $p=0.88$

5/6: $F(1,2628)=0.003$, $p=0.96$

4/6: $F(1,2628)=0.24$, $p=0.62$

Uninformative: 0/6: $F(1,2628)=0.06$, $p=0.81$

6/6: $F(1,2628)<0.001$, $p=0.98$

5/6: $F(1,2628)=0.02$, $p=0.88$

4/6: $F(1,2628)=0.12$, $p=0.73$

Reliably correct: 0/6: $F(1,2628)=4.72$, $p=0.03$

6/6: $F(1,2628)=4.86$, $p=0.03$

5/6: $F(1,2628)=3.48$, $p=0.06$

4/6: $F(1,2628)=0.26$, $p=0.61$

<u>Third-order social information</u>

Similar others, reliably incorrect versus uninformative. 0/6: $F(1,2628)=0.13$, $p=0.72$

6/6: $F(1,2628)=3.41$, $p=0.07$

5/6: $F(1,2628)=4.76$, $p=0.03$

4/6: $F(1,2628)=4.50$, $p=0.03$

Similar others, reliably incorrect versus correct: 0/6: $F(1,2628)=5.91$, $p=0.02$

6/6: $F(1,2628)=8.86$, $p=0.002$

5/6: $F(1,2628)=7.30$, $p=0.007$

4/6: $F(1,2628)=0.88$, $p=0.35$

Similar others, uninformative vs reliably correct: 0/6: $F(1,2628)=9.74$, $p=0.002$

6/6: $F(1,2628)=22.10$, $p<0.001$

5/6: $F(1,2628)=22.31$, $p<0.001$

4/6: $F(1,2628)=10.52$, $p=0.001$

Different others, reliably incorrect versus uninformative: 0/6: $F(1,2628)=1.95$, $p=0.16$

6/6: $F(1,2628)=5.62$, $p=0.02$

5/6: $F(1,2628)=5.03$, $p=0.02$

4/6: $F(1,2628)=1.89$, $p=0.17$

Different others, reliably incorrect versus correct: 0/6: $F(1,2628)=0.04$, $p=0.84$

6/6: $F(1,2628)=0.83$, $p=0.36$

5/6: $F(1,2628)=1.09$, $p=0.30$

4/6: $F(1,2628)=1.10$, $p=0.29$

Different others, uninformative versus reliably correct: 0/6: $F(1,2628)=1.88$, $p=0.17$

6/6: $F(1,2628)=10.14$, $p=0.01$

5/6: $F(1,2628) =11.89$, $p=0.0006$

4/6: $F(1,2628)=6.74$, $p=0.009$

Just correct signals, game against nature

Second order

Reliably incorrect. 0.6: $F(1,2628) = 0.07$, $p=0.79$)

6/6: $F(1,2628)=0.63$, $p=0.43$

5/6: $F(1,2628)=0.75$, $p=0.39$

4/6: $F(1,2628)=0.56$, $p=0.45$

Uninformative: 0/6: $F(1,2628) = 7.52$, $p=0.006$

6/6: $F(1,2628) = 17.98$, $p<0.001$

5/6: $F(1,2628)16.26$, $p<0.001$

4/6: $F(1,2628)=5.73$, $p=0.02$

Reliably correct. 0/6: $F(1,2628) = 5.61$, $p=0.02$

6/6: $F(1,2628)=8.91$, $p=0.003$

5/6: $F(1,2628)=8.41$, $p=0.004$

4/6: $F(1,2628) =3.47$, $p=0.06$

Third order

Similar others, reliably incorrect vs uninformative. 0/6: $F(1,2628)=8.47$, $p=0.004$

6/6: $F(1,2628)=13.18$, $p=0.0003$

5/6: $F(1,2628)=11.62$, $p=0.0007$

4/6: $F(1,2628)=4.16$, $p=0.04$

Similar others, reliably incorrect vs correct. 0/6: $F(1,2628)=14.46$, $p=0.0001$)

6/6: $F(1,2628)=31.09$, $p<0.001$

5/6: $F(1,2628)=32.60$, $p<0.001$

4/6: $F(1,2628) =21.56$, $p<0.001$

Similar others, uninformative vs reliably correct. 0/6: $F(1,2628)=0.16$, $p=0.69$

6/6: $F(1,2628)=1.62$, $p=0.20$

5/6: F(1,2628)=2.15, p=0.14

4/6: F(1,2628) = 2.32, p=0.13

Different others, reliably incorrect vs uninformative. 0/6: F(1,2628)=0.04, p=0.84

6/6: F(1,2628)=0.29, p=0.59

5/6: F(1,2628) =0.33, p=0.56

4/6: F(1,2628)=0.22, p=0.64

Different others, reliably incorrect vs correct. 0/6: F(1,2628)=0.5, p=0.48.

6/6: F(1,2628)=2.98, p=0.08

5/6: F(1,2628)=3.41, p=0.06

4/6: F(1,2628)=2.56, p=0.11

Different others, uninformative vs reliably correct. 0/6: F(1,2628)=1.55, p=0.21)

6/6: F(1,2628)=11.43, p=0.0007

5/6: F(1,2628)=13.51, p=0.0002

4/6: F(1,2628)=9.35, p=0.002

<u>Just incorrect signals, game against nature</u>

<u>Second order</u>

Reliably incorrect. 0/6: F(1,2628)=0.54, p=0.46

6/6: F(1,2628)=1.40, p=0.24

5/6: F(1,2628)=1.35, p=0.25

4/6: F(1,2628)=0.81, p=0.37

Uninformative. 0/6: F(1,2628)=1.60, p=0.21

6/6: F(1,2628)=6.98, p=0.008

5/6: F(1,2628)=8.50, p=0.004

4/6: F(1,2628)=5.76, p=0.02

Reliably correct. 0/6: $F(1,2628)=0.10$, $p=0.76$

6/6: $F(1,2628)=0.47$, $p=0.49$

5/6: $F(1,2628)=1.36$, $p=0.24$

4/6: $F(1,2628)=2.93$, $p=0.09$

<u>Third order</u>

Similar others, reliably incorrect vs uninformative. 0/6: $F(1,2628)=10.85$, $p=0.001$

6/6: $F(1,2628)=32.17$, $p<0.001$

5/6: $F(1,2628)=33.57$, $p<0.001$

4/6: $F(1,2628)=18.00$, $p<0.001$

Similar others, reliably incorrect vs correct. 0/6: $F(1,2628)=0.37$, $p=0.54$

6/6: $F(1,2628)=9.13$, $p=0.003$

5/6: $F(1,2628)=15.70$, $p<0.001$

4/6: $F(1,2628)=21.79$, $p<0.001$

Similar others, uninformative vs reliably correct: 0/6: $F(1,2628)=0.20$, $p=0.65$

6/6: $F(1,2628)=0.89$, $p=0.34$

5/6: $F(1,2628)=2.60$, $p=0.11$

4/6: $F(1,2628) = 6.78$, $p=0.009$

Different others, reliably incorrect vs uninformative. 0/6: $F(1,2628)=1.37$, $p=0.24$

6/6: $F(1,2628)=3.38$, $p=0.07$

5/6: $F(1,2628)=2.93$, $p=0.09$

4/6: $F(1,2628)=0.70$, $p=0.40$

Different others, reliably incorrect vs correct: 0/6: $F(1,2628)=0.90$, $p=0.34$

6/6: $F(1,2628)=3.96$, $p=0.05$

5/6: $F(1,2628)=5.02$, $p=0.03$

4/6: $F(1,2628)=3.50$, $p=0.06$

Different others, uninformative vs reliably correct. 0/6: F(1,2628)=0.25, p=0.62

6/6: F(1,2628)=1.76, p=0.18

5/6: F(1,2628)=2.49, p=0.11

4/6: F(1,2628)=2.18, p=0.14

Appendix 11: Linear combinations comparing the social-learners' ability to respond optimally across each level of the social information, for the coordination game.

Note that the bias to respond optimally to reliably-similar signals only exists on trials where the majority of demonstrators chose % and not for 0/6, and so this bias is less pronounced in the coordination game than the game against nature.

<u>All signals, coordination game</u>

<u>Second order</u>

Reliably incorrect. 0/6: F(1,2372)= 0.25, p=0.62

6/6: F(1,2372)=4.59, p=0.03

5/6: F(1,2372)=5.53, p=0.02

4/6: F(1,2372)= 4.74, p=0.03

Uninformative: 0/6: F(1,2372)=3.39, p=0.07

6/6: F(1,2372)=5.28, p=0.02

5/6: F(1,2372)=3.59, p=0.06

4/6: F(1,2372)=0.46, p=0.50

Reliably correct. 0/6: F(1,2372)=0.12, p=0.73

6/6: F(1,2372)=5.77, p=0.02

5/6: F(1,2372) =7.65, p=0.006

4/6: F(1,2372)=7.03, p=0.008

Third order

Similar, reliably incorrect vs uninformative. 0/6: $F(1,2372)=2.17$, $p=0.14$

6/6: $F(1,2372)=5.07$, $p=0.02$

5/6: $F(1,2372)=4.50$, $p=0.03$

4/6: $F(1,2372)=2.23$, $p=0.14$

Similar others, reliably incorrect vs correct. 0/6: $F(1,2372)=0.78$, $p=0.38$.

6/6: $F(1,2372)=10.8$, $p=0.001$.

5/6: $F(1,2372)=13.49$, $p=0.0002$

4/6: $F(1,2372)=10.05$, $p=0.002$

Similar others, uninformative vs reliably correct. 0/6: $F(1,2372)=4.48$, $p=0.03$.

6/6: $F(1,2372)=33.51$, $p<0.001$.

5/6: $F(1,2372)=38.87$, $p<0.001$

4/6: $F(1,2372)=25.12$, $p<0.001$

Different others, reliably incorrect vs uninformative. 0/6: $F(1,2372)=0.04$, $p=0.83$.

6/6: $F(1,2372)=7.87$, $p=0.005$

5/6: $F(1,2372)=10.05$, $p=0.002$

4/6: $F(1,2372)=9.58$, $p=0.002$

Different others, reliably incorrect vs correct. 0/6: $F(1,2372)=0.06$, $p=0.80$

6/6: $F(1,2372)=1.34$, $p=0.24$

5/6: $F(1,2372)=2.01$, $p=0.16$

4/6: $F(1,2372)=2.57$, $p=0.11$

Different others, uninformative vs reliably correct. 0/6: $F(1,2372)=0.21$, $p=0.64$.

6/6: $F(1,2372)=2.58$, $p=0.11$

5/6: $F(1,2372)=2.89$, $p=0.09$

4/6: $F(1,2372)=2.30$, $p=0.13$

Just correct signals, coordination game

Second-order

Reliably incorrect. 0/6: $F(1,2372) = 4.55$, $p=0.03$)

6/6: $F(1,2372) = 0.63$, $p=0.43$

5/6: $F(1,2372)<0.001$, $p=0.99$

4/6: $F(1,2372)=1.77$, $p=0.18$

Uninformative. 0/6: $F(1,2372) = 11.00$, $p=0.0009$

6/6: $F(1,2372) = 15.54$, $p<0.001$

5/6: $F(1,2372)=12.46$, $p=0.0004$

4/6: $F(1,2372)=3.93$, $p=0.05$

Reliably correct: 0/6: $F(1,2372) = 11.15$, $p=0.0009$

6/6: $F(1,2372) = 21.05$, $p<0.001$

5/6: $F(1,2372)=20.27$, $p<0.001$

4/6: $F(1,2372)= 11.51$, $p=0.0007$

Third order

Similar, reliably incorrect vs uninformative. 0/6: $F(1,2372)=20.19$, $p<0.001$

6/6: $F(1,2372)=15.56$, $p<0.001$

5/6; $F(1,2372)=9.77$, $p=0.002$

4/6: $F(1,2372)=0.67$, $p=0.41$

Similar, reliably incorrect vs correct. 0/6: $F(1,2372)=25.06$, $p<0.001$)

6/6: $F(1,2372)=42.73$, $p<0.001$

5/6: $F(1,2372)=41.05$, $p<0.001$

4/6: $F(1,2372)=22.17$, $p<0.001$

Similar, uninformative vs reliably correct. 0/6: $F(1,2372)=0.46$, $p=0.50$

6/6: $F(1,2372)=5.98$, $p=0.01$

5/6: $F(1,2372)=7.44$, $p=0.006$

4/6: $F(1,2372)=7.57$, $p=0.006$

Different, reliably incorrect vs uninformative. 0/6: $F(1,2372)=5.12$, $p=0.03$

6/6: $F(1,2372)=0.27$, $p=0.60$

5/6: $F(1,2372)=0.21$, $p=0.65$

4/6: $F(1,2372)=3.67$, $p=0.06$

Different, reliably incorrect vs correct. 0/6: $F(1,2372)=6.89$, $p=0.009$

6/6: $F(1,2372)=2.24$, $p=0.13$

5/6: $F(1,2372)=0.33$, $p=0.57$

4/6: $F(1,2372)=1.14$, $p=0.29$

Different, uninformative vs reliably correct. 0/6: $F(1,2372)=1.01$, $p=0.32$)

6/6: $F(1,2372)=9.25$, $p=0.002$.

5/6: $F(1,2372)-9.32$, $p=0.002$

4/6: $F(1,2372)=5.78$, $p=0.02$

Just incorrect signals, coordination game

Second order

Reliably incorrect. 0/6: $F(1,2372)=0.24$, $p=0.62$

6/6: $F(1,2372)=0.64$, $p=0.42$

5/6: $F(1,2372)=1.28$, $p=0.26$

4/6: $F(1,2372)=2.07$, $p=0.15$

Uninformative. 0/6: $F(1,2372)=8.65$, $p=0.003$

6/6: $F(1,2372)=29.84$, $p<0.001$

5/6: $F(1,2372)=30.73$, $p<0.001$

4/6: $F(1,2372)=15.77$, $p<0.001$

Reliably correct. 0/6: $F(1,2372)=0$, $p=0.99$

6/6: $F(1,2372)=1.53$, $p=0.22$

5/6: $F(1,2372)=2.19$, $p=0.14$

4/6: $F(1,2372)=2.19$, $p=0.13$

Third order

Similar, reliably incorrect vs uninformative. 0/6: $F(1,2372)=18.65$, $p<0.001$

6/6: $F(1,2372)=49.84$, $p<0.001$

5/6: $F(1,2372)=48.34$, $p<0.001$

4/6: $F(1,2372)=27.29$, $p<0.001$

Similar, reliably incorrect vs correct. 0/6: $F(1,2372)=18.31$, $p<0.001$

6/6: $F(1,2372)=16.64$, $p<0.001$

5/6: $F(1,2372)=13.89$, $p=0.0002$

4/6: $F(1,2372)=6.37$, $p=0.01$

Similar, uninformative vs reliably correct. 0/6: $F(1,2372)=2.96$, $p=0.09$

6/6: $F(1,2372)=0.81$, $p=0.37$

5/6: $F(1,2372)=0.35$, $p=0.56$

4/6: $F(1,2372)=0.01$, $p=0.91$

Different, reliably incorrect vs uninformative. 0/6: $F(1,2372)=0$, $p=0.99$

6/6: $F(1,2372)=1.55$, $p=0.21$

5/6: $F(1,2372)=2.13$, $p=0.14$

4/6: $F(1,2372)=2.31$, $p=0.13$

Different, reliably incorrect vs correct. 0/6: $F(1,2372)=5.90$, $p=0.02$

6/6: $F(1,237)=7.45$, $p=0.006$

5/6: $F(1,2372)=4.26$, $p=0.04$

4/6: $F(1,2372)=0.16$, $p=0.69$

Different, uninformative vs reliably correct. 0/6: $F(1,2372)=5.62$, $p=0.02$

6/6: $F(1,2372)=5.15$, $p=0.02$

5/6: $F(1,2372)=2.24$, $p=0.13$

4/6: $F(1,2372)=0.02$, $p=0.88$

Appendix 12: Logistic regressions predicting whether the social learners chose their social learner optimum during the game against nature in response to correct and incorrect similarity signals.

Predictors included: (i) the centered proportion of demonstrators who chose the demonstrator optimum, (ii) dummies for each combination of similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, and (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose the demonstrator optimum. Robust standard error clustered on social learner.

Parameter	Estimate (game against nature, correct signals)	Estimate (game against nature, wrong signals)
Intercept	0.083 (0.080) 95% CI [-0.07, 0.24]	0.110 . (0.065) 95% CI [-0.02, 0.24]
Centred proportion of demonstrators choosing the demonstrator optimum	-0.460 (0.318) 95% CI [-1.08, 0.16]	1.204 *** (0.303) 95% CI [0.61, 1.80]
Reliably incorrect- different dummy [indicates different and is correct with 0.1 probability]	-0.149 (0.524) 95% CI [-1.17, 0.88]	-0.021 (0.129) 95% CI [-0.27, 0.23]

Uninformative-same dummy [indicates same and is correct with 0.5 probability]	-0.172 (0.212) 95% CI [-0.59, 0.24]	-0.032 (0.216) 95% CI [-0.45, 0.39]
Uninformative-different dummy [indicates different and is correct with 0.5 probability]	-0.076 (0.183) 95% CI [-0.43, 0.28]	0.052 (0.212) 95% CI [-0.36, 0.47]
Reliably correct-same dummy [indicates same and is correct with 0.9 probability]	-0.020 (0.166) 95% CI [-0.16, 0.40]	-0.956 (0.582) 95% CI [-1.59, 0.50]
Reliably correct-different dummy [indicates different and is correct with 0.9 probability]	0.122 (0.144) 95% CI [-0.16, 0.40]	-0.134 (0.551) 95% CI [-1.21, 0.94]
Centred proportion of demonstrators choosing optimum X reliably incorrect-different dummy	-1.023 (2.037) 95% CI [-5.01, 2.96]	-0.541 (0.535) 95% CI [-1.59, 0.50]
Centred proportion of demonstrators choosing optimum X uninformative-same dummy	3.256 *** (0.952) 95% CI [1.39, 5.12]	-4.161 *** (0.930) 95% CI [-5.98, -2.34]
Centred proportion of demonstrators choosing optimum X uninformative-different dummy	-0.241 (0.681) 95% CI [-1.57, 1.09]	-1.931 * (0.863) 95% CI [-3.62, -0.24]
Centred proportion of demonstrators choosing optimum X reliably correct-same dummy	4.141 *** (0.875) 95% CI [2.43, 5.85]	-4.228 (2.825) 95% CI [-9.75, 1.30]
Centred proportion of demonstrators choosing optimum X reliably correct-different dummy	1.576 * (0.618) 95% CI [0.37, 2.79]	-4.154 (2.594) 95% CI [-9.23, 0.92]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

For blocks that gave incorrect similarity information, the social learners were more likely to respond optimally as more demonstrators did (centred proportion of demonstrators choosing the demonstrator optimum effect). This effect reflected the omitted category of reliably incorrect - similar dummy. As a reliably incorrect signal of similarity was only correct with a probability of 0.1, then smart social learners should have treated this signal as if it were always incorrect. Reliably incorrect-similar signals can be flipped to understand that the learners were more likely to be playing a different game version to the demonstrators. Indeed, the social learners followed the minority of the demonstrators under both reliably incorrect-similar and reliably correct-different signals (Figures 5A and 5F), suggesting that they understood this complexity. Of course, the strategy of following the minority on reliably incorrect-similar blocks would only work if the similarity signal was in fact incorrect. The probabilities implemented meant that the social learners would have played the same game version as the demonstrators on approximately 10% of blocks with these reliably incorrect-similar blocks. In this case, following the minority would have made the learners answer sub-optimally.

When the blocks gave incorrect similarity signals, the social learners were less likely to choose their social learner optimum as more demonstrators did for blocks with uninformative-similar and uninformative-different signals. The learners did not treat an uninformative signal as if it were always incorrect. As the uninformative signal rendered the similarity information arbitrary, we expected the social learners just to choose the strategy that they preferred on these blocks. Figure 5 reveals that the social learners responded at chance towards the uninformative-different signals (Figure 5E), suggesting that they understood these signals to be arbitrary. Interestingly, the social learners instead copied the majority around uninformative-similar others (Figure 5B). This strategy led to the social learners answering sub-optimally whenever this signal of similarity happened to be incorrect. Taken together, this pattern may imply that some social learners arbitrarily committed to a strategy of copying similar others– even if they had uninformative signals– to improve their chances of answering optimally on the off-chance that the uninformative signal happened to be correct. That is, they may have followed a social-learning rule to 'copy similar others'.

When the blocks gave correct similarity signals (middle column of Table 4), the social learners were significantly more likely to answer optimally as more demonstrators chose their demonstrator optimum for the following signals: uninformative-similar, reliably correct-similar, and reliably correct-different. The social learners treated these signals as if they were always correct, which allowed them to respond optimally to these signals when the blocks did provide correct similarity information. This strategy made sense for the reliably correct signals as, based on the probabilities implemented in-game, a reliably correct signal of gave correct information about similarity or difference on approximately 90% of the blocks. The social learners followed the majority on uninformative-similar signals (Figure 5B), though this strategy only allowed the learners to answer optimally on approximately half the blocks when uninformative signals happen to give correct similarity information.

Appendix 13: Logistic regressions predicting whether the social learners chose their social learner optimum during the coordination game in response to correct and incorrect similarity signals.

Predictors included: (i) the centered proportion of demonstrators who chose the demonstrator optimum, (ii) dummies for each combination of similarity and reliability information, minus the omitted category of reliably incorrect- similar signals, and (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose the demonstrator optimum. Robust standard error clustered on social learner.

Parameter	**Estimate (coordination game, correct signals, full data)**	**Estimate (coordination game, incorrect signals, full data)**
Intercept	0.127 . (0.067) 95% CI [0, 0.26]	0.016 . (0.065) 95% CI [0-0.11, 0.14]
Centred proportion of demonstrators choosing the demonstrator optimum	0.008 (0.224) 95% CI [-0.43, 0.45]	1.603 *** (0.288) 95% CI [1.04, 2.17]
Reliably incorrect-different dummy [indicates different and is correct with 0.1 probability]	1.389 * (0.674) 95% CI [0.07, 2.71]	0.179 (0.146) 95% CI [-0.11, 0.46]
Uninformative-same dummy [indicates same and is correct with 0.5 probability]	-0.497 * (0.205) 95% CI [-0.90, -0.10]	-0.137 (0.200) 95% CI [-0.53, 0.26]
Uninformative-different dummy [indicates different and is correct with 0.5 probability]	-0.306 . (0.184) 95% CI [-0.67, 0.05]	-0.041 (0.198) 95% CI [-0.43, 0.35]

Reliably correct-same dummy [indicates same and is correct with 0.9 probability]	-0.129 (0.183) 95% CI [-0.49, 0.23]	0.309 (0.292) 95% CI [-0.26, 0.88]
Reliably correct-different dummy [indicates different and is correct with 0.9 probability]	-0.109 (0.137) 95% CI [-0.38, 0.16]	0.964 (0.672) 95% CI [-0.35, 2.28]
Centred proportion of demonstrators choosing optimum X reliably incorrect-different dummy	-4.14 . (2.217) 95% CI [-8.48, 0.19]	0.010 (0.525) 95% CI [-1.03, 1.03]
Centred proportion of demonstrators choosing optimum X uninformative-same dummy	3.894 *** (0.833) 95% CI [2.26, 5.52]	-4.925 *** (0.831) 95% CI [-6.55, -3.30]
Centred proportion of demonstrators choosing optimum X uninformative-different dummy	0.138 (0.682) 95% CI [-1.19, 1.47]	-0.437 (0.853) 95% CI [-2.11, 1.23]
Centred proportion of demonstrators choosing optimum X reliably correct-same dummy	5.575 *** (0.936) 95% CI [3.75, 7.40]	-7.315 *** (1.651) 95% CI [-10.54, -4.09]
Centred proportion of demonstrators choosing optimum X reliably correct-different dummy	1.413 ** (0.548) 95% CI [0.34, 2.48]	-5.983 ** (2.147) 95% CI [-10.18, -1.78]

The asterisks denote the level of significance of our p values, with the following key:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

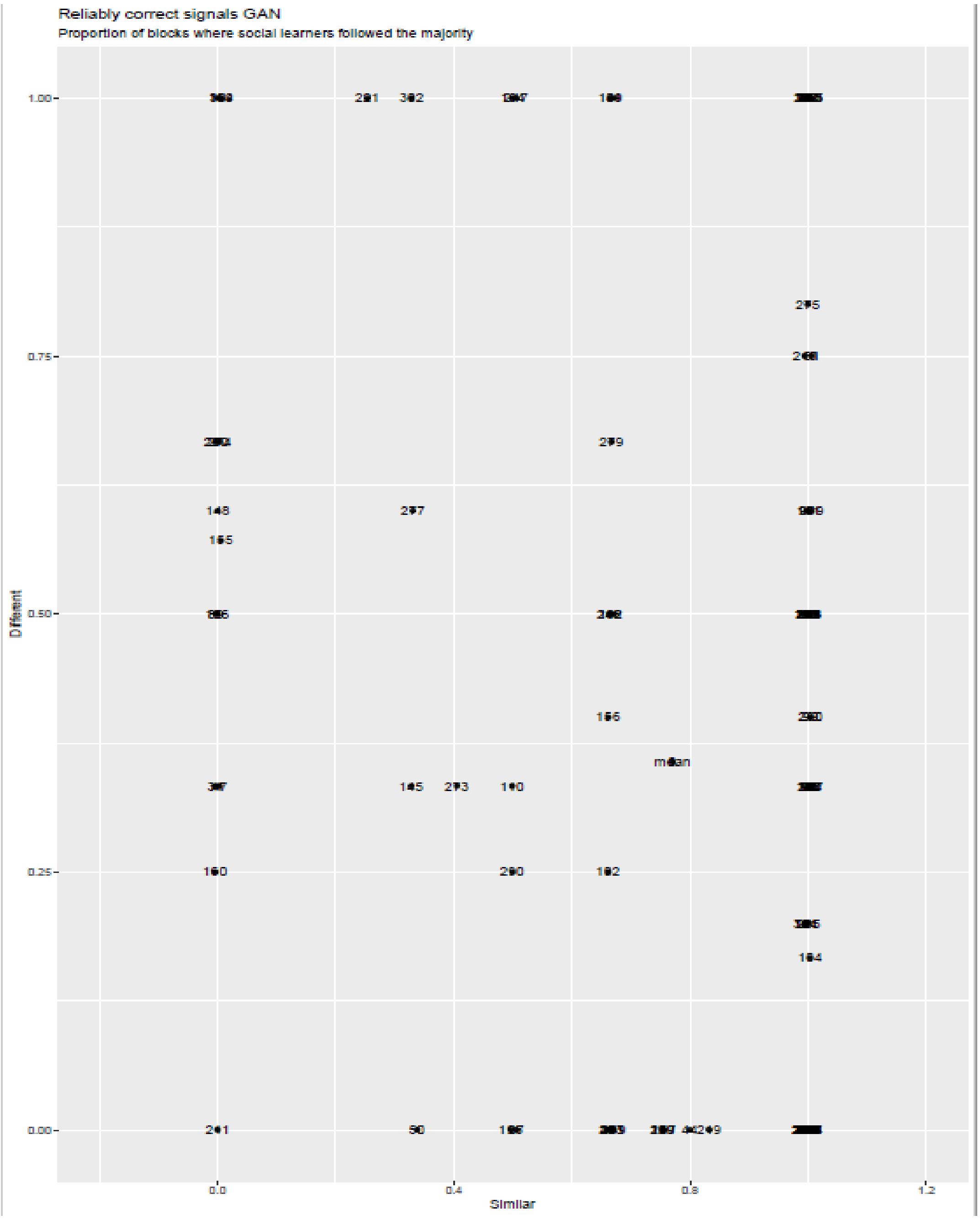
When the blocks gave incorrect similarity information, then the social learners were more likely to coordinate on the social learner optimum as more demonstrators did (Centred proportion of demonstrators choosing the demonstrator optimum effect). This effect reflected the omitted category of reliably incorrect- similar dummy. This effect shows that the social learners responded to reliably incorrect signals of similarity as if these were always incorrect. This makes sense as the probabilities implemented in-game meant that reliably incorrect signals gave false information on approximately 90% of blocks.

When the blocks gave incorrect similarity information, the social learners were less likely to coordinate on their optimal option as more demonstrators coordinated on the demonstrator optimum for the following signals: uninformative-similar, reliably correct-similar and reliably-correct different. Interestingly, when the blocks gave correct social information (middle column Table 5), then the social learners were more likely to coordinate on the social learner optimum as more demonstrators did for the same dummies: uninformative-similar, reliably correct-similar, and reliably correct-different. Taking these results together, the social learners treated these signals as if they were always correct. This led them to answer optimally whenever the blocks happened to provide correct similarity information, though led the social learner to answer sub-optimally whenever the block gave incorrect similarity information. Treating the reliably correct signals as if they are always correct makes sense as the probabilities implemented meant that these blocks gave correct signals of similarity or difference approximately 90% of the time. The social learners also responded to an uninformative signal of similarity as if this signal were always correct, though uninformative signals only provided correct similarity information on approximately half the blocks.

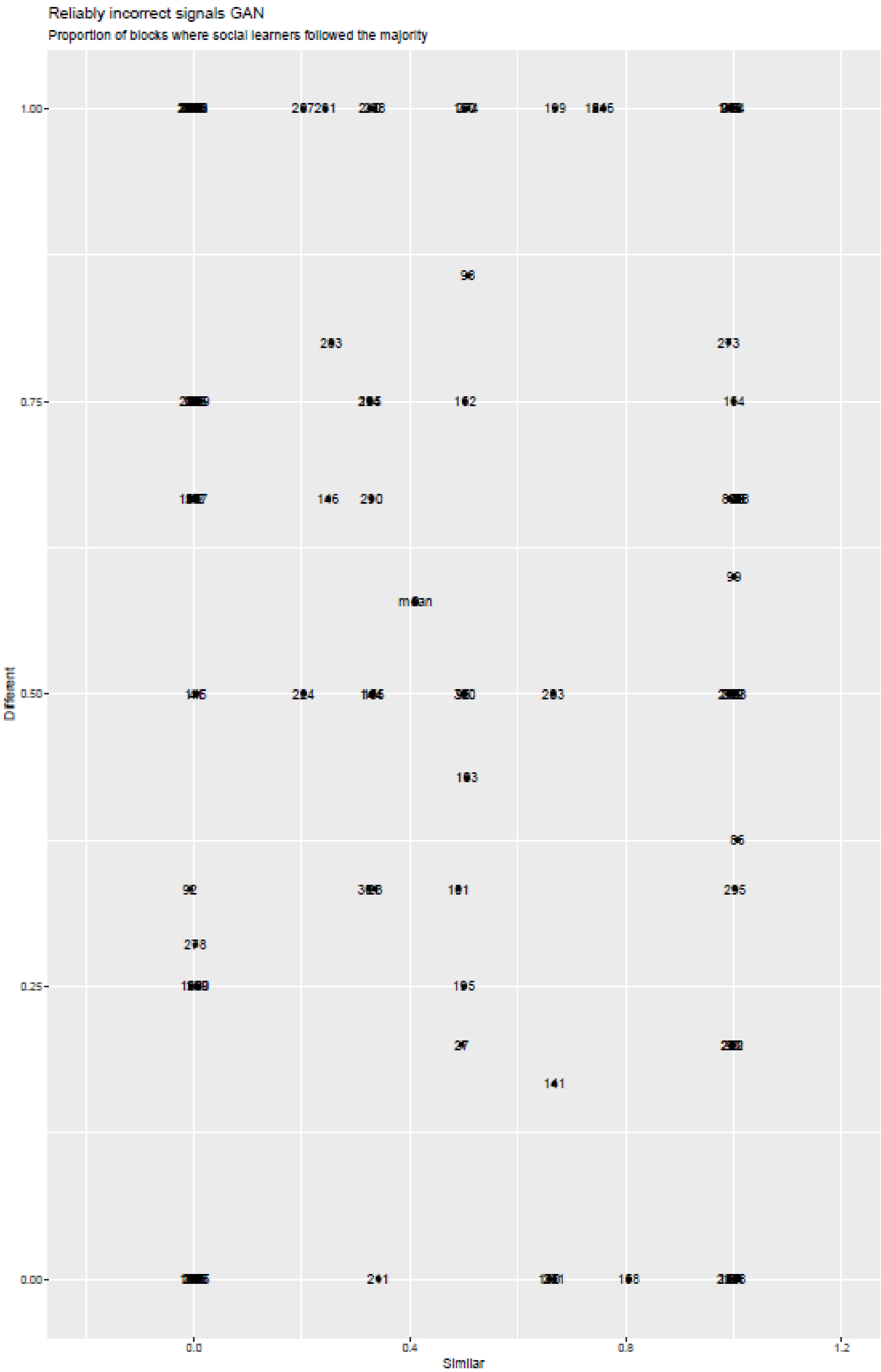
Appendix 14: The scatterplots, showing that the participant IDs often change as a function of the signal they see.

This is to show individual level variation in each of the learner's chosen strategies. As these IDs are hard to see, we also repeat this analysis as a heatmap in appendix 15.

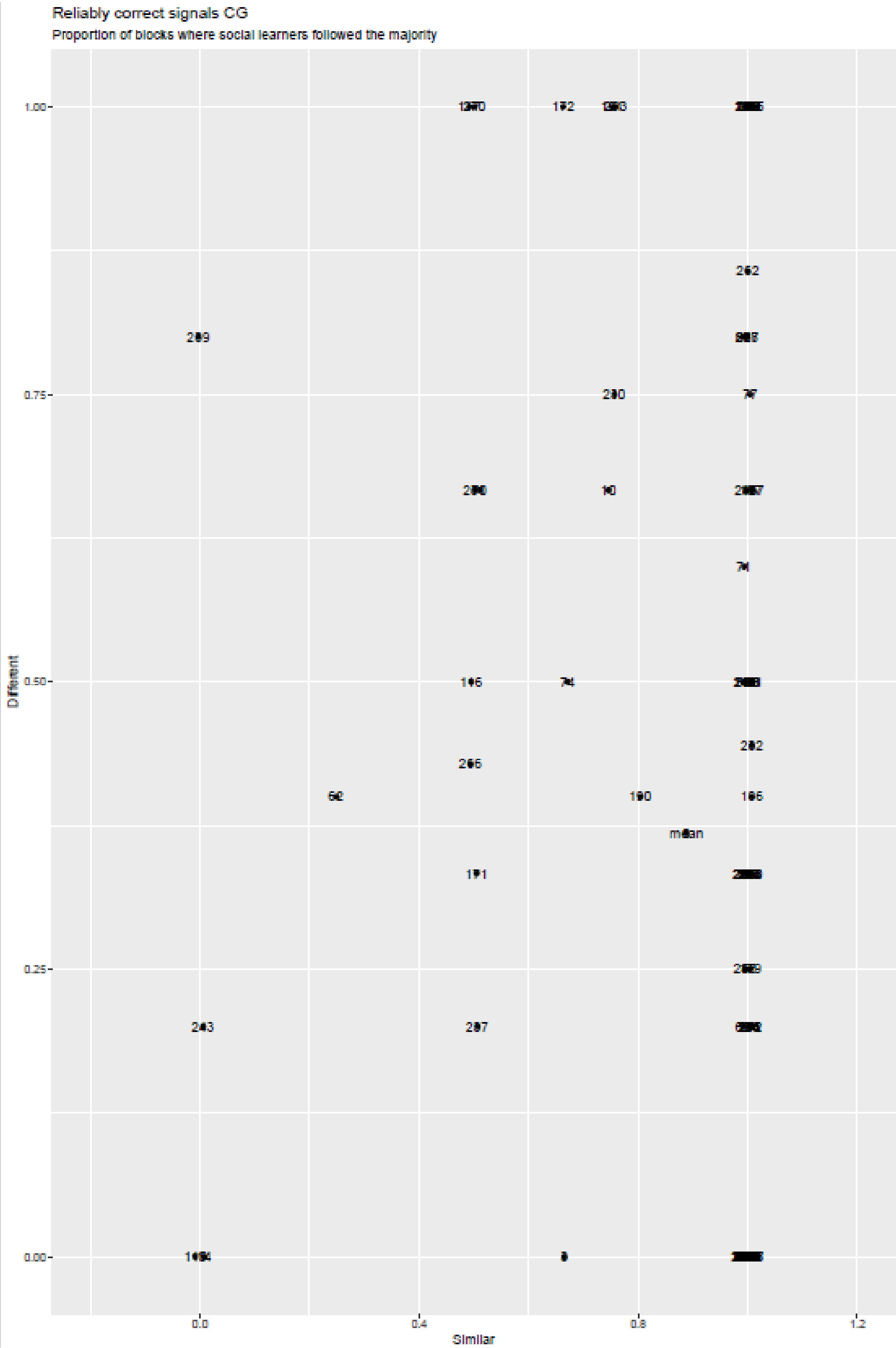
Appendix 14i) Game against nature, reliably correct



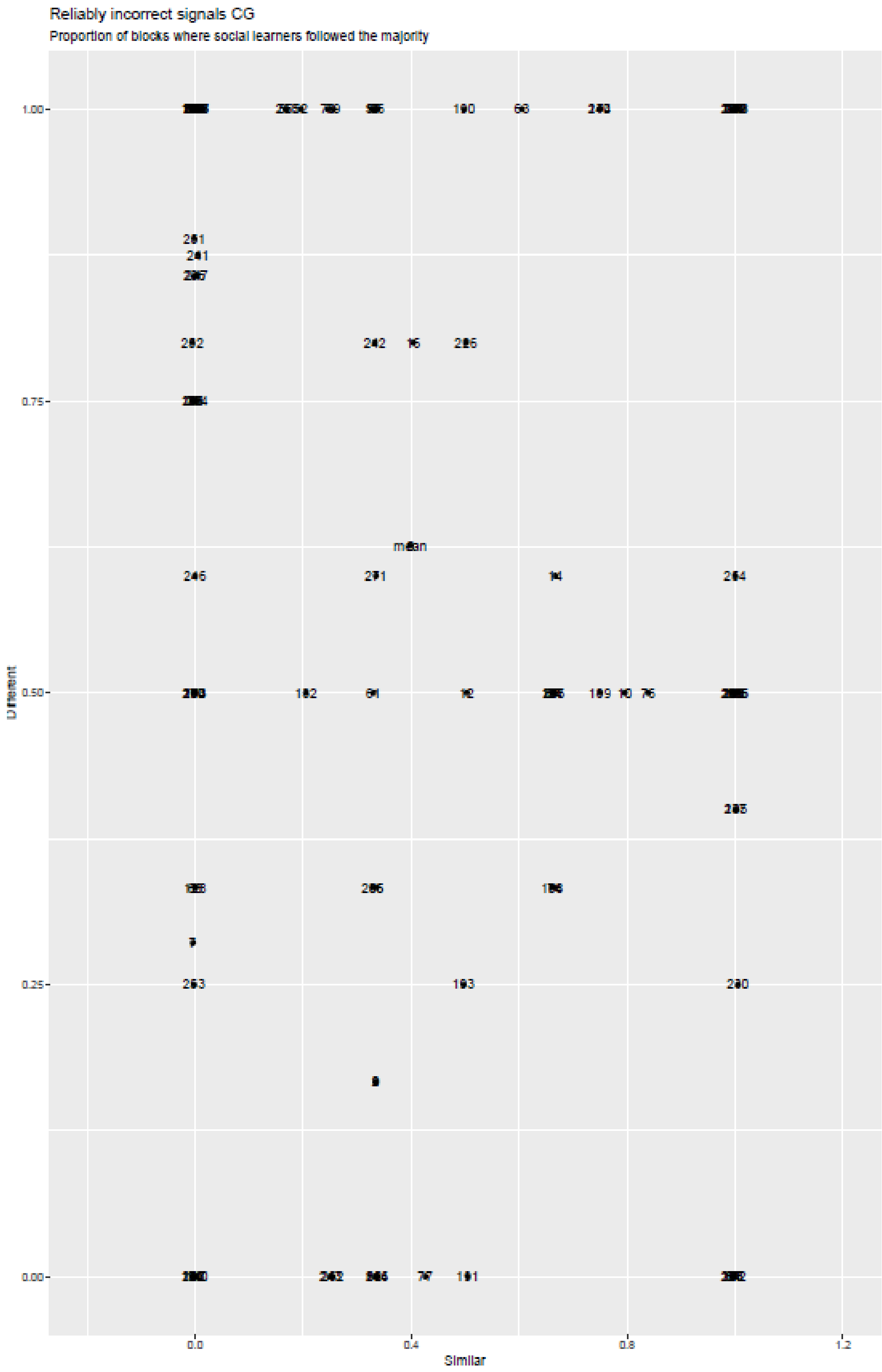
Appendix 14ii) Game against nature, reliably incorrect



Appendix 14iii) Coordination game, reliably correct

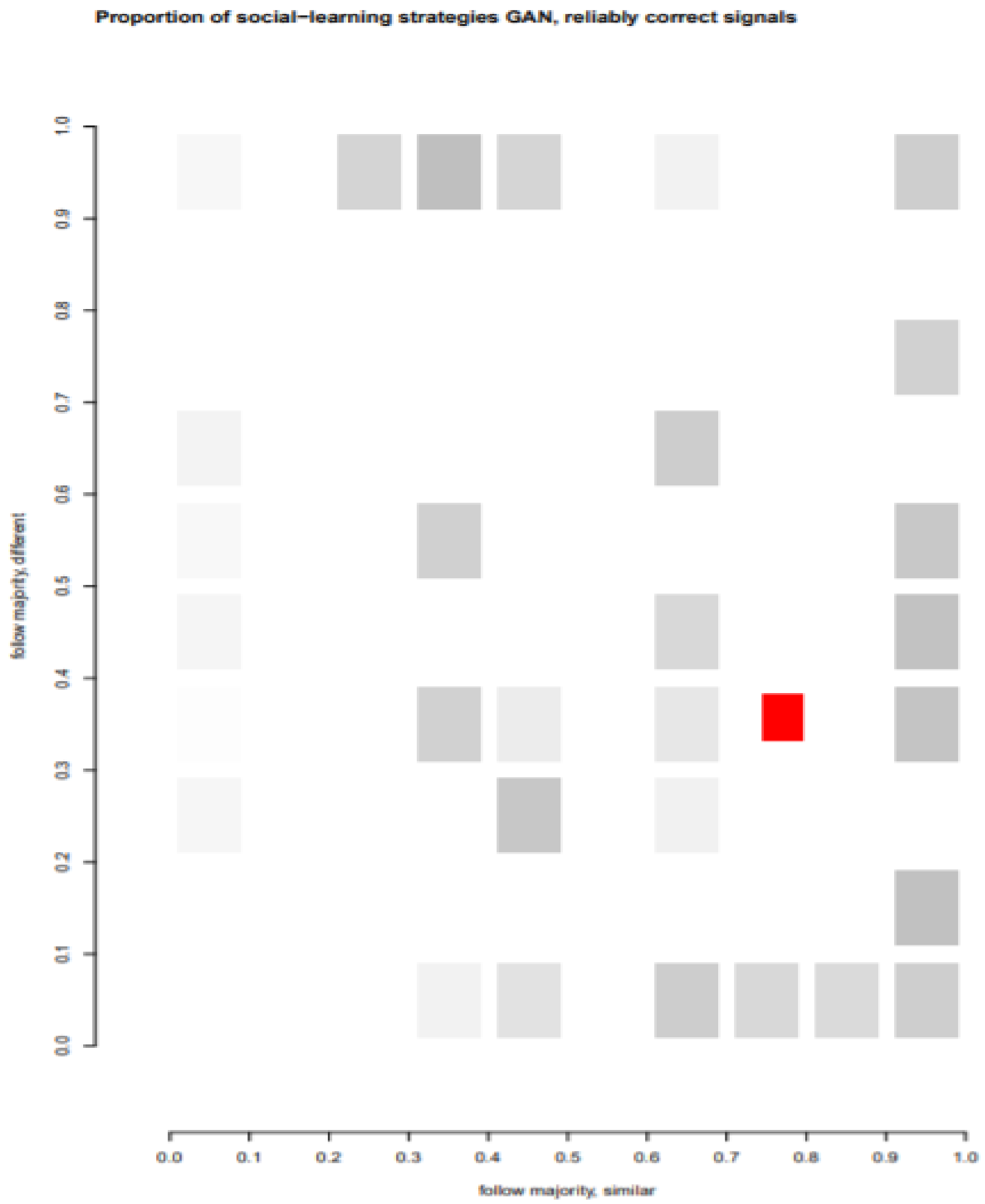


Appendix 14iv) Coordination game, reliably incorrect



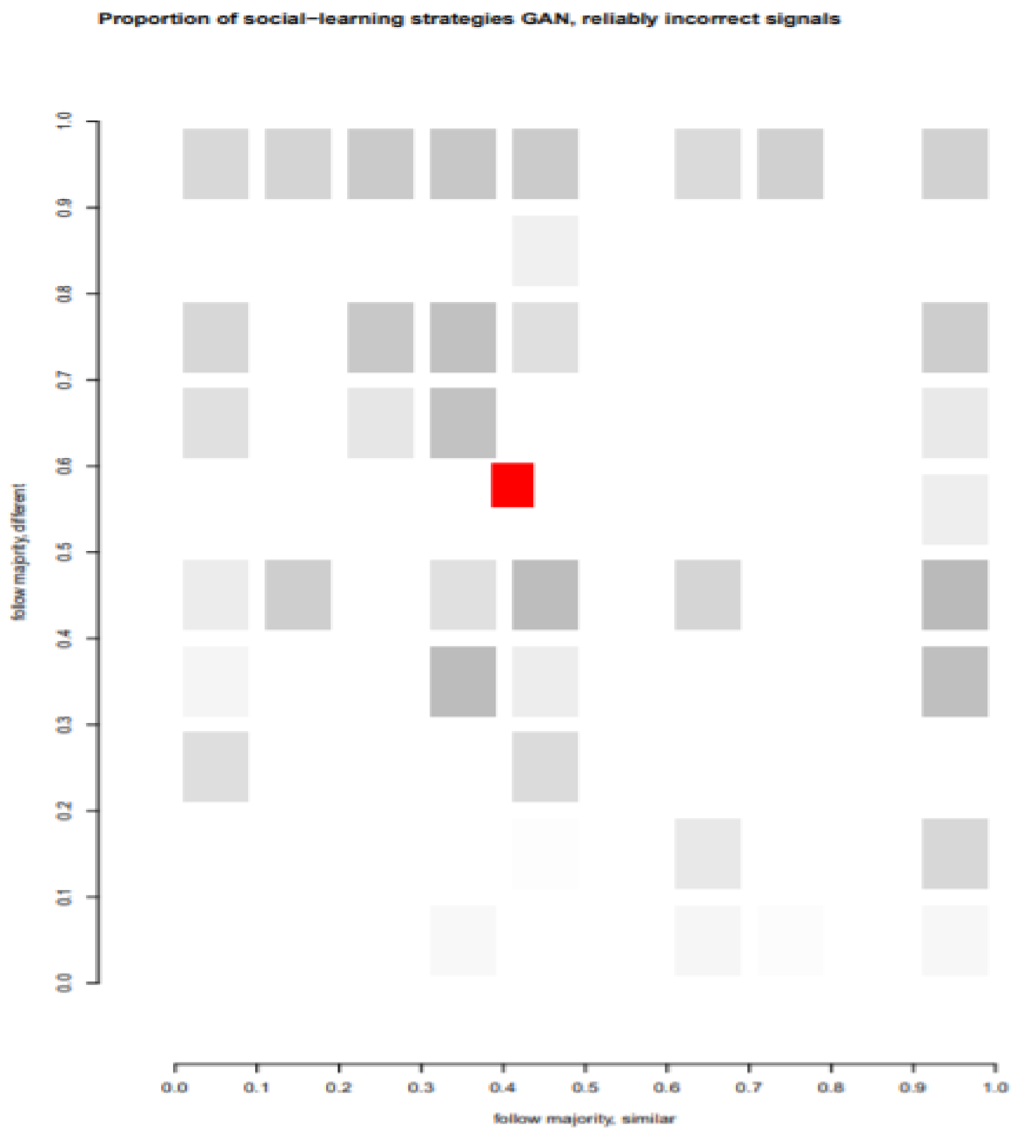
Appendix 15: The heatmaps show the social learner's chosen strategies as a proportion of blocks where they follow the majority of similar others (on the x axis) by the proportion of blocks where they follow the majority of different others (on the y axis). This is to show individual level variation in each of the learner's chosen strategies. The x axis gives 0-1 proportion of all social learners who followed the majority of groups identified as similar, while the y axis gives the 0-1 proportion of all social learners who followed the majority of groups identified as different. The darkness of the grey indicates frequency, with darker colours denoting a denser patch. Any white areas show that none of the learners employed this strategy combination. Finally, the red square denotes the social learning strategy averaged over all participants.

Appendix 15i) Game against nature, reliably correct



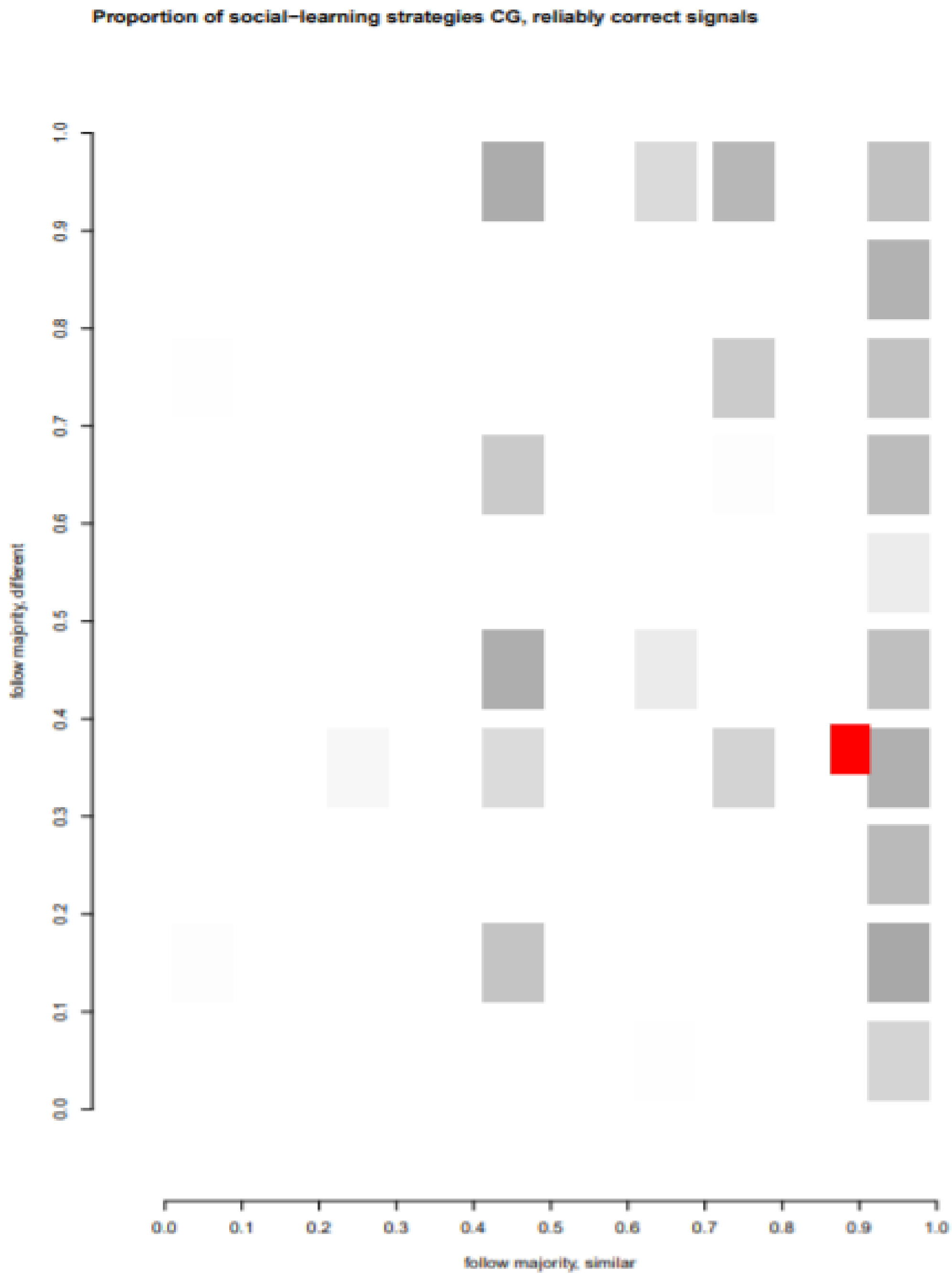
Note that there is a cluster along the top (i.e., these learners always follow the majority of different others, but are less sure when it comes to responding to similar others). There is also a cluster at the side (i.e., these learners always follow the majority of similar others, but they are less sure when it comes to different others). There is also a cluster along the bottom (i.e., these people copy the minority of different others but seem less sure for similar others). There are also some less differentiated strategies in the middle. The average social learning strategy is to copy the majority of similar others with a probability of ~0.7, and to follow the majority of different others with a probability of ~0.3. This perhaps makes sense: in order to choose their optimum, the social learners should copy the majority of similar others and the minority on different others, and so the average strategy is a part adjustment to this. Note that the optimal strategy is given by the bottom right square: an optimal learning strategy is to always copy the majority of reliably similar others but copy the minority of reliably different others. This square is shaded in fairly darkly, suggesting that a few social learners manage to answer optimally.

Appendix 15ii) Game against nature, reliably incorrect



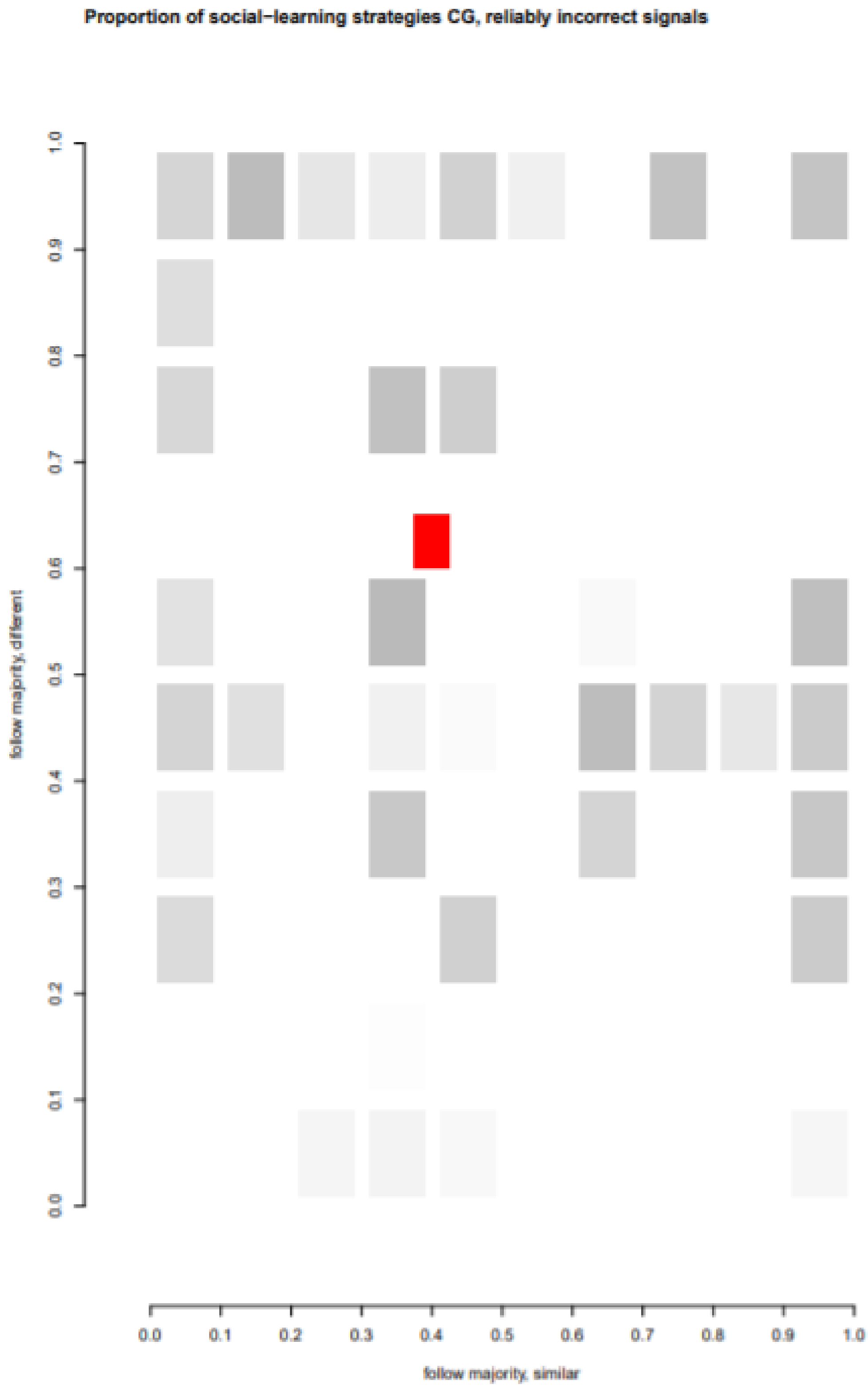
There is definitely a cluster at the top (who always copy the majority of different others with a reliably incorrect signals) and a cluster at the right-hand side (who always copy the majority of similar others with reliably incorrect signals). There also seems to be a cluster at the side, who copy the minority of similar others but are unsure for different others. The optimal square in this case is the top left (always follow the minority of similar others, but majority of different others, with reliably incorrect signals). This is filled in and so some do answer optimally. However, there is a broader range of strategies for the reliably incorrect signals and so the average strategy cancels out at chance and perhaps gives the impression that learners adjust to the reliably incorrect signal less than they actually do.

Appendix 15iii) Coordination game, reliably correct



This graph is perhaps the least spread out. The optimal strategy for a reliably correct signal is to follow the majority of similar others but the minority of different others. We find that this is shaded and so some do answer optimally. There is a cluster along the right-hand side: some learners always copy the majority of similar others but show more variation in the response to different others. There is also a small cluster in the top right: some participants always copy the majority of reliably different others, and trend towards following the majority of similar others. There is also a strip at around 0.4 of following the majority of similar others, where the participants respond differing to the different signal. Although the participants may seemingly respond anywhere to reliably different others, note that all learners avoid following the minority of similar others. This is why the left side of the graph is blank. Thus, following the minority may only be used on a coordination game in response to reliably correct signals of difference or reliably incorrect signals of similarity.

Appendix 15iv) Coordination game, reliably incorrect



The responses to reliably incorrect signals for the coordination game is perhaps the most varied. The optimal solution is to follow the minority of similar others and the majority of different others for a reliably incorrect signal (given by the top left square). As we can see, some of the learners do show optimal strategies. More broadly, we have three edges: some participants copy the majority of similar others, but there is more noise for different others; a cluster at the top where some always copy the majority of different others but are less differentiated for similar others and a cluster on the left where some participants follow the minority of similar others but are less differentiated for different others with reliably incorrect signals. There is then a range of adjustments in the middle.

It may be tempting to conclude that those participants in the middle are ignoring social information and answering optimally. However, we asked the participants to (anonymously) report whether they actually used the social information. Those that denied using social information are not necessarily those in the middle. Some participants take an undifferentiated strategy in response to three orders of social information on purpose.

Chapter 4:
When will you learn to help? Social learners can flexibly respond to a range of social information when learning to cooperate.

Aysha Bellamy
Ryan McKay
Charles Efferson
In preparation for publication

Word count: 12,231 excluding references and appendices

Abstract

Conformity can uphold cooperative or uncooperative group norms. It is unclear how individuals adjust their decision to conform, or use other learning strategies, to information about the group from whom they learn when adapting cooperative norms. This study investigated how flexibly 104 participants learned to play a prisoner's dilemma when uncertain as to whether the group from who they learned upheld a cooperative or uncooperative norm. The participants saw social information including (i) the frequency of choices made by the group from whom they learn; (ii) whether these members were identified as having learned in a similar or different environment to the participants and (iii) the reliability of this similarity information. Similarity information could be reliably correct, uninformative, or reliably incorrect. These reliability values reflected cases where those with similar traits to oneself did not uphold similar cooperative values. The participants processed frequency information and similarity information. The participants responded differently to reliably incorrect signals of similarity. Otherwise, they did not process reliability information. This asymmetric adjustment may be to counteract free riders who infiltrate cooperative groups by pretending to be similar to its' members. Overall, the participants' choice of learning strategies was flexible – to a degree – when learning to cooperate.

Keywords: Cooperation, frequency-dependent social learning strategies, conformity, cultural evolution

1. Introduction

Human societies are uniquely cooperative (Henrich et al., 2001). Cooperation involves accepting a cost in order to benefit another (Chudek et al., 2013). It may be difficult to envision how cooperation arose if individuals act to maximise their own fitness only (Boyd & Richerson, 2007; Colman, 2006). Despite this, individuals have consistently been found to donate substantial amounts of money to strangers who they are unlikely to meet again (Fehr et al., 2002; Fehr & Fischbacher, 2005; Sally, 1995). Whilst concerns have been raised over extrapolating participants performances in laboratory settings to naturalistic environments (Burton-Chellew et al., 2016; Price, 2008), it is also likely that human groups are extremely cooperative (Henrich et al., 2001; 2006). For example, our uniquely advanced tools and cultural systems are likely to be underpinned by our ability to cooperate and share information or resources with others (Dean et al., 2014; Henrich, 2015; Henrich & Muthukrishna, 2021).

There is a debate over how extreme levels of cooperation arose in humans (Chudek et al., 2013). Some posit that living in tight ancestral groups would have created a selection pressure towards cooperation that still drives behaviour today (Price, 2008). Others cite a norm psychology effect, whereby the pressure to adhere to social norms in highly cohesive groups could have led to extreme levels of cooperation (Gintis et al., 2008). Regardless of how cooperation emerged, a large body of literature has focused on the importance of conformity when learning to uphold established cooperative norms (Henrich & Boyd, 2001; Molleman et al., 2013a; 2013b). Conformity may be important to maintain cooperation as this strategy removes a focus on the payoff of one’s behaviour (Szolnoki & Perc, 2015). Thus, individuals may uphold even costly levels of cooperation via conformity (Henrich & Boyd, 1998).

Note that the precise definition of conformity has been subject to a lengthy debate in the gene-culture coevolutionary literature (Kendal et al., 2018; Lachlan et al., 2018; Morgan et al., 2019; van Leeuwen & Haun, 2014). For the sake of clarity, throughout this chapter I define conformity as a social learning strategy where the individual *disproportionately* copies the majority of a *group*. To illustrate with an example, if 75% of a social group are cooperative, then a conformist learner will also cooperate with a probability greater than 0.75. This definition is supported by a recent body of literature (Efferson et al., 2016; Lachlan et al., 2018; Morgan et al., 2019).

An individual could conform to the majority behaviour of a group without understanding the full intricacies of the behaviour being upheld. This is important as societies commonly differ in terms of which domain they cooperate in (Chudek & Henrich, 2011). For example, some groups cooperate when hunting but not when building shelter or vice versa. Even in a certain domain, there may be uncertainty over whether a behaviour is cooperative or uncooperative. For example, hunting. It may sometimes be cooperative to hunt more and donate food to others in the group (Hill, 2002). It may sometimes be cooperative to hunt less so as to not over deplete game (Safin et al., 2015). The rules governing cooperation may be complex and arbitrary. Individuals who wish to uphold the same cooperative (or uncooperative) norm as a social group need only conform to the majority behaviour, without needing to fully understand whether they uphold cooperation or not (Mesoudi, 2018). To reflect this uncertainty, I conduct a social learning study where it is not certain whether the majority of the group from who the participants learn are being cooperative or uncooperative.

Past research suggests that individuals can flexibly choose between conformity and other social learning strategies when learning asocial skills (Mesoudi et al., 2016) and social norms (Shutts et al., 2010). This study seeks to expand these findings by investigating the flexibility of the individual's decision to conform when learning cooperative norms. Conditional on frequency information, I ask when an individual will conform, or follow the majority in a linear fashion, or copy the minority, in response to three pieces of social information about the group from whom they can learn (un)cooperative behaviours.

The first level of social information that I investigate is frequency-dependent social information, or the number of others in the group who cooperate. This will henceforth be referred to as first-order social information. Even following a simple conformity rule, such as to 'always copy the majority', by definition involves responding to frequency information as the individual must track which behaviour is being displayed by the majority (Efferson et al., 2016). Any participant should therefore use first-order social information to inform her choices.

Second, I investigate similarity information which is henceforth referred to as second-order social information. The participants in the present study are informed that the members of the group from whom they can learn made their cooperative decisions in a similar or different decision-making environment to themselves. Individuals who share similar decision-making environments were likely to belong to the same social group throughout the ancestral past (Fischer, 2009). The ultimate goal of a complex social learner may be to learn from those who share a similar decision-making environment, as these individuals are likely to share similar optima (Efferson et al., 2008a).

However, it may have been hard to directly observe whether one shares a similar environment to another in the ancestral past. Those from the same social group are identifiable as they share ethnic markers (Efferson et al., 2008a; McElreath et al., 2003). Thus, looking 'similar' to others may become a useful shorthand to help us identify and learn from our own in-group members (Richerson et al., 2016; Wood et al., 2013). There is a tendency for previous research to focus on similarity based on an observable trait, such as age (Shutts et al., 2010) or gender (Salali et al., 2015). Indeed, evidence suggests that children as young as five prefer to learn cooperative norms from demonstrators who look similar to themselves (House et al., 2013; Salali et al., 2015).

However, there are times when those that look *similar* to oneself are not necessarily those who share similar decision-making environments to oneself. To illustrate with an example, imagine an individual has recently migrated to a new group and does not understand the rules regarding cooperation. She should conform to the cooperative behaviours of the majority of this new group, even if they look different to herself, to ensure she upholds the correct local norms (Deffner et al., 2020; Henrich, 2015). This is why I directly communicate whether the participants share a similar or different decision-making environment to the group from whom they learn in the current study. It was also important to test the most straightforward case of the similarity signal– by directly conveying this information to participants– as this is the first study to test three orders of social information when the participants learn to cooperate.

Finally, I investigate reliability information which is henceforth referred to as third-order information. The similarity information in this study can be reliably correct, uninformative, or reliably incorrect. Reliably correct signals may reflect cases when we are relatively assured of our similarity to others. This may be because the individual has direct knowledge as to who shares the same cooperative norms as herself (Molleman et al., 2014), or because the individuals bases her perceived similarity to others on traits that are readily observable such as age (Jiménez & Mesoudi, 2019) or gender (Efferson et al., 2016).

Reliably incorrect signals may occur when signals which suggest that someone shares the same cooperative intentions to oneself are faked (Sosis et al., 2007). For example, picture a cooperative group who are known to share their abundant wealth equally amongst all members. Now picture that this group are identifiable as they all wear red clothing. It may be quite easy for an outsider to acquire a piece of red clothing and infiltrate the group. This outsider has no intention of cooperation. Instead, she wears the red clothing as a faked signal of her cooperative intentions while actually intending to exploit donations from others. Of course, this example may create a kind of 'arms race' in the lengths that cooperative groups will go to, to prevent free rider infiltration. Perhaps the cooperative group members then start to wear red clothing that is a particular shade, which is difficult for outsiders to acquire. Any individuals who are then found to be wearing an incorrect shade of red are likely to be free riders who are trying to infiltrate the group and exploit resources. That is, the off-shade clothing of the free riders essentially becomes a *reliably incorrect* signal, helping the cooperative group member to identify any individuals around her who are unlikely to share her cooperative intentions (Stein et al., 2021).

Note that this example also highlights the necessity of three orders of social information. Picture an individual member of the cooperative group who is learning to uphold her group's stringent cooperative criteria. Perhaps she initially follows a rule to 'cooperate whenever the majority of others do'. This is a rule based on first-order social information only, but it may prove insufficient. The individual realises that those ingroup members who wear red are much more cooperative than the other individuals who live in nearby groups. Thus, it becomes necessary for the individual to expand her rule to 'cooperate whenever the majority of others who also wear red do'. While those who wear the same red clothing may *look* similar to the focal learner, observable traits of similarity are useful in as far as they imply a shared decision-making environment (Efferson et al., 2016). This social learning rule operates on both first and second-order social information.

However, the infiltration of free riders makes another extension of this rule necessary. The individual then follows a rule to ‘cooperate whenever the majority of others who also wear red do, provided that their clothing is this exact shade of red’. This is now a rule which works on first-order (frequency-dependent), second-order (similar clothing) and third-order social information (reliably correct shading). Of course, the prevalence of free riders could escalate these rules even further. For example, painful and elaborate signals of group membership may be used by extremely cooperative groups to discourage free rider infiltration (Richerson et al., 2016; Sosis et al., 2007).

If such scenarios were frequent enough throughout the ancestral past, then it would pay for the individual to track any reliably incorrect signals of similarity (Smaldino et al., 2018). The reliably incorrect signal is as informationally meaningful as a reliably correct signal in the current study. Reliably incorrect signals may provide a meaningful cue as to who is *faking* group membership. While both signals are informationally meaningful, there may be a trade-off in processing these signals. For example, reliably correct signals may be easier to process than reliably incorrect signals as the latter involves a mental negation (Cutmore et al., 2015). Those that display reliably incorrect signals, such as those free riders wearing the incorrect shade of red in the example above, are also likely to be rarer than individuals who express themselves with reliable and meaningful group markers (McElreath et al., 2003; Richerson et al., 2016). Perhaps it will be harder for individuals to process the reliably incorrect signal in the current study as it was rarer throughout the ancestral past for psychological systems to be needed that could decode such signals.

Note that the learners cannot adjust meaningfully to the uninformative signal. Uninformative signals render the similarity information at chance level of being correct. Thus, the participants cannot calculate with any degree of certainty whether they are making decisions in a similar or different decision-making environment to the group from whom they learn. In this situation, no single social learning strategy can be expected to uphold the same cooperative norm as the group from whom the participant learns better than any other (Efferson et al., 2016). This removes the incentives to the learners' behaviour and thus the participants may choose randomly on these trials, or they may follow an arbitrary personal preference for certain social learning strategies.

To summarise, I investigate whether the participants base their decision to conform or use another social learning strategy on (i) the frequency of choices made by the group from whom they could learn; (ii) whether this group learned in a similar or different decision-making environment to the participants and (iii) the reliability of this similarity signal. This study is novel in testing the participants' response to three orders of social information simultaneously. While it may intuitively make sense for the participants to respond to all social information, there may be trade-offs as it is cognitively demanding to respond to multiple signals (Krafft et al., 2021; Muthukrishna et al., 2016). While it is difficult to make exact predictions, the participants' strategies may broadly fall into one of four categories:

(i) The participants adapt their strategies to all social information. They adjust their behaviour to first-order information (frequency-dependent information), second-order information (similarity information) and third-order information (reliability information).

(ii) Individuals adjust to first-order, second-order and third-order information but there may be a trade-off in social learner cognition which leads to asymmetric adjustments. That is, the individuals may find some types of social information easier to respond to than others. For example, they may find reliably correct signals easier to process than reliably incorrect signals.

(iii) The participants may adjust their strategies to first-order information (frequency-dependent information) and second-order information (similarity information) but may not be able to adjust to third-order social information (reliability information).

(iv) The participants adjust to first-order information (frequency-dependent information) but find both second and third-order social information too complex to process. The participants may just follow a rule to ‘always copy the majority’.

2. Method

2.1 Participants

6 sessions were conducted at the EconLab in Royal Holloway, University of London. 140 participants were recruited (mean age = 20.29, $\sigma = 2.79$, males = 41). This sample met the minimum sample size suggested by an ex-ante power calculation (see pre-registration at OSF | prereg_SD_15mar.docx). Twelve of these participants were international students not from the UK. These international students reported being from Italy (N=2), France, Russia (N=2), India, Kazakhstan, Saudi Arabia, Nigeria, Malaysia, Indonesia, and Japan.

These 140 participants were divided into one of two roles. 36 participants played as demonstrators, who developed a preference to cooperate or defect throughout this study via their own trial-and-error. This number came from assigning 6 demonstrators per session. This was deemed sizeable enough as a social group to provide meaningful social information (Efferson et al., 2016). The remaining 104 participants across all sessions played as social learners who could only base their decisions on social information. While splitting participants into separate roles may seem simplistic, this design crucially allowed the experimenters to draw clean causal inferences regarding the role of social information without the inferential challenges that usually hold (Angrist, 2014; Manski, 2000).

2.2. The prisoner's dilemma

We investigated the learning of cooperative norms with a prisoner's dilemma (Axelrod, 1980; Holt & Roth, 2004; Sally, 1995). In this game, the participant must choose to cooperate or defect with an anonymous partner with whom she has been paired. The logic is this. If she defects when her partner cooperates, then she receives a higher payoff than if she had cooperated. If she defects when her partner defects, then she still does better than if she had cooperated. Defecting when her partner cooperates gives the highest payoff, called the free rider payoff. Cooperating when her partner defects gives the lowest payoff, called the sucker payoff. As defection is the best strategy at the individual level, then it is said to be the Nash equilibrium (Holt & Roth, 2004). However, mutual cooperation gives a higher payoff than mutual defection and thus the game represents a dilemma.

Note that labelling the behavioural options as 'cooperate' or 'defect' may introduce a social desirability effect (Price, 2008). To avoid this, I asked participants to choose between two arbitrary symbols (@ or %). Note that this game still worked as a prisoner's dilemma in the current study. This is because a prisoner's dilemma is any payoff matrix where one option is always expected to give the higher payoff at the individual level (equivalent to defection, or the Nash equilibrium) though the pair of participants would do better if they both coordinated on the other option (equivalent to mutual cooperation) (Molleman & Gächter, 2018).

To illustrate, figure 1 displays how @ or % may align to playing 'cooperate' or 'defect' respectively. Take game left in figure 1, for example. Choosing % when the other person chooses @ gives the focal participant the highest payoff, or free rider payoff. Choosing @ when the other chooses % gives the focal participant the sucker payoff. This makes % the Nash equilibrium. Coordinating on the Nash equilibrium results in fewer points (145) than coordinating on the @ option (235), which is associated with cooperation. These games thus conform to the prisoner's dilemma payoff matrix. Game right instead ties the @ option to the Nash equilibrium while % denotes cooperation. These two games allowed the option associated with cooperation to change randomly between blocks.

	Game Left	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 145** Partner's expected points: 145	**Your expected points: 280** Partner's expected points: 100
You choose @	**Your expected points: 100** Partner's expected points: 280	**Your expected points: 235** Partner's expected points: 235

	Game Right	
	Partner chooses %	Partner chooses @
You choose %	**Your expected points: 235** Partner's expected points: 235	**Your expected points: 100** Partner's expected points: 280
You choose @	**Your expected points: 280** Partner's expected points: 100	**Your expected points: 145** Partner's expected points: 145

Figure 1. The payoff matrix of the prisoner's dilemma shown to the participants. The Nash equilibrium is % in game left and @ in game right. Text in **bold** represents the focal participant's choices.

Figure 1 denotes the expected points. All payoffs were influenced by a random disturbance (M =0, SD = 20 points) which was independently drawn for each period and independently drawn for each participant. I told the participants about this noise including the mean and standard deviation. They were also told that this disturbance meant that they could earn more or less points than expected for their choices. Thus, the participants were aware that they were making decisions under uncertainty (as per Molleman & Gächter [2018]; see appendix 1). This noise is important as, even if cooperation seems to be the most straightforward strategy, there may be other factors affecting our certainty over which behaviours are cooperative (Chudek & Henrich, 2011).

2.3 Materials

The participants read an instructional booklet and answered some pre-game questions to assess their understanding of the game prior to playing it (see appendix 1). The participants played the game at individually screened PCs to ensure anonymity. The game was ran via Z-Tree version 3.5 (Fischbacher, 2007; see appendix 2).

All participants completed a post-experimental survey via Z-Tree (see appendix 3). This asked for key demographics such as age, gender, and country of residence. The social learners also answered questions about their social learning preferences (see Supplementary Materials).

2.4 Procedure and Design

Each of the 6 sessions contained a group of 16-30 participants. The numbers were kept even, as the prisoner's dilemma is played in pairs (Capraro, 2013; Szolnoki & Perc, 2019). After checking the participants' understanding, the computer randomly selected 6 participants to play as demonstrators (labelled 'Type A' in-game). The remaining participants played as social learners (labelled 'Type B' in-game). Participant types stayed the same throughout the session.

Each session consisted of 22 blocks containing 4 periods. The computer began each block by randomly assigning the participants to play in pairs. The pairs were constrained so that the demonstrators could only play with other demonstrators and the social learners could only play with other social learners. The pairings could change randomly between blocks but not across the periods within one block. Finally, the computer randomly assigned the demonstrators to play game left or game right. Note that this decided which option (% or @) was associated to the Nash equilibrium payoff. The option associated with the Nash equilibrium could change randomly between blocks but not within the periods making up one block.

The demonstrators played first in each block. They did not know whether they were playing game left or game right, though they could see the payoff matrices for both games throughout (see figure 2). The demonstrators could learn which option (@ or %) represented cooperation and which represented defection based on the immediate points-based feedback that they received from their and their partners' choices. The disturbance in points helped to mirror a realistic level of uncertainty over which behaviour was considered cooperative.

The demonstrators choices' in the final period of the block formed the social information that was shown to the social learners. I chose only the final period, as by then the demonstrators should play cooperation or defection as based on their preferences which they could learn via trial-and-error across each block. The social learners only played once in the final period of each block. They could not see the feedback from their decisions and instead could only base their decisions on social information. This included: (i) the frequency of demonstrators who had chosen @ or % by the final period of a block (first-order social information); (ii) a signal identifying that the social learners played the same or different game to the demonstrators (similarity information, or second-order social information) and (iii) whether this similarity signal was reliably correct, uninformative, or reliably incorrect (reliability information, or third-order social information).

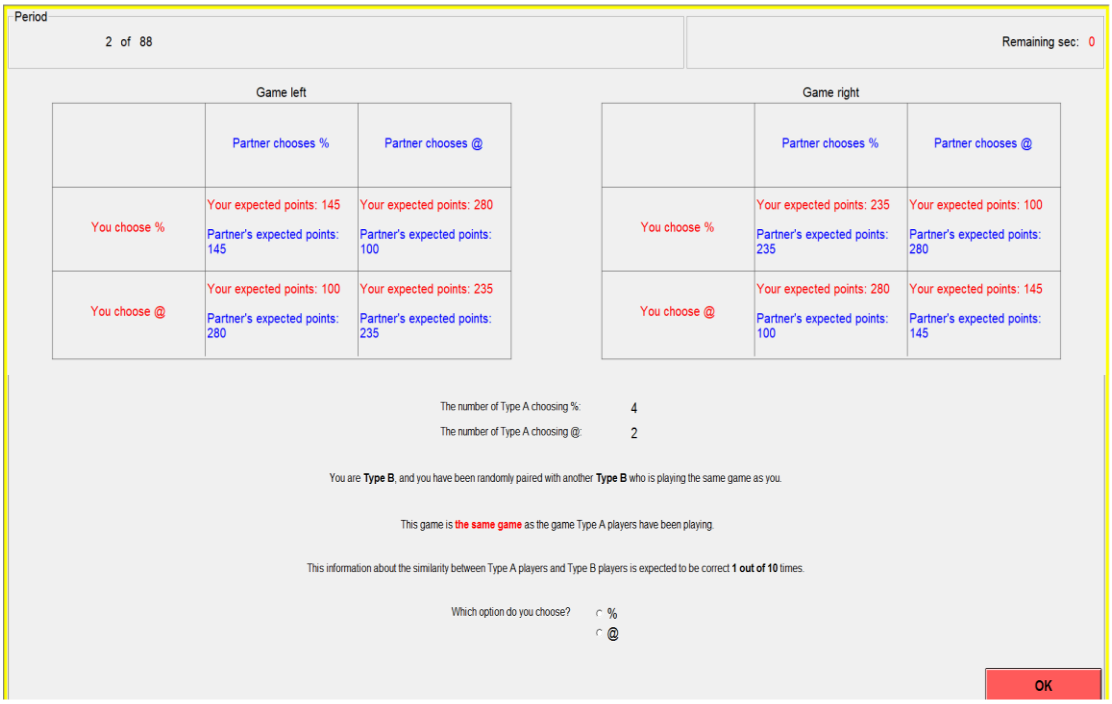


Figure 2. A screenshot of the typical round played by the social learners. Note that Type A participants were demonstrators and Type B participants were social learners.

I avoided the term 'demonstrator' or 'social learner' in case it led the participants respond to the task in certain ways. The top half of the screen reminds the participants of the expected payoffs for the prisoner's dilemma. The bottom half of the screen provides social information regarding: (i) the frequency-dependent social information (i.e., the number of demonstrators who chose @ or %); (ii) similarity information (i.e., the signal telling social learners that they played the same or different game to the demonstrators) and (iii) the reliability information.

Note that frequency-dependent information was not manipulated but emerged endogenously in each game, as this represented the actual choices made by the demonstrators who played first in each block. By using arbitrary symbols, I could investigate some realistic uncertainty about which behaviour is cooperative. There are so many domains that different social groups cooperate in, and many arbitrary rules regulating cooperation, that this may be difficult for individuals to learn (Chudek et al., 2013). Indeed, conformity is useful as one does not have to understand whether one upholds a cooperative or uncooperative norm. One simply conforms to what the majority of the group are doing in order to match their cooperative (or uncooperative) preferences (Szolnoki & Perc, 2015; Yang & Tian, 2017). To reflect this uncertainty, I merely show participants the number of demonstrators who chose @ and % in each round, rather than translating this to 'cooperate' or 'defect'.

This study had a 2X3 within-subjects design. I manipulated both the similarity information (2 levels: same or different) and the reliability information shown to the participants (3 levels: reliably correct, uninformative, or reliably incorrect). The similarity signal was conveyed in relation to the game being played. A 'similar' signal implied that the social learners and the demonstrators played the same game. The social learners should therefore choose the same option as the majority of demonstrators if they wished to uphold the same cooperative (or uncooperative) norm as this group. For example, both the demonstrators and social learners played game left and so both should choose @ to cooperate. A ‘different’ signal implied that the social learners played a different game to the demonstrators. If the social learners made decisions in a different environment to the demonstrators, then they should choose the opposite option to the majority in order to uphold the same cooperative (or uncooperative) norm as this group. For example, the demonstrators played game left where @ represents cooperation, but the social learners played game right where % represents cooperation.

The reliability information was conveyed as probabilities. A reliably correct signal would be correct 9 in 10 times, an uninformative signal would be correct 5 in 10 times and a reliably incorrect signal would be correct 1 in 10 times. Note that a third of social learners were assigned to a ‘reliably correct’ group, a third to an ‘uninformative’ group and a third to a ‘reliably incorrect’ group at the start of each block. The computer then assigned the social learners to play game left or game right and then tracked whether this game matched the one assigned to the demonstrators. The similarity signal shown to the social learners was probabilistically decided by this reliability group. For example, take those assigned to the reliably correct group. Suppose these social learners played game left and the demonstrators had also played game left. The social learners would then see a signal informing them that they played the ‘same game’ as the demonstrators with a probability of 0.9 or would see a signal informing them that they played a ‘different game’ to the demonstrators with the remaining probability. This ensured that the reliability signal was informationally meaningful.

The participants then completed the survey measuring their social learning preferences (see supplementary materials). Participants were debriefed and paid individually based on their points gathered across the session. This was converted to pay at a rate of 1150 points to £1, plus a show-up fee of £4 for each participant. On average, participants earned £14.31 (SD = £1.02) for the session, with pay ranging from £12.02 to £17.60.

This study was self-certified via the Royal Holloway University of London's School of Psychology ethical criteria (see appendix 4). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

2.5. Analysis

Prior to determining the flexibility of the social learners' decision to conform, I needed to check that the demonstrators provided meaningful social information. If the demonstrators did learn to play the game meaningfully, then there should have been a preference towards cooperation or defection across the periods within the sessions. Section 3.1 focuses on whether the demonstrators provided meaningful yet varied social information.

The analysis of interest can be split into two steps. First, I investigated how flexible the social learners were when choosing a frequency-dependent social learning strategy. That is, I investigated whether the learners could adjust their behaviour to all three pieces of social information. I investigated this with a general linear mixed model, with the proportion of social learners who chose % as the dependent variable and participant ID and session ID as random effects. Fixed effects include: (i) the centered proportion of demonstrators who chose % on each block, (ii) dummies representing each combination of the similarity and reliability information and (iii) interactions between these dummies and the centered proportion of demonstrators who chose %. These five dummies were: reliably incorrect – different signals; uninformative – similar signals; uninformative – different signals; reliably correct – similar signals and reliably correct – different signals. The omitted category was the reliably incorrect – similar signals. Note that the similarity and reliability information were combined into dummies rather than analysing the social information as three separate factors to avoid three-way interactions that may be difficult to interpret in a large regression model (see appendix 5 for script). Section 3.2 focuses on the flexibility of the social learners' strategies.

While this first analysis shows how *flexible* the learners are, it does not show whether they have a *social learning preference.* If the learners are affected by trade-offs, then they make asymmetric adjustments and copy the choices of certain groups more readily than others. To investigate whether the social learners made asymmetric adjustments in their choices, my third analysis ran a general linear mixed model with the proportion of social learners who cooperated as the dependent variable. The fixed effects include: (i) the centered proportion of demonstrators who cooperated on their final period; (ii) the same similarity and reliability dummies as above and (iii) interactions between these dummies and the centered proportion of demonstrators that cooperated, with the same random effects as bove. Section 3.3 focuses on whether the social learners displayed a preference to cooperate when learning from certain social groups.

Both models were repeated twice. The second model included control variables. Characteristics such as age and gender can influence one's conformity preferences (Mesoudi et al., 2015; Wen et al., 2019). Note that the regressions reported in section 3 only display the model (with or without controls) that had the best fit of the data according to the ANOVA comparison.

Note that these methods, materials, and sampling plan was preregistered on Open Science Framework [OSF | prereg_SD_15mar.docx]. I kept the exact analysis plan exploratory. This was to allow for any novel investigations as no previous studies have tested third-order complexity when learning to cooperate.

2.6. Predicted social learning strategies

The social learners could only adjust meaningfully to the trials that provided informative signals. The uninformative signal rendered the similarity information at chance level of being correct. The social learners would thus have no guarantee that choosing any one frequency-dependent social learning strategy would make them more or less likely to uphold the same cooperative norm as the group of demonstrators from whom they learned. Thus, the social learners could only adjust meaningfully to reliably incorrect and reliably correct signals of similarity or difference.

Even though the social learners were uncertain over whether the majority of demonstrators cooperated or defected, it is important to remember that the payoffs of one's decisions in a prisoner's dilemma are conditional on a partner. The social learners should be aware that the majority of demonstrators made choices conditioned on what they would expect their partner to do. Thus, the social information from the demonstrators meaningfully reflects the cooperative– or uncooperative– norms of the majority of that group.

Although it is difficult to predict the social learners' strategies, I outline the social learning strategies that the social learners should use if they wish to uphold the same cooperative– or uncooperative– norm as the group from whom they learn. In cases where the social learners played the same game as the demonstrators, then the same symbol aligned with cooperation for both participant types. Note that the social learners were likely to be playing the same game as the demonstrators under both a reliably correct signal of similarity and a reliably incorrect signal of difference. The social learners could follow the majority of demonstrators under both these signals in order to uphold the same cooperative (or uncooperative) norm as this group.

When the social learners played a different game to the demonstrators, then they made decisions in a different environment. Thus, the opposite option (% or @) would be tied to cooperation for the social learners compared to the demonstrators. The social learners could choose the *opposite* option to the majority of demonstrators to uphold the same cooperative or uncooperative norm. Alternatively, they could choose the *same* option as the majority of demonstrators to uphold the opposite cooperative or uncooperative norm to this group.

Of course, if the signal of difference was related to social characteristics rather than decision-making environments then some individuals may coordinate on the opposite norm to the group just to appear different (Chudek & Henrich, 2011; Smaldino & Jones, 2021). As the group only differed in terms of the option associated with the cooperative payoff however, then I expected that the social learners may still try to uphold the same cooperative or uncooperative norm as the group from whom they learn. Thus, I might expect that the social learners follow the minority of demonstrators who made decisions in the different game. Note that the social learners were likely to be playing a different game to the demonstrators under both a reliably correct signal of difference and a reliably incorrect signal of similarity. Astute social learners would be aware that these are equivalent conditions. Table 1 depicts the social learning strategies that the social learners may use if they are trying to uphold the same cooperative (or uncooperative) norm as the group from whom they learn.

Table 1. The strategies that would allow the social learners to uphold the same cooperative or uncooperative norm as the group from whom they could learn, when seeing one of the four informationally meaningful trials.

		Third-order social information	
		Reliably incorrect	**Reliably correct**
Second-order social information	**Similar**	Follow the minority	Follow the majority
	Different	Follow the majority	Follow the minority

3. Results

3.1: Did the demonstrators prefer to cooperate or defect?

The demonstrators' choices formed the frequency-dependent social information shown to the social learners. It was therefore important that the demonstrators had a clear preference (whatever this preference might actually entail) by the end of the block. This is similar to how one would decide whether or not to cooperate with another person based on one's own experiences. If the demonstrators had simply answered randomly, then there would be no way for the social learners to choose a strategy that could meaningfully map onto cooperative (or uncooperative) norms.

Table 2 confirms that the demonstrators formed a preference over the periods within a block. The significant predictor in this regression, 'finalPeriodDummy', indicates that the demonstrators were significantly less likely to cooperate when playing in the final period of each block than any other period (estimate=-0.253, z=-2.96, p=0.003). This shows that the demonstrators developed a preference to defect by the final period of a block based on trial-and-error.

We also had to check for a preference for one arbitrary symbol over the other, I ran a ‘percentageAsCooperate’ dummy in this model (table 2). This dummy indicated whether % happened to be associated to a cooperative payoff. The significant estimate reported in table 2 suggests that the demonstrators were *less* likely to cooperate whenever the % symbol happened to be tied to the cooperative payoff (estimate=-0.456, z=-6.16, p<0.001). This suggests an arbitrary preference in some of the demonstrators to choose the opposite symbol (@) throughout. Alternatively, perhaps the demonstrators were less likely to cooperate when the % symbol was tied to this option, as this symbol may have had connotations of only taking a small percentage of the total earnings. However, I note that past research has used such arbitrary symbols without issue, supporting their use here (Efferson et al., 2016).

Of course, it is then necessary to confirm that the demonstrators provided a range of social information across the blocks and did not arbitrarily choose one option throughout. To check for this meaningful variation, I created a histogram plotting the demonstrators’ choices over the blocks within the game (figure 3). The most frequent outcome was for a block to end with half the demonstrators choosing to cooperate and half choosing to defect. This may appear random, though it should be noted that the histogram has a positive skew. Put simply, this means that the demonstrators defected more often than they cooperated. This suggests that the demonstrators had a preference to defect.

Moreover, 39.90% of the demonstrators chose to cooperate by the average final period. Importantly, this number was significantly different to the 50% of the group that would be expected to cooperate by chance (one-sample t-test: t(21) = -2733.9, p<0.001). At the level of the pair, mutual defection was the most common. Mismatches were also common, as the demonstrators who had cooperative preferences may have been paired with demonstrators who had defection preferences. Mutual cooperation was rarer (see appendix 5). This variation aligns to past research which shows a range of cooperative or uncooperative preferences across participants (Li et al., 2021; Szolnoki & Perc, 2012), though my study may have less mutual cooperation than typical (see Discussion, Section 4; McAuliffe et al., 2019). For the demonstrators in this study, defection was clearly the most common preference. The demonstrators thus provided meaningful– but varied– social information about their (un)cooperative preferences.

Table 2. A general linear mixed model, modelling whether the demonstrators cooperated by the final period of the blocks. Fixed effects include a dummy for the final period and % as the cooperative option, with participant ID and session ID as fixed effects, with Wald's 95% confidence intervals.

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.05	0.09	0.55	0.58	-0.12	0.22	1.05
Final period dummy	-0.25	0.09	-2.96	0.003	-0.42	-0.09	0.78
% as cooperate	-0.46	0.07	-6.16	<0.001	-0.60	-0.31	0.63

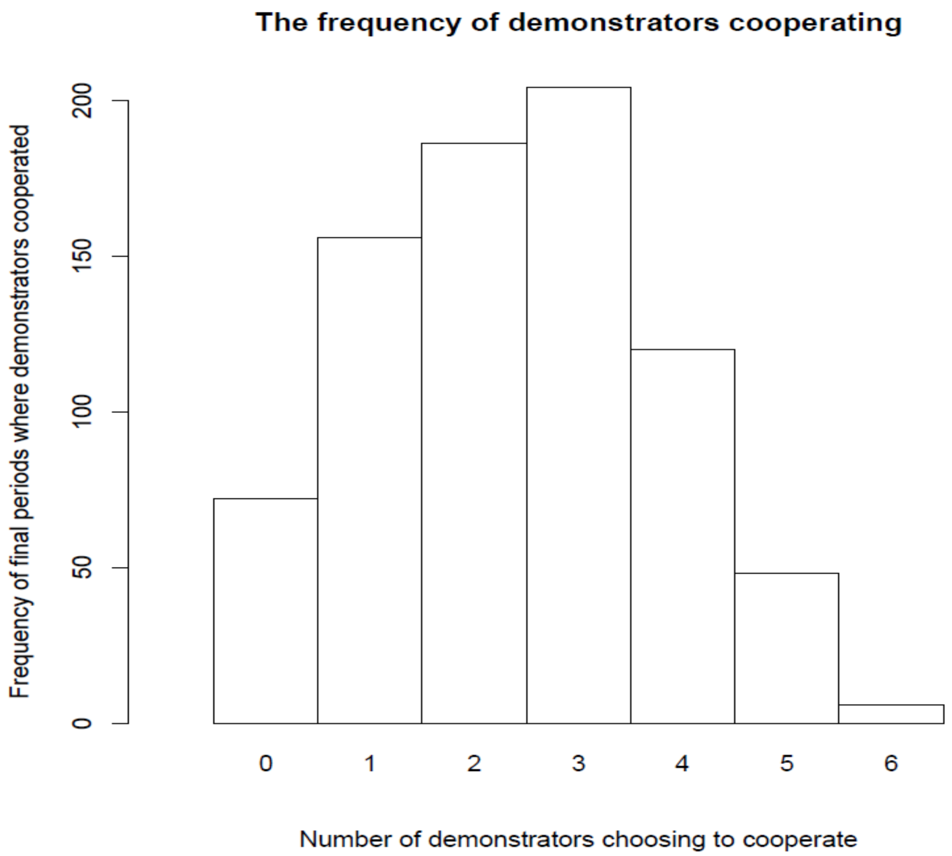


Figure 3. A histogram displaying the number of demonstrators who chose to cooperate distributed across the final periods of all blocks in the prisoner's dilemma.

3.2. Did the social learners flexibly adjust their strategies to each level of the social information presented?

To visualise the social learners' chosen strategies, figure 4 plots the proportion of social learners who chose % as a function of the number of demonstrators who chose % for each combination of the similarity and reliability signal. This allows us to compare the social learners' chosen frequency-dependent social learning strategies across each level of the similarity and reliability information.

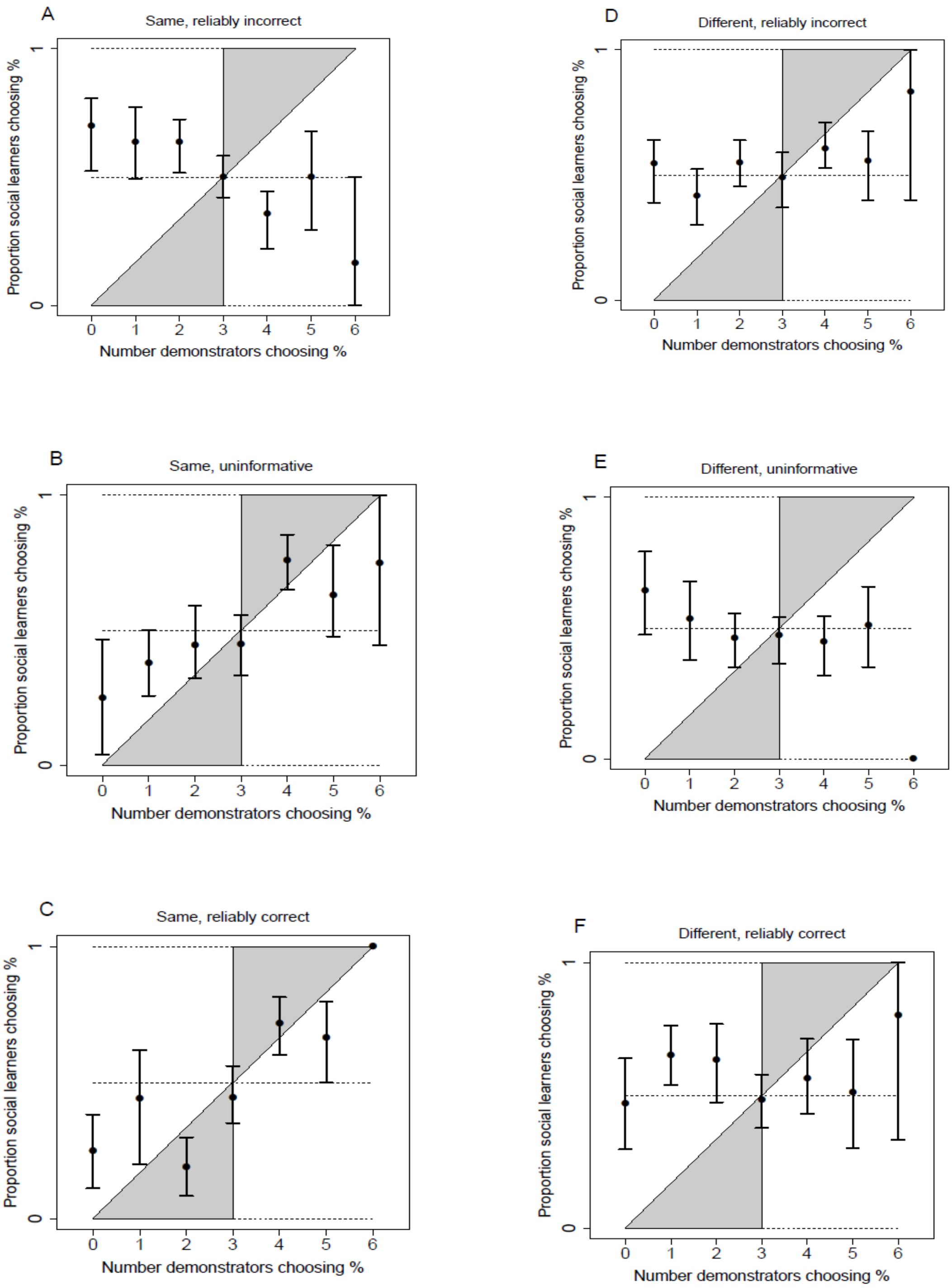


Figure 4. The proportion of social learners who chose % based on the number of demonstrators who chose %. The different panels show the social learners' responses to frequency-dependent social information by each level of the second and third-order social information, for the prisoner's dilemma. The error bars give the 95% bootstrapped confidence interval clustered on the social learners, to reflect the

multiple observations gathered per learner. The regions shaded in grey depict where the social learners' data would fall if they used conformity, while the dashed lines give points of reference for the proportion of learners choosing % at 0, 0.5, and 1.

For clarity going forward, we refer to signals where the demonstrators and the social learners were suggested to have played the 'same game' as each other as cases where the social learners learned from groups of similar others. Signals suggesting that they played a 'different game' to the demonstrators are referred to as learning from 'different others'. The social learning strategies depicted in figure 4 show a response to first-order social information. The social learners responded differently to blocks where the majority of demonstrators chose % compared to blocks where the minority of demonstrators chose %, for similar others at least (left-hand column of figure 3). This shows that the social learners responded to first-order social information as expected. To understand how the social learners responded to similarity information, one need only compare the side-by-side panels in Figure 4. The social learners used a distinct strategy in response to similar versus different others throughout. They processed second-order social information.

To understand how the social learners respond to reliability information, one need only compare the panels in the same column of figure 4. Take the social learners' response to similar others in the left column. The social learners copied the majority of the group with uninformative (figure 4B) and reliably correct signals of similarity (figure 4C), though these strategies did not reach the S-shaped curve typical of true conformity. The preference to follow the majority was closer to being linear than matching the *disproportionate* tendency with which I define conformity (Efferson et al., 2016). Social learners followed the minority of a group of similar others with reliably incorrect signals instead (figure 4A). This broadly matches the social learning strategies that were predicted in table 1, suggesting that the social learners did process second and third-order social information.

However, the social learners' responses to the reliability signal associated to groups of different others varied around chance (figures 4D-F). The social learners may thus have experienced an asymmetric adjustment in that they appeared to only process third-order social information when learning from groups of similar others. The logistic regression in table 3 investigates whether the social learners chose % based on all social information available, in order to further understand whether there was an asymmetric adjustment in the social learners' strategies.

Table 3. The general linear mixed model, modelling whether social learners chose %. Fixed effects included (i) the centred number of demonstrators who chose % on their final period, (ii) each combination of the similarity and reliability information, minus the omitted category of reliably incorrect– similar signals, and (iii) the interactions between each of these dummies and the centered proportion of demonstrators who chose %. I centered the proportion of demonstrators so that any block where 3/6 demonstrators chose % became the omitted category of the regression, and thus were reflected in the intercept. The random effects include participant ID and session ID. I also include Wald's confidence interval for each estimate. See appendix 6 for the full model, with control variables.

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.06	0.11	0.52	0.60	-0.16	0.27	1.06
Centred proportion of demonstrators choosing %	-1.79	0.44	-4.09	<0.001	-2.65	-0.93	0.17

Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	0.08	0.15	0.57	0.57	-0.20	0.37	1.09
Uninformative-same dummy [signal indicates different and is correct with 0.5 probability]	0.08	0.16	0.47	0.64	-0.24	0.39	1.08
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	-0.11	0.15	-0.78	0.43	-0.40	0.17	0.89
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability]	-0.10	0.16	-0.63	0.53	-0.41	0.21	0.90
Reliably correct-different dummy [signal indicates	0.16	0.15	1.09	0.28	-0.13	0.45	1.17

different and is correct with 0.9 probability]							
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	2.36	0.59	4.02	<0.001	1.21	3.52	10.64
Centred proportion of demonstrators choosing % X uninformative-same dummy	4.11	0.67	6.16	<0.001	2.78	5.42	60.77
Centred proportion of demonstrators choosing % X uninformative-different dummy	1.18	0.60	1.96	0.05	0.002	2.37	3.27
Centred proportion of demonstrators choosing % X reliably correct-same dummy	4.85	0.68	7.10	<0.001	3.51	6.19	128.21
Centred proportion of demonstrators choosing % X reliably correct-different dummy	1.47	0.61	2.41	0.02	0.28	2.67	4.36

The model in table 3 shows that the social learners became less likely to choose % as more demonstrators did (estimate=-1.793, z=-4.09, $p<0.001$). This estimate reflects the omitted category of reliably incorrect signals from similar others. The social learners followed the minority around reliably incorrect signals from similar others. This was predicted in table 1. If the social learners understood that a reliably incorrect signal of similarity is in fact more likely to indicate that the demonstrators played the different game to themselves, then they understood that they should follow the minority to uphold the same cooperative (or uncooperative) norm as this group.

The analysis reveals that the social learners became more likely to choose % as more demonstrators chose this under the following dummies: reliably incorrect–different, uninformative–same, uninformative-different, reliably correct–same and reliably correct–different dummies. Note that a reliably incorrect signal of difference is equivalent to a reliably correct signal of similarity and so it was expected that the social learners would follow the majority under both these blocks (table 1). The social learners showed a complex understanding of these signals.

The social learners' response to the reliably correct signal of difference did not match what I predicted in table 1, however. Flexible social learners would have realised that a reliably incorrect signal of similarity and a reliably correct signal of difference were equivalent in terms of social information and so I expected the social learners to follow the minority under both blocks. Instead, the social learners' responses seem to vary at chance around reliably different others (figure 4F). Perhaps the social learners did not understand that the reliably correct signal of difference was equivalent to a reliably incorrect signal of similarity. Alternatively, the decision of whether to conform around reliably different others may not be clear cut (see section 4).

To further test the social learners' flexibility, I performed linear combinations to investigate any differences in the social learners' preference to choose % across the similarity and reliability signals. I focused this analysis on blocks where the majority of the demonstrators had chosen % ($\geq$4), or none had. The logic is this. The social learners who respond flexibly to multiple pieces of social information would have a different preference to this frequency information as based on the different values of the similarity and reliability information. The social learners who instead followed a rule to 'always follow the majority' would choose % whenever the majority of the demonstrators had.

These linear combinations reveal that the social learners showed a significantly distinct response to similar versus different others at each level of the reliability signal (see appendix 7). This shows that the social learners processed second-order social information. The linear combinations also reveal an asymmetric adjustment to the reliability information. The social learners used a significantly distinct social learning strategy around similar others with reliably incorrect signals in comparison to similar others with both uninformative and reliably correct signals. There was no difference between the social learners' responses to uninformative versus reliably correct signals of similarity. The social learners showed a distinct response to reliably incorrect versus uninformative signals of difference, but otherwise there were no differences in the social learners' responses to any other reliability signals shown alongside groups of different others. The social learners processed third-order social information in as far as they showed a unique response to those with reliably incorrect signals of social information.

To summarise, the social learners processed second-order social information, responding differently to groups of similar versus different others. There was, however, an asymmetric adjustment to third-order social information. The social learners identified a reliably incorrect signal of similarity and responded to this in a different way as they did to the uninformative and reliably correct signals of similarity. The social learners may have responded differently to a group of different others with reliably incorrect signals, but this effect was less pronounced than for reliably incorrect signals from similar others. This is an interesting asymmetric adjustment in the social learners' chosen strategies, but this analysis cannot show whether the social learners were more likely to uphold cooperative norms when learning from groups with certain social signals. Section 3.3 now addresses whether the social learners upheld the same cooperative (or uncooperative) norm as the group of demonstrators from whom they learned.

3.3. Did the social learners choose to cooperate whenever the demonstrators did?

The uninformative signal rendered the similarity information at chance likelihood of being correct and so the social learners could only cooperate or defect at chance in response to these signals. The social learners could only develop a meaningful preference towards upholding the same cooperative (or uncooperative) norms as the group from whom they learned in response to blocks with informationally meaningful signals. The logistic regression in table 4 investigates whether the social learners cooperated as based on the number of demonstrators cooperating and the similarity and reliability information.

Table 4. The general linear mixed model, modelling whether the social learners chose to cooperate during the prisoner's dilemma. Fixed effects included: (i) the centered proportion of demonstrators who chose to cooperate on their final periods, (ii) dummies for each combination of the similarity and reliability information, minus the omitted category of reliably incorrect– similar signals, and (iii) interactions between each of these dummies and the centered proportion of demonstrators who cooperated. Random effects included participant ID and session ID, with Wald's 95% confidence interval for each parameter. See appendix 8 for full model with control variables.

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.03	0.10	0.30	0.76	-0.17	0.24	1.03
Centred proportion of demonstrators choosing cooperate	1.40	0.44	3.17	0.002	0.54	2.27	4.07
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	-0.02	0.15	-0.11	0.91	-0.32	0.28	0.98
Uninformative-same dummy [signal indicates different and is correct with 0.5 probability]	0.11	0.16	0.69	0.49	-0.21	0.43	1.12
Uninformative-different dummy [signal indicates	-0.03	0.15	-0.17	0.87	-0.32	0.27	0.97

different and is correct with 0.5 probability]							
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability]	0.03	0.17	0.20	0.84	-0.29	0.36	1.03
Reliably correct- different dummy [signal indicates different and is correct with 0.9 probability]	-0.09	0.15	-0.59	0.55	-0.38	0.20	0.92
Centred proportion of demonstrators choosing % X reliably incorrect- different dummy	-0.94	0.61	-1.55	0.12	-2.13	0.25	0.39
Centred proportion of demonstrators choosing % X uninformative- same dummy	-0.32	0.66	-0.48	0.63	-1.61	0.98	0.73
Centred proportion of demonstrators choosing % X uninformative- different dummy	-0.62	0.62	-1.01	0.31	-1.83	0.59	0.54

Centred proportion of demonstrators choosing % X reliably correct-same dummy	1.34	0.71	1.89	0.06	-0.06	2.73	3.82
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-1.35	0.61	-2.23	0.03	-2.54	-0.16	0.26

The logistic regression in table 4 shows that the social learners cooperated whenever all the demonstrators had (estimate=1.403, z=3.170, p=0.002). Note that this estimate reflects the omitted category for reliably incorrect signals from similar others. By following the minority of the group of similar others with reliably incorrect signals (figure 4A), then the social learners managed to uphold cooperation whenever the majority of demonstrators had cooperated. That is, the social learners were seemingly aware that they should choose the opposite option to groups of similar others with reliably incorrect signals, in order to uphold the same cooperative norm as this group.

Interestingly, the social learners were less likely to cooperate as more demonstrators did when responding to a reliably correct signal from different others, in comparison to the baseline of reliably incorrect signals from similar others (estimate= -1.353, z= -2.227, p=0.03). This could suggest that the lack of response towards reliably different others in Figure 4F is significantly distinct to the minority-based learning strategy towards unreliably similar others in Figure 4A. Despite the fact that these two signals are informationally equivalent, the social learners do not show a definitive strategy towards reliably different others, suggesting that learning cooperative norms from this group may represent a special case.

An interesting caveat of my design was that the similarity signal was only correct with a certain probability. For example, take a block where the social learners were told that they were playing the same game as the demonstrators and that this information was only correct 1 time out of 10 (i.e., a reliably incorrect signal of similarity). In this case, the signal may correctly inform the social learner that she played the same game as the group from whom she learned with a probability of 0.1, though it was more likely that this signal was incorrect and that the social learner played a different game to the demonstrators.

To investigate whether the social learners understood this complexity, I repeated the above analysis for blocks where the similarity information happened to be correct versus blocks where the similarity information happened to be incorrect only (see appendix 9). Let’s start with the blocks that happened to give correct similarity information only. The social learners cooperated whenever the majority of demonstrators had, provided that the blocks gave uninformative or reliably correct signals of similarity. This suggests that the social learners treat an uninformative and reliably correct signal of similarity as if they always provided correct information. Let us now turn to blocks which happened to give incorrect signals of similarity. The social learners were more likely to cooperate whenever more demonstrators did, under blocks with reliably incorrect signals of similarity only. The social learners were less likely to cooperate as more demonstrators did under every other signal combination. Taken together, these analyses show that the social learners treat an uninformative and reliably correct signal of similarity as if they were always correct but treat a reliably incorrect signal as if it was always incorrect. This suggests that the social learners understand the reliability signals given alongside similar others at least.

To further understand the social learners' preferences, I performed linear combinations between the social learners' desire to cooperate at each level of the similarity and reliability information (see appendix 10). I performed these linear combinations on blocks where the majority of the demonstrators had cooperated ($\geq$4), or none had. Social learners who used the social information flexibly should have cooperated whenever the majority of demonstrators had cooperated, regardless of the signals. If there were trade-offs in the amount of social information that the social learners could process, then the social learners may have found it easier to uphold the same cooperative norms as certain social groups compared to others. These linear combinations show that the social learners were significantly more likely to cooperate when learning from a group of similar others with reliably correct signals than they were when learning from a group of different others with reliably correct signals. This may be due to the random strategy at the aggregate level seen in response to the reliably correct signals from different others.

To summarise, the social learners were the most likely to cooperate when learning from similar others with reliably correct signals (figure 4C). The random response to reliably correct signals of different others meant that the social learners were less likely to uphold the same cooperative behaviours as this group in comparison to reliably similar others. They also treated a reliably incorrect signal of similarity as if the similarity information was always false.

3.4. Summary of the social learning strategies

The social learners responded to first-order social information (the frequency of choices made by the demonstrators) and to second-order social information (the similarity information). However, they showed an asymmetric adjustment as the social learners were more likely to uphold the *same* cooperative norm as the demonstrators when the group had learned in a (reliably) similar decision-making environment to themselves, as opposed to a (reliably) different decision-making environment. The social learners found it easier to copy the majority choice around reliably similar others than to follow the minority around reliably different others.

The social learners did process third-order social information (the reliability of the similarity signals) though only to an extent. They processed reliability information in as far as they could identify reliably incorrect signals from similar others and follow the minority choice of this group. Although the social learners did process reliably incorrect signals from different others distinctly, this effect was not as strong. To summarise, second-order social information was clearly important when learning to cooperate. The social learners most readily learned cooperative norms from groups of similar others with reliably correct signals. They processed third-order social information enough to avoid conforming to groups of similar others with reliably incorrect signals.

4. Discussion

This study investigated the flexibility with which individuals chose frequency-dependent social learning strategies when learning to cooperate under conditions of uncertainty. The social learners flexibly adjusted to both first-order (frequency-dependent) and second-order (similarity) information, though they showed an asymmetric response to third-order (reliability) information. This supports prediction (ii) of the introduction. The social learners adjusted to all levels of the social information but there was a trade-off which resulted in asymmetric adjustments. The social learners processed the reliability information when learning from similar others but not different others.

Perhaps the social learners can process second-order social information due to a high rate of intergroup contact which may have exposed individuals to the cooperative norms of different social groups throughout the ancestral past (Boyd & Richardson, 1985 (chapter 7); Efferson et al., 2008a). However, there was a trade-off when processing second-order social information. The social learners more readily copied the majority of similar others than they follow the minority around different others. This corroborates past research which shows that individuals learn cooperative norms from demonstrators who are similar to themselves (House et al., 2013; Salali et al., 2015). This preference for similarity extends to the members of an entire group, too.

While social learners displayed a preference to learn from groups of similar others, it is important to acknowledge that the operationalization of the similarity signal in the current study directly communicated whether the social learners shared a 'similar' or 'different' decision-making environment to the group from whom they learned. Previous work instead focused on similarity on observable traits (House et al., 2013; Molleman et al., 2019; Salali et al., 2015; Shutts et al., 2010). Note that it is important to copy the cooperative behaviours of those that look similar to oneself as far as looking similar may be a short-hand for those that learn to cooperate in a similar environment to oneself (McElreath et al., 2003; Richerson et al., 2016). I directly told the participants about which groups shared similar decision-making environments–

even if the individual was unsure as to whether these norms are cooperative or uncooperative– to remove the need for participants to calculate their similarity to others and test the three orders of social information as straightforwardly as possible.

However, it could be that the social learners display a stronger preference to copy those that *appear* similar rather than those who made decisions in a similar decision-making environment as in the current study. A preference to learn from, and interact with, those that *appear* similar may be shaped by ancestral pressures towards in-group preferences and out-group prejudices (Efferson et al., 2008a; Ihara, 2011; Konrad & Morath, 2012; Vogt et al., 2013). Finding a bias to copy those that *look* similar to oneself could have implications in academic and business groups, as these still tend to consist of white middle- to upper-class individuals which may not represent all individuals (Bell et al., 2021; Chávez & Mitchell, 2020).

Future extensions to this study may therefore wish to consider how similarity on an arbitrary trait– such as an on-screen avatar– affects how the participants learn from others rather than directly telling them about the underlying decision-making environments. This would be similar to how ethnic markers would likely have become signals cueing similar decision-making environments throughout the ancestral past (Efferson et al., 2008a). To incorporate the reliability signals, imagine that on some rounds this avatar remains uninformative but, on other rounds, it becomes linked to the game (left or right) being played such that they become a reliably correct or reliably incorrect signal of who shares a similar decision-making environment to oneself.

The social learners did not blindly copy groups of similar others, however. They processed the reliability information. Social learners followed the majority of a group of similar others with reliably correct signals (figure 4C), though they instead followed the minority around a group of similar others with reliably incorrect signals (figure 4A).

A reliably incorrect signal of similarity implies that the demonstrators learned to make decisions in a different environment to the social learners. In order to uphold the same cooperative– or uncooperative– norm as the demonstrators from whom they learn, then the social learners would know to choose the opposite option to the demonstrators as this was a binary decision-making game (Efferson et al., 2016).

Reliably incorrect signals of similarity may represent faked signals of shared group membership (Stein et al., 2021; Watson-Jones & Legare, 2016). Individuals who faked signals of group membership were often free riders who would not reciprocate any cooperative donations. Of course, a successful cooperative group member should avoid copying these free riders to avoid social punishment herself (Molleman et al., 2019). In the introduction, I walked through an example of a free-rider who tried to infiltrate a cooperative group by wearing red clothing, similar to the cooperative group's members. However, her clothing was an off-shade of red and so may have become a reliably incorrect signal of similarity to the cooperative group's members. Such scenarios may have been common enough in the ancestral past to select for social learner cognition that is flexible enough to account for reliably incorrect signals of similarity. Put simply, cooperative groups were at risk of infiltration by free riders who pretend to be similar to the cooperative group's members throughout the ancestral past. In order to counter this infiltration, it was important that individuals could identify reliably incorrect signals of similarity as these were likely to be faked by free riders.

While the social learners adjusted to reliably incorrect signals from different others distinctly to uninformative signals, their responses to the reliability signals given alongside groups of different others were less pronounced at the aggregate level than their response to the reliability signals from similar others (see figure 4). It could have been a fairly common scenario that free riders faked similarity to cooperative group members in order to elicit a resource donation (Sosis et al., 2007). Therefore, our cognition could have evolved to compensate for this. The inverse scenario, where a person with similar cooperative intentions to oneself pretends to be different, was perhaps less likely to occur. After all, a mismatch on a cooperative interaction is likely to carry serious fitness consequences. Cooperating when no one else does could result in a substantial loss of resources (Capraro, 2013) while free riding in an otherwise cooperative group is likely to attract negative attention in the form of ostracism, reputation loss, or social punishment (Molleman et al., 2019; Price & Johnson, 2011; van den Berg et al., 2012). As it was risky to clash with the cooperative or uncooperative actions of others in the social group, then it is unclear whether those who uphold similar cooperative values to oneself would ever have pretended to be different.

It is demanding to process multiple sources of social information simultaneously (Muthukrishna et al., 2016). If there was an uneven selection pressure for individuals to be able to process the reliability of similarity signals– but not different signals– in the ancestral past, then it would make sense for the social learners to process the reliability signals alongside similar but not different others in the current study. This finding thus supports past research suggesting that social learner cognition is flexible but not fully so (Efferson et al., 2016; Krafft et al., 2021).

Whilst I reference the importance of the ancestral past (Mercier & Morin, 2019), it must be noted that this thesis cannot conclude whether this bias to learn from reliably similar others was genetic and/or cultural in nature. The preference to learn from reliably similar others may be socially learned, as the participants are unlikely to have direct experience with those that pretend to be different to themselves and so they may not have formed a learning rule to respond to this scenario throughout their daily lives (Heyes, 2016).

It was unclear whether this asymmetric trade-off occurred as all participants adjusted asymmetrically to the signals, or whether some participants were completely flexible in their choices of frequency-dependent social learning strategies whilst others merely followed a rule-of-thumb to always 'follow the majority' or 'minority' which would cancel out to an asymmetric adjustment at the group level. To investigate the individual variation in social learning strategies, I created heatmaps and scatterplots to envision the participant's response to both reliably correct and reliably incorrect (see appendix 11-12) signals. These plot the proportion of blocks where the social learners follow the majority of similar others by the blocks where the social learners followed the majority of different others (note that the proportion who follow the minority is simply the remainder). These plots reveal two interesting observations. First, the scatterplots reveal that each participants' social learning strategies changed between reliably correct and reliably incorrect signals. Tht is, all social learners did try to adjust to third-order social information.

However, few did so optimally. The heatmaps show the social learning strategies across the entire space of third-order social information. The optimal response to reliably correct signals would be to always follow the majority of similar others, and to never follow the majority of different others. A few participants did this. For reliably incorrect signals, one should never copy the majority of similar others but always copy the majority of different others. Again, a few participants did this. But most participants answered sub-optimally as they lie elsewhere in the social learning strategy space.

Interestingly, I compare the participants who partially adjusted to both the similar and different signals to those that reported ignoring social information in the end-survey (see supplementary material). I found that the participants in the middle of the social learning strategy space, who did not commit to a full strategy, were not necessarily the ones who ignored social information. This suggests that this partial adjustment is a meaningful– if suboptimal– strategy. This finding corroborates past research which suggests that a subset of participants will never conform even when conformity would have been optimal (Efferson et al., 2008b; McElreath et al. 2008) or will always conform even when conformity would have been suboptimal (Goeree & Yariv, 2015). I extend this finding of suboptimal strategies to learning cooperative norms in a Prisoner's Dilemma.

Another observation suggesting that the social learners were not fully flexible when processing multiple sources of social information comes from their response to reliably correct signals from different others (figures 4F). The social learners seemingly chose one of the two options randomly in response to these blocks. Astute social learners would have realised that a reliably incorrect signal of similarity provides equivalent social information to a reliably correct signal of difference and so it was expected that the participants would follow the minority of groups displaying both signals. However, there was a significant difference in the social learning strategies being used on these blocks, which lead to different rates of cooperation across the blocks. The social learners did follow the minority for blocks with reliably incorrect signals of similarity but respond at chance to reliably correct signals of difference.

It is perhaps less clear cut what one should do in response to social information from reliably different others. Some may oppose the cooperative norms of different others on principle (Chudek & Henrich, 2011). On the other hand, some individuals may assimilate cooperation from out-group members and attempt to kickstart this in their own group if they see evidence that the outgroup members are getting better payoffs from cooperation (Burton-Chellew et al., 2012; 2015). Of course, this strategy may only be possible with access to group-wide payoffs which I did not include in the current study. This was because I wished to focus on the flexibility of frequency-dependent social learning strategies, such as conformity. Without payoff information, the social learners may not have felt comfortable committing to a definitive strategy around groups of reliably different others. Further research should investigate how flexible the social learners are when responding to groups of reliably different others with access to payoff information, too.

Another reason as to why the social learners may have found it difficult to commit to a definitive social learning strategy in response to reliably correct signals from different others may have been due to the social information presented in this study. The social learners could only follow the majority or minority choices made by the demonstrators in terms of the arbitrary symbols, @ and %. Due to the design of the game, the social learners could not know for sure which of these symbols represented ‘cooperation’ or ‘defection’. For example, imagine that a social learner was told that 6/6 demonstrators chose % and that she was reliably similar to these demonstrators. The social learner could not know for sure whether both her and the demonstrators played game left, or whether they both played game right. In the former case, all 6 demonstrators chose % as they wished to defect. In the latter case however, all 6 demonstrators chose % as they wished to cooperate. That is, the social learner could not know for sure whether the group of demonstrators from whom she learned were upholding a cooperative or an uncooperative norm.

This design was important for two reasons. First, labelling options as ‘cooperate’ or ‘defect’ could create a social desirability effect where the participants chose to cooperate to please the experimenter (Price, 2008). Second, there is a wealth of domains in which a group may cooperate (Chudek & Henrich, 2011). For example, some groups cooperate to go hunting but not to build shelter, while other groups cooperate to build shelter but not to hunt (Chudek & Henrich, 2011; Chudek et al., 2013). This uncertainty becomes even greater when one considers that it is often unclear which behaviour would be cooperative in a given domain. For example, the decision to hunt. Sometimes, it is cooperative to hunt more and share one's food with the group (Hill, 2002). Other times, it is cooperative to hunt less so as to not over-deplete natural resources (Safin et al., 2015). The arbitrary nature of cooperative rules could introduce confusion to individuals who are just learning the cooperative norm of a social group, such as recent migrants (Mesoudi, 2018) or young children (Legare, 2019). As these individuals may be uncertain over the cooperative norms of the group, then it would make sense to just conform to the majority of the group without fully understanding what this behaviour entails (Szolnoki & Perc, 2015). That is, by not clearly aligning the arbitrary symbol to a 'cooperate' or 'defect' payoff, then this game reflected a natural level of uncertainty in the social learners.

Of course, the social learners may have expressed a different use of social information if they were aware that the majority of demonstrators 'cooperated' or 'defected'. Perhaps when these labels are available, the social learners may try to uphold cooperation regardless of what the demonstrators did due to a social desirability effect. Or, if my intuition is correct that the three pieces of social information about the group from whom one learns is important, then I would instead expect to see that the social learners adjust to frequency, similarity, and reliability information regardless of whether the demonstrators transmitted a preference to cooperate or defect. It would be useful for future research to make this comparison. For the purposes of my current study however, I wished to investigate how social learners adjusted their strategies to social information *about* the group from whom they learn without necessarily knowing whether they learn to uphold a *cooperative* or *uncooperative* behaviour.

A limitation of this study's design that could also be addressed by future research is the structure of the prisoner's dilemma. The blocks in each session repeated for four periods. All participants knew that cooperative interactions were finite and had an end-point. This could explain why most demonstrators had a preference to defect by the final period of the block. Participants do not continue cooperation if they are aware that this round is the end-point of an interaction with someone (McAuliffe et al., 2019). It would be interesting to model a repeated dilemma, where each round of interaction repeats with a probability, *p*. This could introduce some blocks where one-shot interactions are common and some blocks where repeated interactions are common. The opportunity for cooperative interactions to repeat, and the uncertainty over whether they do in fact repeat, may have been necessary for costly levels of human cooperation to have emerged in the ancestral past (Delton et al., 2011; Krasnow & Delton, 2016).

To summarise, the social learners showed a complex yet asymmetric adjustment to social information when learning to cooperate under uncertainty. They based their strategies on both first-order (frequency) and second-order (similarity) information. There was an asymmetric adjustment to third-order information. The social learners processed reliability enough to follow the minority around groups of similar others with reliably incorrect signals. A reliably incorrect signal of similarity may imply that there are free-riders who are only pretending to have similar cooperative intentions to oneself in order to exploit one's resources. The social learners did not seemingly process the reliability signal alongside different others. There was perhaps less pressure to do this in the ancestral past as individuals were unlikely to pretend to be different to each other. After all, social information is demanding to process, and social learner cognition is unlikely to be fully flexible. Indeed, these findings show that the social learners' decision to conform or use other social learning strategies was flexible to a third-order strategy space, though there was an upper limit to social learner flexibility, when learning to cooperate.

5. References

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98-108. https://doi.org/10.1016/j.labeco.2014.05.008.

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, *24*(1), 3-25.https://doi.org/10.1177/002200278002400101.

Bell, M. P., Berry, D., Leopold, J., & Nkomo, S. (2021). Making Black Lives Matter in academia: a black feminist call for collective action against anti-blackness in the academy. *Gender, Work & Organization*, *28*(1), 39-57. DOI: 10.1111/gwao.12555.

Boyd, R., & Richerson, P. J. (1985). Frequency-dependent Bias and the Evolution of Cooperation. In R. Boyd & P.J. Richerson (Eds.) *Culture and the Evolutionary Process*, (p 204-241). Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (2007). Cultural adaptation and maladaptation: Of Kayaks and Commissars. In S.W. Gangestad & J.A. Simpson (Eds.) *The evolution of the mind: Fundamental questions and controversies* (p327-332). New York: Guilford Press.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences*, *113*(5), 1291-1296. https://doi.org/10.1073/pnas.1509740113.

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142678. https://doi.org/10.1098/rspb.2014.2678.

Burton-Chellew, M. N., & West, S. A. (2012). Pseudocompetition among groups increases human cooperation in a public-goods game. *Animal Behaviour*, *84*(4), 947-952. https://doi.org/10.1016/j.anbehav.2012.07.019.Capraro, V. (2013). A model of human cooperation in social dilemmas. *PloS one*, *8*(8), e72427. https://doi.org/10.1371/journal.pone.0072427.

Chávez, K., & Mitchell, K.M.W. (2020). Exploring bias in student evaluations: Gender, race, and ethnicity. *PS: Political Science & Politics*, *53*(2), 270-274. doi:10.1017/S1049096519001744.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*(5), 218-226. https://doi.org/10.1016/j.tics.2011.03.003.

Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation, and the shaping of human social psychology. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 425–457). US: The MIT Press.

Colman, A. M. (2006). The puzzle of cooperation. *Nature*, *440*(7085), 744-745. DOI: https://doi.org/10.1038/440744b.

Cutmore, T. R., H., Halford, G. S., Wang, Y., Ramm, B. J., Spokes, T., & Shum, D. H.K. (2015). Neural correlates of deductive reasoning: An ERP study with the Wason Selection Task. *International Journal of Psychophysiology*, *98*(3:Part 1), 381-388. https://doi.org/10.1016/j.ijpsycho.2015.07.004.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: a comparative perspective. *Biological Reviews*, *89*(2), 284-301. doi: 10.1111/brv.12053.

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, 108(32), 13335-13340. DOI: https://doi.org/10.1073/pnas.1102131108.

Efferson, C., Lalive, R., Cacault, M. P., & Kistler, D. (2016). The evolution of facultative conformity based on similarity. *PloS one*, *11*(12), e0168551. https://doi.org/10.1371/journal.pone.0168551.

Efferson, C., Lalive, R., & Fehr, E. (2008a). The coevolution of cultural groups and ingroup favoritism. *Science*, *321*(5897), 1844-1849. DOI: 10.1126/science.1155805.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008b). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, *29*(1), 56-64. https://doi.org/10.1016/j.evolhumbehav.2007.08.003.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature*, *13*(1), 1-25.

Fehr, E. & Fischbacher, U. (2005). Human Altruism – Proximate Patterns and Evolutionary Origins. *Analyse & Kritik*, *27*(1), 6-47. https://doi.org/10.1515/auk-2005-0101.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178. https://doi.org/10.1007/s10683-006-9159-4.

Fischer, I. (2009). Friend or foe: subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General*, *138*(3), 341-350. DOI: https://doi.org/10.1037/a0016073.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*(2), 241-253. https://doi.org/10.1007/s11211-008-0067-y.

Goeree, J. K., & Yariv, L. (2015). Conformity in the lab. *Journal of the Economic Science Association*, *1*(1), 15-28. https://doi.org/10.1007/s40881-015-0001-7.

Henrich, J. (2015). *The secret of our success*: *How culture is driving human evolution, domesticating our species and making us smarter.* USA: Princeton University Press.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and human behavior*, *19*(4), 215-241. https://doi.org/10.1016/S1090-5138(98)00018-X.

Henrich, J, & Boyd R. (2001). Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology, 208*(1), 79– 89. https://doi.org/10.1006/jtbi.2000.2202.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, *91*(2), 73-78. DOI: 10.1257/aer.91.2.73.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. & Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767-1770. DOI: 10.1126/science.1127333.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes, C. (2016). Blackboxing: social learning strategies and cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150369. https://doi.org/10.1098/rstb.2015.0369.

Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature*, *13*(1), 105-128. https://doi.org/10.1007/s12110-002-1016-3.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization, 10*(4), 15-41. https://doi.org/10.1080/00208825.1980.11656300.

Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*(12), 3999-4002. https://doi.org/10.1073/pnas.0308738101.

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B.S., McElreath, R. & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, *110*(36), 14586-14591. https://doi.org/10.1073/pnas.1221217110.

Ihara, Y. (2011). Evolution of culture-dependent discriminate sociality: a gene–culture coevolutionary model. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1566), 889-900. https://doi.org/10.1098/rstb.2010.0247.

Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(20), 1-12. https://doi.org/10.1057/s41599-019-0228-7.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Konrad, K. A., & Morath, F. (2012). Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology*, *306*, 61-67. https://doi.org/10.1016/j.jtbi.2012.04.013.

Krafft, P. M., Shmueli, E., Griffiths, T. L., Tenenbaum, J. B., & Pentland, A. (2021). Bayesian collective learning emerges from heuristic social learning. *Cognition*, *212,* 104469. https://doi.org/10.1016/j.cognition.2020.104469.

Krasnow, M. M., & Delton, A. W. (2016). Are humans too generous and too punitive? Using psychological principles to further debates about human social evolution. *Frontiers in psychology*, *7*, 799. https://doi.org/10.3389/fpsyg.2016.00799.

Lachlan, R.F., Ratmann, O. & Nowicki, S. (2018). Cultural conformity generates extremely stable traditions in bird song. *Nature Communications,* 9(2417). https://doi.org/10.1038/s41467-018-04728-1.

Legare, C. H. (2019). The development of cumulative cultural learning. *Annual Review of Developmental Psychology*, *1*, 119-147. https://doi.org/10.1146/annurev-devpsych-121318-084848.

Li, X., Molleman, L., & van Dolder, D. (2021). Do descriptive social norms drive peer punishment? Conditional punishment strategies and their impact on cooperation. *Evolution and Human Behavior*, *42*(5), 469-479. DOI: https://doi.org/10.1016/j.evolhumbehav.2021.04.002.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives*, *14*(3), 115-136. DOI: 10.1257/jep.14.3.115.

McAuliffe, W. H.B., Burton-Chellew, M. N., & McCullough, M. E. (2019). Cooperation and Learning in Unfamiliar Situations. *Current Directions in Psychological Science*, *28*(5), 436-440. https://doi.org/10.1177/0963721419848673.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3515-3528. https://doi.org/10.1098/rstb.2008.0131.

McElreath, R., Boyd, R., & Richerson, P.J. (2003). Shared norms and the evolution of ethnic markers. *Current anthropology*, *44*(1), 122-130. https://doi.org/10.1086/345689.

Mercier, H., & Morin, O. (2019). Majority rules: how good are we at aggregating convergent opinions?. *Evolutionary Human Sciences*, *1*, e6. doi:10.1017/ehs.2019.6.

Mesoudi, A. (2018). Migration, acculturation, and the maintenance of between-group cultural variation. *PloS one*, *13*(10), e0205573. https://doi.org/10.1371/journal.pone.0205573.

Mesoudi, A., Chang, L., Dall, S. R.X., & Thornton, A. (2016). The evolution of individual and cultural variation in social learning. *Trends in ecology & evolution*, *31*(3), 215-225. https://doi.org/10.1016/j.tree.2015.12.012.

Mesoudi, A., Cheng, L., Murray, K., & Lu, H. J. (2015). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20142209. https://doi.org/10.1098/rspb.2014.2209.

Molleman, L., & Gächter, S. (2018). Societal background influences social learning in cooperative decision making. *Evolution and Human Behavior, 39*(5), 547-555. https://doi.org/10.1016/j.evolhumbehav.2018.05.007.

Molleman, L., Kölle, F., Starmer, C., & Gächter, S. (2019). People prefer coordinated punishment in cooperative interactions. *Nature human behaviour*, *3*(11), 1145-1153. https://doi.org/10.1038/s41562-019-0707-2.

Molleman, L., Pen, I., & Weissing, F. J. (2013a). Effects of conformism on the cultural evolution of social behaviour. *PloS one*, *8*(7), e68153. DOI: https://doi.org/10.1371/journal.pone.0068153.

Molleman, L., Quiñones, A. E., & Weissing, F. J. (2013b). Cultural evolution of cooperation: the interplay between forms of social learning and group selection. *Evolution and Human Behavior*, *34*(5), 342-349. https://doi.org/10.1016/j.evolhumbehav.2013.06.001.

Molleman, L., Van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications*, *5*(3570), 1-9. https://doi.org/10.1038/ncomms4570.

Morgan, T. J.H., Acerbi, A., & Van Leeuwen, E. J.C. (2019). Copy-the-majority of instances or individuals? Two approaches to the majority and their consequences for conformist decision-making. *PloS one*, *14*(1), e0210748. https://doi.org/10.1371/journal.pone.0210748.

Muthukrishna, M., Morgan, T. J.H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, *37*(1), 10-20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004.

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Social Justice Research*, *21*(2), 228-240. DOI: https://doi.org/10.1007/s11211-008-0064-1.

Price M.E., Johnson D.D.P. (2011) The Adaptationist Theory of Cooperation in Groups: Evolutionary Predictions for Organizational Cooperation. In: G. Saad (eds) *Evolutionary Psychology in the Business Sciences* (pp 95-133). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-92784-6_5.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E.K., Naar, N., Newson, L., Ross, C., Smaldino, P.E., Waring, T.M. & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*, e30. doi:10.1017/S0140525X1400106X.

Safin, V., Arfer, K. B., & Rachlin, H. (2015). Reciprocation and altruism in social cooperation. *Behavioural processes*, *116*, 12-16. https://doi.org/10.1016/j.beproc.2015.04.009.

Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior*, *36*(2), 86-94. https://doi.org/10.1016/j.evolhumbehav.2014.09.004.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and society*, *7*(1), 58-92. DOI: https://doi.org/10.1177/1043463195007001004.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental science*, *13*(4), 599-610. DOI: 10.1111/j.1467-7687.2009.00913.x.

Smaldino, P. E., Flamson, T. J., & McElreath, R. (2018). The evolution of covert signaling. *Scientific reports*, *8*(4905), 1-10. DOI: https://doi.org/10.1038/s41598-018-22926-1.

Smaldino, P., & Jones, J. (2021). Coupled dynamics of behaviour and disease contagion among antagonistic groups. *Evolutionary Human Sciences, 3*, E28. doi:10.1017/ehs.2021.22.

Sosis, R., Kress, H. C., & Boster, J. S. (2007). Scars for war: Evaluating alternative signaling explanations for cross-cultural variance in ritual costs. *Evolution and Human Behavior*, *28*(4), 234-247. https://doi.org/10.1016/j.evolhumbehav.2007.02.007.

Stein, D. H., Hobson, N. M., & Schroeder, J. (2021). A sacred commitment: How rituals promote group survival. *Current Opinion in Psychology*, *40*, p114-120. https://doi.org/10.1016/j.copsyc.2020.09.005.

Szolnoki, A., & Perc, M. (2012). Conditional strategies and the evolution of cooperation in spatial public goods games. *Physical Review E*, *85*(2), 026104. DOI:https://doi.org/10.1103/PhysRevE.85.026104.

Szolnoki, A., & Perc, M. (2015). Conformity enhances network reciprocity in evolutionary social dilemmas. *Journal of The Royal Society Interface*, *12*(103), 20141299. https://doi.org/10.1098/rsif.2014.1299.

Szolnoki, A., Perc, M. (2019). Seasonal payoff variations and the evolution of cooperation in social dilemmas. *Scientific Reports 9*(12575). https://doi.org/10.1038/s41598-019-49075-3.

van den Berg, P., Molleman, L., & Weissing, F. J. (2012). The social costs of punishment. *Behavioral and Brain Sciences, 35*(1), 42-43. DOI:10.1017/S0140525X11001348.

Van Leeuwen, E. J.C., & Haun, D. B.M. (2014). Conformity without majority? The case for demarcating social from majority influences. *Animal Behaviour*, *96*, 187-194. https://doi.org/10.1016/j.anbehav.2014.08.004.

Vogt, S., Efferson, C., & Fehr, E. (2013). Can we see inside? Predicting strategic behavior given limited information. *Evolution and Human Behavior*, *34*(4), 258-264. https://doi.org/10.1016/j.evolhumbehav.2013.03.003.

Watson-Jones, R. E., & Legare, C. H. (2016). The social functions of group rituals. *Current Directions in Psychological Science*, *25*(1), 42-46. https://doi.org/10.1177/0963721415618486.

Wen, N. J., Clegg, J. M., & Legare, C. H. (2019). Smart conformists: Children and adolescents associate conformity with intelligence across cultures. *Child development*, *90*(3), 746-758. DOI: 10.1111/cdev.12935.

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, *33*(4), 341-356. https://doi.org/10.1016/j.dr.2013.08.002.

Yang, H. X., & Tian, L. (2017). Enhancement of cooperation through conformity-driven reproductive ability. *Chaos, Solitons & Fractals*, *103*, 159-162. https://doi.org/10.1016/j.chaos.2017.06.005.

Data Transparency

For the, supplementary material, raw data and all analysis scripts please visit this link: OSF | When will you learn to help? Social-learning strategies remain somewhat flexible to a range of social information when learning to cooperate.

6. Appendix

Appendix 1: The instructions booklet shown to participants, including pre-game questions to check understanding and answer sheet (for experimenter use only).

Welcome! You are invited to participate in a study for approximately **1 ½ hours.** You can earn points during this study, which will be converted to money at the following rate:

1150 points = £1

You will also be paid a show-up fee of **£4** on top of the money you earn. The choices participants make during the study will be anonymous. This means you will not be able to identify the specific participants in the room who make certain choices, and none of the participants will be able to trace your choices back to you.

Please do not communicate with the other participants. If you have questions, or need to withdraw, then please raise your hand and tell the researcher.

Please read this instruction sheet **carefully**. You will then answer some questions to check that you have understood the study. We will not be able to proceed until everyone answers **all** questions correctly. You will also respond to a brief **survey** after the main study.

The study:

To begin, the computer will randomly choose six of you to be Type A Participants. Others will be Type B Participants. As explained later, your type will determine how often you make choices and the information you have when you do so. The study lasts for 88 periods. We will divide these 88 periods into 22 blocks of 4 periods each. Type A Participants will choose every period, which means they will make 88 choices. Type B Participants will only make a choice in the final period of each block, which means that Type B Participants will make 22 choices. Don't worry. Though Type A and Type B Participants do not make the same number of choices, they will have exactly the same opportunity to earn points. We will explain this in detail later. The upper left-hand corner of your screen will have a counter that displays the current period you are in.

The games:

At the beginning of each block of 4 periods, the computer will randomly pair you with another participant of the same type to play a game. A Type A Participant will always be paired with another Type A, and a Type B Participant will always be paired with another Type B participant. Every time you play, both you and your partner must choose one of two options, either option "%" or option "@".

Specifically, there are two games, which we call "Game Left" and "Game Right". At the beginning of each block, the computer will randomly pick which game Type A Participants play and which game Type B Participants play. The computer decides this completely randomly, giving four possible combinations, which are all equally likely to occur (each with a 1 in 4 probability). The four possibilities are:

(i) both types play Game Left,

(ii) Type A Participants play Game Left, Type B Participants play Game Right,

(iii) Type A Participants play Game Right, Type B Participants play Game Left,

(iv) both types play Game Right.

Note that we do not tell you if you are playing Game Left or Game Right. The following tables show you how your points will depend on the choices you make, for each of the games that you might play.

Game left

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 145 Partner's expected points: 145	Your expected points: 280 Partner's expected points: 100
You choose @	Your expected points: 100 Partner's expected points: 280	Your expected points: 235 Partner's expected points: 235

Game right

	Partner chooses %	Partner chooses @
You choose %	Your expected points: 235 Partner's expected points: 235	Your expected points: 100 Partner's expected points: 280
You choose @	Your expected points: 280 Partner's expected points: 100	Your expected points: 145 Partner's expected points: 145

As you can see, one option is worth more points and will therefore result in you earning more money, on average, if you pick this option. The option (% or @) that is worth the most points is different depending on whether you are playing Game Left or Game Right.

You can see that your payoff depends on what your **partner** chooses. Likewise, your partner's payoff depends on what you choose. You can expect to earn a different amount of points on trials where you both you and your partner choose the same option, as compared to trials where you choose a different option to your partner's.

Lastly, points will also be affected by forces outside of your control, as in real life. The tables above show the expected points you will earn, but random shocks will be applied to these values. This random shock has a **mean of 0** and a **standard deviation of 20 points and** will be added onto the expected value shown in the table**.** This means that these random shocks will sometimes lead you to earn **more points** than those shown in the table, and sometimes they will lead you to earn **less points** than those shown in the table. Remember that these random shocks can also lead your **partner** to earn **more OR less** points than those shown in the tables.

For example, assume you are playing Game Left. If your partner chooses %, then you can expect to earn 145 points if you choose %, or 100 points if you choose @. If your partner chooses @, then you can expect to earn 280 points for choosing % or 235 points for choosing @. This means that, if both you and your partner choose %, then you can each expect to earn 145 points. If you both choose @ however, then you can each expect to earn 235 points. The random shock will be applied to your points however so you may earn **more or less** than these expected values.

Similarly, assume you are playing Game Right. If your partner chooses %, then you can expect to earn 235 points if you choose %, or 280 points if you chose @. If your partner chooses @, then you can expect to earn 100 points if you choose %, or 145 points for choosing @. This means that, if both you and your partner choose @, then you can each expect to earn 145 points. If you both choose % however, then you can each expect to earn 235 points. The random shock will be applied to your points however so you may earn **more or less** than these expected values.

In summary, the points that you earn will depend on four things.

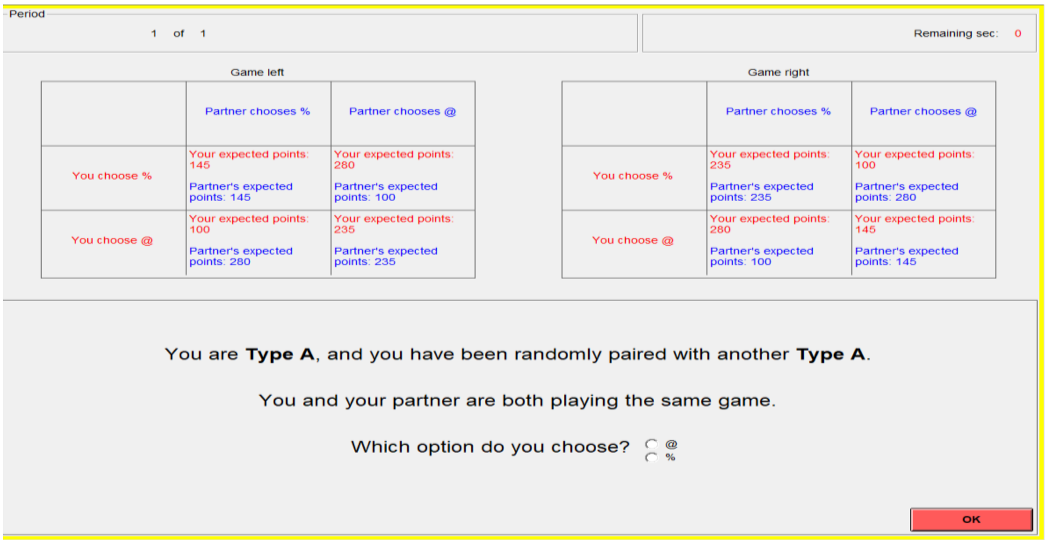
8. **The game being played:** Game Left or Game Right.
9. **The option you choose:** Option % or Option @.
10. **The option your partner chooses:** Option % or Option @.
11. **The random shock:** Random shocks are added to an expected payoff. Random shocks are independent of each other. This means that on some periods the shock will lead you to earn more points than expected, and on some periods they will lead you to earn less points than expected.

IMPORTANT: The computer chooses the games being played at the beginning of each block. The games being played can change from one block to the next, and you will be paired with a different partner of the same type as you in each new block.

Type A vs. Type B Participants

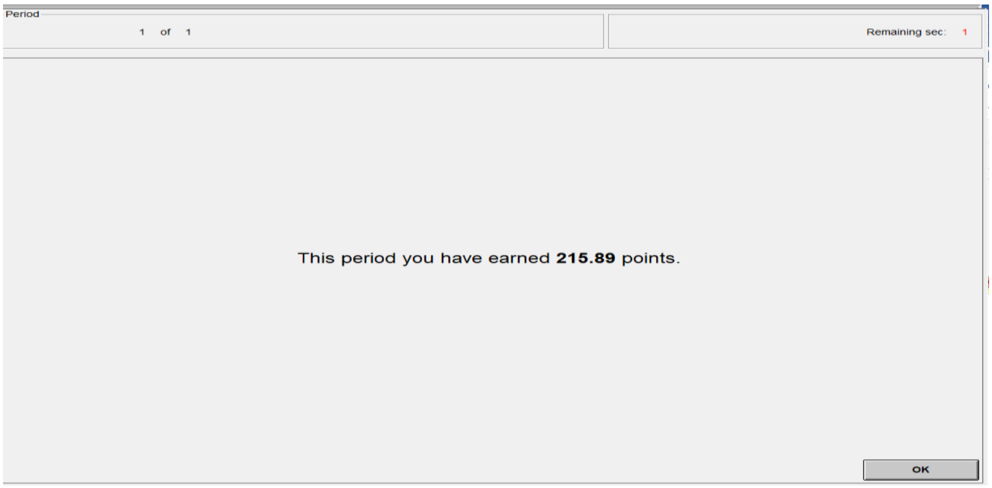
As explained above, before the study begins, the computer will randomly select six participants to be Type A participants. Others will be Type B participants. Your type will not change.

Type A Participants -> Type A Participants choose every period, and each Type A Participant immediately sees the points he or she earns after making a choice. Here is an example of a choice screen for a Type A Participant:

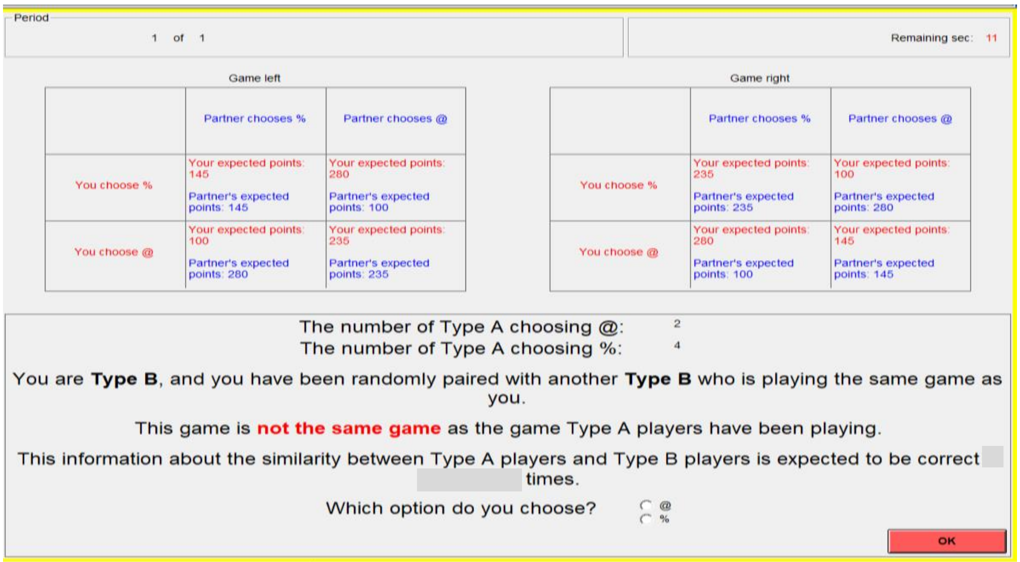


In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This means that you should pay close attention when choosing.

After making a choice and receiving a payoff, each Type A participant will immediately learn how many points he or she earned based on their choice, their partner's choice and a random shock. Here is an example of a feedback screen that a Type A Participant might see:



Type B Participants -> Type B Participants do NOT choose every period. Instead, they only choose in the final period of every block. In earlier periods of a block, Type B Participants simply wait. Here is an example choice screen a Type B Participant might see in the final period of a block:



In this example immediately above, Option @ is listed first. Note, however, that the option listed first can change randomly from one period to the next. That is, you will sometimes see % listed first instead. This means that you should pay close attention when choosing.

Importantly, when Type B Participants make a choice in the final period of a block, they will see the following information.

- The number of Type A Participants who chose option % and the number who chose option @ in the final (fourth) period of the block.
- A SIGNAL indicating if Type B Participants are playing the SAME game as Type A Participants or a DIFFERENT game (Game Left or Game Right). **IMPORTANTLY, you and your partner will always see the same signal.** Remember that the computer decides which game you play completely randomly. It does this at the start of each block by flipping a virtual coin to determine which game all Type A Participants play. It then flips a separate virtual coin to determine the game that all Type B Participants play.
  As you can see, this means that Type B Participants **may or may not** be playing the same game as Type A Participants.
- The probability that the above signal is correct. The signal indicating if Type A Participants play the same game as Type B Participants is not always correct. We do not tell you what this probability is here, but you will see it on the screen in **bold** every time you make a choice. The screen shot provides an example in which we have blurred out the probability that the signal is correct. **IMPORTANTLY, the probability the signal is correct may change from one block to the next, so please pay attention every time you make a**

choice. The probability will ALWAYS be the same for you and your partner.

The following table summarizes the relationship between games and signals:

	Type A Participants play Game Left	**Type A Participants play Game Right**
Type B Participants play Game Left	Correct signal: SAME Incorrect Signal: DIFFERENT	Correct signal: DIFFERENT Incorrect Signal: SAME
Type B Participants play Game Right	Correct Signal: DIFFERENT Incorrect Signal: SAME	Correct Signal: SAME Incorrect Signal: DIFFERENT

After making a choice, the Type B Participants will receive four separate pay-offs based on their choice, their partner's choice and the game they are playing. Four separate random shock values will be added to these pay-offs. This means that, even though Type B Participants make fewer choices than Type A Participants, they have the exact same number of opportunities to earn points.

Type B Participants will not see the points they earn after choosing. Instead, they will only see the total number of points earned across all their blocks at the very end of the study.

***

Final instructions

Once you have played the last block of the game, you will see a short survey you should then complete. This asks about how you played the game and for some demographics. Then, please wait until the researcher calls your seat number to receive your payment. Your earnings will not be told to any other participant.

Now turn the page to sign the consent form and answer the 10 pre-game questions. Raise your hand to alert the researcher when you are finished. Everyone must answer **ALL** pre-game questions correctly before we can begin.

Please keep these instructions to refer back to during the study.

If you have any further questions, feel free to contact the lead PhD researcher (Aysha Bellamy) at: **pejt007@live.rhul.ac.uk**

Comprehension questions

Please answer the following multiple-choice questions, by circling your chosen answer. Everyone must answer all 10 questions correctly before we can begin. You may use your instructions to help you:

***Q1*: Which of the following statements is true, in regards to the number of choices that each participant (Type A and B) makes?**

a) Type A participants make fewer choices than Type B participants.

b) Type A participants make 4X as many choices as Type B participants, but both types of participant have the same opportunities to earn points.

c) Type A Participants make 4X as many choices as Type B Participants, and thus earn 4X as much.

***Q2:* Which of the following statements is true, in regards to the number of points that you should earn if you choose @ and your partner chooses %?**

a) In game Left, I should get 280 points for choosing @ and my partner should get 100 points for %. In Game Right, I would get 100 points for choosing @ and they would get 280 points for choosing %, though a random shock will be applied to mine and my partner's points.

b) In game Left, I should get 100 points for choosing @ and my partner should get 280 points for %. In Game Right, I would get 280 points for choosing @ and they would get 100 points for choosing %, though a random shock will be applied to mine and my partner's points.

c) I will always get 100 points.

***Q3*: Which of the following statements is true, in regards to the partner you are assigned to play the game with?**

a) I get paired with a new partner who is the same type as me in every block.

b) I will play with the same partner throughout the whole study.

c) I get paired with a new partner who is a different type from me in every block.

***Q4:* Which of the following statements is true, in regards to the things that may affect the points that you can earn?**

a) My points earned depend only on which game I am playing, and which option I choose.

b) My expected points depend on **both** the game I am playing and which option me and my partner choose, though the points that I can earn will also be affected by a random shock.

c) My expected points depend only on which game I am playing, and which option I choose, though the points that I can earn will also be affected by a random shock.

***Q5:* Which of the following statements is true, in regards to whether you are playing the same game (Left or Right) as other participants?**

a) I always play the same game as my partner, though Type A and Type B Participants may play the same game or different games.

b) I play a different game to my partner, though Type A and Type B Participants play the same game.

c) All participants play the same game.

***Q6*: Which of the following statements is true, in regards to the feedback that Type A Participants receive?**

a) Type A participants see no information

b) Type A participants see the points made by other Type A participants, but not their own points.

c) Type A participants see their own points, but not the points of other Type A participants.

***Q7:* If you are a Type B Participant, and the signal tells you that you're playing a different game from Type A Participants, is this information necessarily correct?**

a) This information is always correct.

b) This information will sometimes be correct, with a certain probability, and sometimes incorrect, with the remaining probability.

c) This information is never correct.

***Q8:* If you are a Type B Participant, when will you see the probability that the signal (same game as Type A or different game) is correct?**

a) At the very end of the game.

b) After I make each choice.

c) It will be with the information on-screen before I make my choice.

***Q9:* If you are a Type B Participant, you and your partner will always see the same signal indicating whether you're playing the same game as Type A Participants.**

a) True

b) False

***Q10:* If you are a Type B Participant, the probability that this signal (see Q9) is correct will always be the same for you and your partner.**

a) True

b) False

Please now raise your hand and alert the researcher, who will check your answers to the multiple-choice comprehension questions.

Summary for participants (Experimenter will read aloud just after everyone has answered all comprehension questions correctly)

- Welcome to the main portion of the study! Today's session will consist of 88 periods divided into 22 blocks of 4 periods each. When we begin, the computer will randomly assign you to play as a Type A or Type B Participant.

- At the beginning of each block, you will be assigned to play with a partner of the same type as you for all periods in the block.

- For each choice, you will choose between two options (% or @). The points that you can receive will be based on your choices, and the choices of your partner. That is, you can expect to earn a different amount of points on trials where both you and your partner choose an option at the same time, as compared to trials where you choose a different option to your partner.

- The option that is expected to be worth the most depends on whether you are playing Game Left or Right. You do not know which of these games you are playing. The game can change between blocks, but not within blocks. You and your partner will always be playing the same game.

- Type A Participants choose every period and see the points earned immediately after every choice.

- Type B Participants only choose in the final period of each block. Before choosing, they will see (i) how many Type A Participants have chosen @ or %, (ii) a signal indicating whether they are playing the same game (Left or Right) as Type A Participants, and (iii) the probability that this signal is correct. If you are a Type B participant, you and your partner will always see the same information. The information may change, however, from one block to the next, so please pay attention to the on-screen information.

- If you have any questions, please ask them now. Remember not to talk to your fellow participants during the game.

Please now fill in the consent form, if you still wish to take part in the study.

Consent form

The gene-culture coevolution of group decisions

By Aysha Bellamy and Dr. Charles Efferson

Your data will be anonymous and stored electronically in accordance with the Data Protection Act 1988 by the EconLab at Royal Holloway University of London. Your data will be used as part of a PhD thesis in the Psychology Department. The data used will not be personal and will be anonymised. Your personal data cannot be traced back to you and is kept separate to the information you provide on-screen. No harm is anticipated from this study and your participation will be rewarded with money proportional to your performance in-game. You have the right to withdraw participation at any point and still be paid your show-up fee.

Now, please answer the following questions (by circling yes or no):

- Have you read the information sheet about the study?

  Yes No

- Do you understand what you will be doing during this study?

  Yes No

- Have you had the chance to ask questions?

  Yes No

- Have you got satisfactory answers to your questions?

  Yes No

- Do you understand that participation is entirely voluntary, and you

  Yes No

  may leave the study at any time, without giving a reason, and still

  receive the show-up fee?

- Do you understand that your data will remain anonymous and be stored

  Yes No

  securely and confidentially?

- Do you understand that only anonymous data may be shared with third

  Yes No

  parties for replication and publication purposes?

- Do you understand that you will be compensated for your time during

  Yes No

  this study, and that the money you will be paid will be based on the

  points that you earn during the game, on top of the show-up fee?

- Can you confirm that you are over the age of 18?

  Yes No

- Do you agree to take part in this study?

  Yes No

Name in block letters:

____

Sign here: ____________________________________ Date: ______________

Note that this consent form will be stored separately from the responses you provide.

Answer sheet (for experimenter use only)

Double-check their answers and if they have one wrong, ask them to try this q again and return. If consistently get one wrong may have to leave.

***Q1*: B**

***Q2:* B**

***Q3*: A**

***Q4:* B**

***Q5:* A**

***Q6*: C**

***Q7:* B**

***Q8:* C**

***Q9:* A**

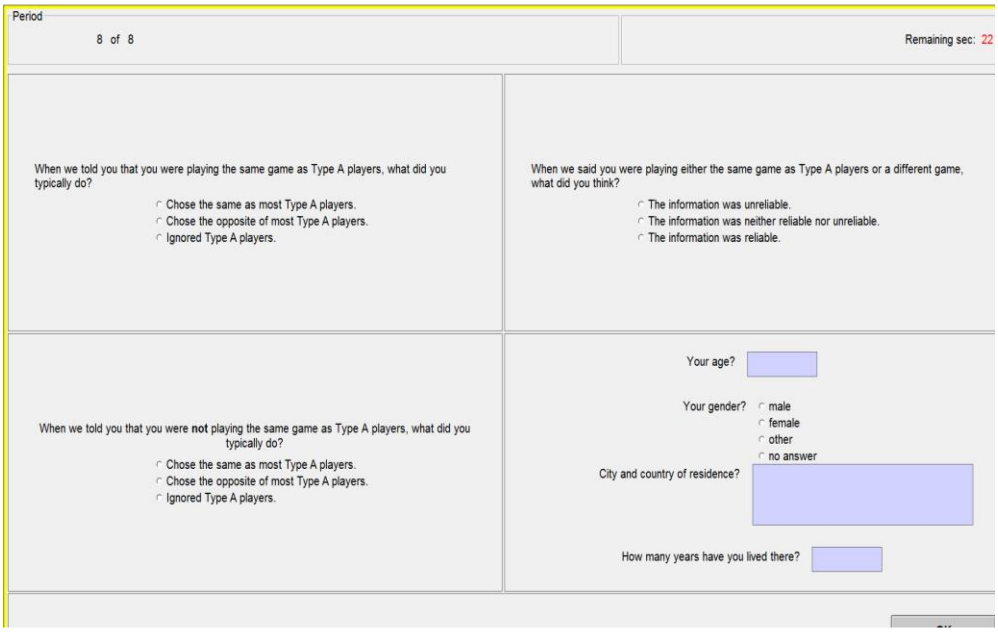
***Q10:* A**

Appendix 2: The script used to run the game via Z-Tree.

Note that the script provided within the Dropbox folder here runs the game for 24 social-learners and 6 demonstrators, though any number of social-learners from 10+ can be accommodated: https://www.dropbox.com/home/Dropbox%20Transfer%20files/Social_Dilemma_Materials

Appendix 3: The end-survey that the social-learners answered on Z-Tree.

See script in Appendix 2 for the code to run this). Note that the demonstrators also provide demographics.



Appendix 4: Ethics materials.

Appendix 4A: The debrief shown to participants after the study.

ID number ___________

Thank you for taking part in **The gene-culture co-evolution of group identities** study via the RHUL EconLab. The data you have provided will be used in my PhD project within the Psychology Department. It will help us to understand if social information biases participants towards certain decisions. Specifically, we are interested in how your similarity to others during the game, and the reliability of this information (1/5/9 in 10 reliable) affected the way that Type B Participants reacted to social information.

What happens now?

You have been paid according to the points you earned plus a show-up fee.

We will keep a copy of your raw data, anonymised by Z-Tree. We will use this data in our analysis, which may be published, though I not identify any one's data specifically. Your unique subject ID codes will ensure that any published data is anonymous (i.e. cannot be traced back to you personally). Your data will be stored securely in accordance with the Data Protection Act 1988.

Thank you!

We are extremely grateful for the time you have given to take part in this study. If you would like any further information about the study, have concerns about your data, or are interested in any of the topics, then please contact me (Aysha Bellamy) with the contact details below. As you leave, **please remember to take this sheet with you.**

Aysha Bellamy : PhD Student in the department of Psychology
Room: Wolfson Building 342/2
Work email: pejt007@rhul.live.ac.uk
Mobile: 07549606411

Appendix 4B: Proof of self-certification ethical approval from Royal Holloway, University of London.

Ethics Review Details

You have chosen to self certify your project.	
Name:	Bellamy, Aysha (2017)
Email:	PEJT007@live.rhul.ac.uk
Title of research project or grant:	The gene-culture co-evolution of group identities.
Project type:	Royal Holloway postgraduate research project/grant
Department:	Psychology
Academic supervisor:	Dr. Charles Efferson
Email address of Academic Supervisor:	Charles.Efferson@rhul.ac.uk
Funding Body Category:	No external funder
Funding Body:	
Start date:	01/11/2018
End date:	01/09/2020

Research question summary:

The theory of gene-culture coevolution suggests that cultural evolution is an important influence on human behaviour. For cultural evolution to take place, one of the key assumptions in this field is that we preferentially learn from those of the same culture as ourselves. In experiments, this should translate to participants preferentially learning from a group of similar others (as similar people are more likely to belong to the same social group). Very few previous studies have investigated this assumption. Those studies that did investigate similarity have not taken into account that our ability to calculate our similarity to others may not always be reliable. These experiments will be the first to investigate a more-realistic, graded similarity-signal, to see how this impacts social-learning style preference. Conformity (or the disproportionate trend to adapt the same behaviour as the majority of a group) is thought to be important in homogenising cultural groups (i.e. making people of the same culture more similar), and in allowing costly levels of cooperation to emerge in human societies. Thus, conformity should be more likely in a group of similar others. Experiment 2 will also address whether conformity can allow cooperation to emerge, by investigating how participants learn from similar and different others during a Social Dilemma task structure.

In summary, the main research question behind these two studies is how people use similarity-information to modify their choice of social-learning strategy. Conformity and cooperation will also be investigated, and these studies will run in a lab in India in the aims of expanding our knowledge of social-learning strategies to those from a diverse range of cultures.

Note that instructions given below will be modified only slightly to accommodate the coordination and social dilemma games.

Research method summary:

Experiments run via Centre for Experimental Social Sciences (CESS) labs in Pune, India. Experiment 1 tests a group of participants en-masse in a best-choice or coordination game, and Experiment 2 tests a Social Dilemma game. Regardless of game-type, participants are assigned to play as individual- or social-learners. The aim of all three games is to choose between two options (@ or %). Individual-learners have four turns in a row, with immediate feedback from their decisions (points). Points are converted to pay with a random shock to reflect exogenous factors influencing decision-making. Individual-learners learn to choose the best option. Social-learners will not see the feedback based on their choices until the end of the game. Instead, they see the number of individual-learners who chose @ or % on their final turn. They also see a sentence telling them that they play the same or different game to individual-learners, and that this signal is likely to be correct only 1/5/9 in 10 of the time (representing an unreliable, chance guess and a reliable signal respectively). The same game means that the same option produces the highest points for individual- and social-learners (e.g. both choose @). At the end of the game, a survey asks social-learners what social-learning strategies they chose and why. Every participant gives key demographics, including age, gender, country of residence and length of time living there. This experiment is ran via Z-Tree, and individuals play at separately-screened computers, anonymising data. CESS provides anonymised files for analysis. The 'game' also involves paper-based instructions and multiple-choice questions to check understanding. The coordination game involves trying to sync your choice to another participant, with whom you cannot interact, whilst Social Dilemmas involve choosing between an individual-maximising strategy (defection) in favour of a strategy which is better when both parties choose it (cooperation).

Risks to participants

Does your research involve any of the below?

Children (under the age of 16),

No

Participants with cognitive or physical impairment that may render them unable to give informed consent,
No

Participants who may be vulnerable for personal, emotional, psychological or other reasons,
No

Participants who may become vulnerable as a result of the conduct of the study (e.g. because it raises sensitive issues) or as a result of what is revealed in the study (e.g. criminal behaviour, or behaviour which is culturally or socially questionable),
No

Participants in unequal power relations (e.g. groups that you teach or work with, in which participants may feel coerced or unable to withdraw),
No

Participants who are likely to suffer negative consequences if identified (e.g. professional censure, exposure to stigma or abuse, damage to professional or social standing),
No

Details,

Design and Data

Does your study include any of the following?

Will it be necessary for participants to take part in the study without their knowledge and/or informed consent at the time?,
No

Is there a risk that participants may be or become identifiable?,
No

Is pain or discomfort likely to result from the study?,
No

Could the study induce psychological stress or anxiety, or cause harm or negative consequences beyond the risks encountered in normal life?,
No

Does this research require approval from the NHS?,
No

If so what is the NHS Approval number,

Are drugs, placebos or other substances to be administered to the study participants, or will the study involve invasive, intrusive or potentially harmful procedures of any kind?,
No

Will human tissue including blood, saliva, urine, faeces, sperm or eggs be collected or used in the project?,
No

Will the research involve the use of administrative or secure data that requires permission from the appropriate authorities before use?,
No

Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?,
No

Is there a risk that any of the material, data, or outcomes to be used in this study has been derived from ethically-unsound procedures?,
No

Details,
Financial inducements are based on in-game performance. These cannot be negative (i.e. participants cannot lose pay) and points are set to be high, allowing for all participants to feel they are good at playing the game. Inducements will also have a random shock applied to them, so it will be made clear to participants that their pay earned in no way reflects how 'smart' they are at playing the game, so this use of financial incentives is not deemed to be unethical.

Risks to the Environment / Society

Will the conduct of the research pose risks to the environment, site, society, or artifacts?,
No

Will the research be undertaken on private or government property without permission?,
No

Risks to Researchers/Institution

Does your research present any of the following risks to researchers or to the institution?

Is there a possibility that the researcher could be placed in a vulnerable situation either emotionally or physically (e.g. by being alone with vulnerable, or potentially aggressive participants, by entering an unsafe environment, or by working in countries in which there is unrest)?,
No

Is the topic of the research sensitive or controversial such that the researcher could be ethically or legally compromised (e.g. as a result of disclosures made during the research)?,
No

Will the research involve the investigation or observation of illegal practices, or the participation in illegal practices?,

No

Could any aspects of the research mean that the University has failed in its duty to care for researchers, participants, or the environment / society?,
No

Is there any reputational risk concerning the source of your funding?,
No

Is there any other ethical issue that may arise during the conduct of this study that could bring the institution into disrepute?,
No

Details,

Declaration
By submitting this form, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Certificate produced for user ID, PEJT007

Date:	05/11/2018 10:11
Signed by:	Bellamy, Aysha (2017)
Digital Signature:	Aysha Bellamy
Certificate dated:	11/5/2018 10:41:50 AM
Files uploaded:	Self-Assessment-851-2018-03-12-14-36-PEJT007.pdf ins_india_BC_5Nov2018.docx

Ethics Self Assessment

Your answers indicate that you do not need ethical approval. If your research includes use of animals as research subjects, you will have been emailed separate guidance which must be followed before you begin your research. Should the circumstances of your research alter in any way please revisit this process to validate your project.

Applicant details

Declaration

By clicking the 'submit form' button, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Project type:	Royal Holloway postgraduate research project/grant
Name:	Bellamy, Aysha (2017)
Email:	PEJT007@live.rhul.ac.uk
Academic supervisor:	Dr. Charles Efferson
Department:	Psychology
Title of research project or grant:	The gene-culture co-evolution of group identities.
Email address of Academic Supervisor:	Charles.Efferson@rhul.ac.uk
Funding Body Category:	No external funder
Funding Body:	

Information about the Research Project

Will the research project involve the use of human participants or human tissue (with or without their knowledge or consent at the time)?, No

Are the results of the research project likely to expose any person or community to physical or psychological harm?, No

Will the research project involve the use of animals as research subjects?, No

Will you have access to personal information that allows you to identify individuals or company confidential information (that is not covered by confidentiality terms within an agreement or by a separate confidentiality agreement)?, No

Does the conduct of the research project present a significant risk to the environment or society?, No

Are there any other ethical issues raised by this research project that in the opinion of the PI require further ethical review?, No

Does the PI believe that the results of this research could reasonably lead to legal action or negative press coverage, for which the PI would require University support?, No

Certificate produced for user ID PEJT007

Certificate dated 3/12/2018 2:36:57 PM

Appendix 5: The analysis scripts used in this paper.

These can be found at [OSF | When will you learn to help? Social-learning strategies remain somewhat flexible to a range of social information when learning to cooperate].

Note that file called *analysisSDgame_app2.R* performs the regressions, *estandBootImitationFunctions_SD.R* performs boot-strapped estimations and *plotEstAndBootImitationFunctions_SD_7point.R* plots the graph in figure 4 accordingly. The *clx*.R* scripts perform the bootstrapped clustering as a function. Finally, the script called *linearCombo_UseSD_21.R* performs all the linear combinations reported in-text and *individualVariance_AnalaysisScatterSD* builds the scatterplots and heatmaps in appendices 11-12. The raw data is also available (see *SDRawData.csv* file), plus the supplementary materials detailing the analysis of the self-reported participant strategies.

Appendix 6: Further analysis for predicting when social learners chose %.

Appendix 6A: The ANOVA comparing the glmms predicting whether the social learners chose %.

These suggest that the model without control variables had the best fit of the data and so this is the model that was reported in-text in Table 3.

Model	DF	AIC	BIC	logLik	Chi squre	Chi	Chi df	Chi prob
Basic	14	3078.5	3158.8	-1525.5	3050.5			
Controls	23	3082.3	3214.2	-1518.2	3036.3	0.028	1	0.868
Controls, No % as cooperation dummy	22	3080.3	3206.5	-1518.2	3036.3	14.155	8	0.078

Appendix 6B: For transparency purposes, we also include the basic model without control variables to predict whether the social learners chose %.

Predictors include (i) the centered proportion of demonstrators who chose %, (ii) dummies for each combination of similarity and reliability signal and (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose %. Robust standard errors clustered on the social learner. Note that the significant betas of the main predictors match those reported in Table 3 in-text. That is, the model is not drastically affected by excluding the control variables and demographics.

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.89	0.54	1.66	0.10	-0.16	1.94	2.44
Centred proportion of demonstrators choosing %	-1.76	0.44	-4.00	<0.001	-2.63	-0.89	0.17
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	0.09	0.15	0.64	0.53	-0.20	0.38	1.10
Uninformative - same dummy [signal indicates different and is	0.09	0.15	0.54	0.59	-0.23	0.40	1.09

correct with 0.5 probability]							
Uninformative - different dummy [signal indicates different and is correct with 0.5 probability]	-0.10	0.15	-0.70	0.48	-0.39	0.18	0.90
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability	-0.09	0.16	-0.59	0.56	-0.41	0.22	0.91
Reliably correct- different dummy [signal indicates different and is correct with 0.9 probability]	0.17	0.15	1.16	0.25	-0.12	0.46	1.19
Centred proportion of demonstrators choosing % X reliably incorrect- different dummy	2.30	0.60	3.81	0.0001	1.12	3.48	9.97
Centred proportion of demonstrators choosing % X	4.11	0.67	6.13	<0.001	2.80	5.42	60.99

uninformative-same dummy							
Centred proportion of demonstrators choosing % X uninformative-different dummy	1.15	0.61	1.89	0.058	-0.04	2.34	3.16
Centred proportion of demonstrators choosing % X reliably correct-same dummy	4.81	0.70	6.91	<0.001	3.44	6.17	122.52
Centred proportion of demonstrators choosing % X reliably correct-different dummy	1.49	0.61	2.43	0.02	0.29	2.69	4.42
% as cooperation	-0.02	0.09	-0.17	0.87	-0.19	0.16	0.98
Block index	-0.0001	0.007	-0.02	0.99	-0.01	0.01	1.0
Female dummy	-0.12	0.12	-1.03	0.30	-0.36	0.11	0.88
Other gender dummy	-0.25	0.42	-0.61	0.54	-1.06	0.56	0.78
Age	-0.02	0.02	-1.12	0.26	-0.07	0.02	0.98
Time in residence	-0.006	0.008	-0.83	0.41	-0.02	0.009	0.99
UK dummy (participant lives in the UK)	-0.21	0.26	-0.81	0.42	-0.73	0.30	0.81
Individualistic dummy (participant lives in	-0.91	0.43	-2.10	0.04	-1.75	-0.06	0.40

individualistic country besides the UK)							
Middle Hofstede dummy (participant lives in a country in the middle of Hofstede's (1980) individualism-collectivism scale)	0.63	0.40	1.57	0.12	-0.16	1.42	1.88

Appendix 7: The linear combinations performed between each level of the similarity and reliability information when predicting whether the social learners chose %.

We run the linear combinations for blocks where the majority of demonstrators chose % ($\geq$4) or blocks where none chose %. See script entitled *linearCombo_UseSD_21.R* at [OSF | When will you learn to help? Social-learning strategies remain somewhat flexible to a range of social information when learning to cooperate] for the code used to calculate these.

Second-order (similar vs different)

- Reliably incorrect signals (0/6: $\chi 2$ (1) = 12.79, p=0.0003)

  6/6: $\chi 2$ (1) = 13.17, p=0.0003

  5/6: $\chi 2$ (1) = 10.94, p=0.0009

  4/6: $\chi 2$ (1) = 6.39, p=0.01

- Uninformative signals (0/6: $\chi2$ (1) = 14.15, p=0.0002)

  6/6: $\chi2$ (1) = 18.26, p<0.0001

  5/6: $\chi2$ (1) = 15.94, p<0.0001

  4/6: $\chi2$ (1) = 10.54, p=0.001

- Reliably correct signals (0/6: $\chi2$ (1) = 30.25, p<0.001)

  6/6: $\chi2$ (1) = 13.22, p=0.0002

  5/6: $\chi2$ (1) = 8.54, p=0.003

  4/6: $\chi2$ (1) = 2.08, p=0.15

<u>Third-order (comparing reliability for similar others)</u>

- Similar, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 31.86, p<0.001)

  6/6: $\chi2$ (1) = 30.03, p<0.0001

  5/6: $\chi2$ (1) = 24.57, p<0.001

  4/6: $\chi2$ (1) = 13.56, p=0.0002

- Similar, reliably incorrect vs correct (0/6: $\chi2$ (1) = 47.77, p<0.001)

  6/6: $\chi2$ (1) = 35.95, p<0.0001

  5/6: $\chi2$ (1) = 27.61, p<0.001

  4/6: $\chi2$ (1) = 12.27, p=0.0005

- Similar, uninformative vs reliably correct (0/6: $\chi2$ (1) = 1.99, p=0.16)

  6/6: $\chi2$ (1) = 0.23, p=0.63

  5/6: $\chi2$ (1) = 0.05, p=0.82

  4/6: $\chi2$ (1) = 0.06, p=0.81

Third-order (comparing reliability for different others)

- Different, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 1.79, p=0.18)

  6/6: $\chi2$ (1) = 5.15, p=0.02

  5/6: $\chi2$ (1) = 5.02, p=0.03

  4/6: $\chi2$ (1) = 4.25, p=0.04

- Different, reliably incorrect vs correct (0/6: $\chi2$ (1) = 3.15, p=0.08)

  6/6: $\chi2$ (1) = 1.08, p=0.30

  5/6: $\chi2$ (1) = 0.65, p=0.42

  4/6: $\chi2$ (1) = 0.13, p=0.72

- Different, uninformative vs reliably correct (0/6: $\chi2$ (1) = 0.20, p=0.66)

  6/6: $\chi2$ (1) = 1.33, p=0.25

  5/6: $\chi2$ (1) = 1.83, p=0.18

  4/6: $\chi2$ (1) = 2.66, p=0.10

Appendix 8: Further analysis for predicting when social learners chose to cooperate.

Appendix 8A: The ANOVA comparing the glmms predicting whether the social learners cooperate.

This suggests that the basic model without control variables had the best fit of the data and so this is the model that was reported in-text in Table 4.

Model	DF	AIC	BIC	logLik	Chi squre	Chi	Chi df	Chi prob
Basic	14	3145.8	3226.1	-1558.9	3117.8			
Controls	23	3157.5	3289.4	-1555.7	3111.5	3.33	1	0.07
Controls, No % as cooperation dummy	22	3158.8	3285.0	-1557.4	3114.8	2.99	8	0.94

Appendix 8B: For transparency purposes, we also include the full model with control variables to predict whether the social learners chose to cooperate. Predictors include (i) the centered proportion of demonstrators who chose to cooperate, (ii) dummies for each combination of similarity and reliability signal and (iii) interactions between each of these dummies and the centered proportion of demonstrators who chose to cooperate and (iv) control predictors. Robust standard errors clustered on the social learner. Note that the significant betas for the main predictors match those reported in Table 4 in-text. That is, adding control variables and demographics does not drastically affect the model. There is a trend for those from countries in the middle of Hofstede's (1980) individualism-collectivism scale to cooperate less than those from collectivist countries (the omitted category). As this tendency did not reach significance however, we will not comment on it further.

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.36	0.41	0.88	0.38	-0.45	1.76	1.44
Centred proportion of demonstrators choosing %	1.33	0.44	3.00	0.003	0.46	2.21	3.80
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	0.004	0.15	0.03	0.98	-0.30	0.31	1.00

Uninformative-same dummy [signal indicates different and is correct with 0.5 probability]	0.12	0.16	0.76	0.46	-0.20	0.44	1.13
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	-0.02	0.15	-0.16	0.87	-0.32	0.27	0.98
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability	0.05	0.17	0.32	0.75	-0.28	0.38	1.06
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.09	0.15	-0.61	0.54	-0.38	0.20	0.91
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	-0.83	0.61	-1.36	0.17	-2.03	0.37	0.43
Centred proportion of	-0.28	0.66	-0.42	0.68	-1.58	1.02	0.76

demonstrators choosing % X uninformative-same dummy							
Centred proportion of demonstrators choosing % X uninformative-different dummy	-0.53	0.62	-0.86	0.39	-1.75	0.69	0.59
Centred proportion of demonstrators choosing % X reliably correct-same dummy	1.43	0.71	2.00	0.045	0.03	2.83	4.17
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-1.36	0.61	-2.22	0.03	-2.55	-0.16	0.26
% as cooperation	0.16	0.09	1.82	0.07	-0.01	0.33	1.17
Block index	-0.002	0.007	-0.24	0.81	-0.02	0.01	1.0
Female dummy	-0.07	0.09	-0.78	0.44	-0.25	0.11	0.93
Other gender dummy	0.23	0.32	0.73	0.46	-0.39	0.86	1.26
Age	-0.01	0.02	-0.64	0.52	-0.04	0.02	0.99
Time in residence	-0.006	0.008	-0.83	0.41	-0.01	0.01	1.0
UK dummy (participant lives in the UK)	-0.15	0.20	-0.74	0.46	-0.54	0.25	0.86
Individualistic dummy	-0.08	0.32	-0.23	0.82	-0.71	0.56	0.93

(participant lives in individualistic country besides the UK)							
Middle Hofstede dummy (participant lives in a country in the middle of Hofstede's (1980) individualism-collectivism scale)	-0.30	0.30	-1.01	0.31	-0.89	0.28	0.74

Appendix 9A: We repeat the analysis predicting whether social learners chose to cooperate but this time we distinguish between blocks which happened to give correct information about similarity or difference (see left-hand column) and blocks which happened to give incorrect information about similarity or difference (see the right-hand column).

The analysis essentially shows us that the social learners will cooperate whenever they see similar others with uninformative or reliably correct signals, provided that these similarity signals are correct (left-hand column). The social learners will cooperate whenever the majority of reliably incorrect-similar others do, though this is because they treat these signals as if they are always incorrect (right-hand column). Note that for all other signals, the social learners are less likely to cooperate whenever the similarity information happens to be incorrect. This may suggest that the social learners only treat blocks with reliably incorrect signals from similar others as if they are always incorrect.

Correct signals:

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.06	0.06	0.10	0.92	-0.12	0.13	1.01
Centred proportion of demonstrators choosing %	0.31	0.25	1.21	0.20	-0.17	0.79	1.37
Reliably incorrect- different dummy [signal indicates different and is correct with 0.1 probability]	0.51	0.29	1.77	0.08	-0.05	1.07	1.66
Uninformative- same dummy [signal indicates different and is correct with 0.5 probability]	0.32	0.18	1.80	0.07	-0.03	0.67	1.38
Uninformative- different dummy [signal indicates different and is correct with 0.5 probability]	-0.15	0.18	-0.82	0.41	-0.50	0.21	0.86
Reliably correct- same dummy	0.06	0.15	0.38	0.70	-0.23	0.35	1.06

[signal indicates different and is correct with 0.9 probability]							
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.15	0.14	0.38	0.70	-0.43	0.12	0.86
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	-0.22	1.12	-0.19	0.85	-2.41	1.98	0.81
Centred proportion of demonstrators choosing % X uninformative-same dummy	2.64	0.73	3.63	0.003	1.22	4.06	14.00
Centred proportion of demonstrators choosing % X uninformative-different dummy	1.19	0.78	1.52	0.13	-0.34	2.73	3.29
Centred proportion of demonstrators choosing % X reliably	2.82	0.63	4.47	<0.001	1.58	4.06	16.81

correct-same dummy							
Centred proportion of demonstrators choosing % X reliably correct-different dummy	0.23	0.53	0.44	0.66	0.81	1.27	1.26

Incorrect signals:

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.02	0.06	0.38	0.70	-0.09	0.14	1.02
Centred proportion of demonstrators choosing %	1.69	0.24	7.14	<0.001	1.23	2.16	5.44
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	-0.10	0.14	-0.71	0.48	-0.37	0.17	0.91
Uninformative-same dummy [signal indicates different and is correct with	-0.26	0.22	-1.19	0.24	-0.69	0.17	0.77

0.5 probability]							
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	0.08	0.15	0.53	0.60	-0.22	0.38	1.08
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability]	-0.92	0.77	-1.19	0.24	-2.43	0.60	0.40
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	0.14	0.22	0.64	0.52	-0.29	0.57	1.15
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	-1.30	0.52	-2.52	0.01	-2.31	-0.29	0.27
Centred proportion of demonstrators choosing % X	-3.83	0.91	-4.21	<0.001	-5.62	-2.04	0.02

uninformative-same dummy							
Centred proportion of demonstrators choosing % X uninformative-different dummy	-1.23	0.59	-2.09	0.04	-2.38	-0.08	0.29
Centred proportion of demonstrators choosing % X reliably correct-same dummy	-8.10	3.74	-2.17	0.03	-15.43	-0.77	0.0003
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-3.80	1.02	-3.73	0.0002	-5.79	-1.80	0.02

Appendix 9B: For transparency, wee repeat the analysis predicting whether social learners chose to cooperate for blocks which happened to give correct information about similarity or difference (see left-hand column) and blocks which happened to give incorrect information about similarity or difference (see the right-hand column), though this time including control predictors.

The control predictors and demographic variables do not significantly affect how likely it is that the social learners will cooperate in the face of correct or incorrect similarity signals. Moreover, the same main predictors are significant in this analysis highlighting that adding control variables to the model does not drastically change the results.

Correct signals:

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.31	0.40	0.76	0.45	-0.49	1.10	1.36
Centred proportion of demonstrators choosing %	1.32	0.25	1.30	0.19	-0.16	0.81	1.38
Reliably incorrect-different dummy [signal indicates different and is correct with 0.1 probability]	0.51	0.29	1.76	0.08	-0.06	1.07	1.66
Uninformative-same dummy	0.34	0.18	1.86	0.06	-0.02	0.69	1.40

[signal indicates different and is correct with 0.5 probability]							
Uninformative-different dummy [signal indicates different and is correct with 0.5 probability]	-0.16	0.19	-0.89	0.37	-0.52	0.19	0.85
Reliably correct- same dummy [signal indicates different and is correct with 0.9 probability	0.07	0.15	0.45	0.66	-0.23	0.36	1.07
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	-0.17	0.14	-1.19	0.23	-0.44	0.11	0.85
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	-0.19	1.22	-0.17	0.86	-2.39	2.01	0.83

Centred proportion of demonstrators choosing % X uninformative-same dummy	2.66	0.73	3.67	0.0003	1.23	4.08	14.24
Centred proportion of demonstrators choosing % X uninformative-different dummy	1.15	0.79	1.46	0.14	-0.39	2.69	3.16
Centred proportion of demonstrators choosing % X reliably correct-same dummy	2.84	0.63	4.49	<0.001	1.60	4.09	17.20
Centred proportion of demonstrators choosing % X reliably correct-different dummy	0.13	0.54	0.25	0.81	-0.92	-1.18	1.14
% as cooperation	0.16	0.09	1.91	0.057	-0.005	0.33	1.18
Block index	0.0006	0.007	0.09	0.93	-0.01	0.01	1.0
Female dummy	-0.08	0.09	-0.88	0.34	-0.26	0.10	0.92
Other gender dummy	0.23	0.32	0.72	0.47	-0.39	0.85	1.26
Age	-0.01	0.02	-0.68	0.50	-0.04	0.02	0.99
Time in residence	-0.01	0.006	-0.19	0.85	-0.01	0.01	1.0
UK dummy (participant lives in the UK)	-0.12	0.20	-0.59	0.55	-0.51	0.27	0.88

Individualistic dummy (participant lives in individualistic country besides the UK)	-0.03	0.32	-0.10	0.93	-0.66	0.60	0.97
Middle Hofstede dummy (participant lives in a country in the middle of Hofstede’s (1980) individualism-collectivism scale)	-0.29	0.30	-0.99	0.32	-0.87	0.29	0.75

Incorrect signals:

	Estimate	Standard error	Z value	P value	Lower CI	Upper CI	Inverse logit (probability)
Intercept	0.25	0.41	0.60	0.55	-0.56	1.06	1.30
Centred proportion of demonstrators choosing %	1.68	0.24	6.93	<0.001	1.20	2.15	5.34
Reliably incorrect- different dummy [signal indicates different and is correct with 0.1 probability]	-0.08	0.14	-0.61	0.54	0.35	0.19	0.92
Uninformative- same dummy [signal indicates different and is correct with 0.5 probability]	-0.27	0.22	-1.21	0.23	0.35	0.17	0.76
Uninformative- different dummy [signal indicates different and is correct with 0.5 probability]	0.08	0.15	0.53	0.59	0.70	0.17	1.08
Reliably correct- same dummy	-0.88	0.78	-1.13	0.26	0.22	0.38	0.41

[signal indicates different and is correct with 0.9 probability							
Reliably correct-different dummy [signal indicates different and is correct with 0.9 probability]	0.14	0.22	0.64	0.52	2.41	0.64	1.15
Centred proportion of demonstrators choosing % X reliably incorrect-different dummy	-1.23	0.52	2.37	0.02	2.89	0.57	0.29
Centred proportion of demonstrators choosing % X uninformative-same dummy	-3.91	0.92	-4.25	<0.001	2.26	-0.21	0.02
Centred proportion of demonstrators choosing % X uninformative-different dummy	-1.14	0.59	-1.93	0.053	5.71	-2.10	0.32
Centred proportion of demonstrators choosing % X reliably	-8.01	3.72	-2.15	0.03	2.30	0.02	0.0003

correct-same dummy							
Centred proportion of demonstrators choosing % X reliably correct-different dummy	-3.71	1.02	-3.63	0.0003	-15.30	-0.72	0.02
% as cooperation	0.12	0.09	1.44	0.15	-5.71	-1.71	1.13
Block index	0.001	0.007	0.19	0.85	-0.05	0.29	1.00
Female dummy	-0.06	0.09	-0.70	0.49	-0.25	0.12	0.94
Other gender dummy	0.30	0.32	0.94	0.35	-0.33	0.93	1.35
Age	-0.008	0.02	-0.70	0.49	-0.04	0.02	0.99
Time in residence	-0.007	0.006	-0.12	0.91	-0.01	0.01	1.0
UK dummy (participant lives in the UK)	-0.11	0.20	-0.53	0.60	-0.51	0.29	0.90
Individualistic dummy (participant lives in individualistic country besides the UK)	-0.02	0.32	-0.07	0.95	-0.66	0.61	0.98
Middle Hofstede dummy (participant lives in a country in the middle of Hofstede's (1980) individualism-	-0.26	0.30	-0.87	0.38	-0.85	0.33	0.77

collectivism scale)	

Appendix 10: The linear combinations performed between each level of the similarity and reliability information when predicting whether the social learners chose to cooperate.

We run the linear combinations twice for blocks where the majority of demonstrators (≥4) cooperate and blocks where none cooperate. See script entitled *linearCombo_UseSD_21.R* at [OSF | When will you learn to help? Social-learning strategies remain somewhat flexible to a range of social information when learning to cooperate] for the code used to calculate these. In text, we report the results for all signals. Note that the analysis for just incorrect vs just correct similarity signals broadly confirms what the collapsed analysis reported in Section 3.3 of the main text reports: the social learners treat both an uninformative and a reliably correct signal of similarity as if these were always correct, but they treat a reliably incorrect signal of similarity as if this was always incorrect.

<u>All signals, social learner cooperate</u>

<u>Second-order (similar vs different)</u>

- Reliably incorrect signals (0/6: $\chi2$ (1) = 2.63, $p=0.10$)

  6/6: $\chi2$ (1) = 1.54, $p=0.21$

  5/6: $\chi2$ (1) = 1.20, $p=0.27$

  4/6: $\chi2$ (1) = 0.65, $p=0.42$

- Uninformative signals (0/6: $\chi2$ (1) = 0.003, p=0.96)

  6/6: $\chi2$ (1) = 0.47, p=0.49

  5/6: $\chi2$ (1) = 0.55, p=0.46

  4/6: $\chi2$ (1) = 0.66, p=0.42

- Reliably correct signals (0/6: $\chi2$ (1) = 14.88, p=0.0001)

  6/6: $\chi2$ (1) = 10.90, p=0.001

  5/6: $\chi2$ (1) = 9.03, p=0.003

  4/6: $\chi2$ (1) = 5.61, p=0.02

Third-order (comparing reliability for similar others)

- Similar, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 0.76, p=0.38)

  6/6: $\chi2$ (1) = 0.01, p=0.91

  5/6: $\chi2$ (1) <0.001, p=0.98

  4/6: $\chi2$ (1) = 0.07, p=0.79

- Similar, reliably incorrect vs correct (0/6: $\chi2$ (1) = 3.92, p=0.048)

  6/6: $\chi2$ (1) = 2.42, p=0.12

  5/6: $\chi2$ (1) = 1.95, p=0.16

  4/6: $\chi2$ (1) = 1.12, p=0.29

- Similar, uninformative vs reliably correct (0/6: $\chi2$ (1) = 7.26, p=0.007)

  6/6: $\chi2$ (1) = 2.46, p=0.12

  5/6: $\chi2$ (1) = 1.68, p=0.20

  4/6: $\chi2$ (1) = 0.58, p=0.44

Third-order (comparing reliability for different others)

- Different, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 0.38, p=0.54)

  6/6: $\chi2$ (1) = 0.15, p=0.70

5/6: $\chi2$ (1) = 0.10, p=0.75

4/6: $\chi2$ (1) = 0.04, p=0.84

- Different, reliably incorrect vs correct (0/6: $\chi2$ (1) = 0.24, p=0.62)

  6/6: $\chi2$ (1) = 0.53, p=0.47

  5/6: $\chi2$ (1) = 0.51, p=0.48

  4/6: $\chi2$ (1) = 0.43, p=0.51

- Different, uninformative vs reliably correct (0/6: $\chi2$ (1) = 1.16, p=0.28)

  6/6: $\chi2$ (1) = 1.25, p=0.26

  5/6: $\chi2$ (1) = 1.09, p=0.30

  4/6: $\chi2$ (1) = 0.77, p=0.38

Appendix 10 (learning to cooperte, just correct signals):

<u>Second-order (similar vs different)</u>

- Reliably incorrect signals (0/6: $\chi2$ (1) = 1.11, p=0.29)

  6/6: $\chi2$ (1) = 0.36, p=0.55

  5/6: $\chi2$ (1) = 0.75, p=0.39

  4/6: $\chi2$ (1) = 1.67, p=0.20

- Uninformative signals (0/6: $\chi2$ (1) = 0.28, p=0.59)

  6/6: $\chi2$ (1) = 3.56, p=0.06

  5/6: $\chi2$ (1) = 3.97, p=0.046

  4/6: $\chi2$ (1) = 4.42, p=0.04

- Reliably correct signals (0/6: $\chi2$ (1) = 9.72, p=0.002)

  6/6: $\chi2$ (1) = 9.98 p=0.002

5/6: $\chi 2$ (1) = 8.73, p=0.003

4/6: $\chi 2$ (1) = 6.15, p=0.01

Third-order (comparing reliability for similar others)

- Similar, reliably incorrect vs uninformative (0/6: $\chi 2$ (1) = 8.81, p=0.003)

  6/6: $\chi 2$ (1) = 12.55, p=0.003

  5/6: $\chi 2$ (1) 11.57, p=0.0007

  4/6: $\chi 2$ (1) = 1.66, p=0.20

- Similar, reliably incorrect vs correct (0/6: $\chi 2$ (1) = 22.71, p<0.001)

  6/6: $\chi 2$ (1) = 13.21, p=0.0003

  5/6: $\chi 2$ (1) = 10.53, p=0.001

  4/6: $\chi 2$ (1) = 5.92, p=0.01

- Similar, uninformative vs reliably correct (0/6: $\chi 2$ (1) = 0.75, p=0.39)

  6/6: $\chi 2$ (1) = 0.09, p=0.76

  5/6: $\chi 2$ (1) = 0.23, p=0.63

  4/6: $\chi 2$ (1) = 0.58, p=0.45

Third-order (comparing reliability for different others)

- Different, reliably incorrect vs uninformative (0/6: $\chi 2$ (1) = 4.07, p=0.04)

  6/6: $\chi 2$ (1) = 0.004, p=0.95

  5/6: $\chi 2$ (1) = 0.10, p=0.75

  4/6: $\chi 2$ (1) = 0.04, p=0.84

- Different, reliably incorrect vs correct (0/6: $\chi 2$ (1) = 2.04, p=0.15)

  6/6: $\chi 2$ (1) = 0.37, p=0.54

  5/6: $\chi 2$ (1) = 0.91, p=0.34

  4/6: $\chi 2$ (1) = 0.31, p=0.58

- Different, uninformative vs reliably correct (0/6: $\chi2$ (1) = 1.29, p=0.26)

  6/6: $\chi2$ (1) = 0.78, p=0.38

  5/6: $\chi2$ (1) = 0.61, p=0.44

  4/6: $\chi2$ (1) = 2.23, p=0.14

Appendix 10 (learning to cooperate, just incorrect signals):

<u>Second-order (similar vs different)</u>

- Reliably incorrect signals (0/6: $\chi2$ (1) = 5.80, p=0.02)

  6/6: $\chi2$ (1) = 4.73, p=0.03

  5/6: $\chi2$ (1) = 1.40, p=0.046

  4/6: $\chi2$ (1) = 2.66, p=0.10

- Uninformative signals (0/6: $\chi2$ (1) = 4.39, p=0.04)

  6/6: $\chi2$ (1) = 5.95, p=0.01

  5/6: $\chi2$ (1) = 5.51, p=0.02

  4/6: $\chi2$ (1) = 4.44, p=0.04

- Reliably correct signals (0/6: $\chi2$ (1) = 0.57, p=0.45)

  6/6: $\chi2$ (1) = 1.55 p=0.21

  5/6: $\chi2$ (1) = 1.63, p=0.20

  4/6: $\chi2$ (1) = 1.74, p=0.19

<u>Third-order (comparing reliability for similar others)</u>

- Similar, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 16.70, p<0.001)

  6/6: $\chi2$ (1) = 13.62, p=0.002

  5/6: $\chi2$ (1) =11.67, p=0.0006

  4/6: $\chi2$ (1) = 7.93, p=0.005

- Similar, reliably incorrect vs correct (0/6: $\chi2$ (1) = 5.20, p=0.02)

  6/6: $\chi2$ (1) = 3.91, p=0.048

  5/6: $\chi2$ (1) = 3.60, p=0.058

  4/6: $\chi2$ (1) = 2.98, p=0.08

- Similar, uninformative vs reliably correct (0/6: $\chi2$ (1) = 1.08, p=0.30)

  6/6: $\chi2$ (1) = 1.18, p=0.28

  5/6: $\chi2$ (1) = 1.13, p=0.29

  4/6: $\chi2$ (1) = 1.03, p=0.31

Third-order (comparing reliability for different others)

- Different, reliably incorrect vs uninformative (0/6: $\chi2$ (1) = 0.21, p=0.65)

  6/6: $\chi2$ (1) = 0.20, p=0.65

  5/6: $\chi2$ (1) = 0.31, p=0.58

  4/6: $\chi2$ (1) = 0.51, p=0.47

- Different, reliably incorrect vs correct (0/6: $\chi2$ (1) = 8.19, p=0.004)

  6/6: $\chi2$ (1) = 2.28, p=0.13

  5/6: $\chi2$ (1) = 1.38, p=0.24

  4/6: $\chi2$ (1) = 0.26, p=0.61

- Different, uninformative vs reliably correct (0/6: $\chi2$ (1) = 6.28 p=0.01)

  6/6: $\chi2$ (1) = 3.13, p=0.08

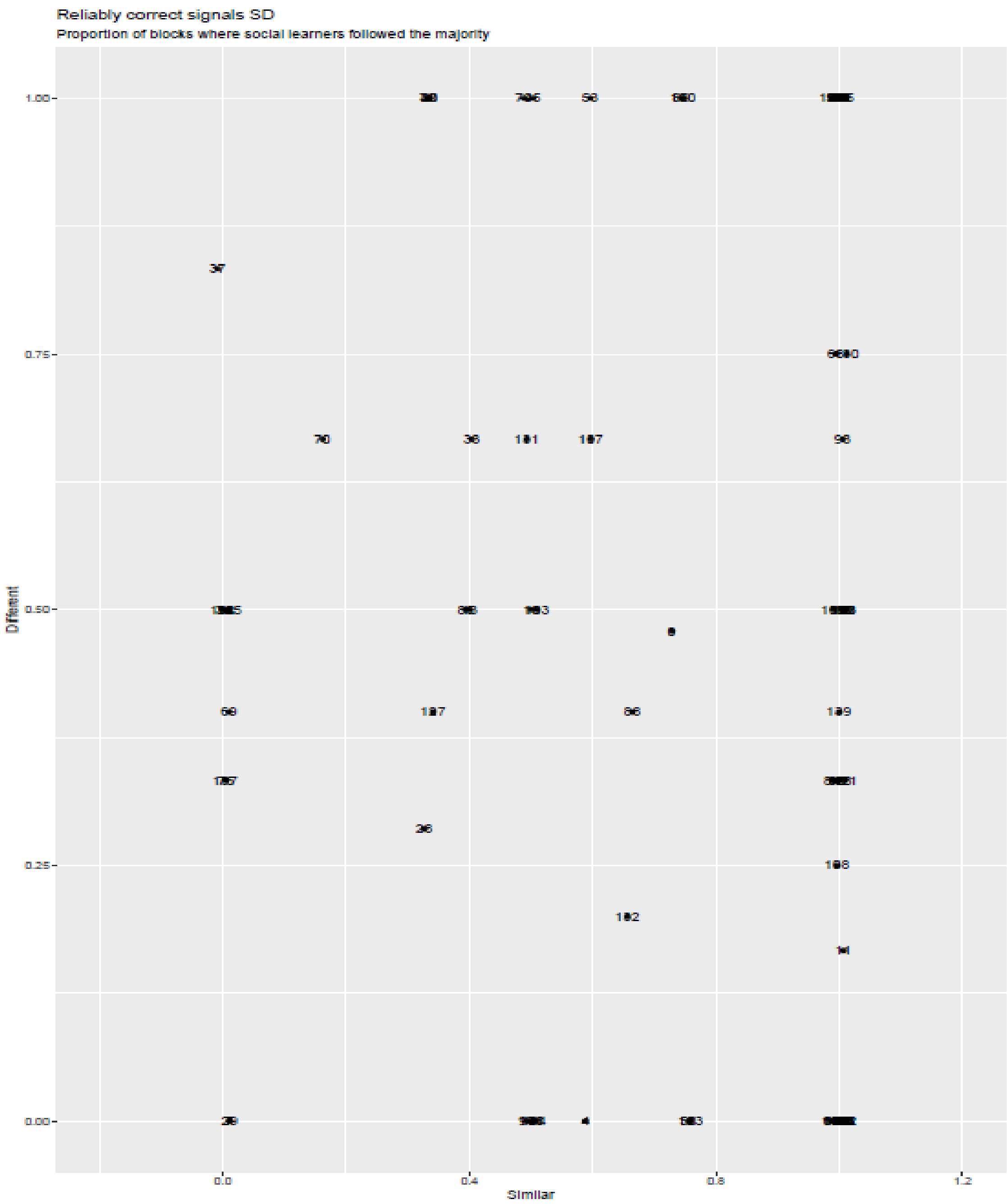
  5/6: $\chi2$ (1) = 2.33, p=0.13

  4/6: $\chi2$ (1) = 1.03, p=0.31

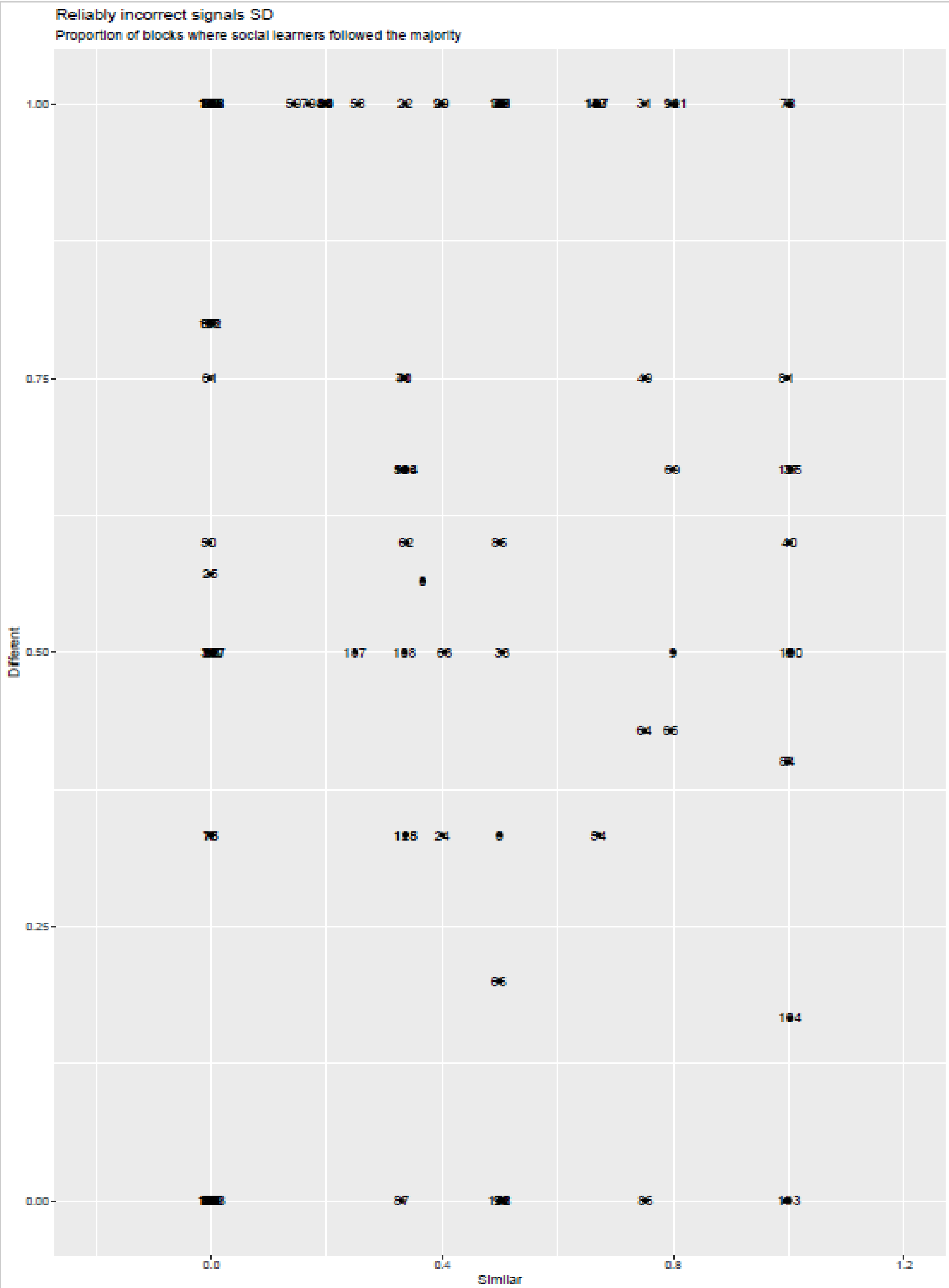
Appendix 11: The scatterplots, showing that the participant IDs often change as a function of the signal they see.

This is to show individual level variation in each of the learner's chosen strategies. As these IDs are hard to see, we also repeat this analysis as a heatmap in appendix 15.

Appendix 11i) reliably correct signals

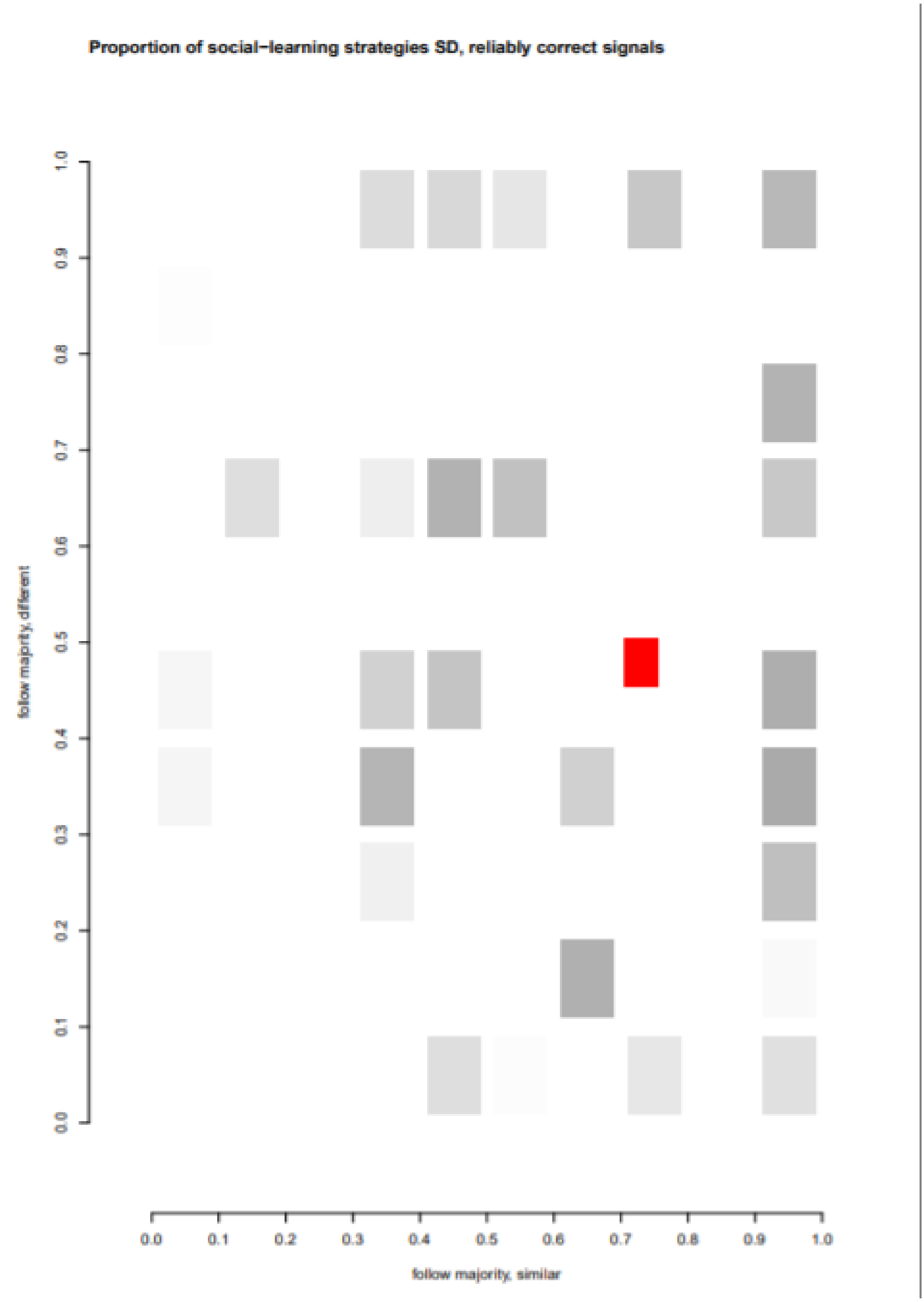


Appendix 11ii) Reliably incorrect signals



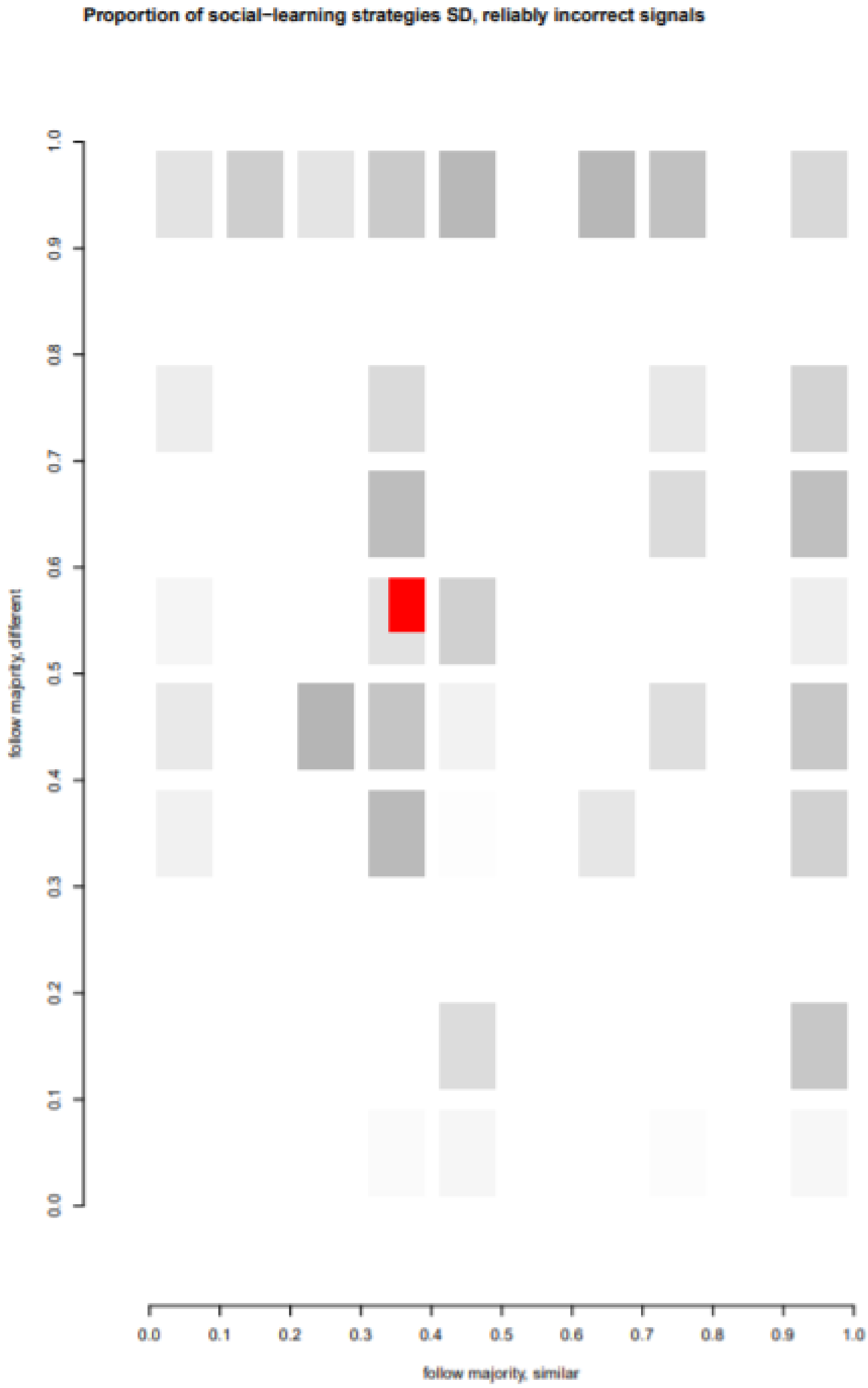
Appendix 12: The heatmaps show the social learner’s chosen strategies as a proportion of blocks where they follow the majority of similar others (on the x axis) by the proportion of blocks where they follow the majority of different others (on the y axis). This is to show individual level variation in each of the learner’s chosen strategies. The x axis gives 0-1 proportion of all social learners who followed the majority of groups identified as similar, while the y axis gives the 0-1 proportion of all social learners who followed the majority of groups identified as different. The darkness of the grey indicates frequency, with darker colours denoting a denser patch. Any white areas show that none of the learners employed this strategy combination. Finally, the red square denotes the social learning strategy averaged over all participants.

Appendix 12i) The heatmap of social learning strategies for reliably correct signals



If the social learners wished to uphold the same cooperative norms as the group from whom they learned, then they should always follow the majority around similar others and follow the minority around different others when seeing reliably correct signals. This would mean that the optimal strategy space would be in the bottom right corner of the heatmap in Figure 12i. As can be seen, some social learners do answer optimally. There also some social learners who also follow the majority all the time. None of the social learners seem to follow the minority around reliably correct signals from similar others. There are some social learning strategies in the middle of the heatmap, suggesting that some social learners partially adjust. Comparing the participant IDs in the centre of the scatterplot from Appendix 11 to the social learners that self-reported ignoring social information (see script in appendix 5), we can see that these social learners who partially adjust are not necessarily the ones that self-report ignoring social information. That is, the partial adjustment may be a meaningful– if suboptimal– strategy.

Appendix 12ii) Heatmap of social learning strategies for reliably incorrect signals



For the reliably incorrect signals, the social learners should have followed the minority of similar others and should always follow the majority of different others in order to uphold the same cooperative norm as the group from whom they learn. This is the equivalent to the top left corner of the graph, which is filled in showing that some social learners did adjust their strategies as we expected. The top right square is also filled, suggesting that some social learners always followed the majority regardless of the signal. Again, there are partial adjustments in the middle of the graph.

Chapter 5:
Jack of all trades: Modular cognition and motivation may underlie our ability to master skills over distinct domains.

Aysha Bellamy
Charles Efferson
In preparation for publication

Word count: 14,858 excluding references and appendices

Abstract

An important aspect of every human society is our ability to master complex skills across various domains. This ability may be underlaid by a series of modular processes, each designed to deal with one recurrent issue that we faced throughout the ancestral past or may be underlaid by a domain-general reasoning ability which functions flexibly across multiple domains. This model compared modular and domain-general psychology to see which was likely to uphold our skill learning across multiple domains. Agents (N = 100) made decisions in two games against nature tasks, which were designed to mirror learning asocial skills in two distinct domains. I allowed both agent cognition and/or motivation to be modular (i.e., specialised to each domain) or domain-general (i.e., flexible across both domains). I found that domain-general psychology can uphold skill learning over two similar domains, where the fitness tied to matching one's skillset to the current environment align over both domains. Modular psychology was instead needed when mastering skills over two distinct domains. Modular cognition was important in driving agent behaviour, though this coevolved with motivation to influence behaviour in a myriad of ways. This suggests that drift may also play a role in skill learning.

Key words: skill learning, domain-general, modular, cognition

1. Introduction

Throughout the ancestral past, individuals had to master skillsets as varied as tool use, hunting, gathering, processing food, and sewing (Henrich, 2004; 2015). It is important to understand the psychological mechanisms that uphold our ability to display such a varied range of skills. Specifically, this model aims to investigate whether the acquisition of asocial skills is likely to be upheld by domain-general or modular psychological mechanisms.

Some researchers argue that the learning of broad skillsets is likely to be underlaid by a domain-general mechanism (Bolhuis et al., 2011). Domain-general cognition consists of a central processor that flexibly oversees decision-making across a variety of environmental inputs (Kan et al., 2013; Vergauwe et al., 2010). Due to its flexibility, a domain-general system may explain our ability to excel in skillsets that are necessary for careers that have only emerged recently in the ancestral past and so are unlikely to be underlaid by a genetically-selected bias (Muthukrishna & Henrich, 2016). For example, the increased demand for computer specialism (Robinson et al., 2020).

This theory contrasts with the dominant view of massive modularity in Evolutionary Psychology (Cosmides & Tooby, 1994). This is the concept that humans have many evolved cognitive processes, each designed to deal with one specific input to reach a certain behavioural output (Cosmides & Tooby, 1994; Pietraszewski & Wertz, 2021). These modules are designed to process recurrent issues faced throughout the ancestral past (Price, 2008), and may be organised hierarchically to underlie complex skill learning (Manoel et al., 2002). The putative cheater-detection module is an example of a task specific module and helps us to identify and avoid others who may break social rules. This module may help participants to enhance their skills on an otherwise difficult 'if-then' logic task, by helping to identify hypothetical cheaters (Cosmides et al., 2010; van Lier et al., 2013).

It is currently unclear whether this ability to learn skills over a variety of domains is underlaid by specialised modular processes (Charbonneau, 2016; Goschke & Bolte, 2012) or by an impressive generic reasoning ability (Bolhuis et al., 2011; Mesoudi, 2011). This chapter seeks to compare modular and domain-general psychology in a population of agents (who learn to display skills across two distinct domains), with the aim of determining which process is likely to uphold our ability to master skills over multiple domains.

Previous research comparing domain-generality versus modular psychology tends to be theoretical (Burke, 2014; Fodor, 1983; Frankenhuis & Ploeger, 2007; Spunt & Adolphs, 2017; Stephen, 2014; Stokes & Bergeron, 2015), or inferred based on observations of participant behaviour (Charbonneau, 2016; Mesoudi, 2011; Miu et al., 2020). However, these findings may be subject to the 'inverse problem' (Deffner et al., 2020). Complex and flexible behaviour does not necessarily evidence complex and flexible cognition, and seemingly complex behaviour can arise from simple cognitive processing throughout the animal kingdom (Amodio et al., 2019). For example, Clever Hans was a horse that could tap out sums with his hoof, but this was likely to be underlaid by operant conditioning rather than complex cognition (Lapuschkin et al., 2019).

To illustrate the inverse problem in reference to skill learning, imagine a gatherer collecting fruit in a field. She sees a spider and immediately withdraws to a safe distance. Her quick reaction may be consistent with a threat-detection module. Being exposed to poisonous animals such as spiders throughout the ancestral past may have created a selection pressure for the development of a module to quickly detect and avoid such animals (Öhman & Mineka, 2001). Alternatively, the agent's reaction may be consistent with a domain-general system, which does not have infinite processing power (Örün & Akbulut, 2019; Vergauwe et al., 2021). These systems may have to up-weigh the processing of survival relevant stimuli. Here, a domain-general system may up-weigh the processing of a potentially poisonous spider.

This chapter take a different approach to behavioural observation. I create an agent-based model where the agent is unsure which behaviour is optimal. I then vary the modularity or domain-generality of agent psychology, and then observe how this affects the agents' ability to master skills across two distinct domains. This mirrors the natural uncertainty that individuals have regarding which behaviour is likely to be optimal when learning a new skill (Mesoudi et al., 2015).

Modules in Evolutionary Psychology must have functional specialisation, working on one input at a time (Pietraszewski & Wertz, 2021). As such, I consider a modular agent to have psychological mechanisms that evolve separately per domain (Cosmides & Tooby, 1994). Based on the definition of domain-general psychology, I consider a domain-general agent to have one psychological system which must remain flexible across both domains (Bolhuis et al., 2011). These skill domains can be thought of as any which would have been beneficial to master throughout the ancestral past. Throughout, I reference the examples of weapon use when hunting (Tomka, 2013) and cooking food (Wrangham, 2009).

Previous research has focused on *cognitive* modularity (Cosmides & Tooby, 1994) though it is my view that *motivation* is just as important. To illustrate with an example, should our motivation to detect and avoid poisonous spiders be the same as our motivation to create and share folktales (Da Silva & Tehrani, 2016)? Some agents seem unmotivated to master skills via their own trial-and-error in certain domains

(Mesoudi, 2008). For example, some individuals rely heavily on the work of others when writing code (Miu et al., 2020), This indicates that motivation is important to skill-learning. I therefore model motivation to be modular (i.e., specialised to each of the two skill domains) or domain-general (i.e., flexible across both domains).

Investigating motivation is important for two reasons. First, previous research has focused on one aspect of cognition or motivation at a time (Delton et al., 2011), typically to avoid psychological polymorphism. Leaving psychological traits to coevolve could produce a wealth of psychological phenotypes that all support similar behaviours at the individual level (Kurzban & Houser, 2005). To avoid psychological polymorphism, previous models have investigated simplistic representations of human nature (Laland, 1993). However, there is a trade-off between the simplicity of the model and the extent to which it captures meaningful – and complex – human behaviour (Kendal et al., 2018). It is my view that fixing either cognition or motivation, and allowing the other to evolve, is an oversimplification. Our cognitive and motivational systems would have evolved in-tandem as we mastered a range of skillsets throughout our ancestral past, and thus this model allows both traits to coevolve.

Second, modelling modularity separately for cognitive and motivational thresholds allows for exploring partly modular agents, between the two extremes of full modularity (Cosmides & Tooby, 1994) and full domain-generality (Bolhuis et al., 2011). For example, a partly modular agent may have modular cognition and domain-general motivation. This agent could reason distinctly about the value of specific behaviours, but only show a generic drive towards mastering any skill. An alternative partly modular agent could have domain-general cognition and modular motivation. They would reason generically about the value of behaviours over many skill domains but have a specific desire to master each skill.

To summarise, I investigate how agents come to master skills in two domains. This model has two novel aims:

1. To investigate how cognition and motivation coevolve to influence our behaviour over multiple skill domains.
2. To investigate whether the psychology underlying our ability to master skills over multiple domains is likely to be fully modular, partly modular, or fully domain-general.

2. Methods

2.1. Model description

The 100 agents play games against nature tasks, in which one behaviour will lead to an optimal payoff. As the agent's behaviour does not affect the fitness of any other agents in her group, these tasks mirror the learning of asocial skills (Molleman & Gächter, 2018).

Agents play two games against nature tasks, which I refer to throughout as domain A and domain B. In each domain, the agent chooses one of two behaviours (0 or 1) and receives a payoff if her behaviour matches the environmental state (0 or 1). The agent may select the optimal behaviour in both, one of neither of the domains. This decision-making environment is realistic, as skills are only useful if employed in the right context. For example, the optimal design of a weapon depends heavily on environment (Mesoudi et al., 2015), and local food must often be processed in specific ways (Henrich, 2015). All model parameters and variables, plus their notation used throughout this paper, are given in Table 1.

Table 1. The notations and variables used throughout the model.

Symbol	Model description
$S_A, S_B \in \{0,1\}$	The environmental state in domain A or B respectively.
$B_A, B_B \in \{0,1\}$	The behaviour that the agent employs in domains A and B respectively.
$x_A, x_B \in (-\infty, +\infty)$	The cue summary drawn from the environment to help the agent decide which state the environment is likely to be in for both domains A and B respectively.
$s_A, s_B \in \{0,1\}$	The state that the agent believes the environment is in for both domains A and B respectively. These can be wrong.
$T_A, T_B \in (-\infty, +\infty)$	The cognitive threshold of evidence needed to believe that the state is 1 in domains A and B. (Note, just T for domain-general agents).
$\alpha_A, \alpha_B \in [0,1]$	The agents' motivation to play behaviour 1 when they believe that the state is 1 for domains A and B respectively. (Note, just α for domain-general agents).
$\beta_A, \beta_B \in [0,1]$	The agents' motivation to play behaviour 1 when they believe that the state is 0 for domains A and B respectively. (Note, just β for domain-general agents).
$p_A, p_B \in \{0.1, 0.5, 0.9\}$	The probability that the state will be 1 in both domains A and B respectively.
$f_{Azero}, f_{Bzero} \in \{0.25, 1, 4\}$	The fitness that the agent gains when her skillsets match the environment when the environmental state is 0. These are modelled separately for domains A and B. These can take a low (0.25), intermediate (1) or high value (4).
$f_{Aone}, f_{Bone} \in \{0.25, 1, 4\}$	The fitness that the agent gains when her skillsets match the environment when the environmental state is 1. These are modelled separately for domains A and B. These can take a low (0.25), intermediate (1) or high value (4).

Let the agents chosen behaviour be denoted by B_A for domain A and B_B for domain B. Both components could take the realisation $\in\{0,1\}$. This behviour will only be beneficial to fitness if it matches the envrironmental state in each domain. The environment in domain A can take one of two states,as represented by a random variable, S_A. The state of the environment in domain B was represented by a random variable, S_B. Both had support $\in\{0,1\}$ across both models.

To illustrate this logic in relation to the game against nature task, imagine that the game in domain A represents the use of a bow-and-arrow. Let $B_A = 1$ represent the decision to use a bow-and-arrow, and $B_A = 0$ represent the decision to not use the bow-and-arrow. That is, $B_A = 1$ always represents the decision to act whilst $B_A = 0$ represents the decision not to act. As this is the first comparison of modular and domain-general decision-making in a population of theoretically-evolving agents, we take the simplest representation of a domain-general system as one that must decide to act or not act over multiple domains (Pietraszewski & Wertz, 2021). This should not be seen as a full and comprehensive comparison of modular or domin-general architecture, but is the first step towards this goal. Let $S_A = 1$ denote an environment where using the bow-and-arrow would be appropriate (e.g., when hunting intermediate sized prey animals from a distance; Tomka, 2013). Cases where S_A and $B_A = 1$ indicate acting successfully. When $S_A = 1$ and $B_A = 0$, the agent failed to act when it would have been optimal (i.e., 'fails to use bow-and-arrow when appropriate').

Let $S_A = 0$ represent an environment where it would be inappropriate to use the bow-and-arrow (e.g., when hunting a larger animal with a tough hide, or in close-contact fighting; Tomka, 2013). S_A and $B_A = 0$ implies that the agent shows flexible decision-making and 'avoids acting when inappropriate'. $S_A = 0$ and $B_A = 1$ implies that the agent 'acts inappropriately'. Figure 1 gives the formal payoff of the game against nature in domain A.

		Environmental state in domain A	
		0 (should not use bow-and-arrow)	**1 (should use bow-and-arrow**
Focal agent's behaviour	**0 (does not use bow-and-arrow)**	$f_{Azero} \in \{0.25,1, 4\}$ (avoids using bow-and-arrow when inappropriate)	0 (fails to use bow-and-arrow when appropriate)
	1 (uses bow-and-arrow)	0 (uses bow-and-arrow inappropriately)	$f_{Aone} \in \{0.25,1, 4\}$ (successful bow-and-arrow use)

Figure 1. The payoff matrix for the game against nature in domain A (use of bow-and-arrow). The payoff of failing to match one's skill to the environment is always set at 0. Thus, the payoff to the agent for matching her skillset to the environment is driven entirely by the size of the f_{Azero} and f_{Aone} parameters.

The payoff of failing to match one's skillsets to the environment is standardised at 0. Of course, in the ancestral past using the wrong weapon when hunting could result in a heavy cost. For example, failing to hit an animal with a bow-and-arrow represents at best a loss of food and at worst a potentially fatal incident with a startled animal (Churchill, 1993). In my model, fitness payoffs were always positive as this allowed for reproduction to be tied to cumulative fitness.

The appeal of matching one's skillsets to the environment is driven entirely by the size of the f_{Azero} and f_{Aone} parameters. These are modelled separately and could take realisations $\in\{0.25,1, 4\}$ respectively. I focus on cases where these parameters take either a low (0.25) or high value (4). Cases where f_{Azero} and f_{Aone} = 0.25 imply there is only a small difference in payoff between using a bow-and-arrow or not. For example, in a friendly athletics demonstration it does not matter too much which equipment one displays. Cases where f_{Azero} and f_{Aone}= 4 instead suggest a strong pressure to use the bow-and-arrow when appropriate. It is paramount to use the bow-and-arrow successfully at a distance to bring down easily startled game, though using it in close range can lead to missed opportunities or injury (Tomka, 2013). Domain B may represent skill learning in either a similar or distinct domain to domain A. For an example of a similar domain, imagine that domain B represents the decision of whether or not to use a spear thrower (which can be used as a projectile weapon from a distance). There would be an overlap between the environments where it would be optimal to hunt with a bow-and-arrow and environments where it would be optimal to hunt with a spear thrower (Hughes, 1998).

As an example of a distinct domain to domain A, imagine that domain B represents cooking. As our diet expanded into a various meats and grains, it became increasingly important for our digestive health to process these foods correctly (Henrich, 2015; Wrangham, 2009). Let B_B = 1 represent the decision to spend more time cooking food and B_B = 0 represent the decision to spend less time cooking food. Let S_B = 1 represent a case where cooking food would be appropriate. For example, meat must be cooked, as the ingestion of raw meat is likely to contain harmful bacteria, and we cannot digest raw meat, having lost a significant portion of our intestinal tract since the discovery of fire (Henrich & Muthukrishna, 2021). Cooking also aids digestibility for plants (Wrangham, 2009). In my model, cases where S_B and B_B = 1 imply that the agent acts successfully (i.e., cooks food when needed). Cases where S_B = 1 and B_B = 0 imply that the agent fails to act successfully (i.e., fails to cook when necessary). This could have adverse fitness effects and can even be fatal (Henrich, 2015; Wrangham, 2009).

Let S_B = 0 represent a food which does not have to be cooked. Cases where S_B and B_B = 0 imply that the agent successfully avoids acting when unnecessary (i.e., she avoids overcooking her food). Cases where S_B = 0 and B_B = 1 imply that the agent acts unnecessarily (i.e., they overcook; for example, some plants emit harmful toxins when overcooked; Stahl et al., 1984). There may be cases where to undercook (f_{Bzero}) or to overcook (f_{Bone}) would not have high fitness consequences (i.e., f_{Bzero} and f_{Bone} = 0.25). For example, steaks can be consumed anywhere from blue to well done and – although individuals may have strong opinions on how they like their steak – it does not hurt to eat a steak that is under or over-cooked to your preference. Cases with high fitness consequences (f_{Bzero} and $f_{Bone} = 4$) instead represent cases where to under or over-cook would be deadly. For example, the cassava root has to be processed to an exact, detailed method or it may make one very ill due to toxins (Henrich, 2015). Figure 2 shows the payoff matrix when domain B represents the decision to cook.

		Environmental state in domain B	
		0 (should not cook food)	**1 (should cook food)**
Focal agent's behaviour	**0 (does not cook)**	$f_{Bzero} \in \{0.25, 1, 4\}$ (successfully avoids over-cooking)	0 (under-cooks food)
	1 (cooks)	0 (over-cooks food)	$f_{Bone} \in \{0.25, 1, 4\}$ (successful cooking)

Figure 2. The payoff matrix for the game against nature in domain B (cooking food). The payoff of failing to match one's skill to the environment is always set at 0. Thus, the payoff to the agent for matching her skillset to the environment is driven entirely by the size of the f_{Bzero} and f_{Bone} parameters.

In order to display a behaviour, the agent first formulates a belief as to the likely value of the environmental state. Let s_A represent the agent's belief about the value of the environmental state in domain A, and s_B represent the agent's belief about the value of the environmental state in domain B. In domain A, the agent either believes that it would be appropriate (s_A=1), or inappropriate to use a bow-and-arrow (s_A=0). In domain B, the agent either believes that she should cook her food (s_B=1) or that cooking the food is unnecessary (s_B=0).

Each agent formulates her belief by independently and randomly drawing a cue summary from the environment. This cue summary can be thought of as the cues that one weighs up when trying to decide which skill would be optimal to master in the current environment. For example, cues suggesting the use of a bow-and-arrow include whether the agent is hunting a medium-sized prey animal with a soft hide. The cue summary is probabilistically – not deterministically – decided by the environmental state and is drawn independently for both domains. When S_A or S_B = 1, the agent draws a cue summary from a normal distribution with a Mμ =1, σ =1. When S_A or S_B = 0, the agent draws a cue summary from a normal distribution with a Mμ = -1, σ =1.

Let the agent's private signal (i.e., cue summary) in domain A be represented by x_A and let the agent's private signal in domain B represented by x_B. These could be any random value from these respective cue distribution curves. These cue summaries could be thought of as the agent's private signal of evidence that the state was 1. Note that these cue summaries were probabilistic but not deterministic. Indeed, the difficulty of distinguishing between these environments was driven by the overlap of the two distribution curves (~31.7%). Whilst the agent could usually form accurate beliefs about the environment, she could occasionally be wrong about the environmental state. Let s_A and s_B to refer to the agent's belief about the environmental state. These could take support $\in\{0,1\}$ and importantly, these could be wrong. Thus, the agents made decisions regarding which skill to master under uncertainty.

Once the agent has drawn a cue summary from the environment (x_A or x_B), she can translate this into her beliefs about the environmental state in each domain (s_A or s_B). The agent does this using her cognitive threshold (the minimum amount of evidence needed to conclude that the state is 1 in each domain). These thresholds can take any positive or negative value. Let the cognitive threshold in domain A be represented by T_A and the cognitive threshold in domain B be represented by T_B. For example, if the agent's cue summary exceeds her cognitive threshold for domain A ($x_A > T_A$), she believes that the state is 1 ($s_A = 1$), and that it is appropriate to use the bow-and-arrow. When this cue summary does not exceed the threshold ($x_A <= T_A$), the agent does not have enough evidence to believe that s_A is 1 and concludes that using the bow-and-arrow would be inappropriate ($s_A=0$). Similarly, in domain B, the agent will believe she has to cook her food when her signal exceeds her cognitive threshold ($x_B > T_B$, $s_B = 1$) or believe that cooking is unnecessary when it does not ($x_B <= T_B$, $s_B = 0$). Of course, the agent's beliefs in each domain may or my not be correct and so they make decisions under uncertainty.

The cue summaries were modelled so that 0 would represent unbiased cognitive thresholds. An agent that evolved positive cognitive thresholds thus needed more evidence from the environment before she would believe that the state was 1. Conversely, an agent who evolved negative cognitive thresholds was less discerning and needed less evidence from the environment to believe that the state was 1.

Agents with modular cognition can reach distinct conclusions about the value of the environmental state in domains A and B because their cognitive thresholds are allowed to evolve separately in each domain (T_A and T_B). For example, this agent could conclude that it is appropriate to use the bow-and-arrow in domain A (s_A=1), and inappropriate to cook the food in domain B (s_B=0). In contrast, domain-general agents can only reason generically about the likelihood of the state being 1 over both domains A and B. This is because I constrain their cognitive threshold so that $T = T_A = T_B$. In any domain, the agent's beliefs about the environment can be incorrect. Thus, agents master asocial skills in conditions of uncertainty.

After forming a belief about the state of the environment, the agent can then only display a skill if she is motivated to do so. Each agent drew a random number form the uninform interval [0,1] and compared this with an internal motivational threshold. Whenever the agent believed that the environmental state was 1 ($s_A = 1$ or $s_B = 1$), then the agent acts (i.e., plays behaviour 1) with a motivation of $Probability_{State\ 1}$ (which I label α for future reference). Whenever this random number exceeded the α threshold, then the agent would play behaviour 1. If this random number did not exceed α, then the agent would play behaviour 0. The probability of not acting (i.e, playing behaviour 0) when the agent believed that the environmental state took the value 1 was therefore given by $1- \alpha$.

To illustrate, imagine that the agent has enough evidence from the environment to conclude that using the bow-and-arrow would be appropriate ($x_A > T_A$, $s_A = 1$). She will then use the bow-and-arrow with a motivation of α_A or fail to use the bow-and-arrow with a motivation of $1- \alpha_A$. Likewise, imagine that the agent has enough evidence from the environment to conclude that she should cook her food ($x_B > T_B$, s_B=1). The agent thus cooks her food with a motivation of α_B or fails to cook with a motivation of $1- \alpha_B$. Thus, α_A and α_B values represent the agent's motivation to play behaviour 1 when she believes that this would be optimal given the environmental state.

Whenever the agent believed that the environmental state was 0 ($s_A = 0$ or $s_B = 0$), then she would act (i.e., play behaviour 1) with a motivation of $Probability_{State\ 0}$ (which I label β for future reference). The probability of not acting (i.e., playing behaviour 0) when the agent believed that the environmental state took the value 0 was therefore given by $1- \beta$. β is therefore the agent's desire to play behaviour 1 when she thought that this was the suboptimal skill to master in her environment.

To illustrate, imagine that the agent's cue summary does not exceed her cognitive threshold ($x_A <= T_A$ or $x_B <= T_B$), and she believes that the bow-and-arrow is inappropriate ($s_A=0$), or that cooking is unnecessary (s_B=0). The agent then uses the bow-and-arrow with a motivation of β_A or fails to use the bow-and-arrow with a motivation of $1 - \beta_A$. Likewise, the agent cooks her food with amotivation of β_B, or does not cook her food with a motivation of $1 - \beta_B$. Thus, β_A and β_B values always represent the agent's motivation to play behaviour 1 when she believes that this would be suboptimal given the environmental state.

Agents with modular motivation can show a distinct desire to use the bow-and-arrow or cook (either successfully or inappropriately). This is because agents with modular motivation can have α_A, α_B, β_A and β_B values evolve separately in each domain. The agents with domain-general motivation can only show a generic desire to play behaviour 1 when they believe that this is optimal, as I constrain their motivation so that $\alpha = \alpha_A = \alpha_B$. Likewise, the agents with domain-general motivation can only show a generic desire to play behaviour 1 when they believe that this was suboptimal, as I constrain their motivation so that $\beta = \beta_A = \beta_B$.

Note that I model modularity separately for the cognitive and motivational thresholds. This gives four possible combinations of agent types:

1. Fully modular agents can reason distinctly about the likelihood of the bow-and-arrow and cooking being appropriate across the two domains (T_A, T_B). They can also show a distinct desire to play behaviour 1 (i.e., use the bow-and-arrow or cook) across the two domains (α_A, α_B, β_A, β_B).
2. Partly modular agents with modular cognition only can reason distinctly about the likelihood of the bow-and-arrow and cooking being appropriate across the two domains (T_A, T_B). However, they can only show a generic desire to play behaviour 1 or behaviour 0 across both domains (α, β). That is, they can only show a generic desire to use the bow-and-arrow and to cook.
3. Partly modular agents with modular motivation only can only reason generically about the likelihood that the state was 1 across both domains (T). They only have a generic belief that using the bow-and-arrow and cooking is appropriate. However, they have a distinct desire to play behaviour 1 in both domains (α_A, α_B, β_A, β_B).
4. Domain-general agents can only reason generically about the likelihood that the state is 1 across both domains (T) and show a generic desire to play behaviour 1 or 0 across both domains (α, β).

I run four models, each eliciting one of the four agent types above. It is not my aim to investigate which agent type is the most successful and therefore the most likely to become evolutionary dominant across the run. Instead, it is my intention to compare the mastering of skillsets, and the endpoints of evolution of the cognitive and motivational thresholds, across the four types of agents.

Cognitive and motivational thresholds are left to endogenously evolve throughout these models. These components decide the agents behaviour (B_A and B_B). The payoff of the agent's behaviour also depends on the environmental state (S_A and S_B) in each domain. The agents receive an exogenous fitness value of 1 on top of the fitness that they accrue via playing the games against nature. This reflects the fitness that the agents may gather from factors besides their ability to master asocial skills.

For every simulation, the population evolved through 20,000 generations. Each agent gained fitness via her behaviour in each domain as dependent on the environmental state. All agents received an exogenous fitness value of 1 on top of the points they earned from the two games. This reflected the agents fitness from behaviours besides skill acquisition. These values were then summed to give a total fitness value per agent. Note that the agents chance of having offspring was directly proportional to their total fitness. For every one of the 100 agents, I translated their total fitness value into a fitness value which was a cumulative proportion of the entire generation. Agents who had more fitness thus had a higher cumulative proportion value. To assign offspring genotype, I simply allowed a random interval between [0,1] which corresponded to this cumulative proportion space. Thus, parental agents with a larger proportional fitness would have a higher chance of having offspring, as the offspring agents were more likely to be sampled from these larger proportional fitness values. Once the offspring's parental agent was identified, the offspring then inherited her parental agent's psychological variables. Note that I allowed cognition and motivation to be inherited independently from different parental agents. This did not mirror sexual reproduction, but was instead intended to reflect how continuous, complex traits such as psychological phenomena are likely to be coded at multiple allele sites.

Agents inherited the cognitive thresholds with a mutation rate of 0.5. As this was a continuous variable with no fixed end points, I assumed mutations to be common, but constrained mutation to only occur in small chunks in each time step ($M\mu$ = 0, σ =

0.00125). Agents inherited their motivational thresholds with a lower mutation rate of 0.05. Mutations occurred in steps of plus or minus 0.01 at each timestep. Note that mutation was constrained so that the motivational thresholds could only take a value between 0 and 1. The agents could never be less than 0%, or more than 100%, motivated to play a certain behaviour. Mutation was modelled independently for the two psychological components. The offspring and parental generation did not co-exist. Offspring agents replaced parental agents entirely once they had inherited their traits, as is typical of a Wright Fisher model.

2.2. Parameters

We ran 100 simulations for each possible combination of parameter space (see appendix 1 for full space). Each simulation consists of a population of 100 agents evolving over 20,000 generations. First, I vary the probability of the environmental state being 1 in domain A (denoted by p_A) and domain B (denoted by p_B). These are modelled separately in each domain and could take the realisations $p_A \in \{0.1, 0.5, 0.9\}$ and $p_B \in \{0.1, 0.5, 0.9\}$ respectively. The likelihood of the bow-and-arrow being appropriate does not depend on the likelihood of cooking being appropriate.

I also vary the f_{Azero}, f_{Bzero}, f_{Aone} and f_{Bone} parameters. These are also modelled separately in each domain. These give the fitness when the agent's skillsets match the environmental state in domains A and B. I focus on cases where matching gives a small (0.25) or large payoff (4). As the full parameter space is too large to investigate here, I instead focus the analyses on the following four combinations of interest:

1. Runs where the agents make decisions in two similar domains (e.g., deciding whether to use a bow-and-arrow or spear thrower). On these runs, both the priors (p_A and p_B = 0.1) and the fitness when the agents' skillsets match the

environment favour behaviour 0 (f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). For example, the individual may engage in close-contact hunting instead of using a bow-and-arrow or spear thrower.

2. Runs where the agents make decisions in two similar domains, but the most common environmental state is not necessarily the best fit for the agent's skillset. On these runs, the priors favour state 1 (p_A and p_B = 0.9) though the fitness when the agents skillset matches the environment favours behaviour 0 (f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). For example, it would be common for the agent to be able to use a bow-and-arrow to hunt deer, but the rarer and more beneficial prize would be to use close-contact weapons to take down larger animals such as mammoth.
3. Runs where the agent makes decisions in two distinct domains. For example, the decision of when it is appropriate to use the bow-and-arrow or to cook. In these runs, the priors and fitness of matching one's skillsets to the environment favour behaviour 0 in domain A but behaviour 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). It is rare that the agents would need to use the bow-and-arrow to take down deer, but it is common that the agents need to cook their food.
4. Runs where the agent makes decisions in two distinct domains, but the most common environmental state is not necessarily the optimal state for the agents to match their skillsets to (p_A=0.1, p_B=0.9, f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). Here, the prior makes state 0 likely to occur in domain A but the fitness tied to matching one's skillsets to the environment favour state 1 of domain A. For example, it may be common for the agent to have the opportunity to set traps for smaller animals, such as rabbit, but the larger prize would be to use the

bow-and-arrow to hunt the deer which are more elusive in this run. The prior makes state 1 in domain B likely to occur, but the fitness tied to matching one's skillset to the environment favours state 0 in domain B. For example, it is common that we have to cook our food as our intestines cannot process raw meat or grains (Henrich & Muthukrishna, 2021; Wrangham, 2009). It is sometimes the case – though more rarely – that overcooking food is dangerous. For example, some plants emit toxins when overcooked (Stahl et al., 1984).

To summrise the model, the agents must decide to 'act' or 'not to act' in two skill domains. To make this decision, each agent must first decide which environmental state the domain is in (does it favour acting [$s_A = 1, s_B = 1$] or not acting? [$s_A = 0, s_B = 0$]). To decide this, each agent has a private signal (x_A, x_B) which can be thought of as the cues from the environment. These cues must exceed the cognitive threshold ($x_A > T_A$, $x_B > T_B$) for the agent to believe that the state is 1 ($s_A = 1, s_B = 1$). This signal can thought of as all the cues that the agent needs to decide that acting is beneficial in each domain (e.g., cues that suggest the arrow is needed is if we hunt intermediate sized prey with soft hide from a distance). If the private signal does not exceed the threshold ($x_A <= T_A$, $x_B <= T_B$), then the agent instead believes that the state is 0 and that they should not act ($s_A = 0, s_B = 0$). These beliefs can be wrong. After formulating these beliefs, the agent must be motivated in order to act. The agent is motivated to act with a probability of α when they believe that the state is 1 (and1-α for not acting), and are motivated to act with a probability of β when they believe that the state is 0 (and 1-β for not acting). Once the agent 'acts' ($B_A = 1, B_B = 1$) or 'does not act' ($B_A = 0, B_B = 0$) in each domain, they then receive fitness based on whether this behviour actually matches the environmental state ($S_A = 0, S_A = 1$, and $S_B = 0, S_B = 1$) across both domains. This will be affected by the probability that the state is 1 (p_A and p_B) and the fitness parameters set in each domain (f_{Aone}, f_{Azero}, f_{Bone}, f_{Bzero}). All agents receive an exogenous fitness value of 1 on top of this fitness. Fitness is then changed to a cumulative proportion value, so that more fit agents are more likely to have more offspring. Reproduction is therefore proportional to fitness. The offspring inherit cognitive (T_A, T_B) and motivational thresholds (α_A, α_B, β_A, β_B) from separate agents with a small rate of mutation. The offspring generation then overwrite the parental agent at each time step, as characteristic of a Wright-Fisher model (Suchow et al., 2017).

Our ultimate aim is to compare how modular and domain-general agents come to decide to act or not act in each skill domain. To achieve this, fully modular agents were coded to have a separate cognitive and motivtional threshold per domain (T_A, T_B, α_A, α_B, β_A, β_B). Partly modular agents either had modular cognition but domain-general motivation (T_A, T_B, α, β) or they had domain-general cognition and modular motivation (T, α_A, α_B, β_A, β_B). Finally, the fully domain-general agents have constrained cognition so that $T = T_A = T_B$, and constrained motivation so that $\alpha = \alpha_A = \alpha_B$ and $\beta = \beta_A = \beta_B$. This matches the definition of domin-general systems as those that must operate to decide to act or not act over multiple decision-making domains simultaneously.

3. Results

Our model converged on a stable psychological architecture across the 20,000 generations over all simulations (see appendix 2). For each parameter combination, I explore the behavioural outcomes of all agent types in the final generation (section 3.1), the psychological architecture of the final generation of agents (section 3.2) plus how cognition and motivation predict fitness for the final generation of agents (section 3.3).

I then compare the fitness of all four agent types to see whether domain-general or modular psychology is likely to account for skill learning (section 3.4) followed by a summary of the key findings (section 3.5).

3.1. Did the agents learn to display the optimal skill in each domain?

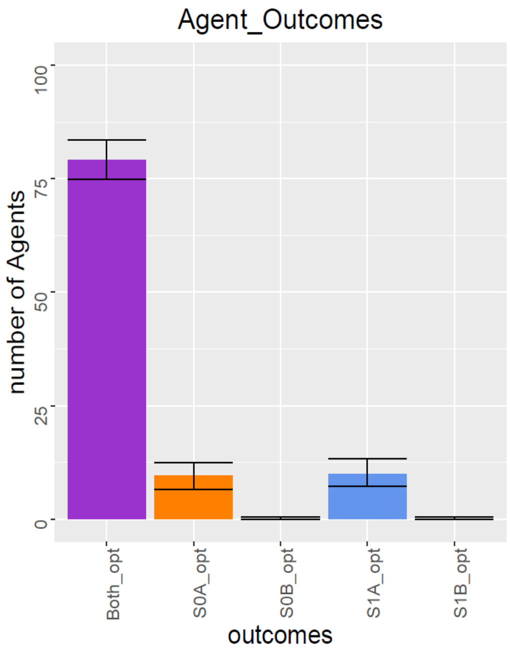
Provided that the agents can track the priors and fitness tied to matching their skillset to the environment, then it would be expected that all agents eventually display the optimal skill in each domain (Mesoudi et al., 2015; Molleman & Gächter, 2018). Figures 3 and 4 display the agents' behavioural outcomes in the final generation. I plot the number of agents who answered both domains correctly and just one correctly. The remaining agents did not answer correctly in either domain.

For runs where the agents made decisions in two similar domains of skills, there was a similar result across all agent types (p_A=0.1, p_B=0.1*;* f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4; figure 3i). The majority of all agents (~80) answered correctly in both domains, with very few agents (~10) answering correctly in state 0 or state 1 of domain A only. When the priors and fitness to matching one's skillset to the environment were consistent across two domains, the agents reached a similar behavioural outcome regardless of their modularity type.

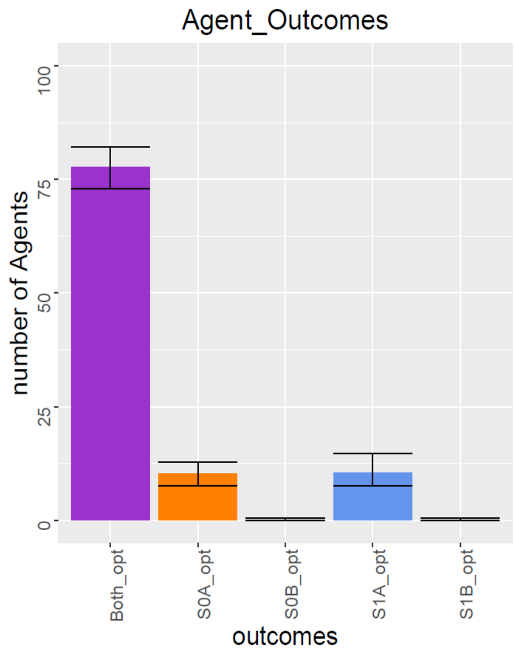
Figure 3ii displays runs where the agents decide over two similar skill domains, though the most common environment is not necessarily the most optimal environment to tailor one's skillsets to (p_A=0.9, p_B=0.9*;* f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). For example, it is common that the agent can hunt deer with a long-distance weapon, though the rarer and more beneficial prize may be to use a different weapon to hunt larger prey. Here, agents with modular motivation only and domain-general agents had ~25 agents answer correctly in both domains, with fewer domain-general agents managing to answer just one domain correctly. Sometimes, these agents aimed to master the rarer skill with the higher payoffs (i.e., using a different weapon to hunt larger prey) and sometimes they would master the most common skill (i.e., using the spear thrower to hunt smaller but commonly-available deer). The fully modular agents were more likely to answer correctly in state 0 of domain B (~25 agents) than answer both domains correctly (~20 agents), suggesting that these agents attempted to display behaviour 0 as this would have a higher payoff, though was less common. To illustrate, fully modular agents would avoid using the spearthrower to hunt the commonly-available deer but instead use a different weapon to hunt larger prey.

The agents with modular cognition only may have performed the best on this run as ~30 agents answered correctly in both domains, with ~20 agents answering correctly in state 1 or state 0 of domain B only. These agents may have used a bow-and-arrow to bring down the commonly-available deer in domain A but would forego the spearthrower in domain B, instead seeking the opportunity to hunt larger prey.

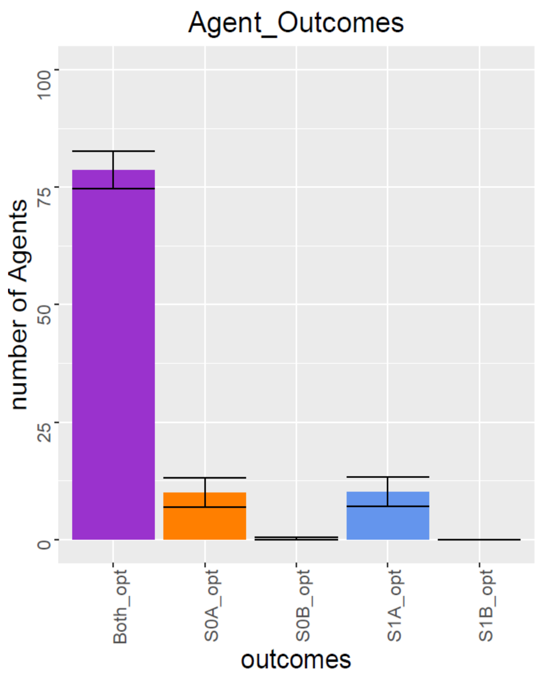
Figure 3i) $p_A=0.1$, $p_B=0.1$; $f_{Aone}=0.25$, $f_{Azero}=4$, $f_{Bone}=0.25$, $f_{Bzero}=4$



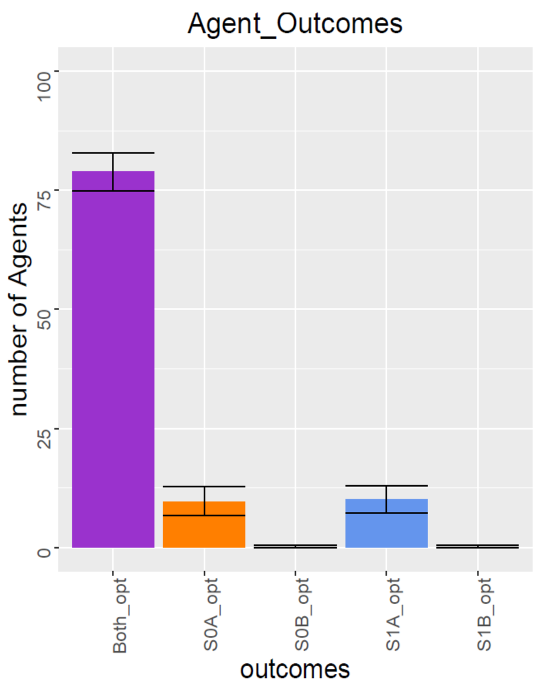
Fully modular



Modular motivation

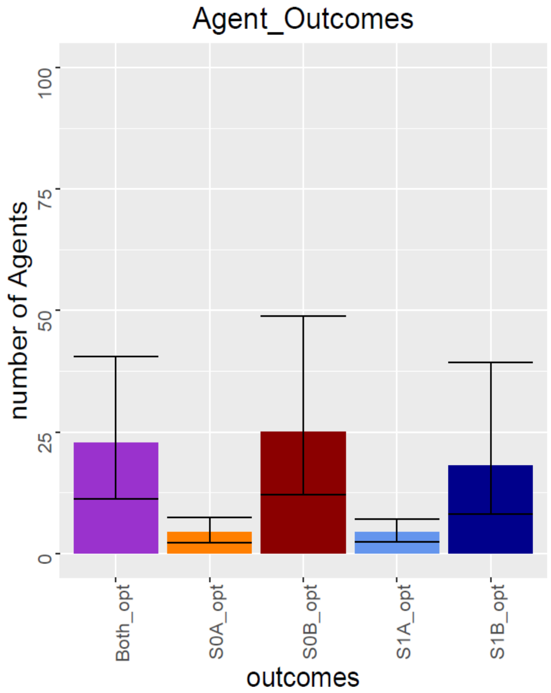


Modular cognition

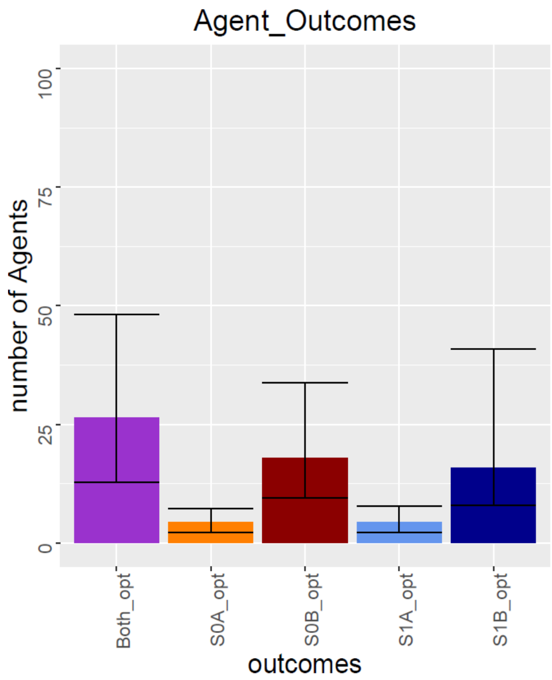


Domain-general

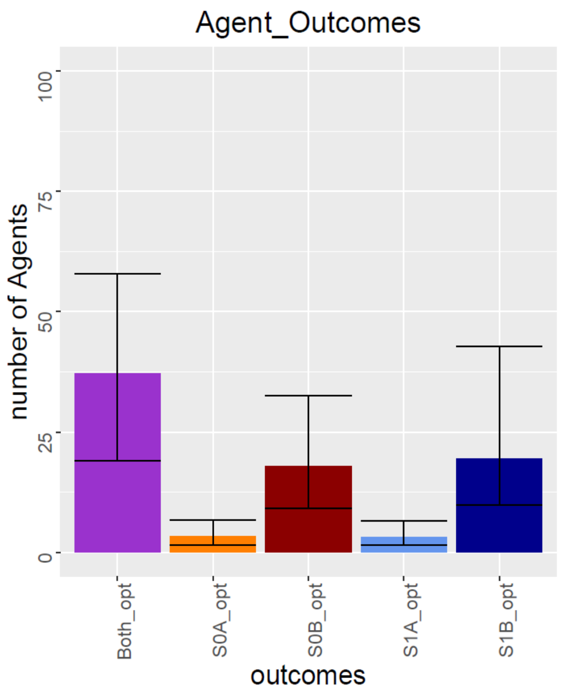
Figure 3ii) p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4



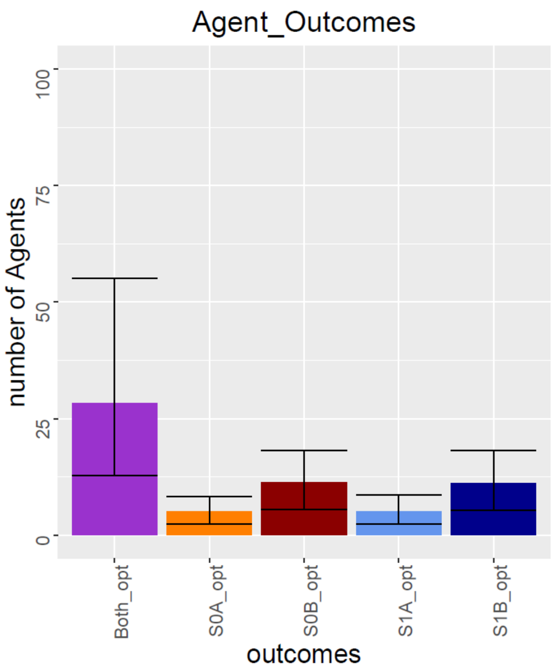
Fully modular



Modular motivation



Modular cognition



Domain-general

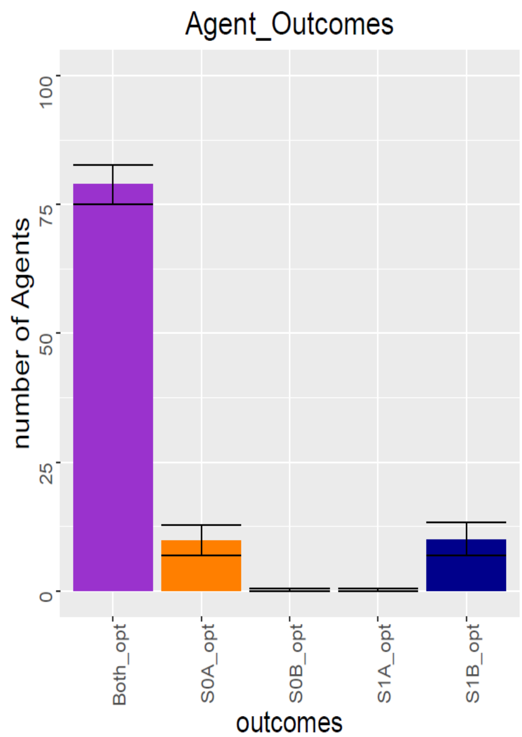
Figure 3. The agent outcomes on runs where the agent made decisions in two similar skill domains. The x axis gives the possible outcomes. Both_opt represents cases where the agent chose the optimal behaviour in both domains, S0A indicates correct answers in state 0 of domain A, S0B in state 0 of domain B, S1A in state 1 of domain A and S1B in state 1 of domain B. Results for the fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents are graphed separately. The graphs display the agents' response to i) runs where the priors and fitness tied to matching one's skillsets to the environment favour state 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4) and, ii) runs where the most common environmental state to match on is not necessarily the state which gives the highest payoff (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

For runs over two distinct domains, most of the modular agents answered optimally across both domains (~80) (p_A*=0.1*, p_B*=0.9;* f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25; see figure 4i). Any agent with a degree of modularity could avoid using the bow-and-arrow to hunt when inappropriate and could cook their food when appropriate. The domain-general agents only had ~50 agents answer optimally in both skill domains, and ~20 agents only answered correctly in state 0 of domain A or state 1 of domain B. The domain-general agents may have sometimes correctly avoided using the bow-and-arrow or may have correctly cooked their food. However, due to having to balance their psychological architecture over two distinct domains, they could not always master both skills optimally.

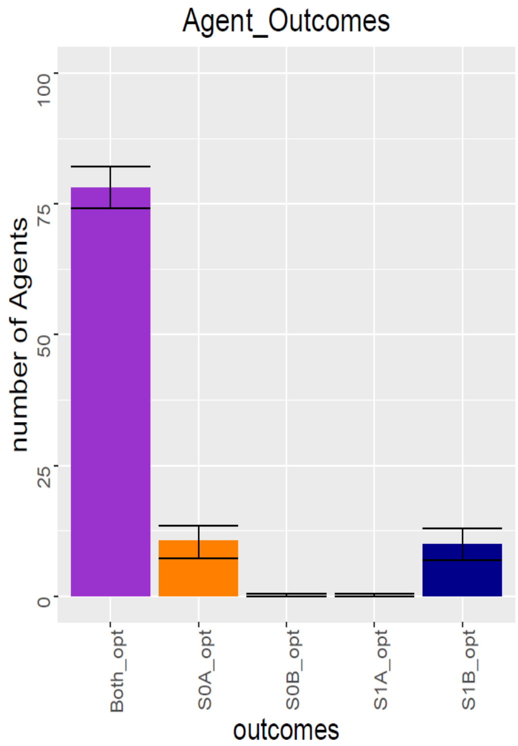
Finally, I move onto a case where the priors and fitness tied to matching one's skillsets to the environment favour two different environmental states over two distinct domains (p_A*=0.1*, p_B*=0.9;* f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4; see figure 4ii). To illustrate the sort of environment that these runs represent in reference to domain A, imagine that it is common for the agent to be able to set traps to catch smaller prey such as rabbits, but a rarer and more prized food source would be to use the bow-and-arrow to hunt deer. In domain B, it is more common that we need to cook food, but it is sometimes the case that overcooking is dangerous (Stahl et al., 1984). Fully modular agents and agents with modular motivation only had similar behaviours. Approximately 20 agents answered correctly in both domains, with ~15-20 agents answering correctly in state 0 of domain A or state 1 of domain B only. The agents with modular cognition did slightly better, with ~40 agents answering correctly in both domains, and ~20 agents answering correctly in state 0 of domain A or state 1 in domain B only. Interestingly, the domain-general agents were the most likely to answer optimally across both domains (~50 agents) out of all agent types.

In sum, when the priors and the fitness tied to matching one's skillset to the environment favoured different skillsets, then all agents tried to adjust their skill to the most common environment rather than the environment with the highest payoffs. To illustrate in reference to domain A, the agents set traps for the common rabbit but do not use the bow-and-arrow to pursue the rarer deer. In domain B, the agents will sometimes overcook as cooking food is often more necessary than not having to cook.

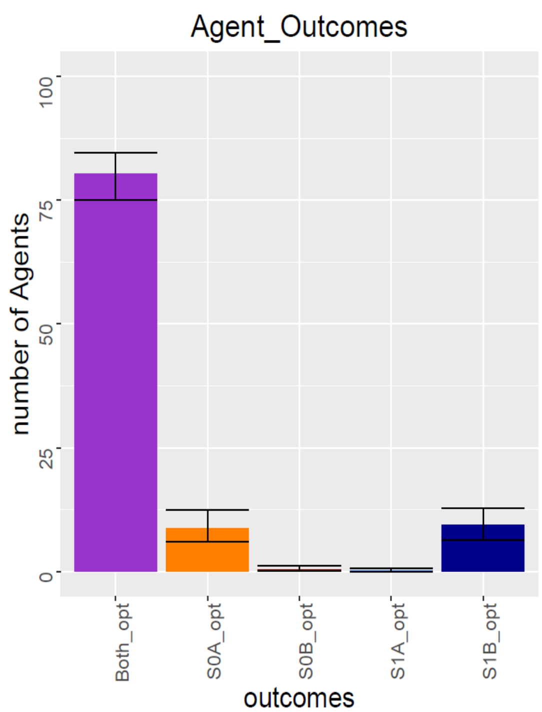
Figure 4i) p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25



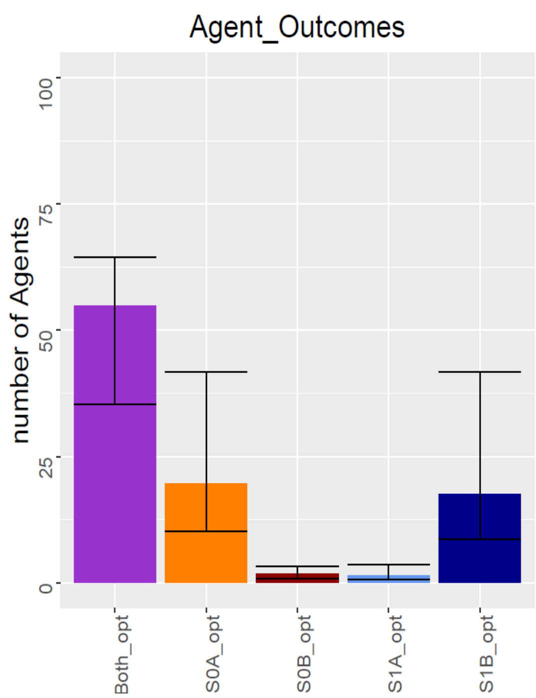
Fully modular



Modular motivation

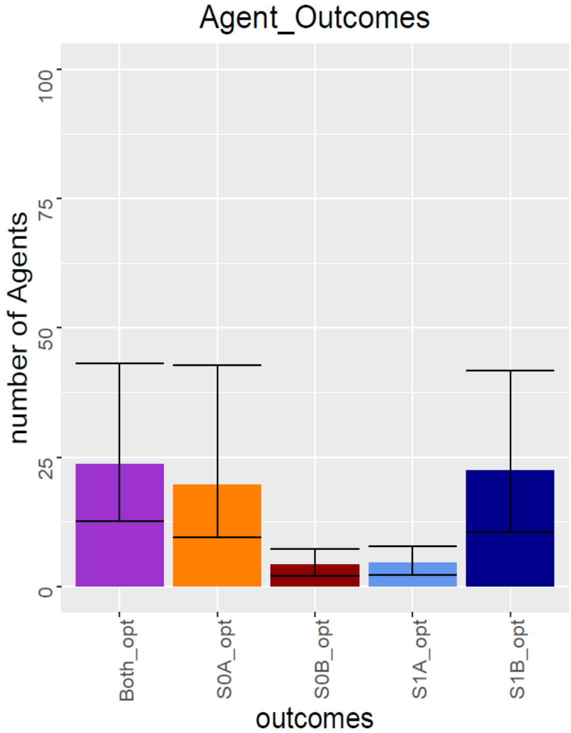


Modular cognition

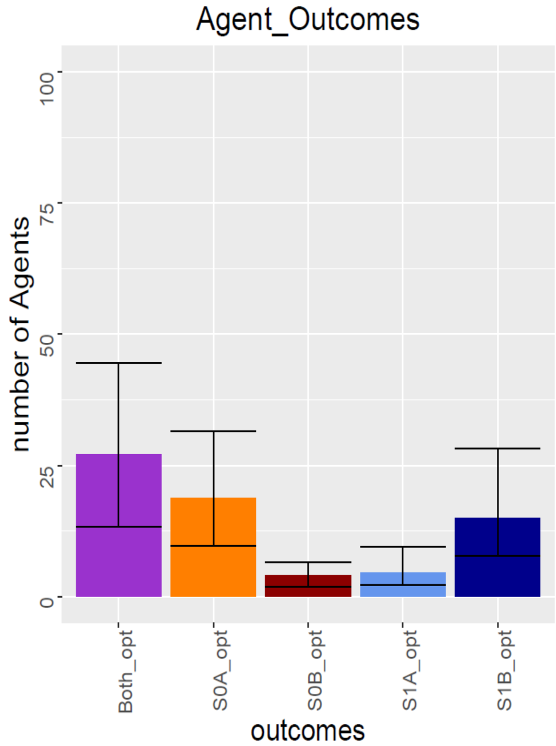


Domain-general

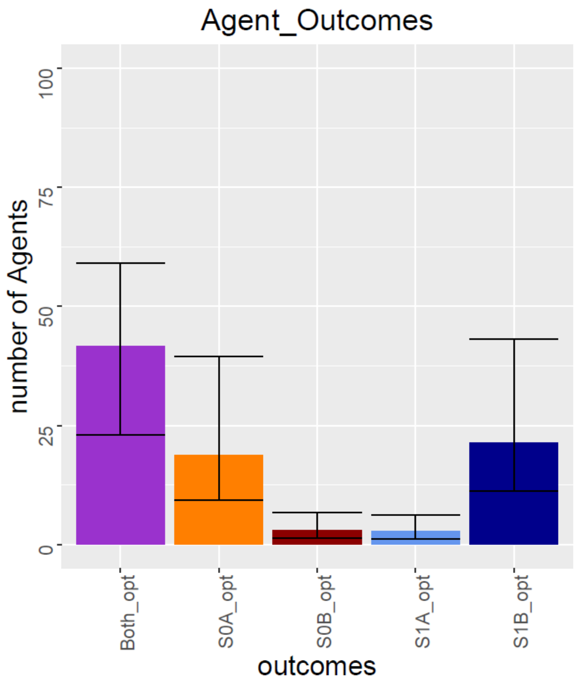
Figure 4ii) p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4



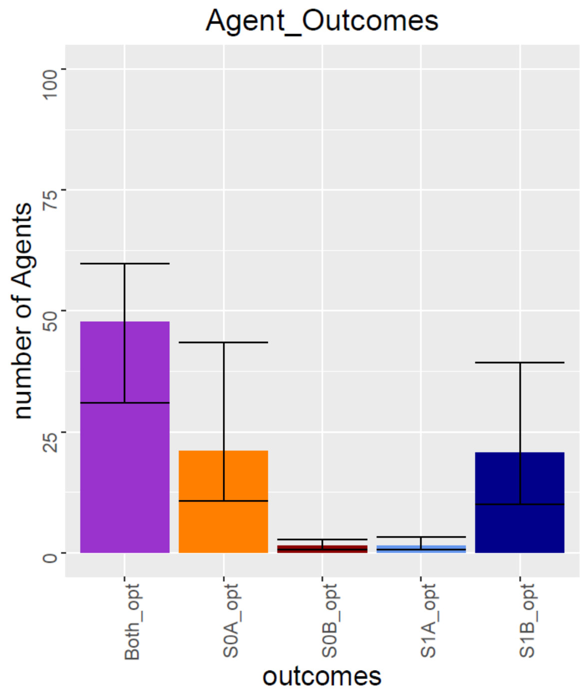
Fully modular



Modular motivation



Modular cognition



Domain-general

Figure 4. The agent outcomes on runs where the agents made decisions in two distinct skill domains. The x axis gives the possible outcomes. Both_opt represents cases where the agent chose the optimal behaviour in both domains, S0A indicates correct answers in state 0 of domain A, S0B in state 0 of domain B, S1A in state 1 of domain A and S1B in state 1 of domain B. Results for the fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents are graphed separately. The graphs display the agents response to i) runs where the priors and fitness tied to matching one's skillset to the environment favoured state 0 in domain A but state 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25) and ii) runs where the most common environment was not necessarily the environment which gave the highest payoff, over two distinct domains (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

3.2: The psychological architecture of the four agent types

I created a series of binned heatmaps to visualise how the agents' cognition and motivation coevolved to drive behaviour (see Figures 5-8). The agent's cognitive thresholds could take any value from - to + ∞. I used the code in appendix 3 to calculate the smallest and largest cognitive thresholds in the final generation of agents, then divided these thresholds into nine equally spaced bins, corresponding to the nine sections of each heatmap in figures 5-8. The agents' α and β motivation thresholds could take any form from [0,1]. I partitioned this interval into ten bins accordingly, {[0, 0.1], (0.1, 0.2], (0.2, 0.3],…(0.9, 1]}. There were two separate bins for α and β motivational thresholds. For every agent, I calculated which cognitive threshold bin they fell into, specifying which of the nine heatmaps to plot their data to. I then calculated which of the motivational threshold bins they fell into. There were 100 bins defined jointly for the α and β values and these corresponded to the exact coordinates of each graph. The α motivational threshold is on the x axis and the β motivational threshold is on the y axis for each heatmap. I did this for all agents across each simulation, and when plotting to the corresponding graph space, I allowed the heatmap colours to track the frequency of agents observed to fall into the same cognitive threshold and motivational threshold bins. Darker colours represent denser patches, with any dark red patches of the graphs approaching 100% of agents in the same psychological architecture space. Meanwhile, lighter patches of the graph show less agents, with faint yellow indicating very few agents have this cognitive and motivational threshold combination. Any white spaces on the graph denotes that none of the agents fit into this exact cognitive and motivational threshold bin combination. On top of plotting the heatmap colours, I also added a black square to each graph. This black square represents the average agent's cognitive threshold and motivational threshold.

Figure 5 shows the psychological architecture of the four agent types on runs where behaviour 0 was the most common and the most fit across both domains (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4; see figure 5). This is the equivalent to runs where both the bow-and-arrow and the spear thrower are rarely needed and provide less calories than mastering a different weapon to hunt larger prey. In these runs, figure 5 reveals that all agents had unbiased cognition between -7 and +7. Here, I use the term 'unbiased cognition' to refer to agents who had cognitive thresholds around 0. As the summary distributions tend to associate negative cue summaries with state 0, and positive cue summaries with state 1, then agents with cognitive thresholds around 0 were likely to perceive the environmental state in each domain without bias. Alternatively, thresholds below 0 made the agent more likely to believe that the state was 1 and thresholds above 0 made the agent more likely to believe that the state was 0. The agents with unbiased cognition (who were likely to accurately perceive the environmental state) had low α and β values. Thus, they were motivated to play behaviour 0 and avoid using both the bow-and-arrow and spear thrower when using these would be inappropriate.

If the agent had a positive cognitive threshold, then their private signal was likely to be below this threshold ($x_A <= T_A$ or $x_B <= T_B$) and thus the agent was likely to believe that the state was 0 (s_A or $s_B = 0$). As the β motivational threshold is used whenever the agent believes that the state was 0, then there was stronger selection acting on this value to be close to 0. As the agents were less likely to believe that the state was 1, then there was a weaker selection acting on α which could take any value (bottom panels of figure 5).

Agents with negative cognitive thresholds were likely to have a private signal that exceeded this threshold ($x_A > T_A$ or $x_B > T_B$; see top panels Figure 5) and so they were likely to believe that the state was 1 (s_A or $s_B = 1$). Note that the agents with negative cognitive thresholds had a *cognitive bias.* I use this term as the agents were likely to develop an erroneous belief about the environmental state in each domain (Efferson et al., 2020). Agents with negative cognitive thresholds had a cognitive bias to believe that the state was 1, despite this run's priors making state 0 more likely. Their motivation compensated for this cognitive bias, however. As there was strong selection acting on α to be 0, then the agent was still unmotivated to use the bow-and-arrow. This compensated for their biased beliefs that the bow-and-arrow was appropriate even when it would be inefficient or dangerous to use. The similar psychological space across all four agent types may account for the similar behavioural outcomes seen in the bar charts of section 3.1.

Now I turn to cases where the agents decide over two similar skill domains, but the most common environment is not the environment which would provide the highest payoff (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4; figure 6). For example, it is common for the agent to be able to use her bow-and-arrow to hunt the small but readily available deer, but the less common target (and a richer source of calories) would be to use a close-range weapon to hunt larger prey such as mammoth. Interestingly, when the priors and the fitness tied to matching one's skillset to the environment favour matching in different states, then all agent types seemingly converged on unbiased cognitive thresholds. Thus, they were likely to perceive the state of the environment accurately.

Fully modular agents, and agents with modular motivation only, had a similar psychological architecture by the final generation of these runs. Agents with negative cognitive thresholds – who were likely to believe that the state was 1 (s_A or s_B = 1) – had a strong selection acting on α to be low to mid, while β took more of a range of values (bottom panels of figures 6i and 6ii). The agents with unbiased cognitive thresholds (middle panels of figures 6i and 6ii) had a strong selection on β to be low, though α took a range of values. The average agent had an α of 0.6 and an average β of 0.2, across both domains. When the agents believed that the state was 0, then they were highly motivated to play behaviour 0. This was the rarer skill with the higher payoffs, and so perhaps it makes sense that the agents were highly-motivated to exploit this skillset whenever it was an option. When the agent believed that the state was 1, then their motivation to play behaviour 1 was weaker. Perhaps this makes sense, as if deer are common, we do not have to be highly-motivated to hunt them with a bow-and-arrow or spear thrower. There will be other opportunities to hunt.

Agents with modular cognition only had a similar psychological architecture by the final generation, but with a tighter cluster on a high α and a low β value. These agents were thus motivated to use the bow-and-arrow or spear thrower when they were likely to encounter the commonly-available deer, though they were also motivated to avoid the bow-and-arrow when given the opportunity to pursue the rarer but larger animals.

Finally, domain-general agents with negative cognitive thresholds had strong selection acting on α to be zero, though β took any value (top panels of figure 6iv). These agents are biased to believe that the state was 1, and thus the lower α value may compensate and ensure that the agents do not miss an opportunity to hunt the larger but rarer animals. Domain-general agents with unbiased thresholds had a high α and low β cluster like the other agent types (middle panels of figure 6iv). Finally, agents with a cognitive bias to believe that the state was 0 (as they had positive cognitive thresholds; see bottom panels of figure 6iv) had a strong selection acting on β to be zero but α could take any value. Overall, the domain-general agents are maximising the chances that they will avoid the bow-and-arrow or spear thrower, in preference of hunting the larger but rarer animals. All the agents are leaning towards mastering the skillset which is rarer but has the higher payoff, rather than mastering the skillset which is common but has a lower payoff.

Figure 5i: Fully modular, domain A. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

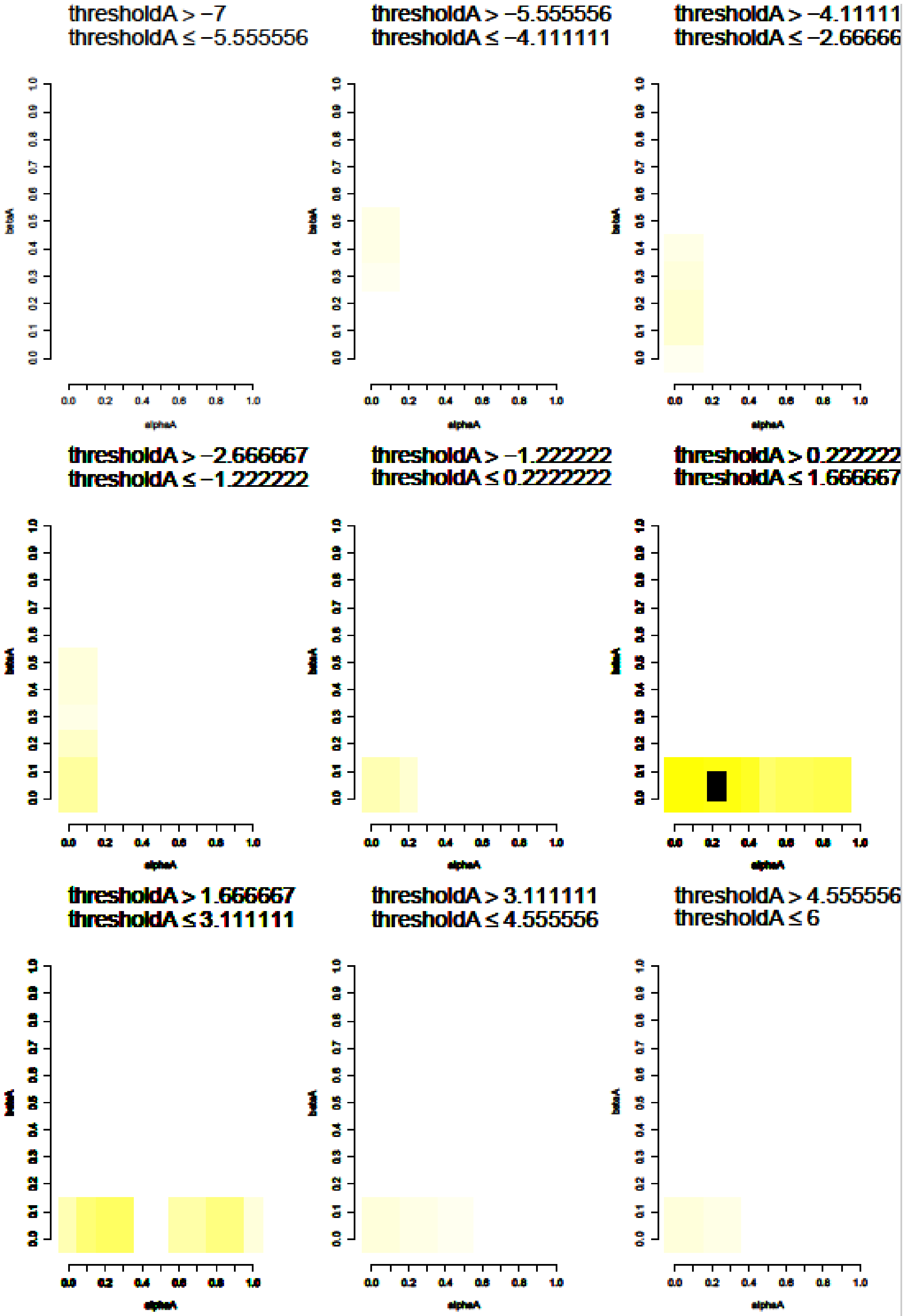


Figure 5i: Fully modular, domain B. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

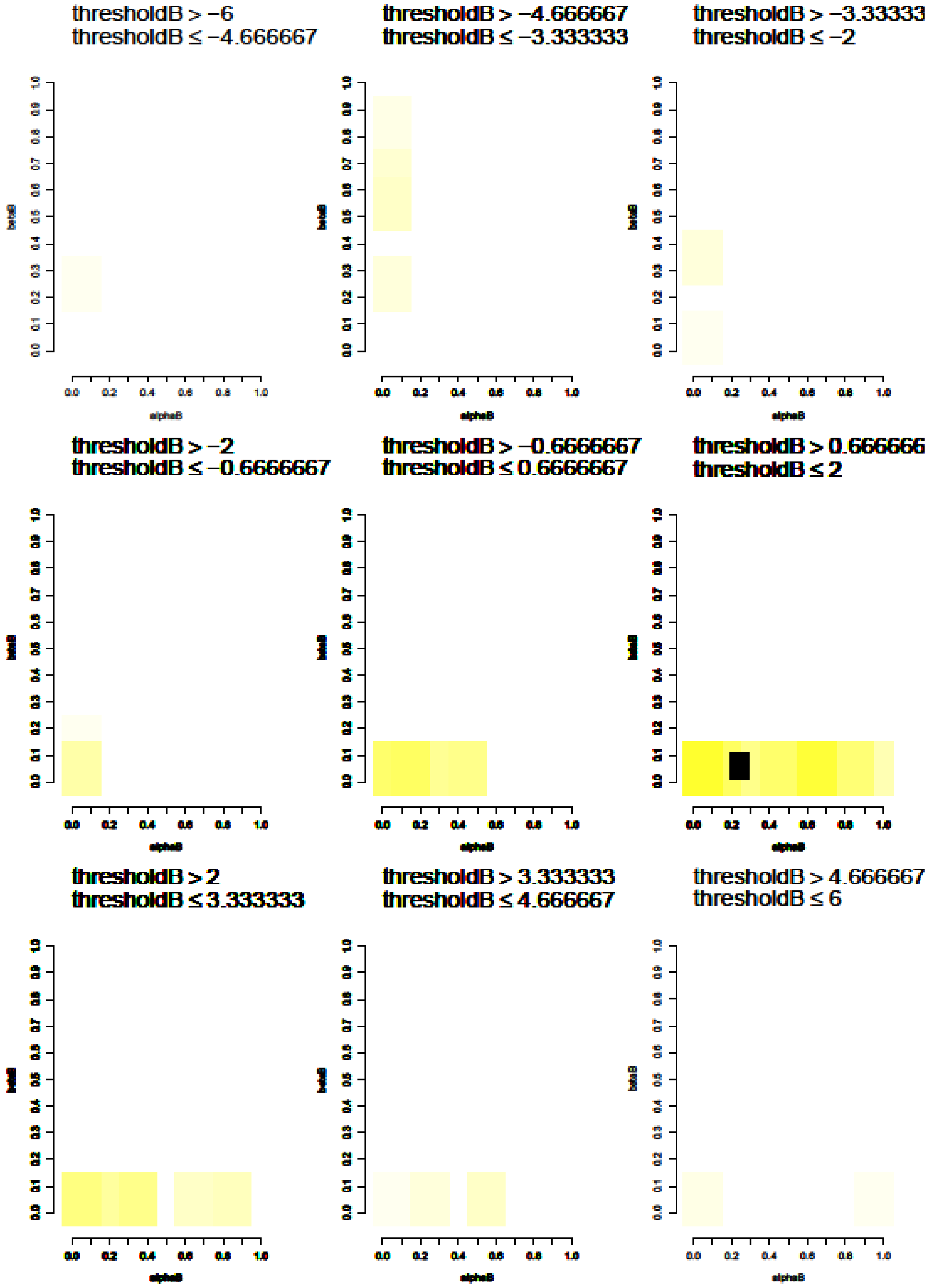


Figure 5ii: Modular motivation, domain A. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

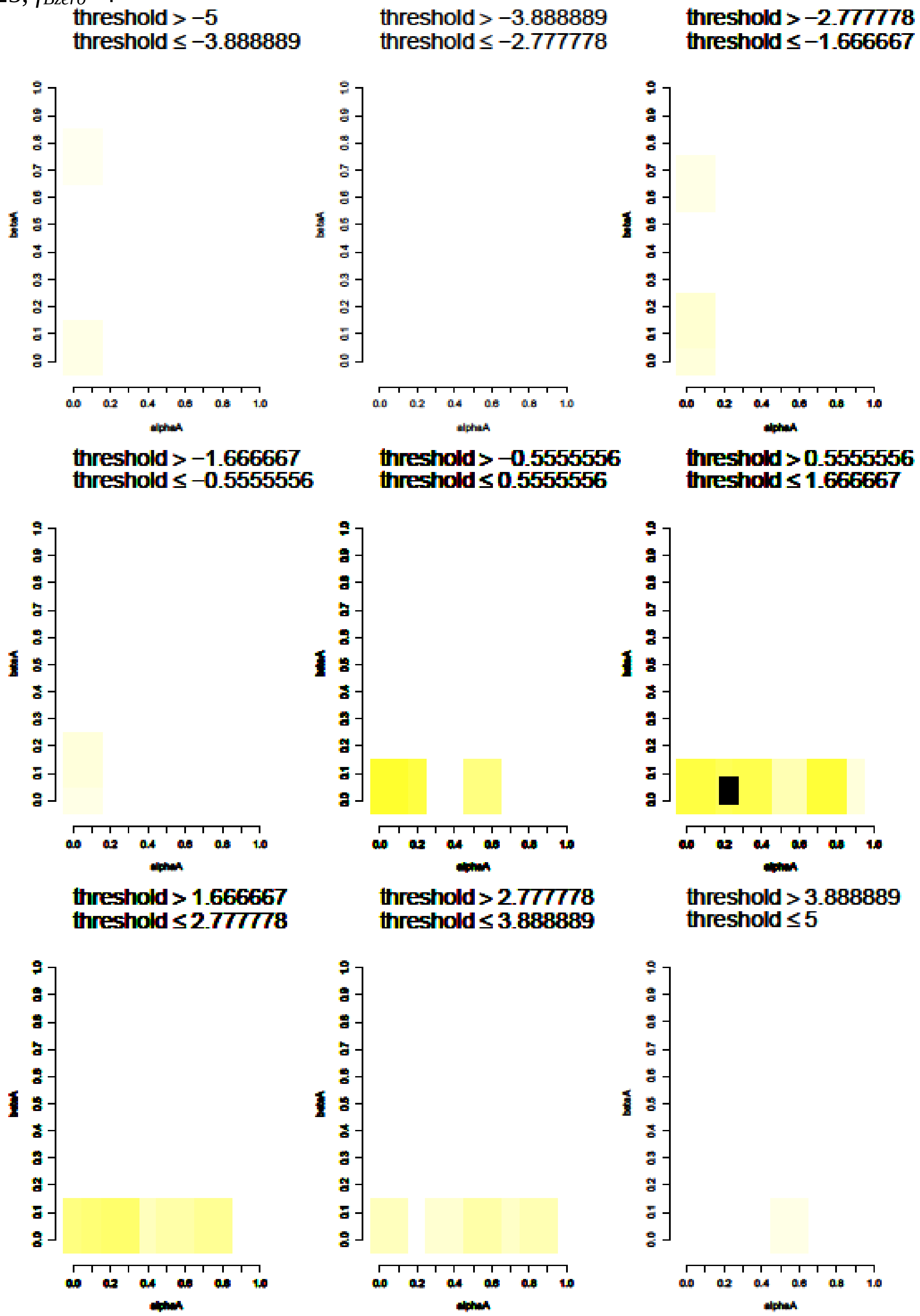


Figure 5ii: Modular motivation, domain B. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

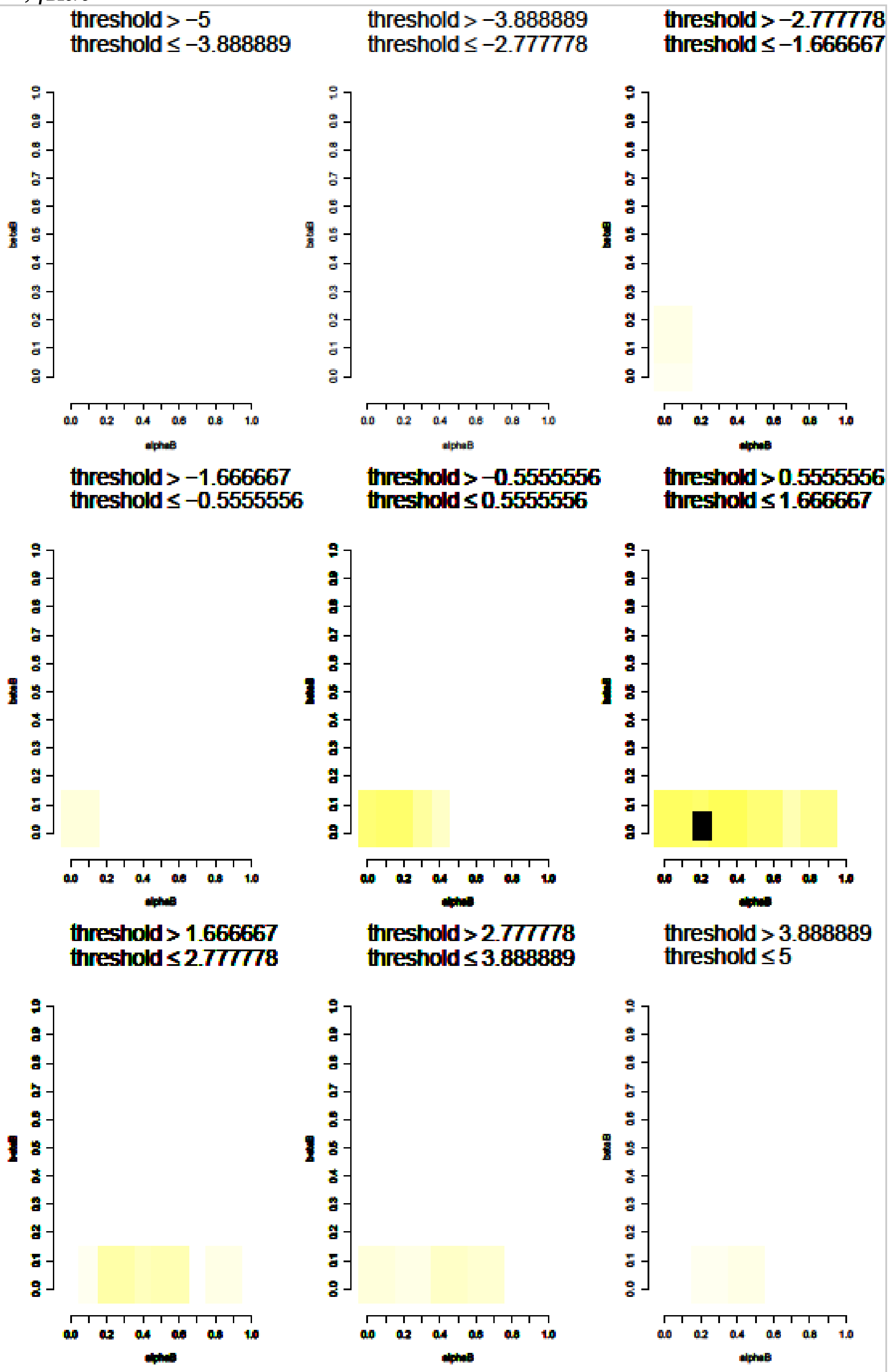


Figure 5iii: Modular cognition, domain A. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

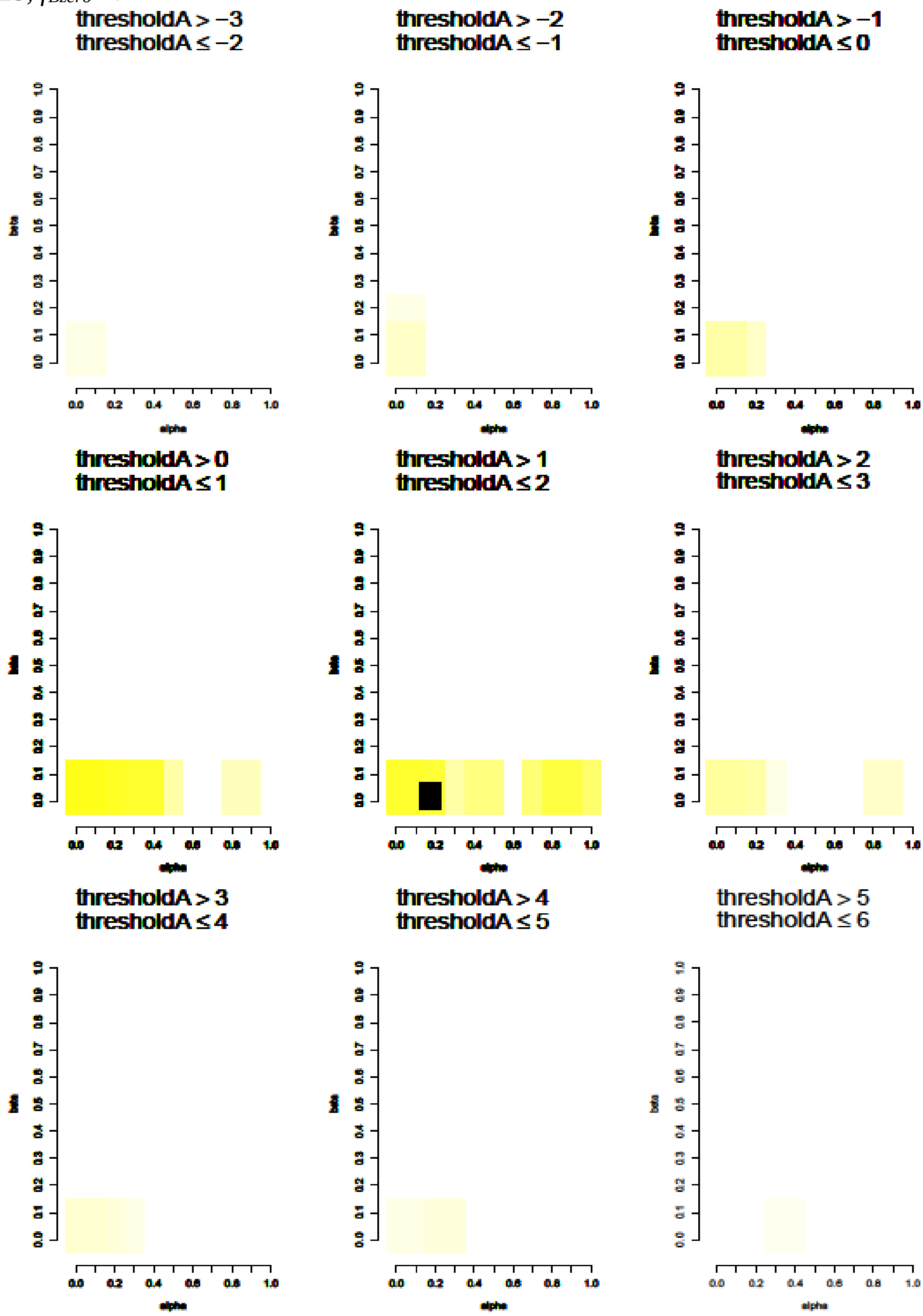


Figure 5iii: Modular cognition, domain B. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

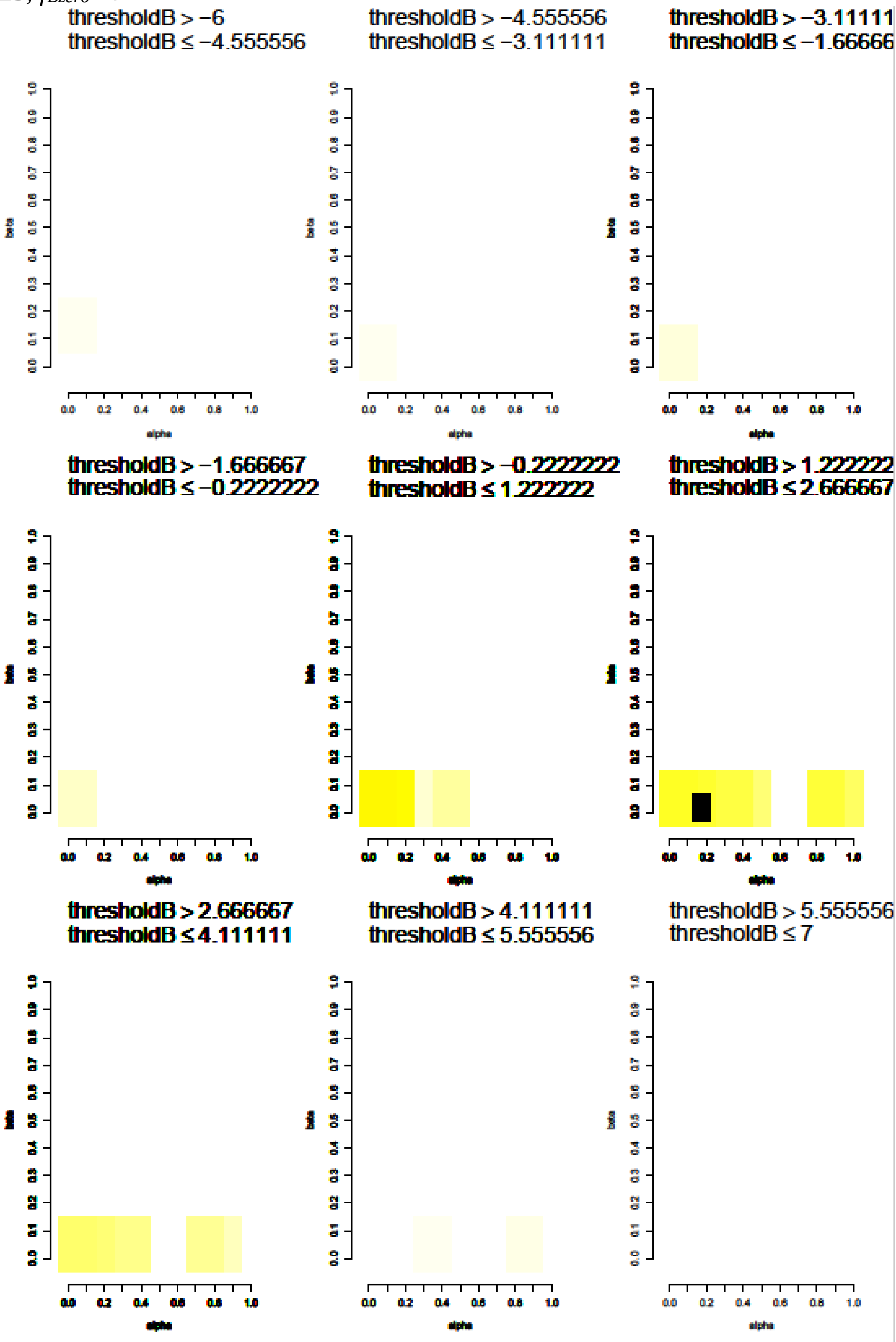


Figure 5iv: Domain-general. p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

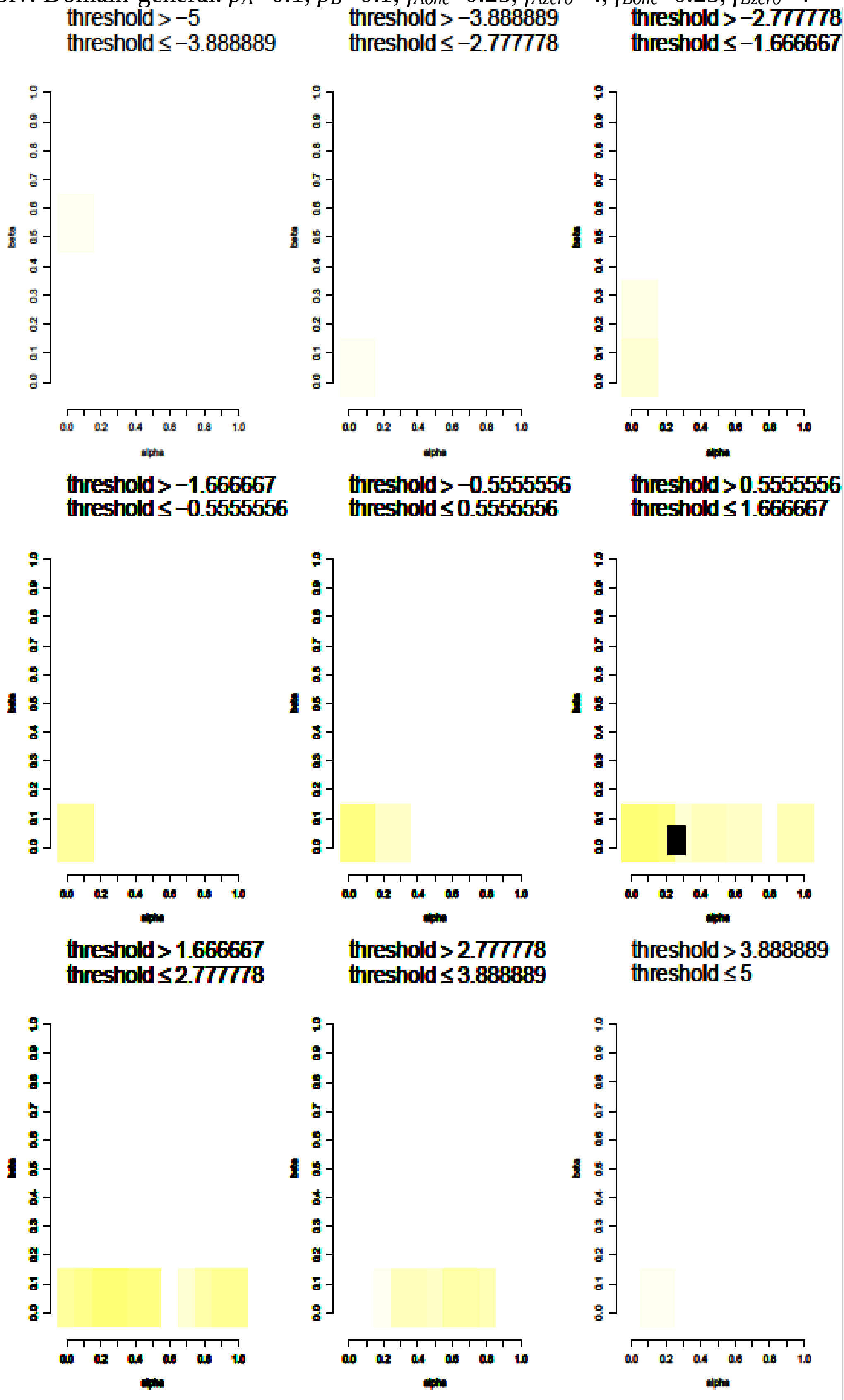


Figure 6i: Fully modular, domain A. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

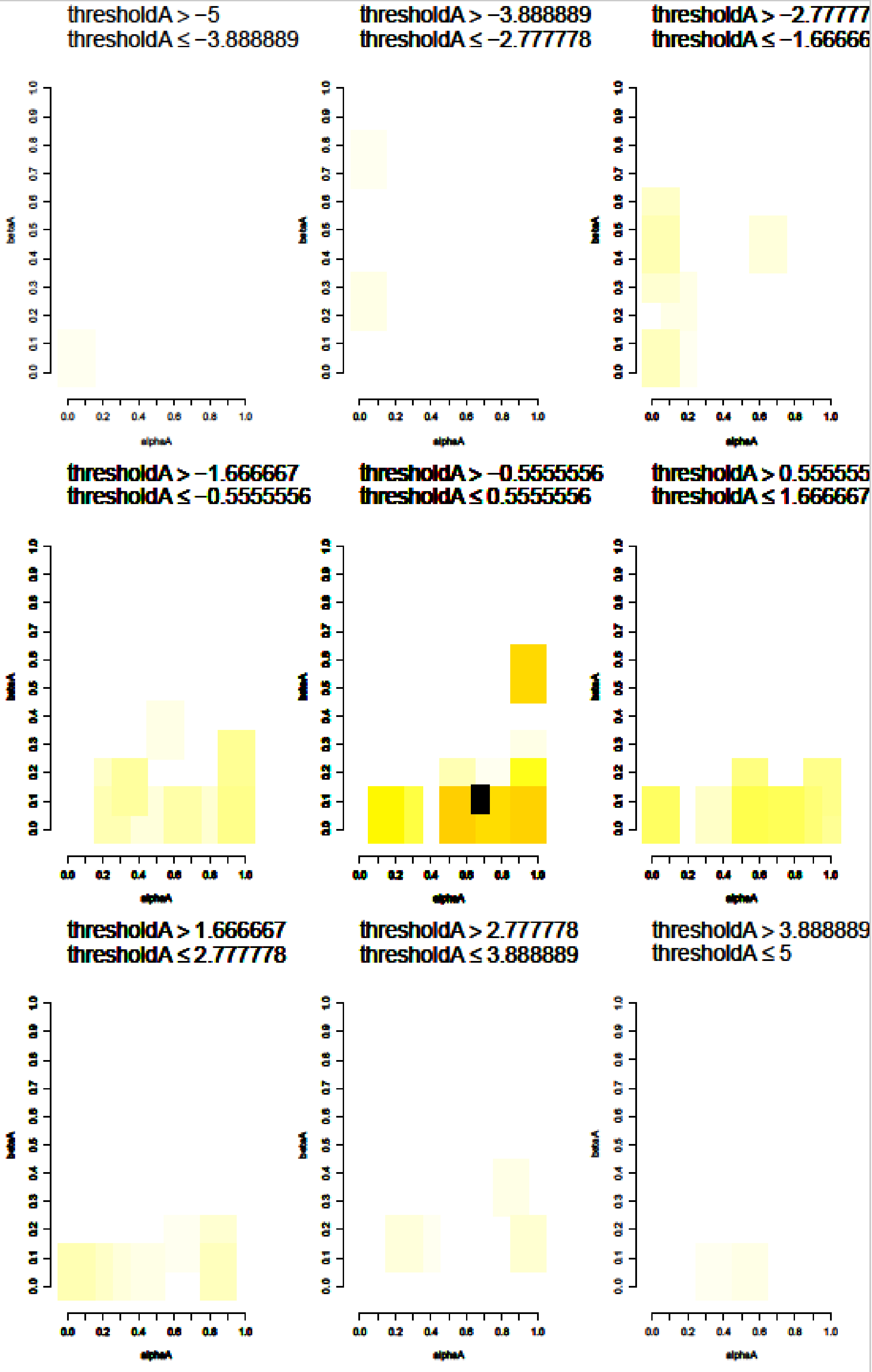


Figure 6i: Fully modular, domain B. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

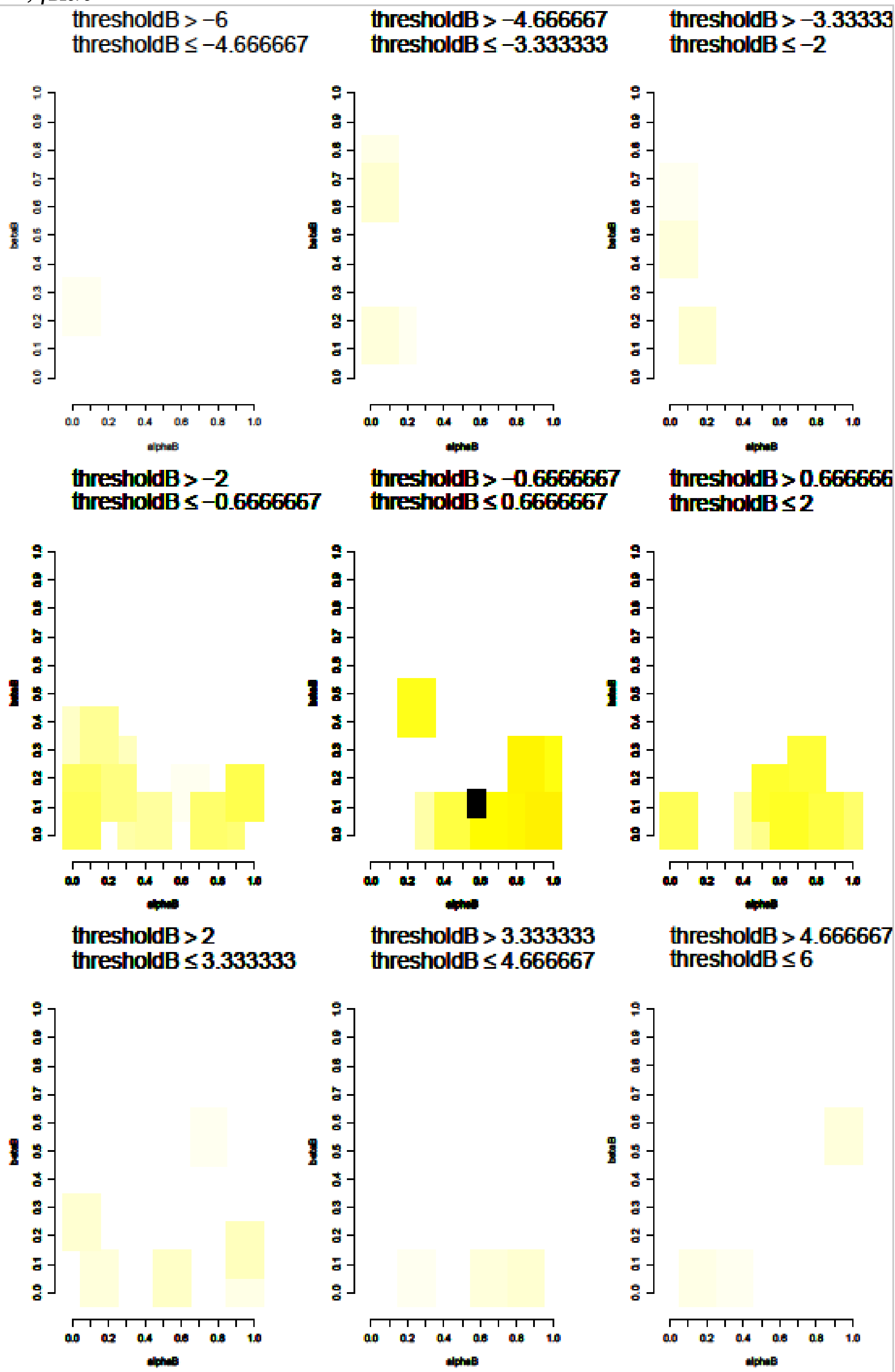


Figure 6ii: Modular motivation, domain A. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

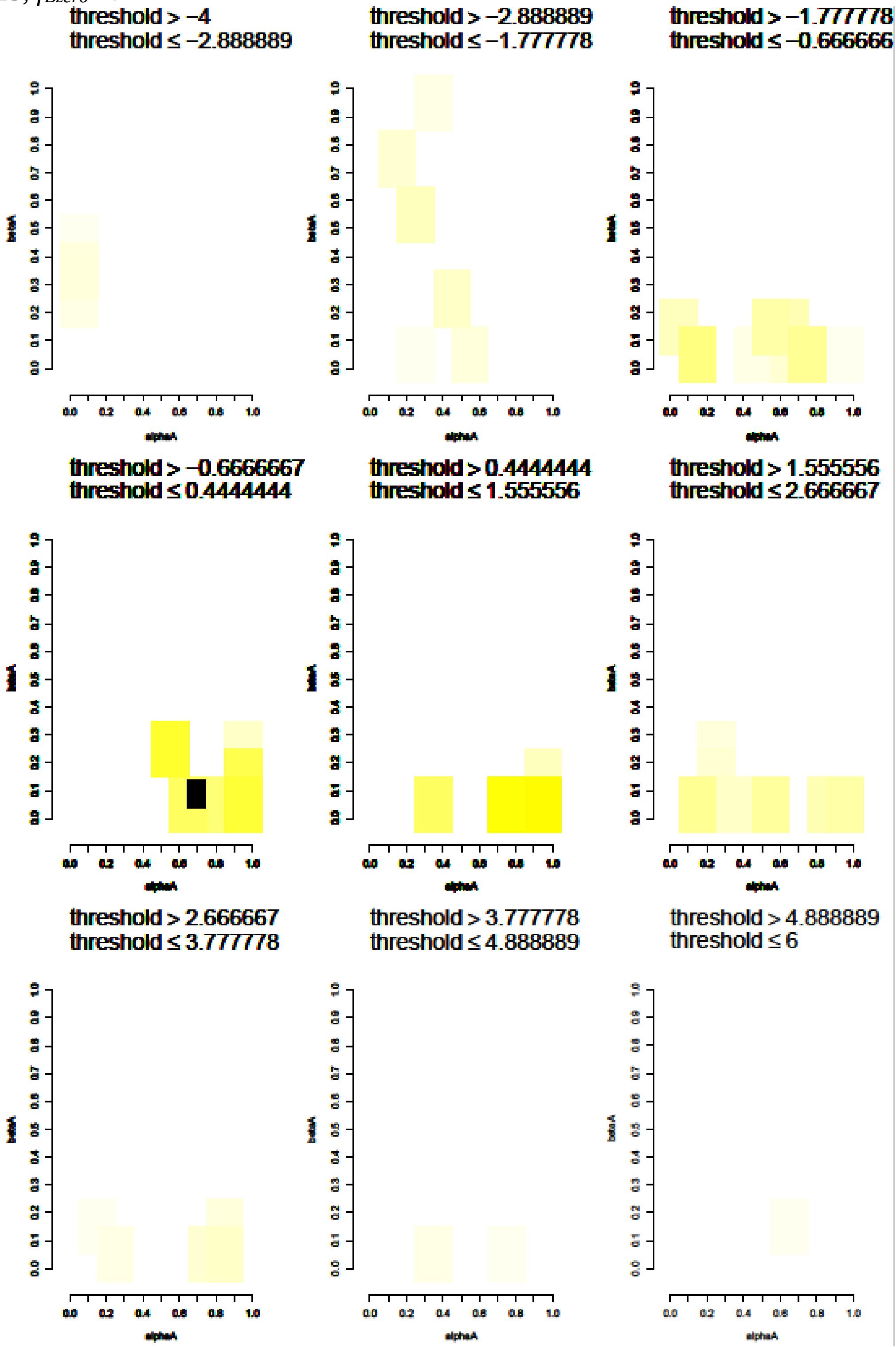


Figure 6ii: Modular motivation, domain B. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

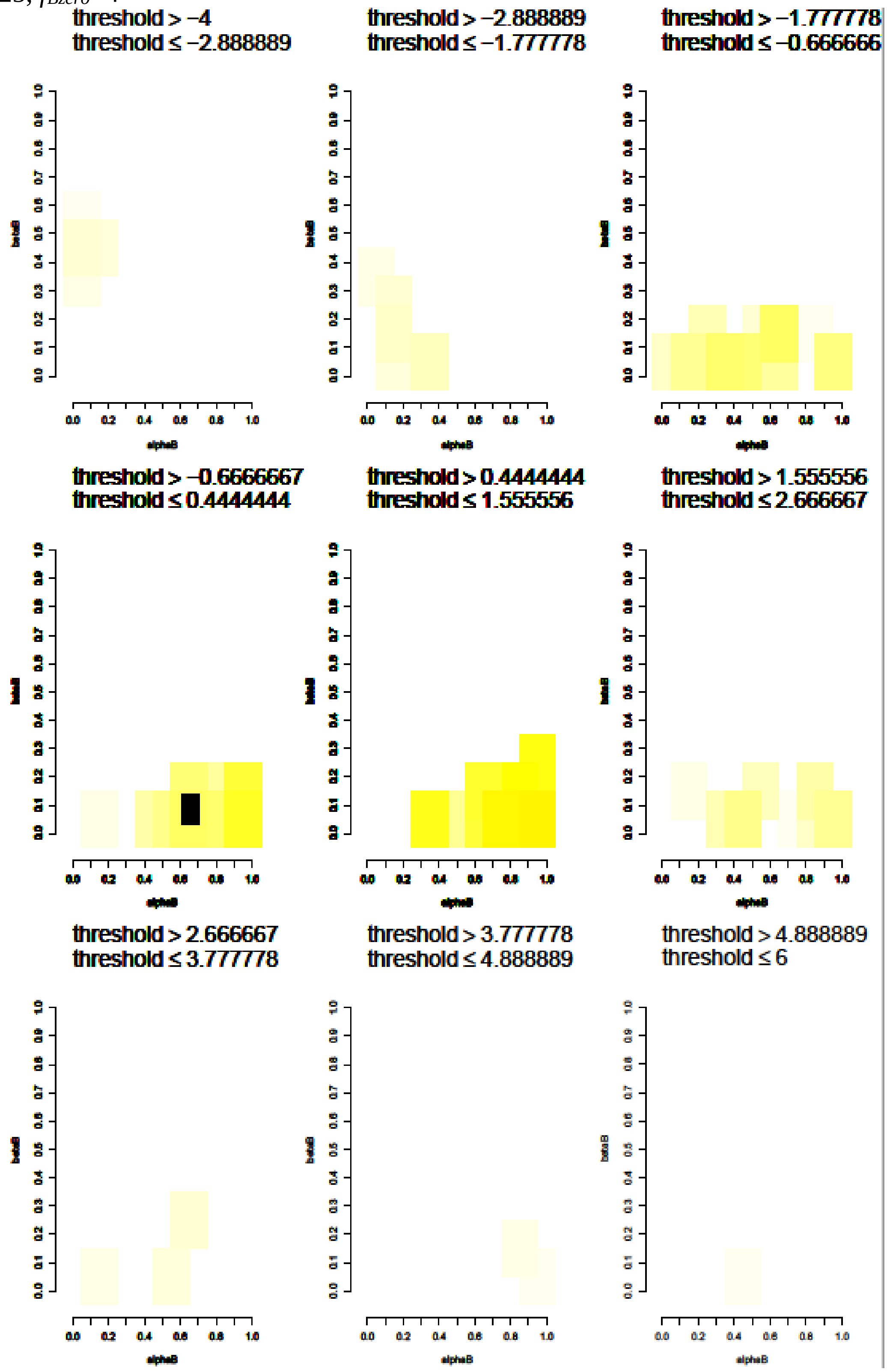


Figure 6iii: Modular cognition, domain A. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

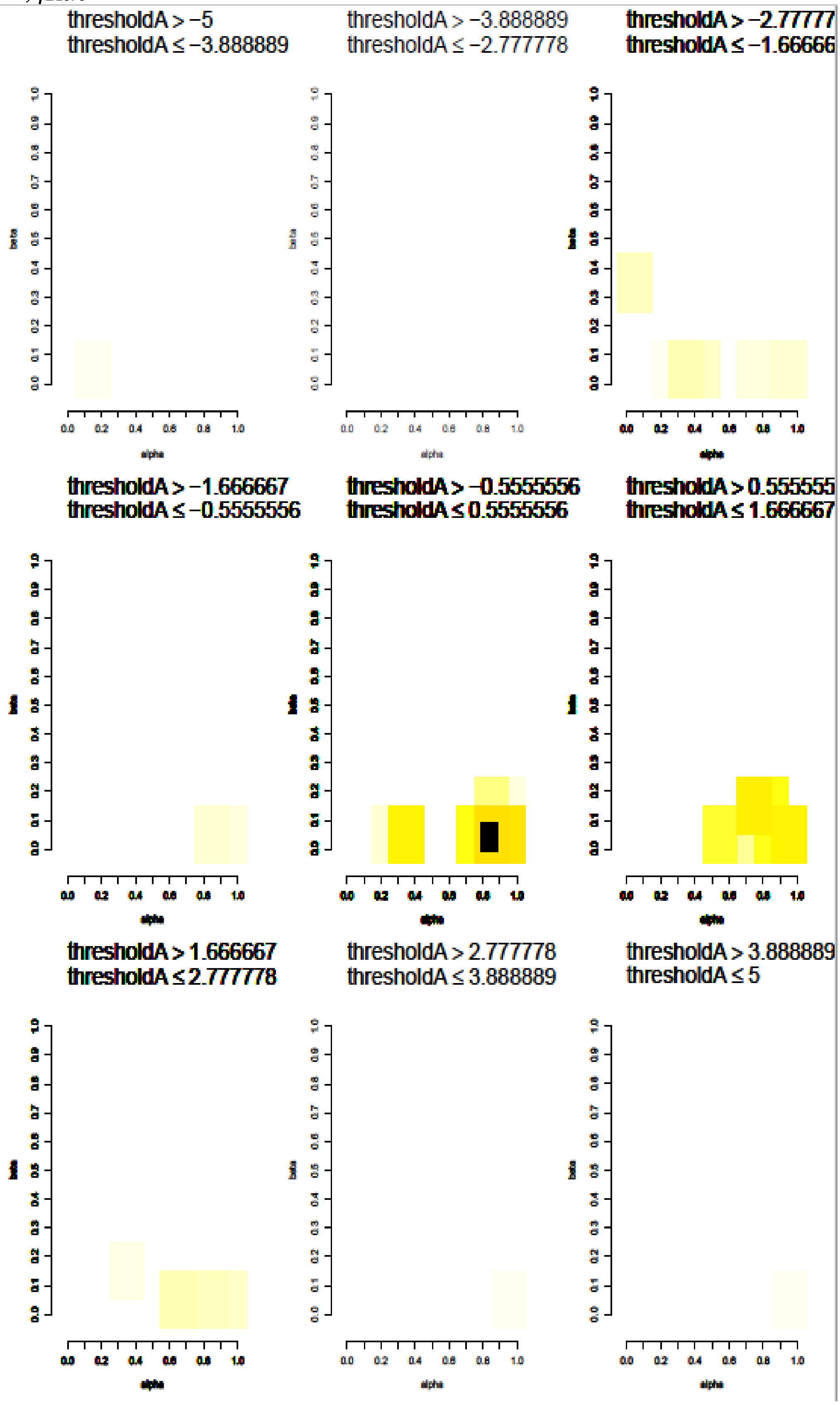


Figure 6iii: Modular cognition, domain B. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

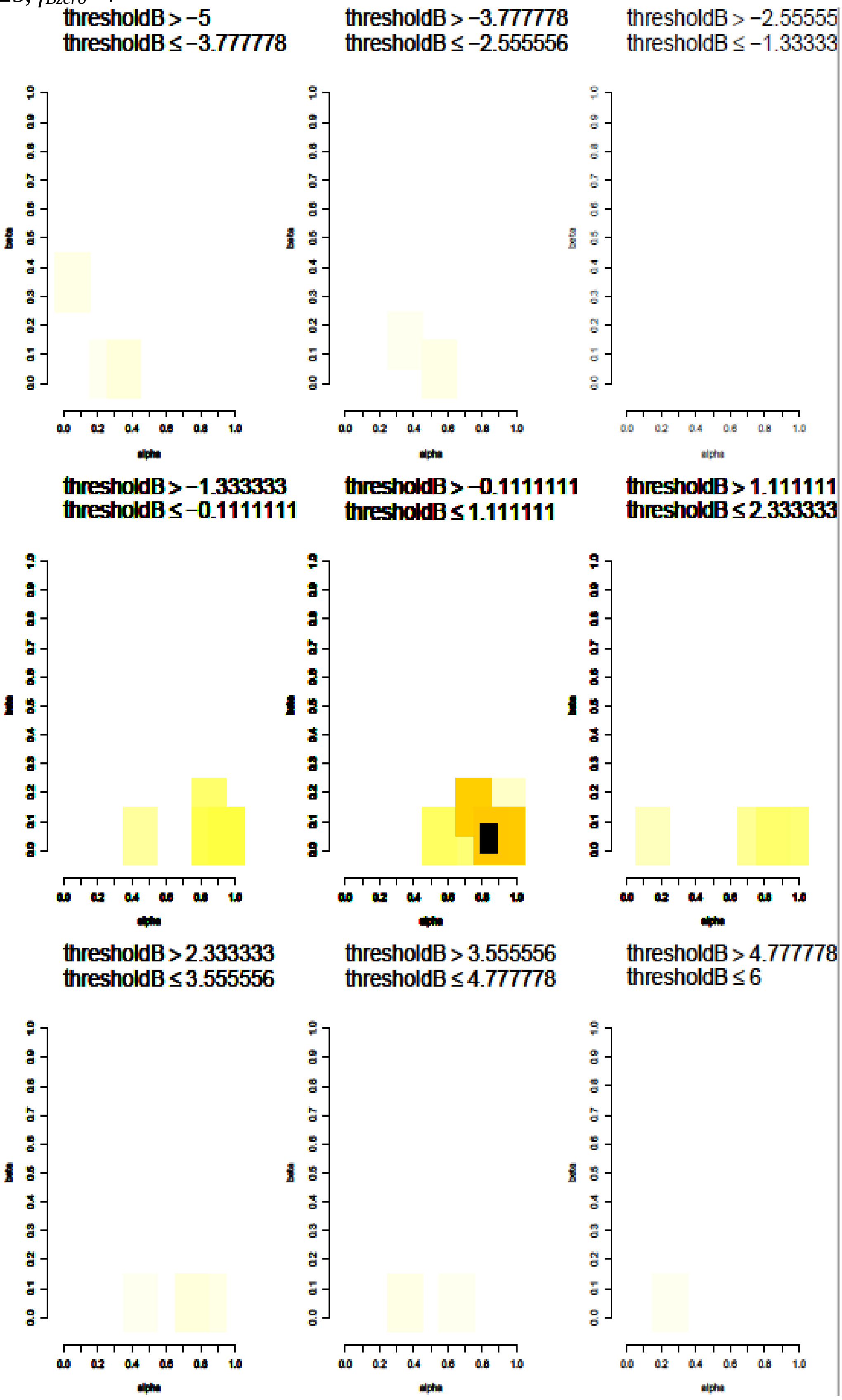


Figure 6iv: Domain-general. p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4

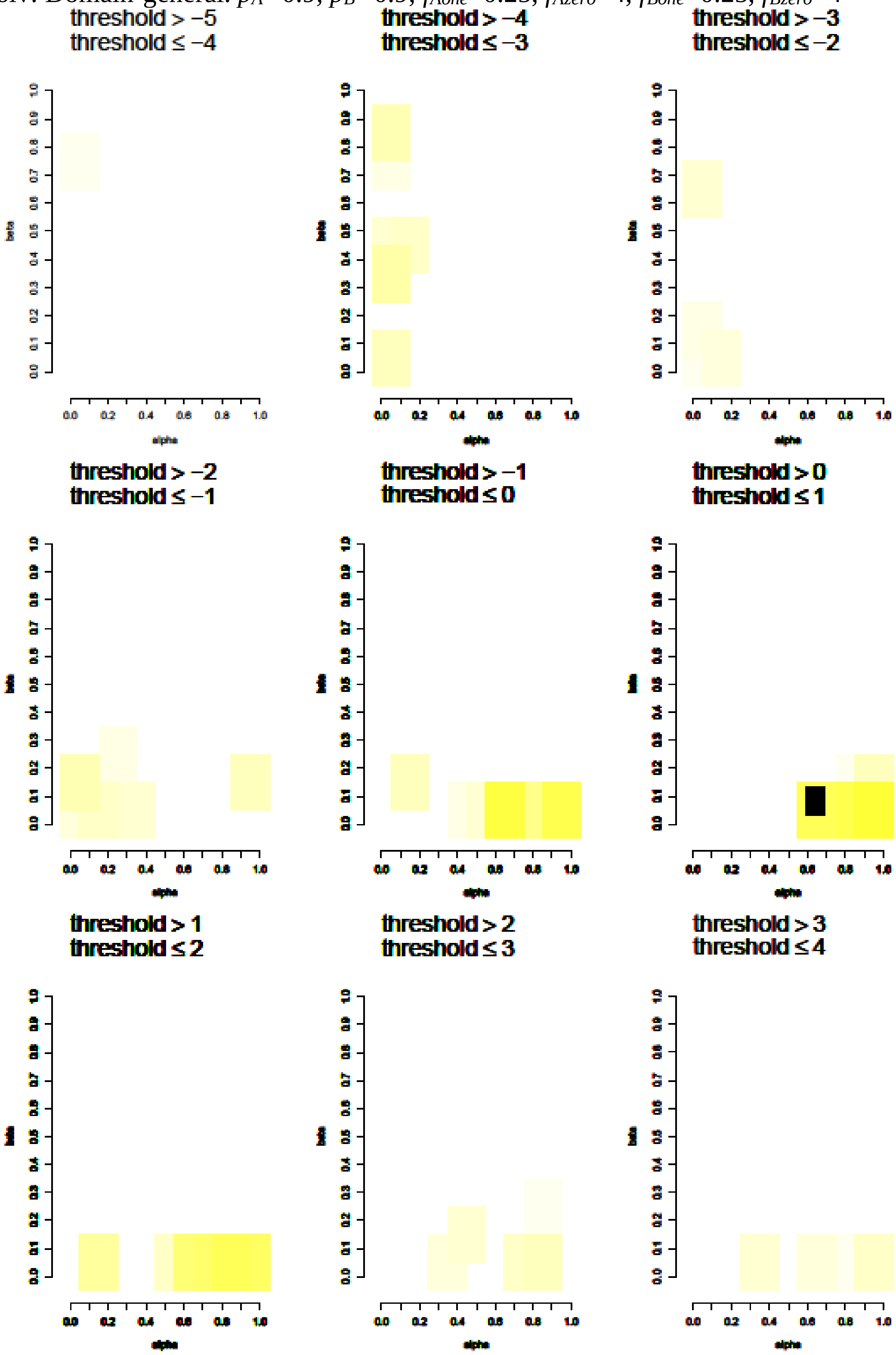


Figure 5 and Figure 6. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) partly modular agents with modular motivation but domain-general cognition; (iii) partly modular agents with modular cognition but domain-general motivation and (iv) fully domain-general agents when deciding in two similar skill domains. Figure 5 gives the results for runs where both the priors and the fitness tied to matching one's skillsets to the environment favour behaviour 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). Figure 6 gives the results for runs where the priors favour state 1 but the fitness tied to matching one's skillsets to the environment favour displaying behaviour 0 (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4).

Next, I consider runs where the agents mastered skillsets over two distinct domains, such as when the bow-and-arrow would be inappropriate and the agents needed to cook their food (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). The fully modular agents and the agents with modular motivation only had a similar psychological space in domain A as they did to the runs where state 0 was favoured in figures 5i and 5ii. That is, the agents were motivated to play behaviour 0 regardless of their cognitive thresholds. The agents were thus motivated to avoid using the bow-and-arrow when this is rarely needed and likely suboptimal.

They had a different psychological space for domain B. Agents with unbiased or negative cognitive thresholds were likely to believe that the state was 1 (s_B=1). In this case, they had a strong selection on α_B to be 1, though the selection acting on β_B was weaker and this threshold could take a range of values. Thus, the agent was motivated to play behaviour 1 and cook their food when this was both common and fit to do so. Agents with positive cognitive thresholds had a cognitive bias to believe that the state was 0 (s_B=0). Despite this, their motivation compensated for this bias as both their α_B and β_B thresholds evolved to be close to 1. Even if the modular agents had a cognitive bias to believe that they did not have to cook their food, they were still highly motivated to cook and so could avoid the negative health consequences associated with eating raw food (Wrangham, 2009). In sum, these modular agents evolved different cognitive and motivational thresholds across domains A and B in order to respond to the different priors and fitness pressures across these domains.

The partly modular agents with modular cognition only could not change their motivational thresholds to the conflicting demands of domains A and B, as they only had one α and β value. In this case, the agents with modular cognition only had a high α and a low β. They were motivated to play behaviour 1 whenever they believed that the state was 1 and were motivated to play behaviour 0 whenever they believed that the state was 0. Interestingly, there were larger shifts in these partly modular agents' cognitive thresholds. As modular cognition was the only modular component that this agent had, these thresholds shifted significantly to meet the contrasting demands of both domains. In domain A, all agents had positive cognitive thresholds (T_A = 0-6, Mμ = 2). This meant that they would believe that the state was 0 when this was likely to be common in domain A. In this case, the low β value allowed them to play behaviour 0 and avoid using the bow-and-arrow when this is less fit and rarely needed. In domain B, these agents had negative cognitive thresholds (T_B =-6 - 0, Mμ = -2). This meant that they believed that the state was 1 when this was likely to be common in domain B. In this case, their high α value allowed them to play behaviour 1. They would cook food when it was both fit to do so and commonly-needed. Thus, these partly modular agents could also tailor their behaviour to the contrasting pressures of both domains (see figure 4i, section 3.1 for behavioural outcomes), though the partly modular agents with modular cognition achieved this via the evolutionary trajectory of their cognitive thresholds while fully modular agents and agents with modular motivation only chiefly achieved this via the evolutionary trajectory of their motivational thresholds.

The domain-general agent also had a high α and a low β, though their cognitive thresholds could not specialise to each domain as they only had one threshold, *T*. They had unbiased thresholds by the final generation (see figure 7iv), and so they tried to play behaviour 1 whenever they believed that the state was 1 or played behaviour 0 whenever they thought that the state was 0. However, this strategy only paid off half the time (see figure 4i, section 3.1). The other half of agents only answered one domain correctly or failed to answer optimally in either domain.

Finally, I consider runs where the agents made decisions in two distinct domains where the most common environmental state was not necessarily the one which provided the highest payoff to the agents for matching their skillsets to (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4; see figure 8). To illustrate, it is common that the agents could lay traps for small animals in domain A, though the rarer but larger prize would be to use the bow-and-arrow to hunt deer. It is common that the agent needs to cook her food in domain B, though on rare occasions there is a higher payoff to avoid overcooking food.

In domain A, the fully modular agents typically had unbiased cognitive thresholds with a high α_A though β_A could take a range of values (see figure 8i). The average α_A was 0.9 and average β_A was 0.4. The fully modular agents were strongly motivated to play behaviour 1 when they believed that the state was 1 in domain A, and were weakly motivated to play behaviour 0 when they believed that the state was 0 in domain A. The fully modular agents were strongly motivated to use a bow-and-arrow to hunt deer when these were rare but a larger source of food. They were less motivated to build traps to catch the small but commonly available prey, perhaps because there would be plenty of opportunities to lay traps again.

The fully modular agents in domain B had a different strategy, however. The agents with unbiased cognition or negative cognitive thresholds– who would believe that the state was 1– had a high α_B though their β_B threshold took a range of values. Agents with positive cognitive thresholds– who would believe that the state was 0– had a low α_B and a high β_B (bottom panels of domain B in figure 8i). These agents were thus likely to play behaviour 1, regardless of their cognitive beliefs. While the fully modular agents tried to use the bow-and-arrow whenever it was rare but had higher payoffs in domain A, they instead merely tried to do the common behaviour in domain B. They always cooked their food as this was commonly needed, despite the fact that undercooking would have higher fitness payoffs on rare occasions.

The partly modular agents with modular motivation only had a similar psychological architecture to the fully modular agents in domain A (see figure 8ii). Both agent types tried to use the bow-and-arrow to hunt the rare deer when this was the option that gave the highest fitness payoffs. Their response to domain B differed, however (see figure 8ii). These partly modular agents had an average α_B of 0.7 and an average β_B of 0.1. They were only somewhat motivated to play behaviour 1 when they believed that the state was 1 in domain B, though they were strongly motivated to play behaviour 0 when they believed tht the state was 0 in domain B. They were somewhat motivated to cook, as cooking is commonly necessary. The agents with modular motivation only were strongly motivated to avoid cooking in the rarer cases where overcooking was dangerous, however. Thus, the fully modular agents would cook in domain B simply as this was commonly needed. The agents with modular motivation only instead tried to avoid overcooking when this was rare but had higher fitness payoffs.

Perhaps both phenotypes are selected for. If it does not matter too much whether the agent has the commonly-needed skill but with lower payoffs, or the rarely-needed skills with higher payoffs, then perhaps a two-strategy equilibrium will exist where the final generation of agents consist of individuals within either psychological space. This would be evidence of psychological polymorphism. By allowing cognition and motivation to coevolve then multiple behaviours may be selected for (Laland, 1993).

To visualise whether psychological polymorphism could have affected the results, I simply adapt the code to create the binned heatmaps in appendix 3 to plot the final generation from one simulation at a time. While different simulations ended up with slightly different psychological architecture, I found no evidence for multiple equilibria. There were no polymorphic states by the final generation of any run. Rather than polymorphism explaining the varied psychological spaces reported in Figure 8, it could instead be that selection pushes the agents' psychological space towards a certain subset of cognitive and motivational thresholds. As the agents within these psychological subsets typically perform equally well in terms of fitness, then perhaps drift will have more of an effect on the exact psychological architecture present in the final generation. Whether the final generation adapted the common skill with lower payoffs, or the rarer skill with higher payoffs, may be down to drift. Drift describes the effect of random sampling errors over the course of evolution (Rorabaugh, 2014). Drift may have acted over these generations to produce two different phenotypes, though both may have been plausibly selected for.

When the priors and the fitness pressures clashed over two distinct domains, the partly modular agents with modular cognition did not experience any great shifts in their cognitive thresholds. Instead, the average agent has unbiased cognition (see figure 8iii). They had a high α and a low β cluster, so they tried to play behaviour 1 whenever they believed that the state was 1 and tried to play behaviour 0 whenever they believed that the state was 0. These agents merely hedged their bets to try to match their behaviour to the environment. The domain-general agents had a similar psychological architecture in the final run, though with more noise (see figure 8iv). When there were contrasting fitness values and prior probabilities across two domains, then the domain-general agents may have struggled to balance their psychological architecture to the contrasting demands of both domains.

Figure 7i: Fully modular, domain A. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

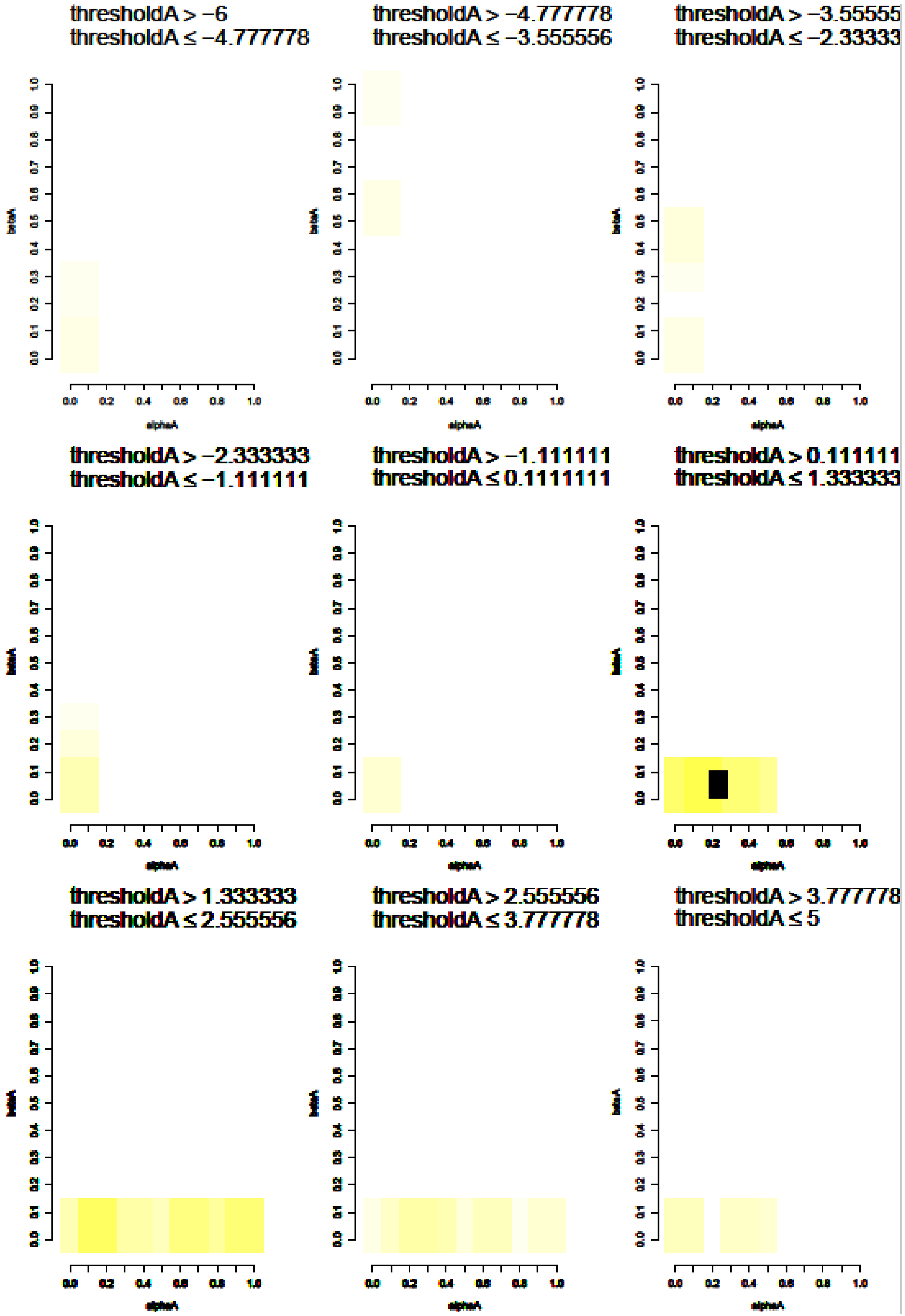


Figure 7i: Fully modular, domain B. $p_A=0.1$, $p_B=0.9$; $f_{Aone}=0.25$, $f_{Azero}=4$, $f_{Bone}=4$, $f_{Bzero}=0.25$

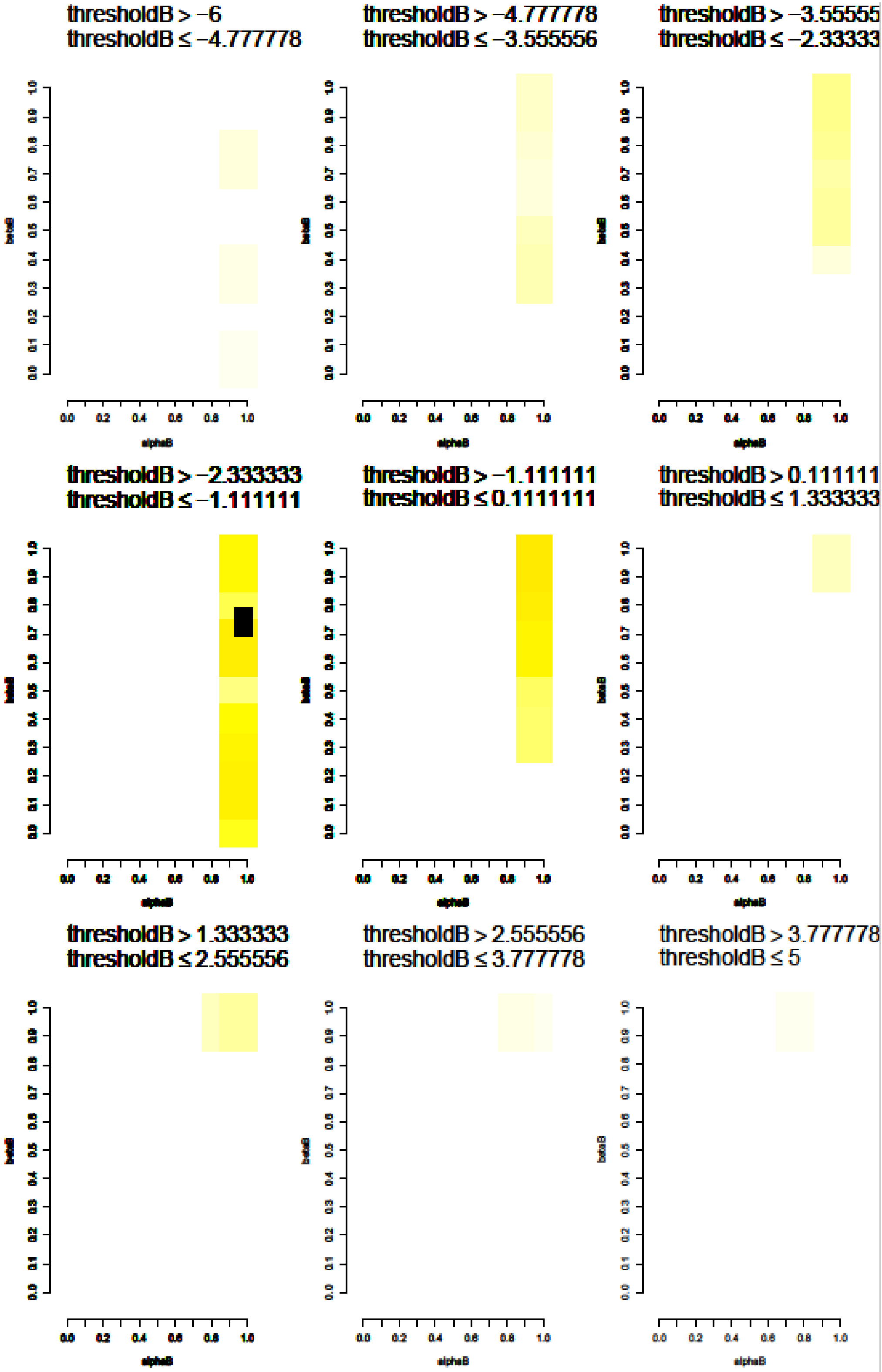


Figure 7ii: Modular, motivation, domain A. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

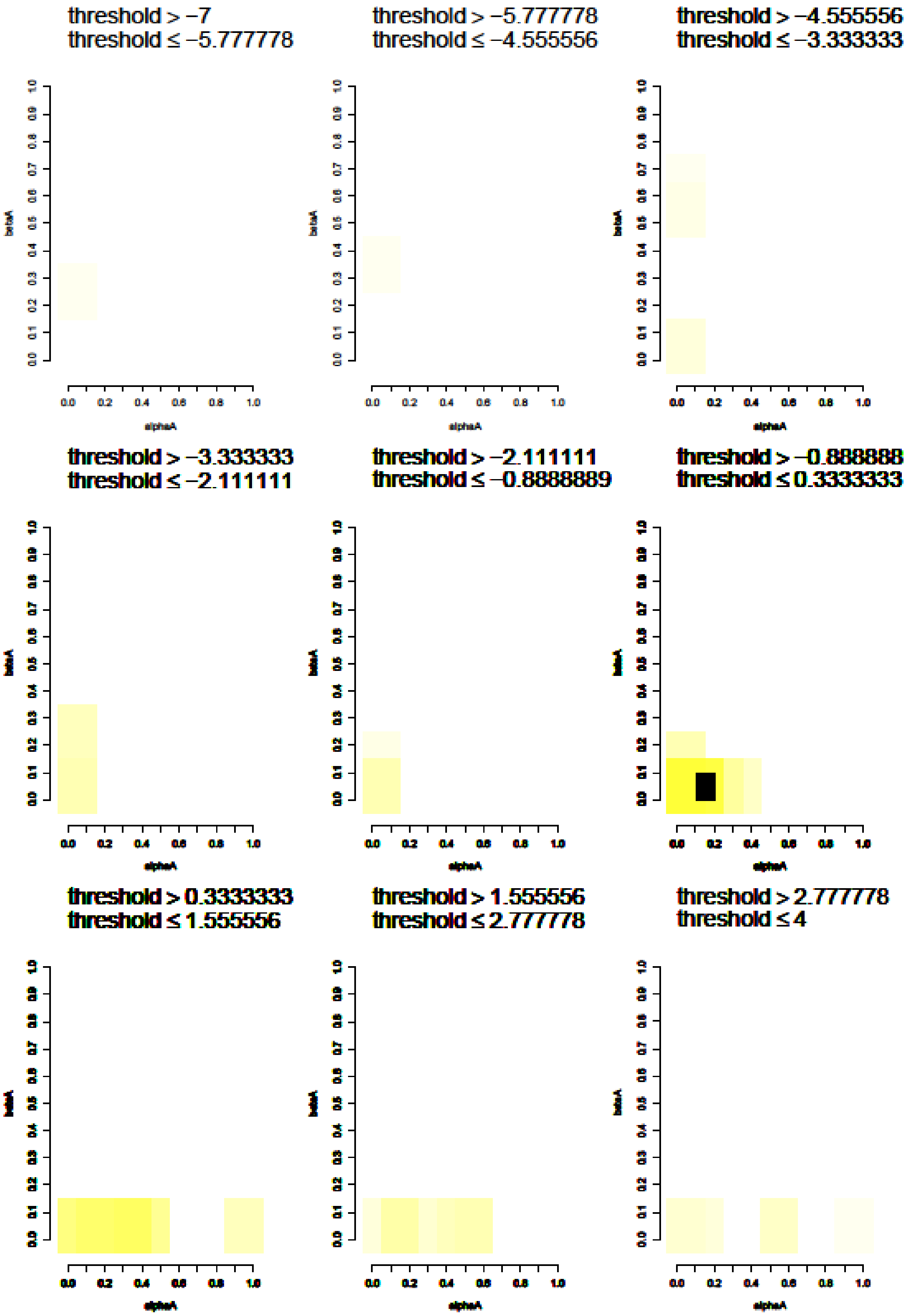


Figure 7ii: Modular, motivation, domain B. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

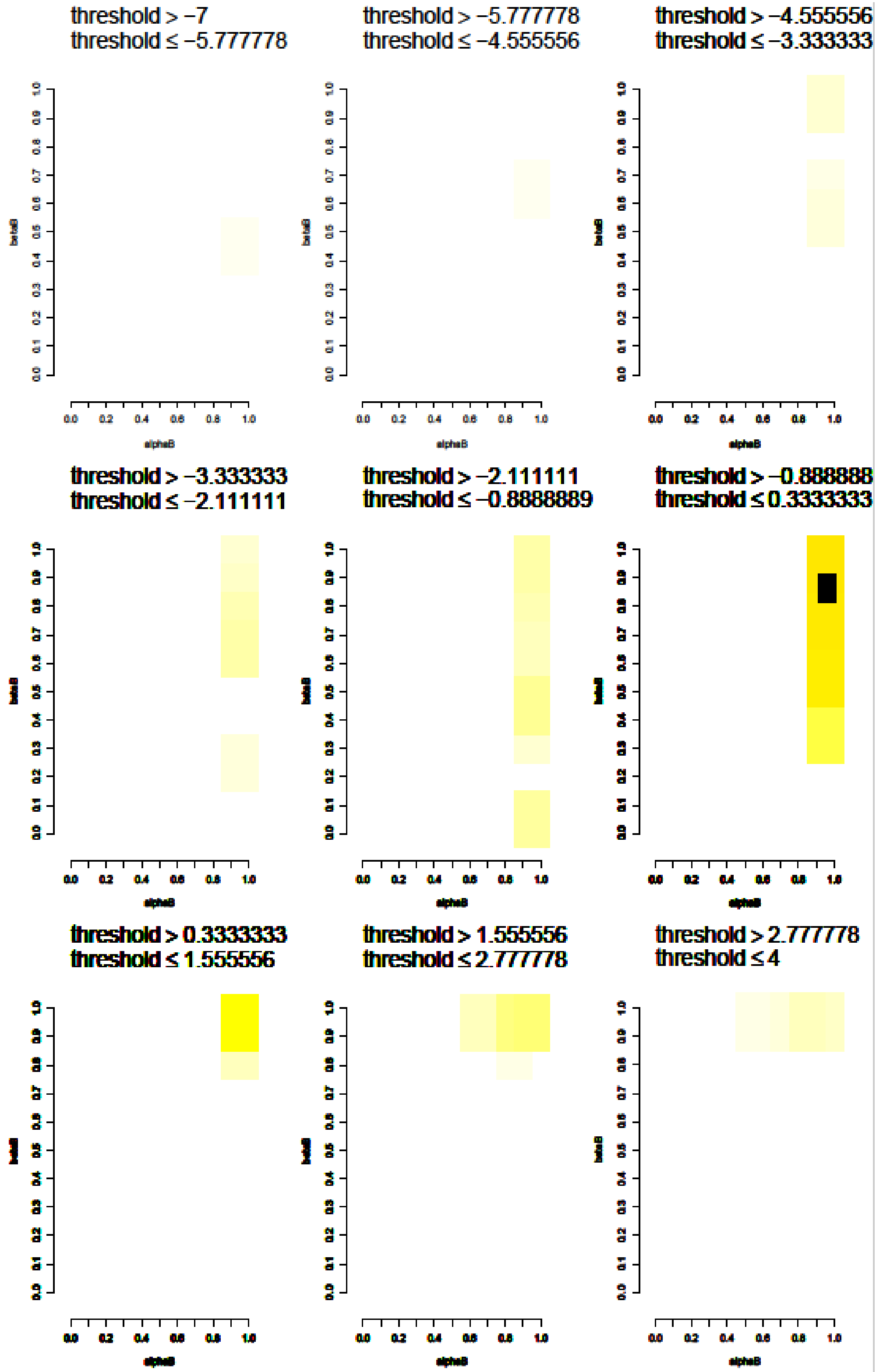


Figure 7iii: Modular, cognition domain A. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

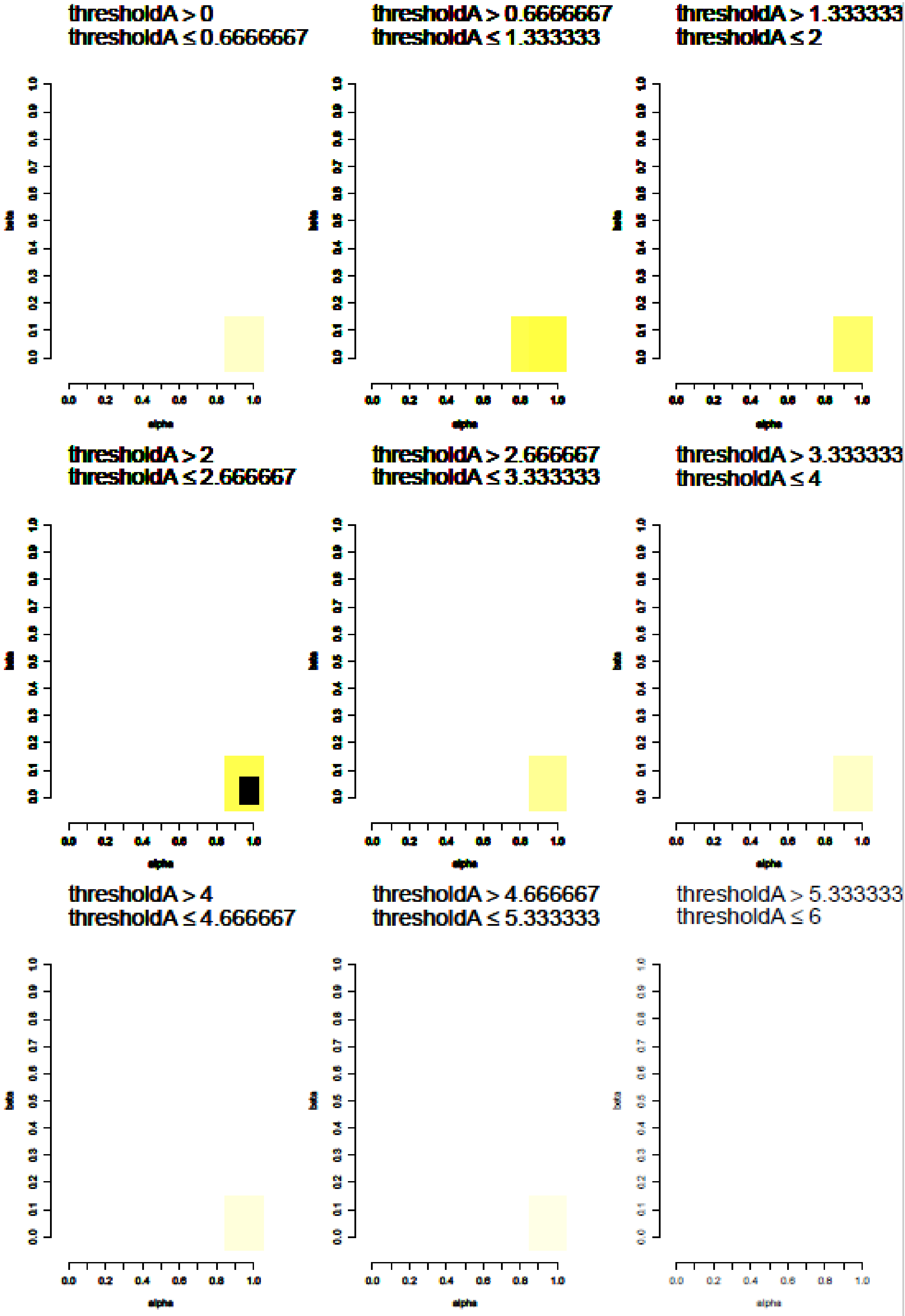


Figure 7iii: Modular, cognition domain B. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

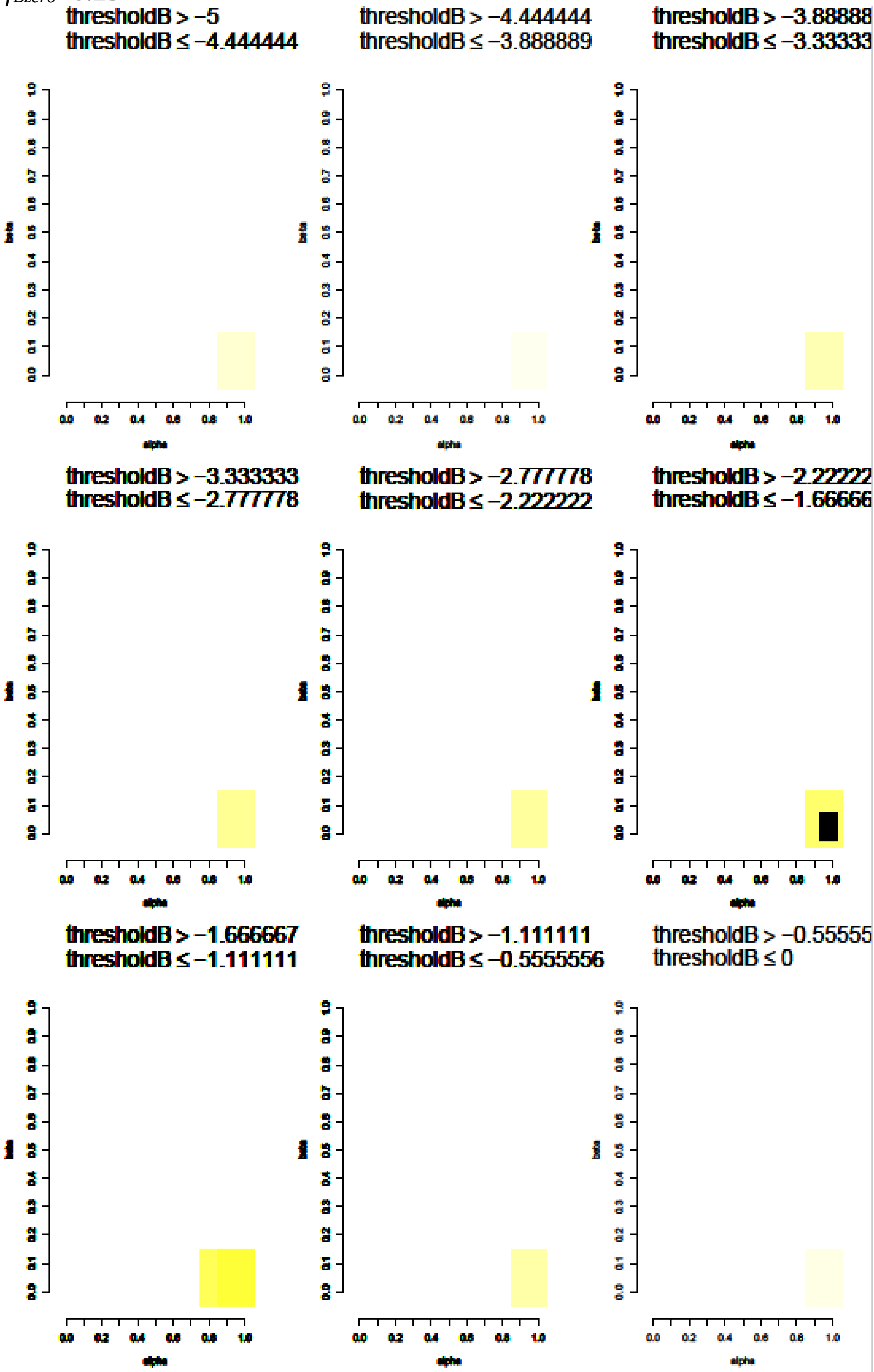


Figure 7iv: Domain-general. p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25

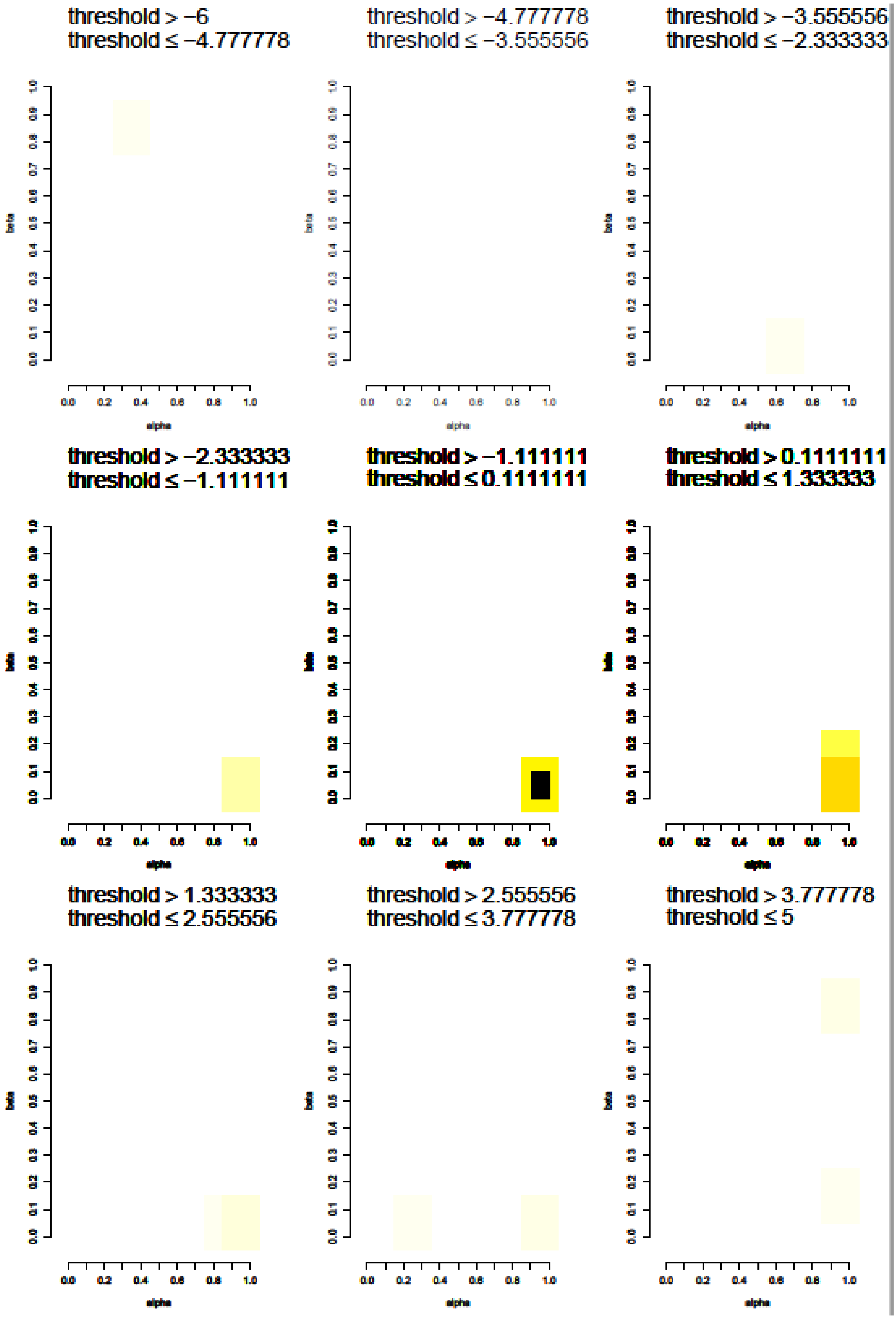


Figure 8i: Fully modular, domain A. $p_A=0.1$, $p_B=0.9$; $f_{Aone}=4$, $f_{Azero}=0.25$, $f_{Bone}=0.25$, $f_{Bzero}=4$

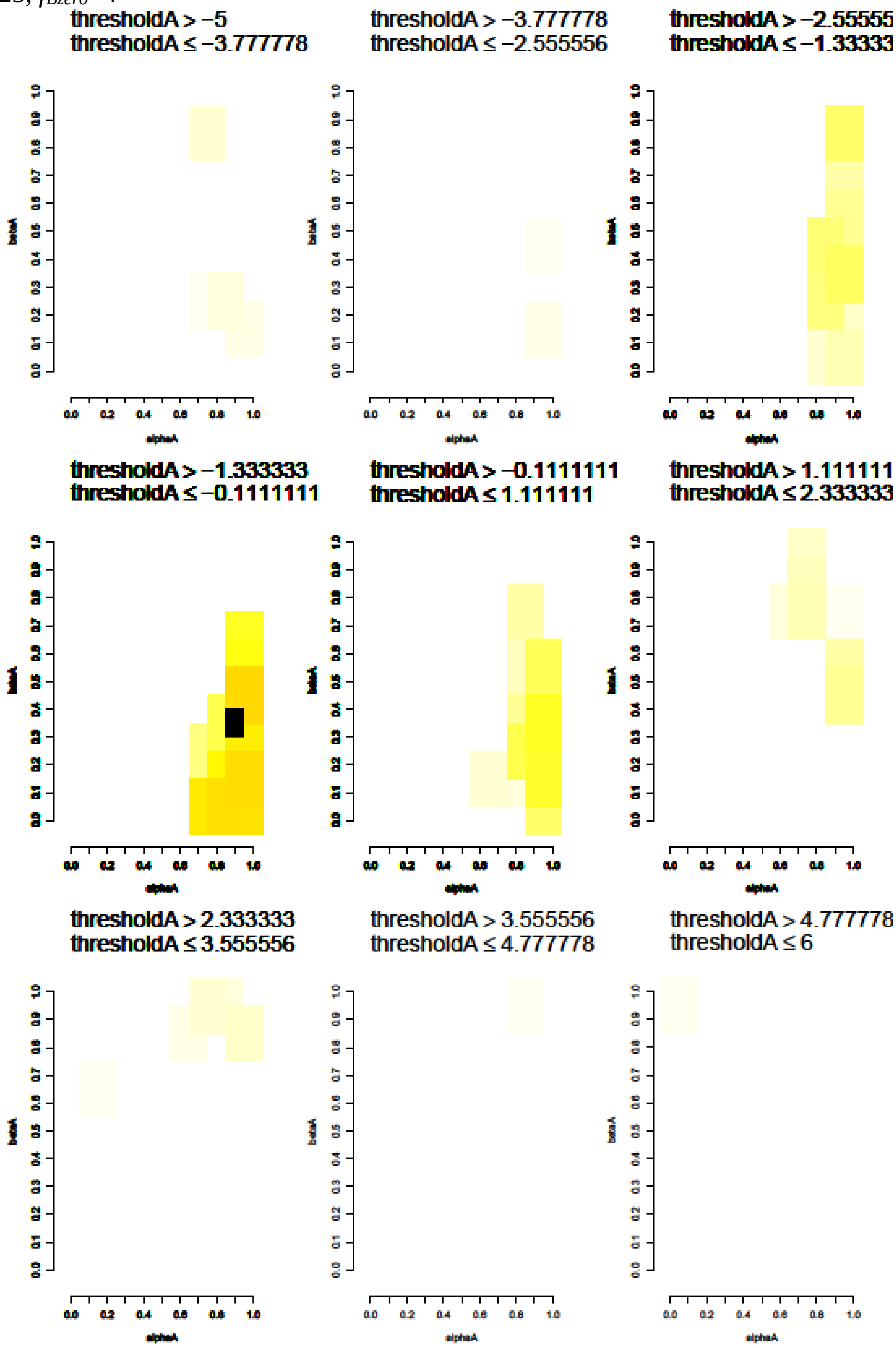


Figure 8i: Fully modular, domain B. p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4

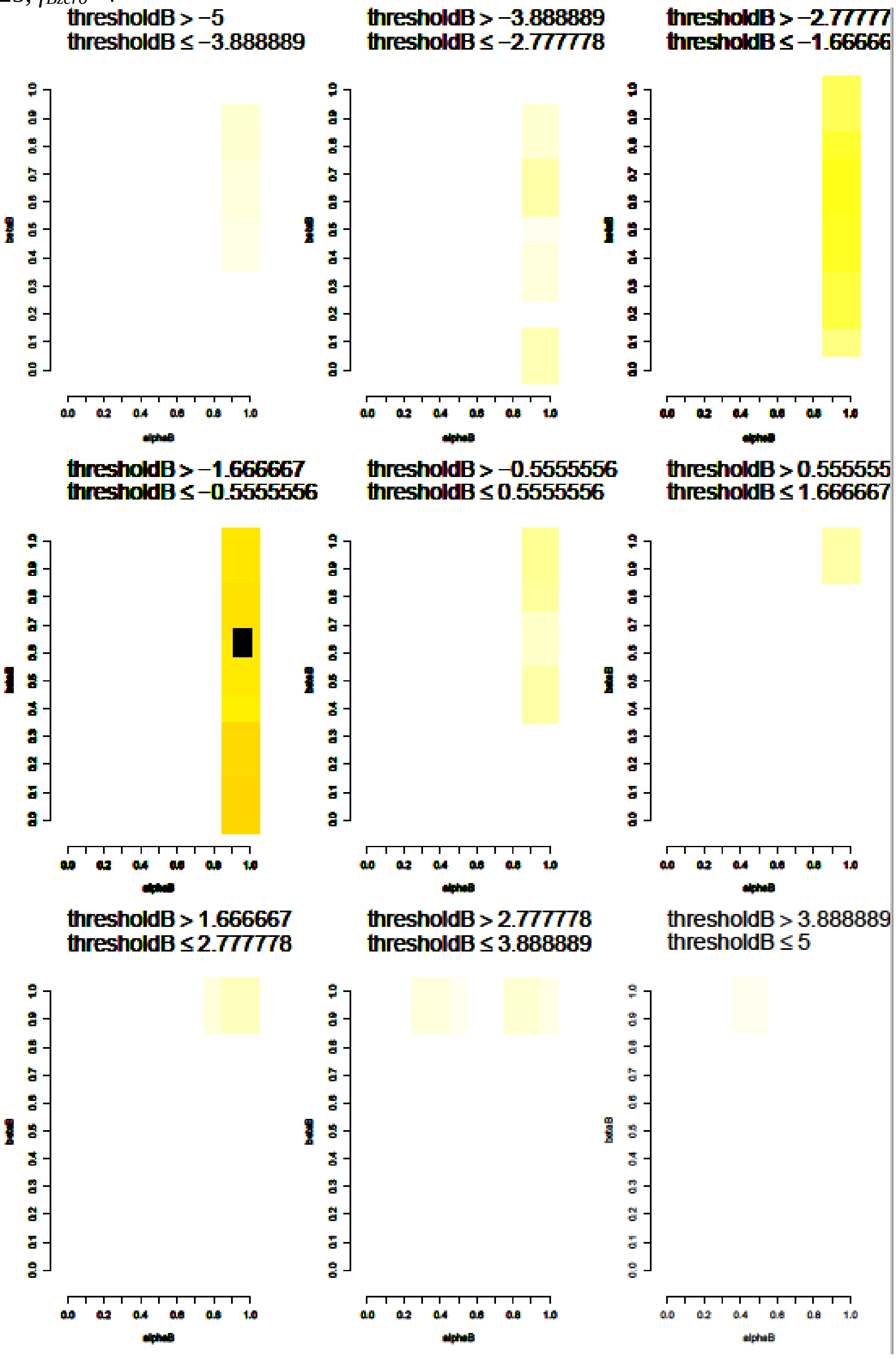


Figure 8ii: Modular motivation, domain A. $p_A=0.1$, $p_B=0.9$; $f_{Aone}=4$, $f_{Azero}=0.25$, $f_{Bone}=0.25$, $f_{Bzero}=4$

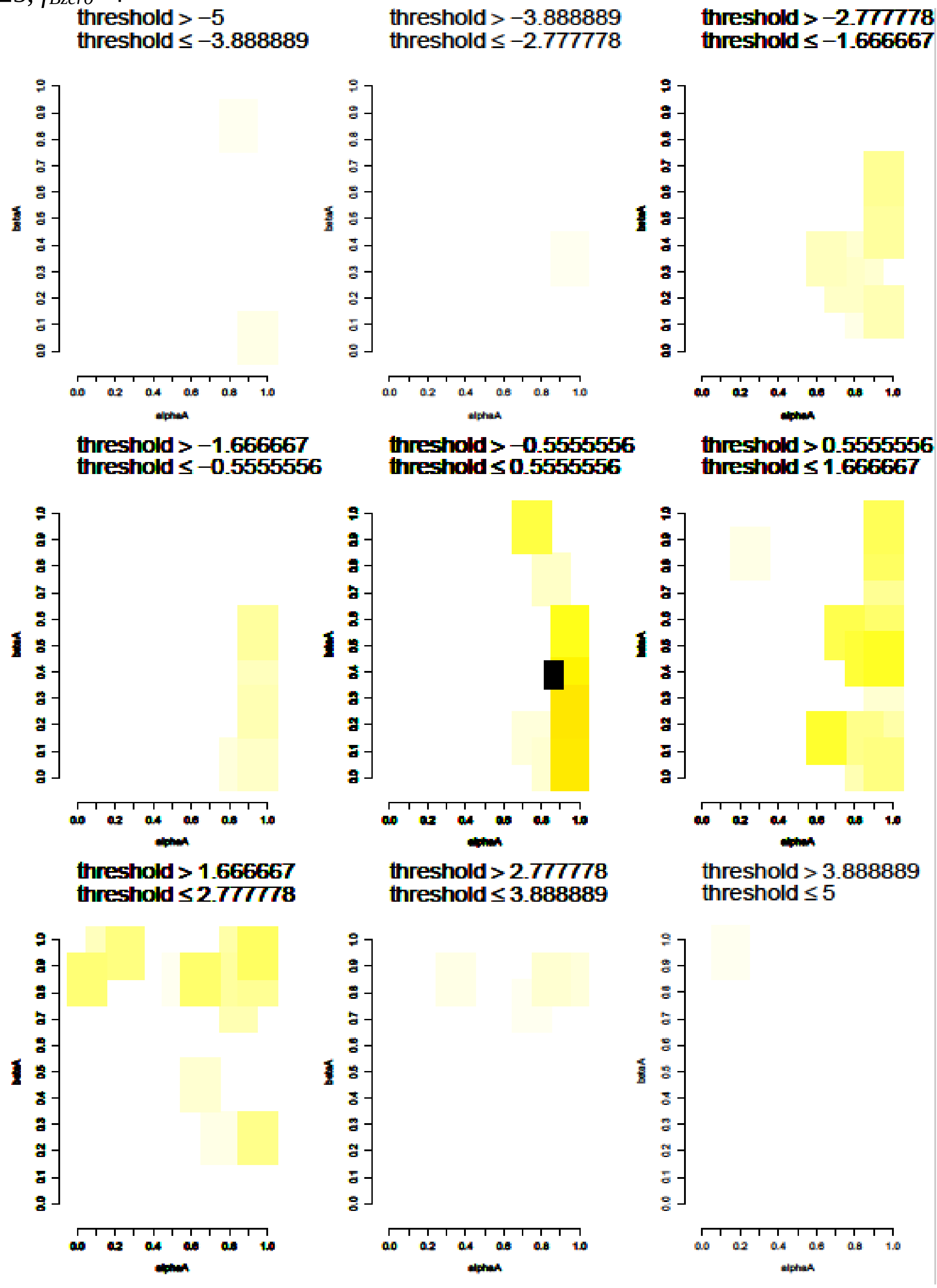


Figure 8ii: Modular motivation, domain B. $p_A=0.1$, $p_B=0.9$; $f_{Aone}=4$, $f_{Azero}=0.25$, $f_{Bone}=0.25$, $f_{Bzero}=4$

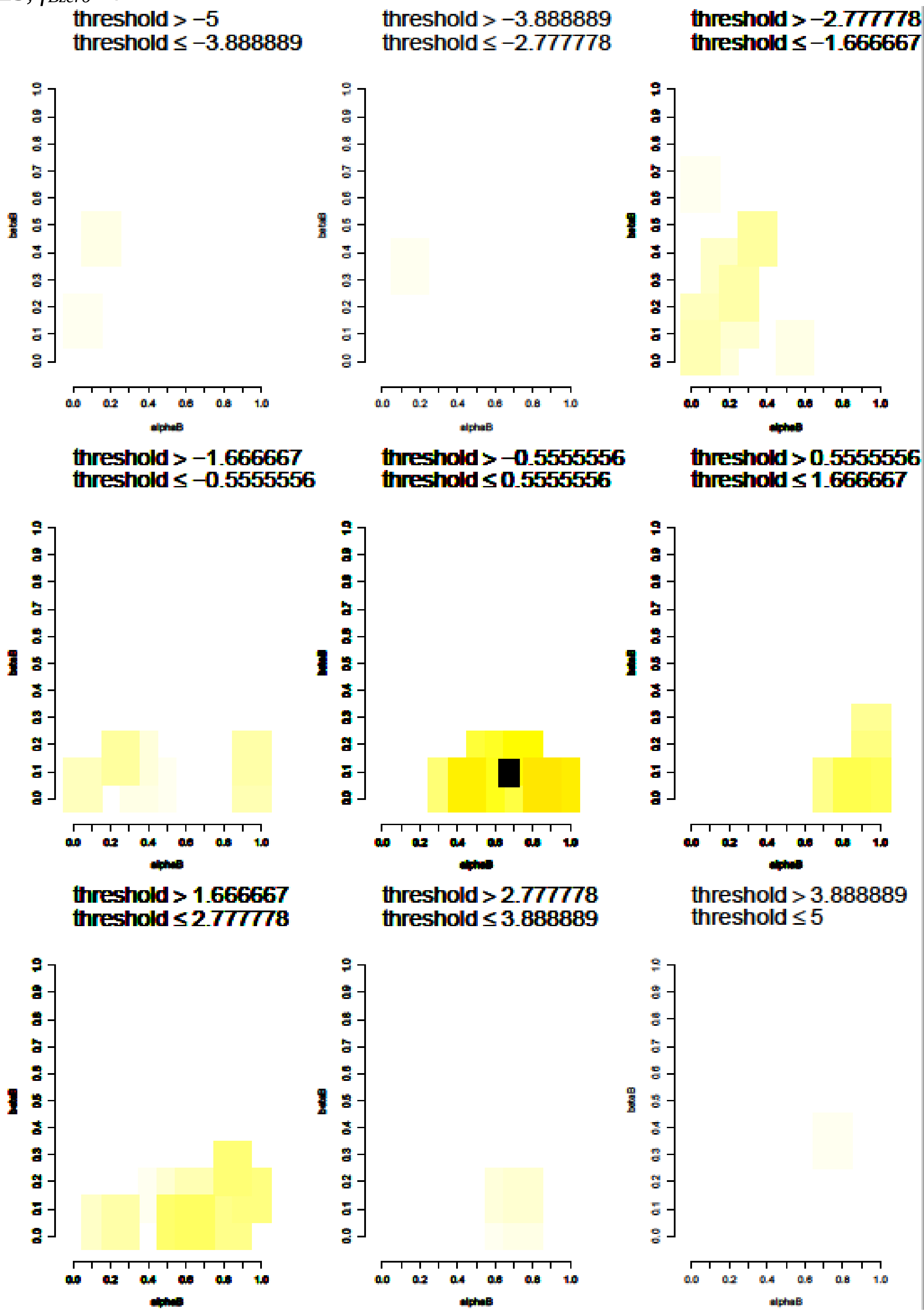


Figure 8iii: Modular cognition, domain A. p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4

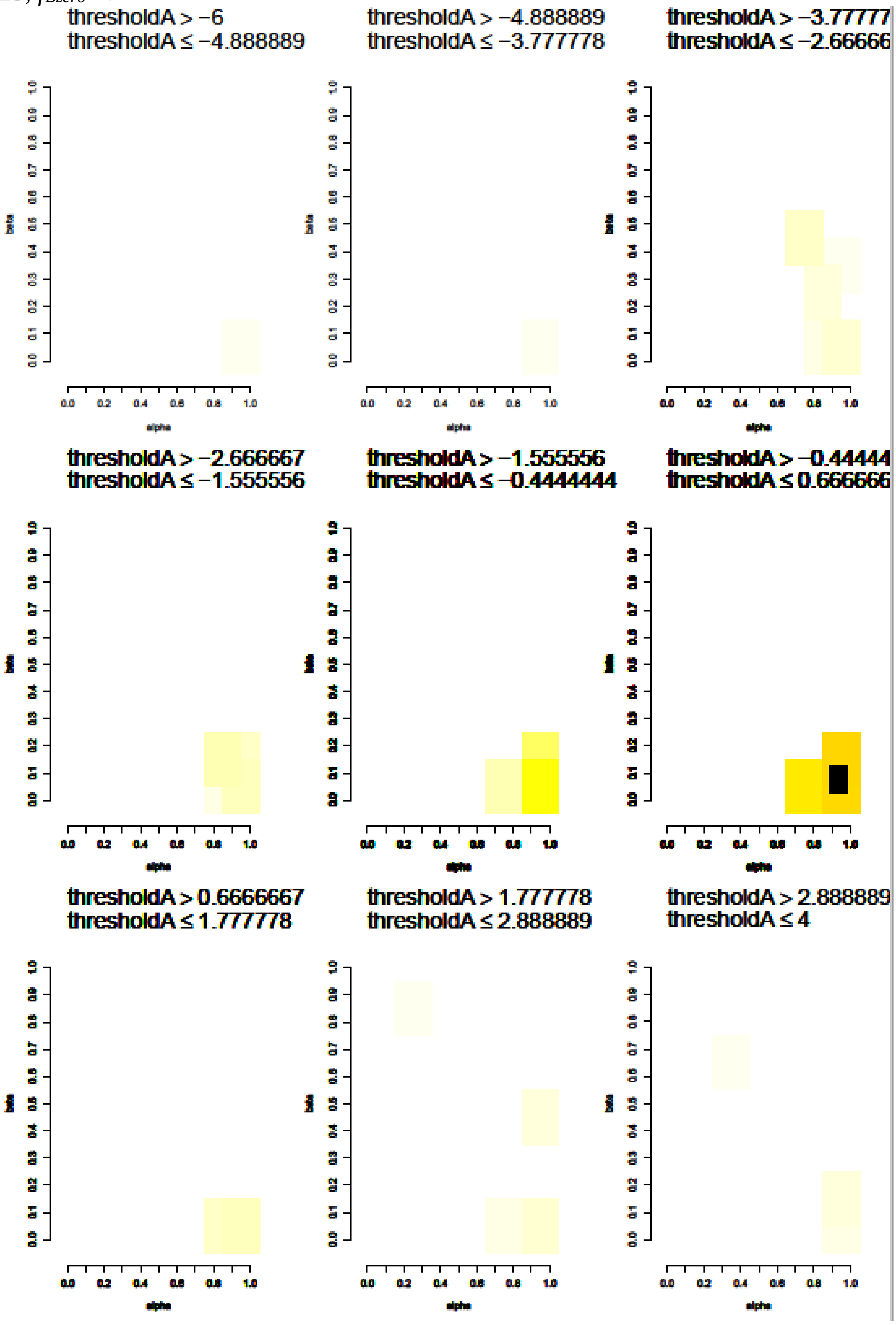


Figure 8iii: Modular cognition, domain B. p_A=0.1, p_B= 0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4

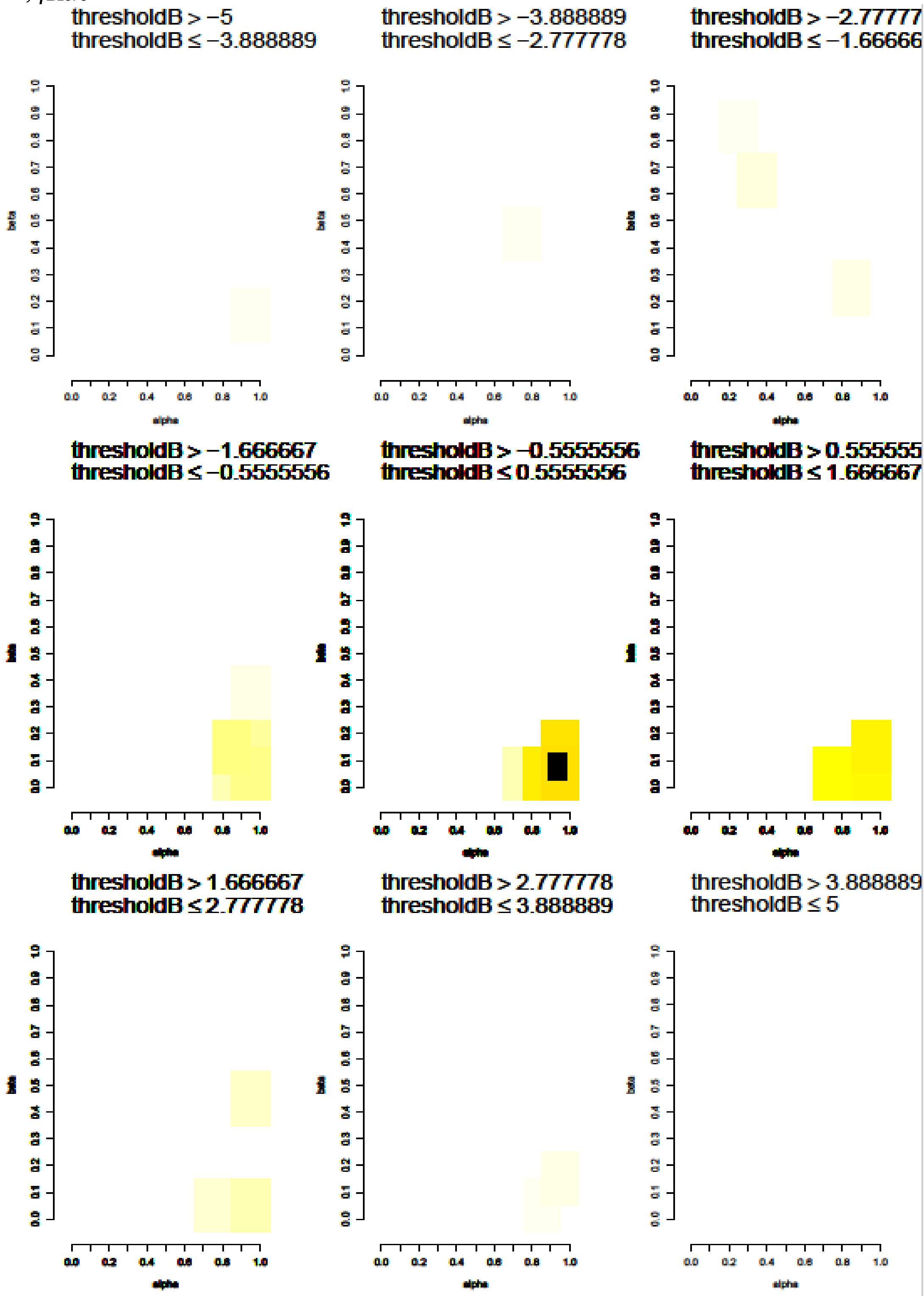


Figure 8iv: Domain-general A. p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4

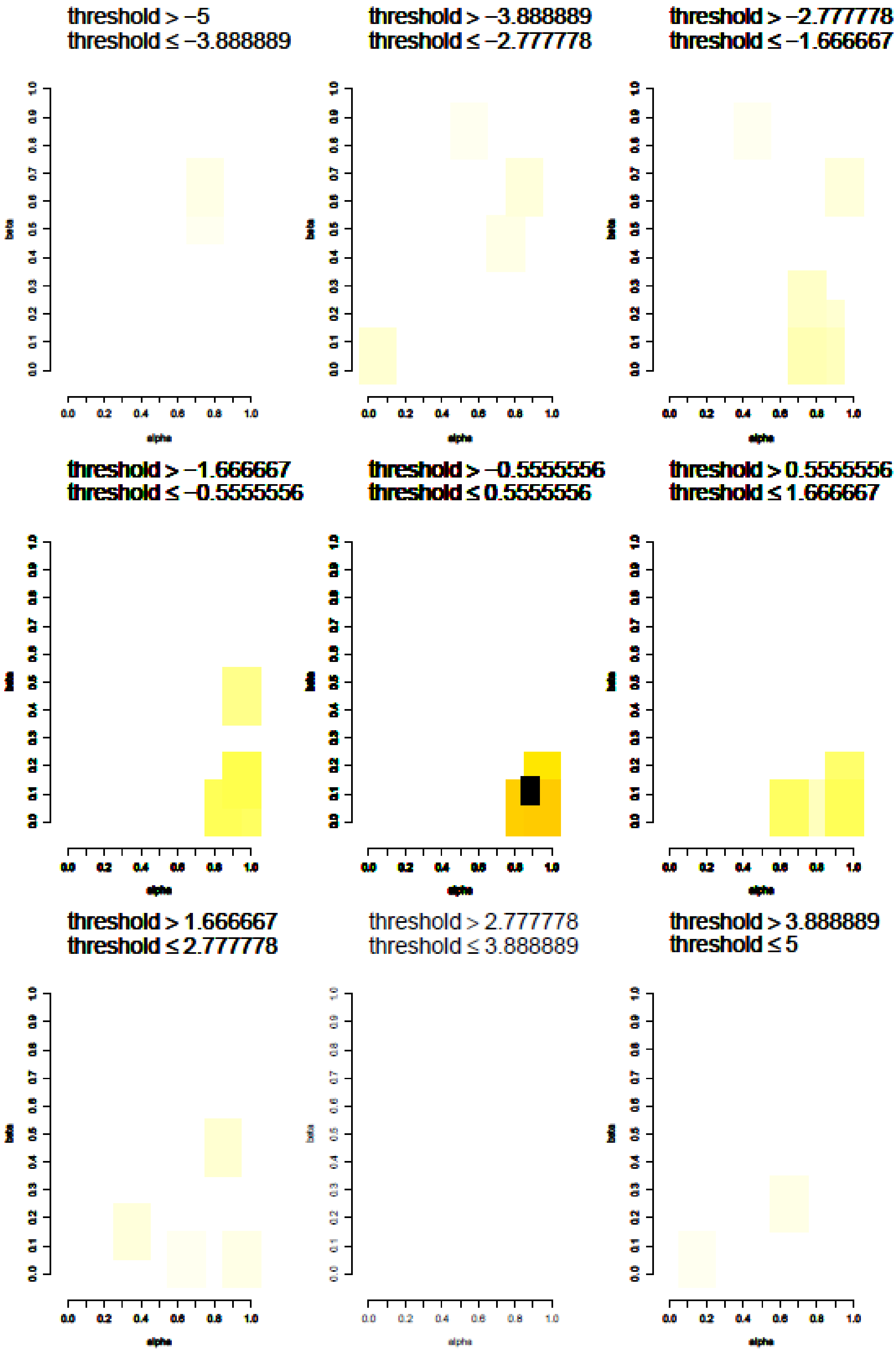


Figure 7 and Figure 8. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) agents with modular motivation only; (iii) agents with modular cognition only and (iv) fully domain-general agents. These are for skill-learning over two distinct domains. Figure 7 gives the results for runs where both the priors and the fitness tied to matching one's skillsets to the environment favour state 0 in domain A but state 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). Figure 8 gives the results for runs where the priors favour state 0 in domain A and state 1 in domain B, but the fitness tied to matching one's skillset to the environment instead favour state 1 in domain A but state 0 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4).

3.3. Did the agents' psychological components affect their fitness?

If both cognition and motivation are influential in driving the skills that agents master, I would expect both components to evolve to affect fitness. Accordingly, I run a regression to predict how cognitive and motivational thresholds affect fitness for the four agent types (see table 2).

Starting with runs where the most common and the most fit behaviour is 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4), then very few components predict fitness. The agents with modular cognition only had a smaller β predicting fitness (estimate=-5.90, p<0.001). The domain-general agents had a smaller α predicting fitness (estimate=-0.24, p=0.002). Both agents were motivated to play behaviour 0. This makes sense as behaviour 0 is both common and fit to play. Note that the domain-general agents had a higher fitness if they were unmotivated to play behaviour 1 when the state was 1. While this may seem counterintuitive, the priors made state 1 very rare and the fitness pressures also made behaviour 1 suboptimal. When in doubt as to whether the bow-and-arrow would be appropriate, the domain-general agents simply avoided using the bow-and-arrow at all. Note that none of the fully modular agents or partly modular agents with modular motivation had psychological components which evolved to affect their fitness. This again suggests that drift has an influential effect in the processes underlying the skills we master (Rorabaugh, 2014).

When the priors favour state 1, but the fitness tied to matching one's environment favours state 0 (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4), then the agents may be torn between displaying a common but less-fit skill, or a rarely needed but high-fit skill. Fully modular agents had a smaller T_B predict fitness (estimate=-0.02, p=0.03) while domain-general agents had a smaller T predict fitness (estimate=-0.03, p=0.003). The more fit agents thus were more likely to believe that the state was 1, perhaps as the priors made this likely. They needed less evidence to suggest bow-and-arrow use when this weapon was commonly needed. All four agents had a larger α value predicting fitness (table 2). All agents were highly motivated to play behaviour 1 when they believed that the state was 1. The most fit agents thus used the bow-and-arrow to hunt the common but small deer.

Interestingly, the fully modular agents were the only agents to have a smaller β_A predict fitness (estimate=-0.30, p=0.02). These agents would play behaviour 0 when they thought that the state was 0 in domain A. The more fit fully-modular agents would avoid using the bow-and-arrow in favour of using close-range weapons to hunt rarer but larger animals. While all agents had a higher fitness if they learned the bow-and-arrow in order to hunt the small but common deer, it was the fully modular agents who also tried to exploit a rarer skillset with a high payoff.

Table 2. The regression results displaying the psychological components that predict fitness for each of the four agent types in runs where the agents decide in two similar skill domains. The fitness tied to matching one's skillset to the environment always favours state 0 (f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4*)*. The results in italic text give the estimates for runs where state 0 was likely (p_A and, p_B = 0.1) and the results in non-italic text give the estimates for runs where state 1 was likely (p_A and p_B = 0.9). Note that agents with domain-general cognition had the same threshold in domain B as they did in A, and agents with domain-general motivation had the same α and β motivational thresholds in domain B as they did in A. For this reason, I arbitrarily block out domain B and focus on domain A for these agent types.

Psychological component left to evolve	**Agent Types**							
	Fully modular agents		**Modular motivation agents**		**Modular cognition agents**		**Domain-general agents**	
Intercept	*8.06* *(0.04)* 1.76 (0.06)	*** *** 	*7.99* *(0.04)* 1.68 (0.06)	*** *** 	*8.19* *(0.03)* 1.62 (0.08)	*** *** 	*8.13* *(0.03)* 1.67 (0.04)	*** ***
Threshold A	*0.006* *(0.01)* -0.02 (0.01)		*0.01* *(0.02)* -0.01 (0.01)		*-0.01* *(0.01)* -0.01 (0.01)		*0.01* *(0.01)* -0.03 (0.01)	 **
Threshold B	*-0.0003* *(0.01)* -0.02 (0.009)	 * 			*-0.02* *(0.01)* -0.009 (0.01)	. 		
***α* motivation A**	*0.06* *(0.09)* 0.14 (0.06)	 ** 	*-0.12* *(0.09)* 0.13 (0.06)	 * 	*-0.02* *(0.09)* 0.36 (0.09)	 *** 	*-0.24* *(0.08)* 0.36 (0.05)	** ***
***α* motivation B**	*-0.01* *(0.08)* 0.14 (0.06)	 * 	*0.11* *(0.10)* 0.19 (0.06)	 ** 				
***β* motivation A**	*-0.18* *(0.27)* -0.30 (0.13)	 * 	*0.01* *(0.20)* -0.11 (0.13)		*-5.90* *(0.98)* 0.17 (0.37)	*** 	*-0.38* *(0.32)* 0.24 (0.12)	 .

***β* motivation B**	*-0.17*	*0.27*	
	(0.17)	*(0.53)*	
	-0.10	-0.08	
	(0.12)	(0.19)	

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p < 0.01)*
** (p< 0.05)*
· (trend: p=0.05 – 0.10 significance)

Table 3 represents two different domains, such as hunting in domain A or cooking in domain B. It is both rare and low payoff to use the bow-and-arrow in domain A and it is both common and fit to cook in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). The partly modular agents with modular cognition had a larger T_A predict fitness (estimate=0.07, p=0.0001). The agents needed more evidence to believe that the bow-and-arrow was optimal in environments where this was rarely needed. The domain-general agents had a smaller cognitive threshold T predicting fitness (estimate=-0.05, p=0.02). These agents were more likely to believe that the state was 1. While this was likely to be beneficial in domain B where state 1 was favoured, this threshold could have led to erroneous beliefs in domain A where state 0 was favoured. By definition, a domain-general agent cannot tailer her thresholds to the contrasting pressures of two domains.

The agents with domain-general motivation had a larger α (domain-general: estimate=2.12, p<0.001; partly modular with modular cognition estimate=3.24, p=0.0002) and a smaller β (domain-general estimate=-1.69, p<0.001; partly modular agents with modular cognition estimate=-4.90, p<0.001) predicting fitness. These agents would play behaviour 1 when they believed that the state was 1, and play behaviour 0 when they believed that the state was 0. These motivational thresholds would help the agents with domain-general motivation to avoid the bow-and-arrow when inappropriate, and to cook when appropriate. Interestingly, the modular motivational components did not evolve to predict fitness (see table 3), despite the large shifts in motivational thresholds for any agents with modular motivation (see figure 7, section 3.2). When cognition and motivation coevolved, it was difficult to pinpoint how exactly each affected fitness. This could suggest that drift, or random noise in each simulation, had a large impact in driving agent fitness and behaviour (Rorabaugh, 2014).

Finally, I consider two distinct domains where the most commonly-needed skill is not the one which gives the highest payoff (p_A=0.1, p_B=0.9*;* f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). For example, it is common to be able to set traps for small prey in domain A, but a rarer prize with higher payoffs would be to use the bow-and-arrow to hunt the larger deer. Similarly, cooking may be commonly-needed in domain B though overcooking can be more dangerous. Modular agents had a larger T_A predict fitness (fully modular: estimate=0.02, p=0.04; partly modular with modular cognition estimate=0.07, p=0.0001). They were more likely to believe that the state was 0 in domain A when this was common. They were likely to believe that the bow-and-arrow was inappropriate. Interestingly, no other cognitive thresholds evolved to affect fitness (see table 3).

Fully modular agents (α_A estimate=0.27, p=0.04; α_B estimate=0.18, p=0.001) and domain-general agents (estimate=0.34, p=0.0002) had larger α values predicting fitness (see table 3). The agents were motivated to play behaviour 1 when they thought that the state was 1. The partly modular agents with modular motivation had a smaller β_B predict fitness (estimate=-0.43, p=0.006) and domain-general agents had a smaller β predict fitness (estimate=-0.21, p=0.02). These agents were motivated to play behaviour 0 when they thought that the state was 0. The agents with modular motivation only displayed this in domain B, where behaviour 0 was rarer but gave higher payoffs. For example, it is common that we need to cook our food (Wrangham, 2009). However, we may occasionally encounter plants which emit toxins when overcooked (Stahl et al., 1984). It seems that the agents with modular motivation only were attempting to cook less in the latter case. These agents tried to master the rarer skillset with higher payoffs, whilst the other agents tried to master the common skillset with lower payoffs. Perhaps both skillsets are viable, but the one we come to master may be influenced by drift.

Table 3. The regression results displaying the psychological components that predict fitness for each of the four agent types in runs where the priors favoured state 0 in domain A and state 1 in domain B (p_A=0.1, p_B=0.9). The results in italic text give runs where the fitness tied to matching one's skillset to the environment favoured matching state 0 in domain A and state 1 in domain B (f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). The results in non-italic text gives the estimates for runs where the fitness tied to matching one's skillset to the environment favours matching in state 1 of domain A but state 0 of domain B (f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}= 4). These clash with the priors. Agents with domain-general cognition had the same threshold in domain B as they did in domain A, and agents with domain-general motivation had the same α and β motivational thresholds in domain B as they did in domain A. For this reason, I arbitrarily block out domain B and focus on domain A for these agent types.

Psychological component left to evolve	Agent Types							
	Fully modular agents		**Modular motivation agents**		**Modular cognition agents**		**Domain-general agents**	
Intercept	*7.52* *(0.54)* 1.57 (0.14)	*** ***	*7.71* *(0.32)* 1.86 (0.10)	*** ***	*4.75* *(0.86)* 1.94 (0.23)	*** ***	*4.49* *(0.26)* 1.61 (0.09)	*** ***
Threshold A	*0.007* *(0.01)* 0.02 (0.01)	 *	*0.03* *(0.02)* -0.01 (0.02)		*0.07* *(0.02)* 0.007 (0.01)	***	*-0.05* *(0.02)* 0.01 (0.01)	*
Threshold B	*-0.02* *(0.01)* -0.008 (0.01)				*-0.01* *(0.02)* -0.010 (0.13)			
***α* motivation A**	*0.04* *(0.09)* 0.27 (0.13)	 *	*-0.07* *(0.09)* 0.03 (0.09)		*3.24* *(0.88)* 0.04 (0.23)	***	*2.12* *(0.27)* 0.34 (0.09)	*** ***
***α* motivation B**	*0.48* *(0.55)* 0.18 (0.06)	 **	*0.37* *(0.32)* 0.06 (0.07)					
***β* motivation A**	*0.11* *(0.17)* -0.05 (0.06)		*0.38* *(0.27)* -0.05 (0.06)	.	*-4.90* *(0.87)* -0.48 (0.19)	***	*-1.69* *(0.23)* -0.21 (0.09)	*** *
***β* motivation B**	*0.02* *(0.08)* -0.05 (0.18)		*-0.07* *(0.13)* -0.43 (0.16)	 **				

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· *(trend: p=0.05 – 0.10 significance)*

3.4: Is there a difference in the fitness of the four agent types?

Section 3.3 addresses my first research aim regarding how cognition and motivation coevolved to predict agent fitness. My analysis of fitness is not complete until I address the second research aim. This model is novel in its aim of comparing the skill-learning behaviour and psychological phenotypes of fully modular, partly modular, and domain-general agents. Here, I continue this analysis by comparing the fitness accrued by each of the 4 agent types by the final generation (see table 4).

Table 4. The regression results displaying any differences in fitness of the four agent types in runs where i) both the priors and fitness tied to matching one's skillset to the environment favour matching in state 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4); ii) runs where the priors favour state 1 but the fitness tied to matching one's skillset to the environment favour matching in state 0 (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4); iii) runs where both the priors and the fitness tied to matching one's skillset to the environment favour state 0 in domain A and state 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25) and iv) runs where the priors favour state 0 in domain A and state 1 in domain B, though the fitness tied to matching one's skillset to the environment favours matching in state 1 of domain A and state 0 of domain B (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). Note that the domain-general agents are the omitted category to which the other three agent types are compared.

Estimates for regression predicting fitness	**Prior probabilities and fitness tied to matching the environment**			
	p_A =0.1, p_B =0.1	*p_A =0.9, p_B =0.9*	*p_A =0.1, p_B =0.9*	*p_A =0.1, p_B =0.9*
	f_{Aone} = 0.25, f_{Azero} = 4, f_{Bone} = 0.25, f_{Bzero}= 4	f_{Aone} = 0.25, f_{Azero} = 4, f_{Bone} = 0.25, f_{Bzero} = 4	f_{Aone} = 0.25, f_{Azero} = 4, f_{Bone} = 4, f_{Bzero} = 0.25	f_{Aone} = 4, f_{Azero} = 0.25, f_{Bone} = 0.25, f_{Bzero} = 4
Intercept	8.06 *** (0.018)	1.91 *** (0.016)	6.43 *** (0.020)	1.88 *** (0.01)
Modular cognition	-0.017 (0.03)	0.012 (0.02)	1.55 *** (0.03)	0.06 ** (0.02)
Modular motivation	-0.06 * (0.03)	-0.04 . (0.02)	1.59 *** (0.03)	-0.02 (0.02)
Fully modular	-0.005 (0.026)	-0.025 (0.022)	1.62 *** (0.03)	0.03 (0.02)

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· (trend: p=0.05 – 0.10 significance)

When the priors and fitness pressures both favoured state 0 (p_A=0.1, p_B=0.1; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4), then the partly modular agents with modular motivation had less fitness than domain-general agents (estimate=-0.06, p=0.02). When mastering two similar skillsets, domain-general psychology seemed to be sufficient and, if anything, modular motivation was a disadvantage.

We now consider runs where both the priors favoured state 1, though the fitness pressures favoured behaviour 0 (p_A=0.9, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=0.25, f_{Bzero}=4). This is when the agent could master two similar skillsets (e.g., the bow-and-arrow and spear thrower) but the most common skillset is not necessarily the one with the highest payoff. The regression revealed no differences in fitness between the domain-general and modular agents. I further broke down these findings with linear combinations to compare the modular agents fitness (see appendix 4). These revealed that the partly modular agents with modular cognition only had a higher fitness than the partly modular agents with modular motivation. A further regression confirmed that agents with modular motivation had less fitness than agents with domain-general motivation on these runs (see appendix 5). This confirms my above interpretation: domain-general psychology is sufficient when mastering two similar skillsets and (if anything) modular motivation is a disadvantage in terms of fitness.

We then consider runs where the priors and fitness pressures favour state 0 in domain A, but state 1 in domain B (p_A=0.1, p_B=0.9; f_{Aone}=0.25, f_{Azero}=4, f_{Bone}=4, f_{Bzero}=0.25). This run is the equivalent of mastering two distinct skillsets. In this run, all modular agent types accrued more fitness than the domain-general agents (fully modular, estimate=1.62, p<0.001; agents with modular cognition only, estimate=1.55, p<0.001; agents with modular motivation only, estimate=1.59, p<0.001). The linear combinations confirmed that the fully modular agents accrued more fitness than partly modular agents with modular cognition only (see appendix 4). A further regression confirmed that agents with modular cognition always had a higher fitness than agents with domain-general cognition, while agents with modular motivation always accrued a higher fitness than agents with domain-general motivation (see appendix 5). Taken together, it thus seems that fully modular psychology is necessary when mastering skillsets in two distinct domains.

Finally, I consider a run where the priors favour state 0 in domain A but the fitness pressures favour behaviour 1, and the priors favour state 1 in domain B though the fitness pressures favour behaviour 0 (p_A=0.1, p_B=0.9; f_{Aone}=4, f_{Azero}=0.25, f_{Bone}=0.25, f_{Bzero}=4). This is the equivalent of mastering two distinct skillsets, where the most common skill to master gives a lower payoff than the rarer skill to master. To illustrate in reference to domain A, there is an abundancy of small prey which we can build traps to catch. However, the larger but rarer deer could be hunted with the bow-and-arrow. To illustrate in reference to domain B, it is common that we need to cook our food. On rare occasions, there would be a higher fitness consequence to avoid overcooking food however (e.g., some plants emit harmful toxins when overcooked; Stahl et al., 1984).

The partly modular agents with modular cognition had a higher fitness than domain-general agents on these runs (estimate=0.06, p=0.007). The linear combinations revealed that both fully modular agents and agents with modular cognition only accrued a higher fitness than the partly modular agents with modular motivation only (see appendix 4). A further regression confirmed that agents with modular cognition always had a higher fitness than agents with domain-general cognition on these runs, though there was no difference in fitness between the domain-general and modular motivational processes (see appendix 5). Taken together, it thus seems that modular cognition is uniquely important when mastering two distinct skillsets, in cases where the priors and the fitness pressures clash. Perhaps modular cognition is needed to weigh up the contrasting fitness pressures of multiple domains. To summarise, domain-general psychology may be sufficient when the agents master skillsets in two similar domains though modular psychology seems to be necessary when the agents master skillsets in two distinct skill domains, with a particular emphasis on modular cognition.

3.5. Summary of the key results

This analysis focused on (i) how our cognition and motivation coevolved to influence the asocial skills that we master and, (ii) whether the decision-making underlying asocial skills was likely to be domain-general, partly modular, or fully modular in nature. This analysis revealed a large range of results which, to aid the discussion, I summarise as three key findings below:

1. Domain-general psychology seems to be sufficient when the agent mastered skillsets over two similar domains, while modular psychology may have been necessary when the agent mastered skillsets over two distinct domains.
2. The partly modular agents with modular cognition only experienced a larger shift in their cognitive thresholds when mastering skills over two distinct domains (section 3.2). Interestingly, the agents with modular cognition outperformed agents with modular motivation when mastering skillsets over two domains, where the most common skill had a lower payoff than a rarer skill. Taken together, this suggests a unique role of modular cognition when agents come to master distinct skillsets.
3. The varied psychological architecture in the final generation of agents (section 3.2), coupled with the fact that few psychological components reliably evolved to predict agent fitness (section 3.3), may suggest that drift had a large impact on the skillsets that we come to master.

4. Discussion

I investigated the psychological systems that likely uphold our ability to master skills over a variety of domains. I found that domain-general psychology was sufficient to explain performance over two similar domains, but modular psychology was necessary when agents mastered skills across two distinct domains.

Domain-general psychology was shown to be important across two similar domains in the current model. Here, I considered two similar domains to be those where certain environments were equally likely to occur and when there were similar fitness payoffs to mastering the correct skillset in each environment. For example, two domains requiring similar weapon-use. Whilst this is just one skill domain, throughout the ancestral past, individuals may also have had to be skilled in multiple other domains (Henrich, 2015; Henrich & Muthukrishna, 2021).

Modular psychology may be best placed to explain performance over such a wide domain of skillsets (Charbonneau, 2016; Goschke & Bolte 2012). Indeed, fully modular agents accrued the most fitness on the runs with two distinct domains. Interestingly, agents with modular cognition consistently accrued more fitness than agents with domain-general cognition and modular motivation (see appendix 4). The partly modular agents with modular cognition also experienced a large shift in their cognitive thresholds.

If modular cognitive processing is important, then these results may support the massive modularity hypothesis of Evolutionary Psychology (Cosmides & Tooby, 1994). To illustrate, let us return to the example from the introduction of the gatherer who sees a spider and quickly withdraws to a safe distance. Her behaviour may be consistent with a threat-detection module which has evolved specifically to process and avoid poisonous animals (Öhman & Mineka, 2001) rather than being evidence of a limited domain-general pool of resources (Vergauwe et al., 2021). While Evolutionary Psychology has been criticised for its limited discussions of culture (Confer et al., 2010; Gangestad et al., 2006), modularity may still be able to account for the wide variation in skillsets that have been observed across cultures. Modules can remain plastic to environmental inputs (Barrett & Kurzban, 2006; Mercado, 2008; Verbeke & Verguts, 2019) and the modules that arise throughout our development are likely shaped by culture (Müller, 2007; Reader, 2006).

Of course, modular motivation was also influential in driving the agent behaviour across two domains. However, there was an inconsistent effect of motivation on agent behaviour and fitness. Consider runs over two distinct domains where the most common behaviour is not necessarily the most fit (see figure 8). In domain B, the fully modular agents would cook food simply as this was commonly needed though this had a lower payoff than selectively undercooking some food, such as plants that can emit harmful toxins (Stahl et al., 1984). Instead, the partly modular agents with modular motivation only tried to undercook when this was rare but had higher payoffs. These agents thus showed two different behaviours, with fully modular agents mastering a common skillset with a lower payoff but the partly modular agents mastering a rarer skillset with a higher payoff. Perhaps this may be due to a path dependence effect, where small differences in the route of evolutionary trajectories differentially affected these agents (Bentley et al., 2004). The modular cognition of fully modular agents may have affected the evolutionary trajectory of their modular motivational thresholds in unexpected ways (Laland, 1993).

Note that there were no cases of psychological polymorphism, or multiple psychological architectures being present, in the final generation of agents across this model's runs (see appendix 3). This suggests that selection was pushing the agents towards a certain subset of psychological space. This is perhaps the most evidence on runs over two distinct domains, where the most commonly needed skillset has a lower payoff than a skillset which is rarely needed (see figure 8). Some agents were selected for that tried to play behaviour 0 when the state was 0 in domain B. These agents master a rarer skillset, but with a higher payoff. For example, these agents stay attuned to the rare event that overcooking is dangerous and try to cook less. Alternatively, some agents are selected for that play behaviour 1. After all, it is more common that we need to cook our food than not (Wrangham, 2009). If either space could be selected for, then the exact psychological architecture of the final generation of agents may depend on drift. Drift describes the random fluctuations in a population which may have large consequences on the evolution of a psychological characteristic further down the generations (Rorabaugh, 2014). Drift has previously been shown to affect cultural evolutionary outcomes in domains such as name preferences and random copying errors in pottery art (Bentley et al., 2004; Billiard & Alvergne, 2018; Rorabaugh, 2014). Further supporting a role of drift in this model, there were very few psychological components which directly predicted agent fitness (see section 3.3).

If drift commonly affects the evolutionary trajectory of skill acquisition, this suggests that future models should explore this process in detail. Interestingly, there is often an overlap in the kind of skills that emerge in different cultural groups who share similar ecological conditions (Henrich, 2004; Richerson & Boyd, 2008). This may suggest that certain factors mediate the effect of drift. While random copying errors can influence the trajectory of cultural evolution (Rorabaugh, 2014), it may be the case that drift can be counteracted by the unique propensity for high-fidelity social learning that is seen across human cultures (Heyes, 2018; Legare, 2017; Muthukrishna et al., 2017; Whiten, 2019). The complex myriad of skills and toolkits that we display is thought to be underlaid by a unique propensity for social learning which produces an advanced cultural repertoire in humans (Henrich, 2015; Mesoudi & Thornton, 2018; Richerson & Boyd, 2008). Future modelling efforts could aim to investigate the space in which such a high-fidelity learning process would likely emerge.

Adding social learning processes to this model may also change the results in regard to the importance of modularity. Social learning may be underlaid by a domain-general reasoning process which allows the learner to craft a multitude of rules regarding what skills to learn and from who (Heyes, 2019). One's social learning ability is often linked to one's ability to engage in trial-and-error learning, which suggests that these learning processes share an underlying mechanism (Mesoudi, 2011; Muthukrishna et al., 2016). Skills can be learned via both trial-and-error (Mesoudi, 2011; Muthukrishna et al., 2016), and social learning (Legare & Nielsen, 2015; McElreath et al., 2018). A mixture of the two processes may be preferred, as trial-and-error can compensate for cases where social learning would lead to suboptimal behaviours (Mesoudi, 2008; Truskanov & Prat, 2018). Taken together, this evidence suggests that asocial skills may be underlaid by domain-general processes, if the model is updated to consider both trial-and-error and social learning.

It is worth noting that this model was simplistic in that the agents needed only to choose between one of two behaviours when mastering a skill. Throughout our ancestral past, it is likely that we would have had to master skills with multiple equilibria which were constantly changing due to environmental fluctuations (Deffner et al., 2020). For example, Mesoudi et al. (2015) employed an arrow designing task, whereby participants could change the material of the arrow, the size of the arrow point or the colour of their arrow (though only the first two affected fitness when hunting). Learning a skill is likely to be more complex in a task where multiple, intertwining factors must be mastered. Perhaps social learning becomes more necessary here (Mesoudi et al., 2015; Muthukrishna et al., 2016; Truskanov & Prat, 2018).

Perhaps the parameter space chosen for my model favoured modularity. This could be due to modelling the prior probability of environmental states, and the fitness when the agent matched their skillset to the environment, separately across both domains. When these variables fluctuate separately, then it perhaps makes sense for a modular psychological system to emerge which supports decision-making over distinct domains. However, it may be the case that certain skillsets become linked over time and thus a domain-general reasoning ability cannot be ruled out (Bolhuis et al., 2011; Heyes, 2019). In my model, I considered the use of a bow-and-arrow when hunting and the decision of when to cook food as two distinct domains. However, these two domains may become linked as having success with a bow-and-arrow when hunting prey implies that the agent must also learn to cook meat.

Finally, perhaps another simplicity of this model is that all agents had to master skills across two domains. Perhaps it is not always necessary for agents to become skilled over multiple domains. Since the agricultural revolution, there has been an intensification in the division of labour across skillsets (Muthukrishna & Henrich, 2016). This is because humans were able to cultivate food to such an extent, that a surplus of supplies allowed individuals within a group to specialise in different relevant skills which could then be exchanged for other resources. For example, some individuals would know how to farm the local crops while other individuals could specialise in a different area, such as finding medicinal plants (Henrich, 2015; Wrangham, 2009). Rather than one agent having to both hunt and cook as in my model, instead imagine a group where some agents specialise in hunting and some other agents specialise in cooking. Interestingly, this division of labour makes our performance in certain skillsets remarkably cooperative (Henrich, 2015). The chef will only cook if the hunter can provide the meat and the hunter will only eat if the chef can cook adequately.

Perhaps modular psychology would only be necessary when an agent has to master a wealth of skillsets across a variety of domains for their own survival. To use the common phrase, perhaps in ancestral environments it was necessary for early humans to be modular in order to become a 'jack of all trades'. However, the intensification of division of labour that emerged approximately 10,000 years ago may have created a different niche for our psychological architecture to evolve to (Laland & O'Brien, 2011; Muthukrsihna & Henrich, 2016). Rather than being a 'jack of all trades', humans may have shifted to become a 'master of one'. Future modelling efforts should investigate how this division of labour may influence the selection pressure for domain-general or modular psychology.

In summary, domain-general psychology was sufficient to underlie the agent's ability to master skills across two similar domains while modular psychology was needed to specialise to the demands of two distinct domains. Modular cognition experienced a large shift to the contrasting demands of multiple distinct domains and was important in driving agent behaviour. Finally, it is worth noting that cognition and motivation coevolved to influence agent behaviour in unique ways which may be affected by drift. Future extensions to this model could make this a more complete account of skill acquisition, for example by adding a period of social learning or by considering a division of labour. However, my main aims of this model were to investigate how cognition and motivation coevolved to influence the emergence of asocial skills, and to highlight whether domain-general, partly modular, or fully modular psychology could uphold these skills. All else being equal, modular psychology may be necessary to uphold the unique wealth of skills that human cultures display.

5. References

Amodio, P., Boeckle, M., Schnell, A. K., Ostojíc, L., Fiorito, G., & Clayton, N. S. (2019). Grow smart and die young: why did cephalopods evolve intelligence?. *Trends in ecology & evolution*, *34*(1), 45-56. https://doi.org/10.1016/j.tree.2018.10.010.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*(3), 628–647. https://doi.org/10.1037/0033-295X.113.3.628.

Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1547), 1443-1450. https://doi.org/10.1098/rspb.2004.2746.

Billiard, S., & Alvergne, A. (2018). Stochasticity in cultural evolution: a revolution yet to happen. *History and philosophy of the life sciences*, *40*(1), 9. https://doi.org/10.1007/s40656-017-0173-y.

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biol*, *9*(7), e1001109. https://doi.org/10.1371/journal.pbio.1001109.

Burke, D. (2014). Why isn't everyone an evolutionary psychologist? *Front. Psychol*. *5*(910), 1-8. doi: 10.3389/fpsyg.2014.00910.

Charbonneau, M. (2016). Modularity and recombination in technological evolution. *Philosophy & Technology*, *29*(4), 373-392. https://doi.org/10.1007/s13347-016-0228-0.

Churchill, S. E. (1993). Weapon technology, prey size selection, and hunting methods in modern hunter-gatherers: implications for hunting in the Palaeolithic and Mesolithic. *Archeological Papers of the American*

Anthropological Association, *4*(1), 11-24. https://doi.org/10.1525/ap3a.1993.4.1.11.

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M. G., Perilloux, C., & Buss, D. M. (2010). Evolutionary psychology: Controversies, questions, prospects, and limitations. *American Psychologist, 65*(2), 110–126. https://doi.org/10.1037/a0018413.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences*, *107*(Supplement 2), 9007-9014. https://doi.org/10.1073/pnas.0914623107.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. Hirschfeld and S.Gelman (Eds.) *Mapping the mind: Domain specificity in cognition and culture* (pp. 84-116). New York: Cambridge University Press.

Da Silva, S. G., & Tehrani, J. J. (2016). Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *Royal Society open science*, *3*(1), 150645. https://doi.org/10.1098/rsos.150645.

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences,* 108(32), 13335-13340. DOI: https://doi.org/10.1073/pnas.1102131108.

Efferson, C., McKay, R., & Fehr, E. (2020). The evolution of distorted beliefs vs. mistaken choices under asymmetric error costs. *Evolutionary Human Sciences, 2*, E27. doi:10.1017/ehs.2020.25.

Fodor, J. A. (1983). *The modularity of mind*: *An essay on faculty psychology*. England, MIT press.

Frankenhuis, W. E., & Ploeger, A. (2007). Evolutionary psychology versus Fodor: Arguments for and against the massive modularity hypothesis. *Philosophical Psychology, 20*(6), 687-710. https://doi.org/10.1080/09515080701665904.

Gangestad, S. W., Haselton, M. G., & Buss, D. M. (2006). Evolutionary foundations of cultural variation: Evoked culture and mate preferences. *Psychological Inquiry, 17*(2), 75-95. DOI: 10.1207/s15327965pli1702_1.

Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual sequences. *Cognitive psychology*, *65*(2), 284-320. https://doi.org/10.1016/j.cogpsych.2012.04.002.

Henrich, J. (2004). Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case. *American Antiquity*, *69*(2), 197-214. doi:10.2307/4128416.

Henrich, J. (2015). *The secret of our success*: *How culture is driving human evolution, domesticating our species and making us smarter*. USA: Princeton University Press.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes C. (2018). Enquire within: cultural evolution and cognitive science. *Philosophical Transaction Royal Society. B 373*: 20170051. http://dx.doi.org/10.1098/rstb.2017.0051.

Heyes, C. (2019). Précis of Cognitive Gadgets: The Cultural Evolution of Thinking. *Behavioral and Brain Sciences, 42*, E169. doi:10.1017/S0140525X18002145.

Hughes, S. S. (1998). Getting to the point: evolutionary change in prehistoric weaponry. *Journal of Archaeological Method and Theory*, *5*(4), 345-408. https://doi.org/10.1007/BF02428421.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637-651. https://doi.org/10.1016/j.cognition.2013.09.001.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences*, *102*(5), 1803-1807. https://doi.org/10.1073/pnas.0408759102.

Laland, K. N. (1993). The mathematical modelling of human culture and its implications for psychology and the human sciences. *British Journal of Psychology*, *84*(2), 145-169. DOI: https://doi.org/10.1111/j.2044-8295.1993.tb02471.x.

Laland, K. N., & O'brien, M. J. (2011). Cultural niche construction: An introduction. *Biological Theory*, *6*(3), 191-202. DOI 10.1007/s13752-012-0026-6.

Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W., & Müller, K. R. (2019). Unmasking Clever Hans predictors and assessing what machines really learn. *Nature communications*, *10*(1096), 1-8. https://doi.org/10.1038/s41467-019-08987-4.

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences*, *114*(30), 7877-7883. https://doi.org/10.1073/pnas.1620743114.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in cognitive sciences*, *19*(11), 688-699. https://doi.org/10.1016/j.tics.2015.08.005.

Manoel, E. D. J., Basso, L., Correa, U. C., & Tani, G. (2002). Modularity and hierarchical organization of action programs in human acquisition of graphic skills. *Neuroscience Letters*, *335*(2), 83-86. https://doi.org/10.1016/S0304-3940(02)01102-3.

McElreath, M. B., Boesch, C., Kühl, H., & McElreath, R. (2018). Complex dynamics from simple cognition: The primary ratchet effect in animal culture. *Evolutionary Behavioral Sciences, 12*(3), 191–202. https://doi.org/10.1037/ebs0000117.

Mercado, E. III. (2008). Neural and cognitive plasticity: From maps to minds. *Psychological Bulletin, 134*(1), 109–137. https://doi.org/10.1037/0033-2909.134.1.109.

Mesoudi, A. (2008). An experimental simulation of the "copy-successful-individuals" cultural learning strategy: Adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*, *29*(5), 350-363. https://doi.org/10.1016/j.evolhumbehav.2008.04.005.

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334-342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001.

Mesoudi, A., Cheng, L., Murray, K., & Lu, H. J. (2015). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20142209. https://doi.org/10.1098/rspb.2014.2209.

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution?. *Proceedings of the Royal Society B*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712.

Miu, E., Gulley, N., Laland, K. N., & Rendell, L. (2020). Flexible learning, rather than inveterate innovation or copying, drives cumulative knowledge gain. *Science Advances*, *6*(23), eaaz0286. DOI: 10.1126/sciadv.aaz0286.

Molleman, L., & Gächter, S. (2018). Societal background influences social learning in cooperative decision making. *Evolution and Human Behavior, 39*(5), 547-555. https://doi.org/10.1016/j.evolhumbehav.2018.05.007.

Müller, G. B. (2007). Evo–devo: extending the evolutionary synthesis. *Nature reviews genetics*, *8*(12), 943-949. https://doi.org/10.1038/nrg2219.

Muthukrishna, M., Doebeli, M., Chudek, M., & Henrich, J. (2017). The Cultural Brain Hypothesis: How culture drives brain expansion, underlies sociality, and alters life history. *bioRxiv*, 209007. doi: https://doi.org/10.1101/209007.

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150192. https://doi.org/10.1098/rstb.2015.0192.

Muthukrishna, M., Morgan, T. J.H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, *37*(1), 10-20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483.

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, *92*, 216-229. https://doi.org/10.1016/j.chb.2018.11.027.

Pietraszewski, D., & A. E., Wertz (2021). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science* DOI: 10.1177/1745691621997113. [preprint].

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Social Justice Research*, *21*(2), 228-240. DOI: https://doi.org/10.1007/s11211-008-0064-1.

Reader, S.M. (2006). Evo-devo, modularity, and evolvability: Insights for cultural evolution. *Behavioral and Brain Sciences*, *29*(4), 361-362. doi:10.1017/S0140525X06399086.

Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. US: University of Chicago press.

Robinson, S., Orsingher, C., Alkire, L., De Keyser, A., Giebelhausen, M., Papamichail, K. N., Shams, P. & Temerak, M. S. (2020). Frontline encounters of the AI kind: An evolved service encounter framework. *Journal of Business Research*, *116*, 366-376. https://doi.org/10.1016/j.jbusres.2019.08.038.

Rorabaugh, A. N. (2014). Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *Journal of Archaeological Science*, *49*, 255-264. https://doi.org/10.1016/j.jas.2014.05.016.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nature Reviews Neuroscience*, *18*(9), 559-567. DOI: https://doi.org/10.1038/nrn.2017.76.

Stahl, A. B., Dunbar, R. I. M., Homewood, K., Ikawa-Smith, F., Kortlandt, A., McGrew, W. C., Milton, K., Paterson, J.D., Poirier, F.E., Sugardjito, J. Tanner, N.M. & Wrangham, R. W. (1984). Hominid dietary selection before fire [and comments and reply]. *Current Anthropology*, *25*(2), 151-168. https://doi.org/10.1086/203106.

Stephen, I. D. (2014). Putting the theory before the data: is "massive modularity" a necessary foundation of evolutionary psychology?. *Frontiers in psychology*, *5*, 1158. https://doi.org/10.3389/fpsyg.2014.01158.

Stokes, D., & Bergeron, V. (2015). Modular architectures and informational encapsulation: A dilemma. *European Journal for Philosophy of Science*, *5*(3), 315-338. https://doi.org/10.1007/s13194-015-0107-z.

Suchow, J. W., Bourgin, D. D., & Griffiths, T. L. (2017). Evolution in mind: Evolutionary dynamics, cognitive processes, and bayesian inference. *Trends in cognitive sciences*, *21*(7), 522-530. https://doi.org/10.1016/j.tics.2017.04.005.

Tomka, S. (2013). The Adoption of the Bow and Arrow: A Model Based on Experimental Performance Characteristics. *American Antiquity*, *78*(3), 553-569. doi:10.7183/0002-7316.78.3.553.

Truskanov, N., & Prat, Y. (2018). Cultural transmission in an ever-changing world: trial-and-error copying may be more robust than precise imitation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743), 20170050. https://doi.org/10.1098/rstb.2017.0050.

Van Lier, J., Revlin, R., & De Neys, W. (2013). Detecting cheaters without thinking: Testing the automaticity of the cheater detection module. *PloS one*, *8*(1), e53827. https://doi.org/10.1371/journal.pone.0053827.

Verbeke, P., & Verguts, T. (2019). Learning to synchronize: How biological agents can couple neural task modules for dealing with the stability-plasticity dilemma. *PLoS computational biology*, *15*(8), e1006604. https://doi.org/10.1371/journal.pcbi.1006604.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource?. *Psychological science*, *21*(3), 384-390. https://doi.org/10.1177/0956797610361340.

Vergauwe, E., von Bastian, C. C., Kostova, R., & Morey, C. C. (2021). Storage and processing in working memory: A single, domain-general resource explains multi-tasking. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000895. [preprint].

Whiten, A. (2019). Conformity and over-imitation: An integrative review of variant forms of hyper-reliance on social learning. *Advances in the Study of Behavior*, *51*, 31-75. https://doi.org/10.1016/bs.asb.2018.12.003.

Wrangham, R. (2009). *Catching fire: How cooking made us human.* UK: Profile Books.

6. Appendices

Appendix 1: The full parameter space of the asocial skills model

The bash file (*runSimulations*.txt)* found at OSF link (see Appendix 3) details the full parameter space controlled by the researcher. Most of these parameters were fixed. The parameters that I varied included:

- probStateAOne and probStateBOne variables (given by p_A and p_B in-text). These took one of 3 values: 0.1, 0.5 and 0.9.
- The fitness when one's skillset matches the environment. Note that if the agent displayed a different behaviour to the environment, then I standardised their fitness at 0. The fitness tied to displaying the skillset that matched the environmental state was given by the f_{Aone}, f_{Azero}, f_{Bone}, and f_{Bzero} parameters in-text. These took one of 3 values: 0.25, 1 and 4.
- The modularity of the agent type. I controlled these with the probFullMod, probTwoCog (for partly modular agents with modular cognition) and probTwoMotiv (for partly modular agents with modular motivation) respectively. These dummies took a 0 or 1 value. For example, if probFullMod was 1, then this run would be for all fully modular agents. If probFullMod was 0, then none of the agents would be fully modular and instead one of the other modularity types would be explored. Note that there is no variable modelled for domain-general agents. I modelled this so that, if probFullMod, probTwoCog and probTwoMotiv were all at 0, then all agents in this run would be domain-general by default. Whilst I consider runs to be unanimous across agent types, an interested researcher could force a split between all 4 agent types (for example, by coding probFullMod, probTwoCog and probTwoMotiv values to take 0.25 value each for an equal

split). Then the agent type that rises to fixation may be evolutionary dominant over the others.

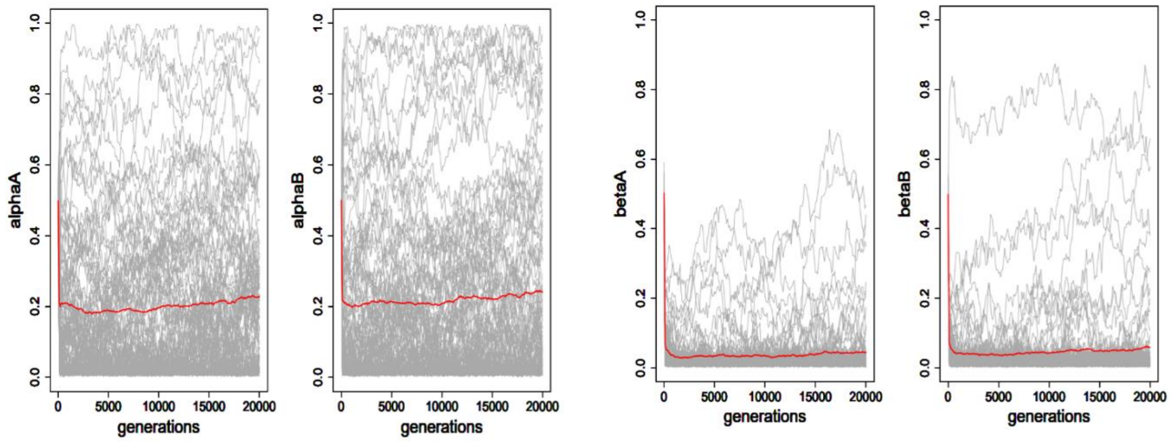
So, for each agent type I varied both the probability of the state being 1 and the fitness tied to matching one's skillset to the environment. Each of these variables took one of 3 values, and I varied 6 such variables. This gave 3^6 variables which gave 729 runs. I ran these for each of the 4 agent types, giving 2916 runs overall. These runs are too extensive to investigate in the appendices and so I merely note that I upload the data to a Dropbox folder repository if of interest but the analysis was only ran for the cases reported above for brevity:

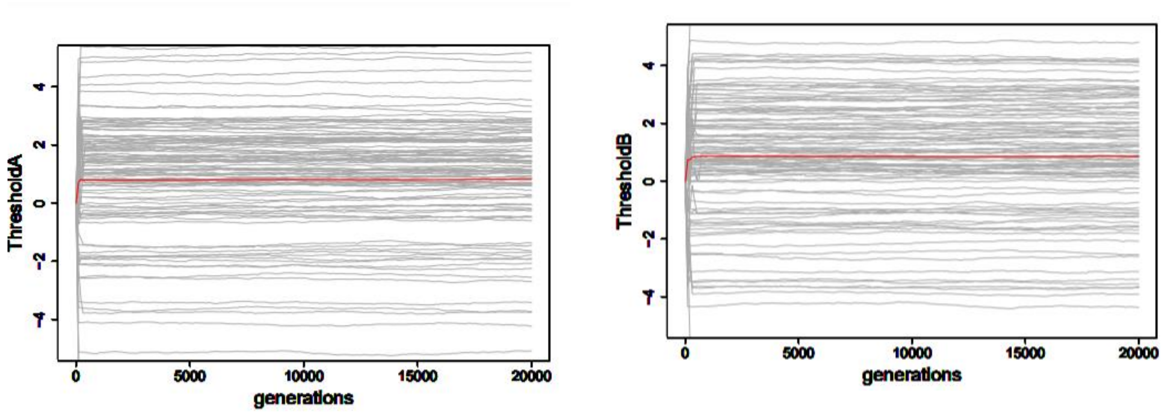
https://www.dropbox.com/sh/ojyhxk6cuks9hu1/AAD8CG_YOA3F3kYVRDw0tq8Ia?dl=0. There are multiple files uploaded, one for each of the parameter combinations. The files called *disagg** contain the data for the final generation of agents. These values were used to create the heatmaps in section 3.2, and the regressions in section 3.3. The file called *popDataAgentTypes** represented the cognitive and motivational values averaged over the four agent types of interest over the generations. This was used to make the line graph in appendix 2. The files called *popDataAgentTypesCorr** give the number of agents answering correctly throughout. This data was used to make the clustered barchart in section 3.1. Finally, the *popData** file gives data aggregated over the agent types but was unused in the analysis.

Appendix 2: The line graphs confirming that these runs converged on a stable psychological architecture.

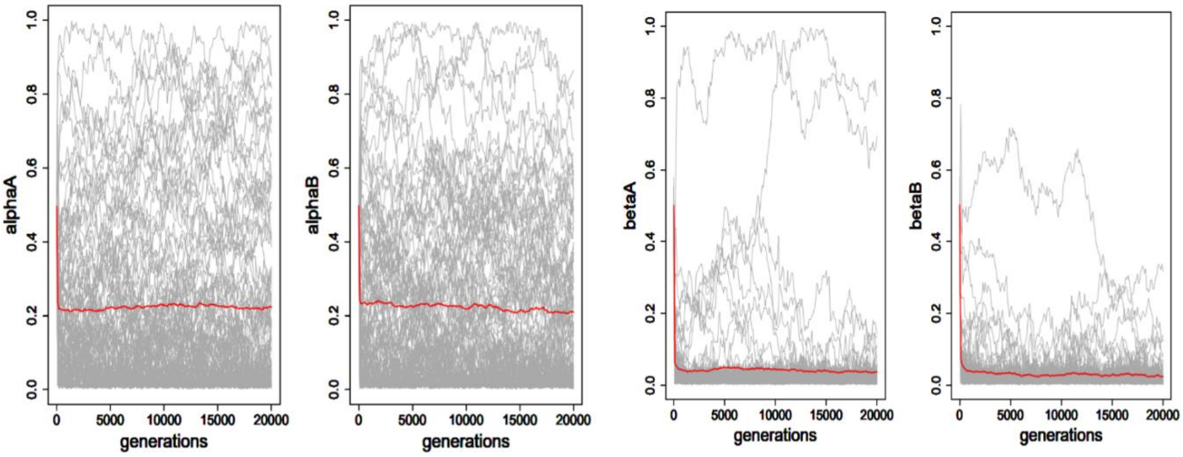
Appendix 2A: The line graphs for runs where the priors and fitness tied to matching skillsets to the environment favour environment 0 ($p_A = 0.1$, $p_B = 0.1$, $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$)

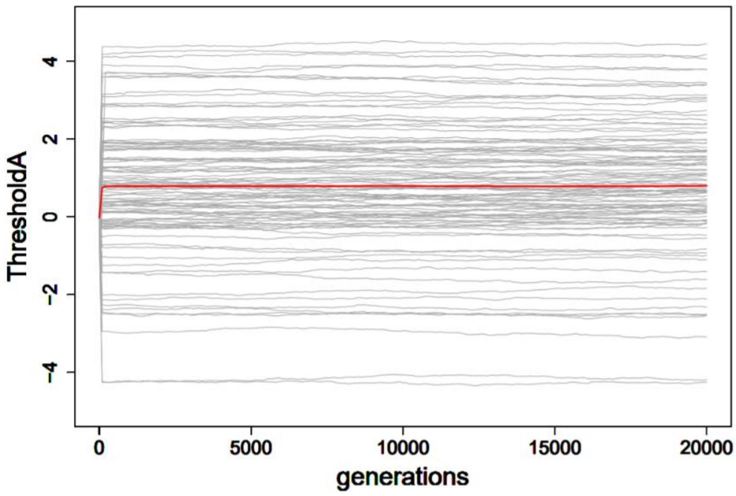
i) Fully modular agents



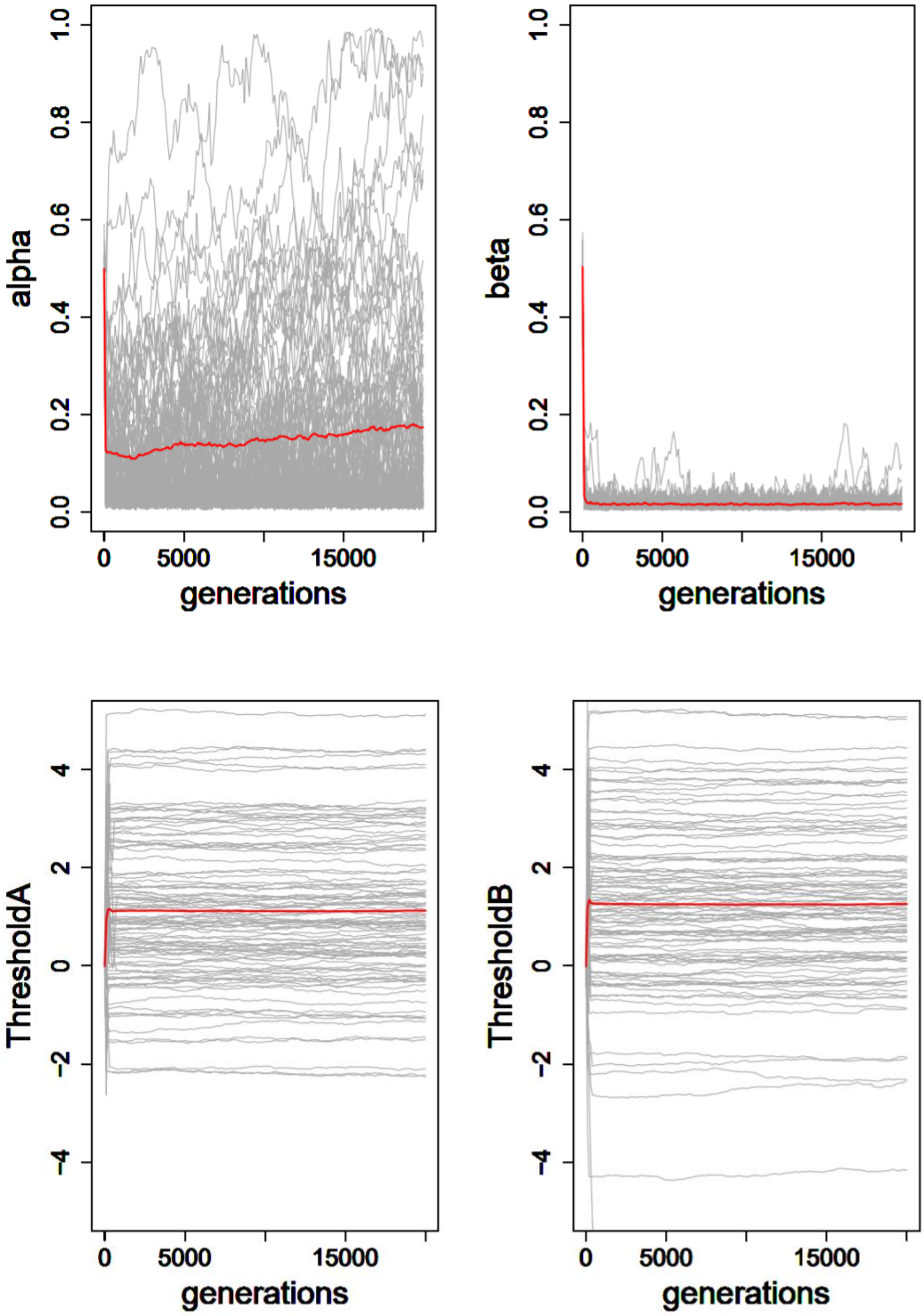


ii) Partly modular agents with modular motivation

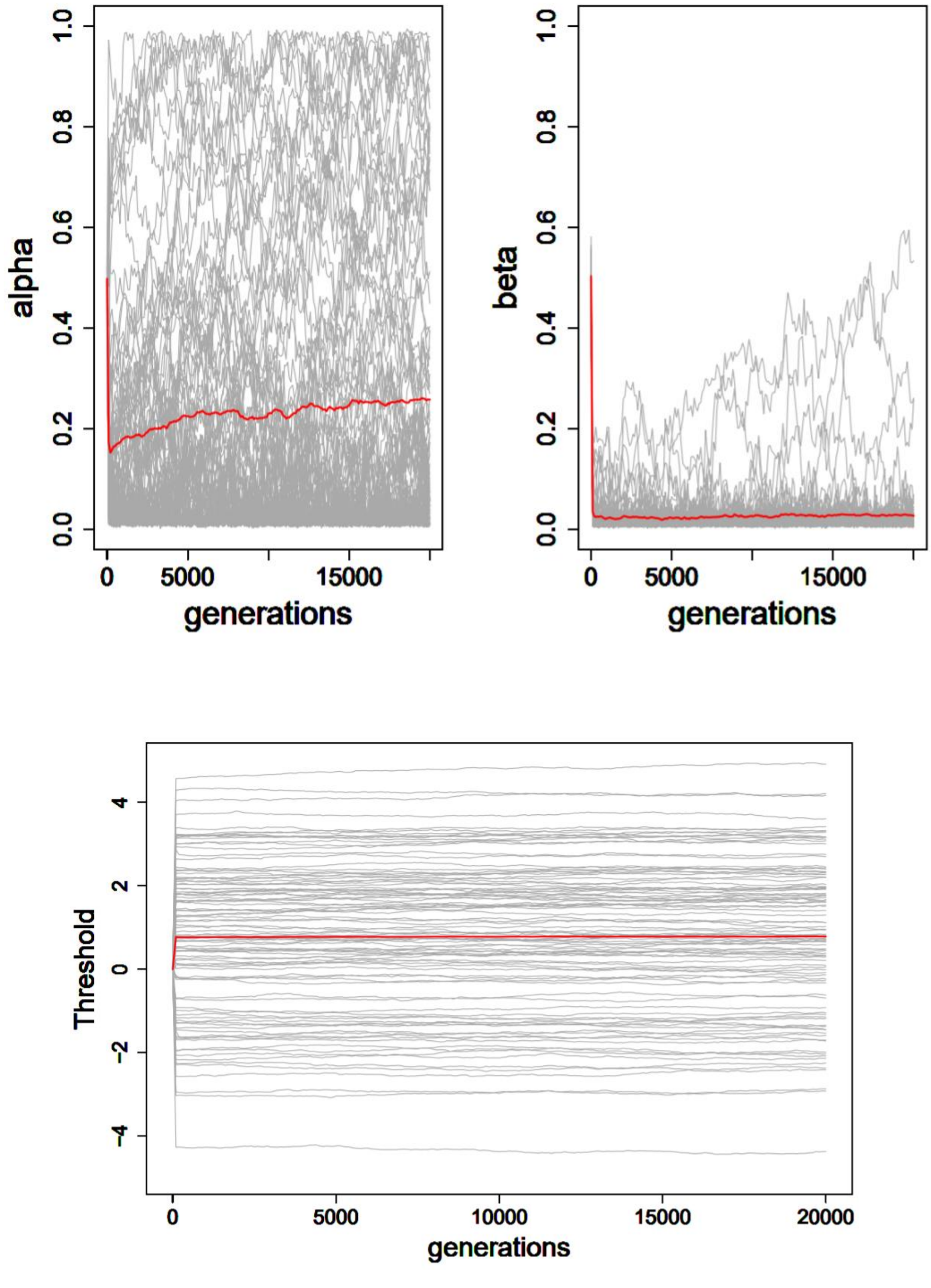




iii) Partly modular agents with modular cognition

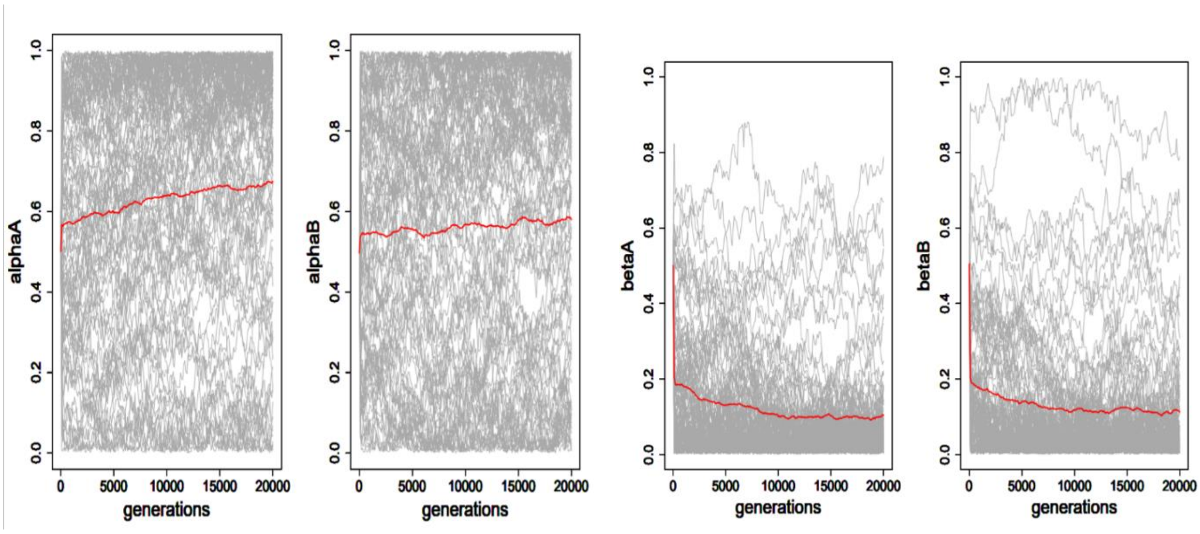


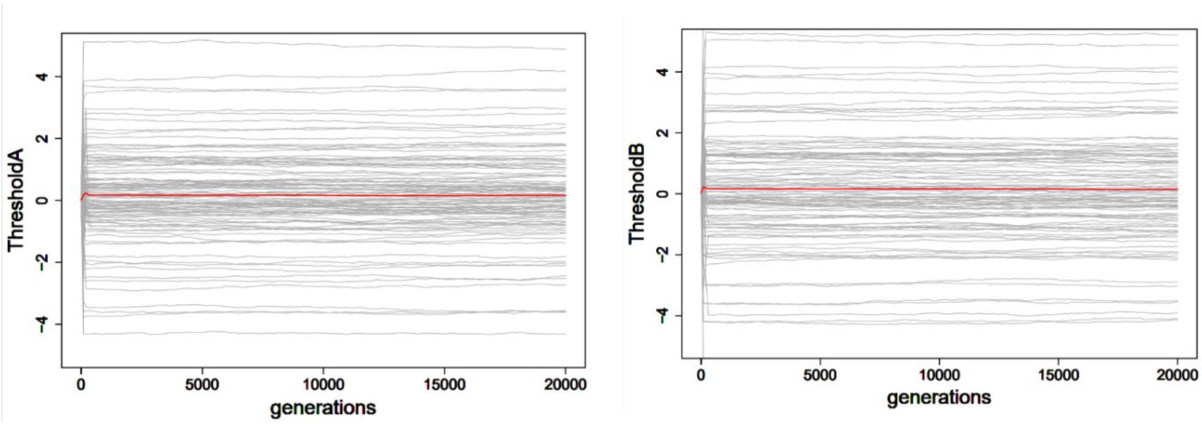
iv) Domain-general agents



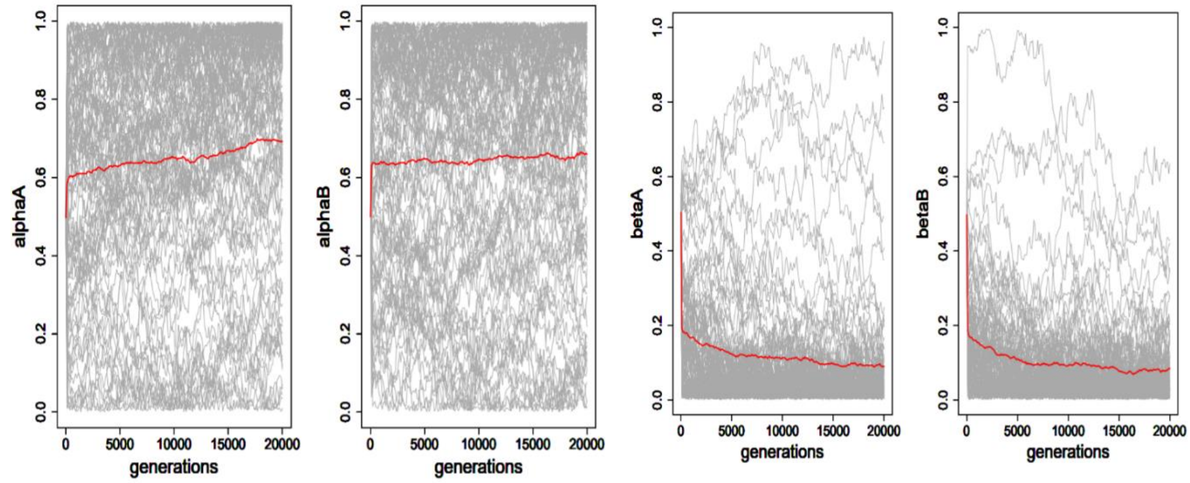
Appendix 2B: The line graphs for runs where the priors favour environment 0 but the fitness tied to matching skillsets to the environment favour environment 1 (p_A =0.9, p_B =0.9, f_{Aone} = 0.25, f_{Azero} = 4, f_{Bone} = 0.25, f_{Bzero} = 4)

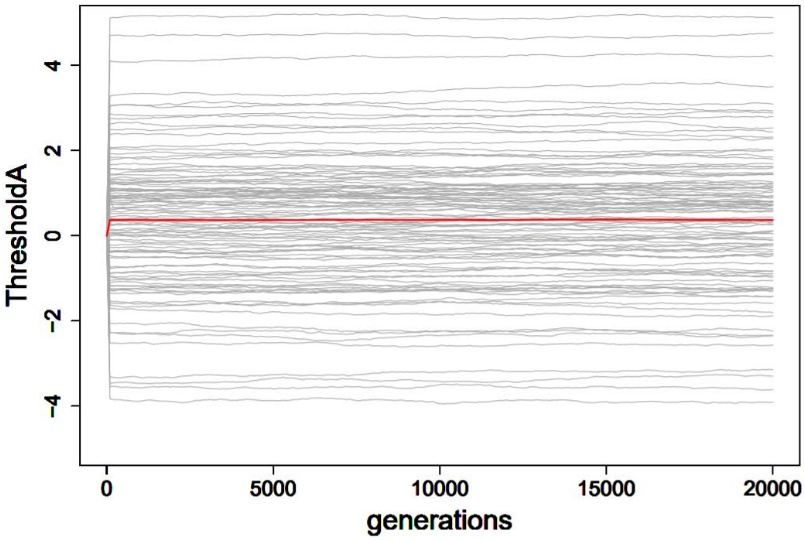
i) Fully modular agents



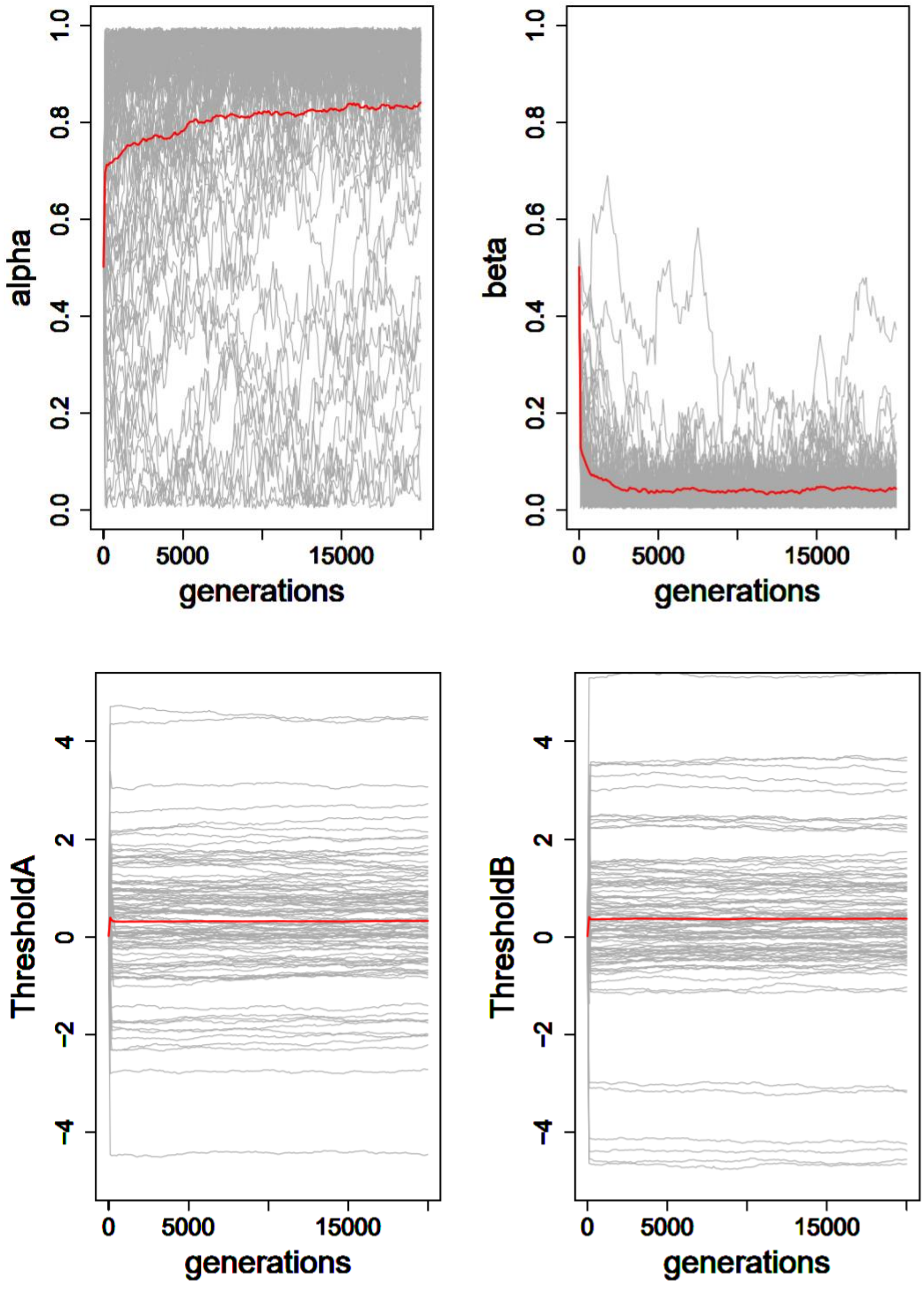


ii) Partly modular agents with modular motivation

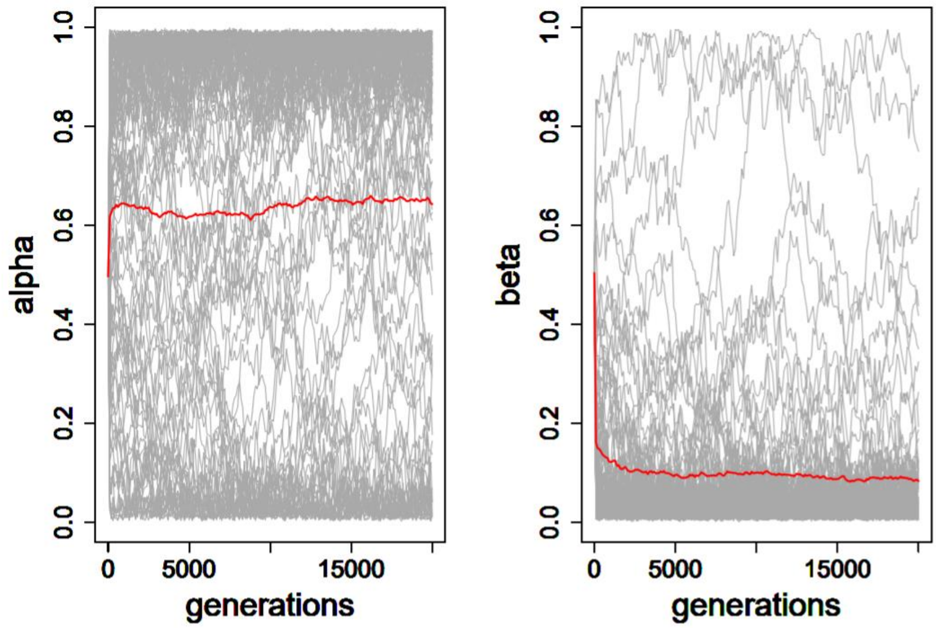


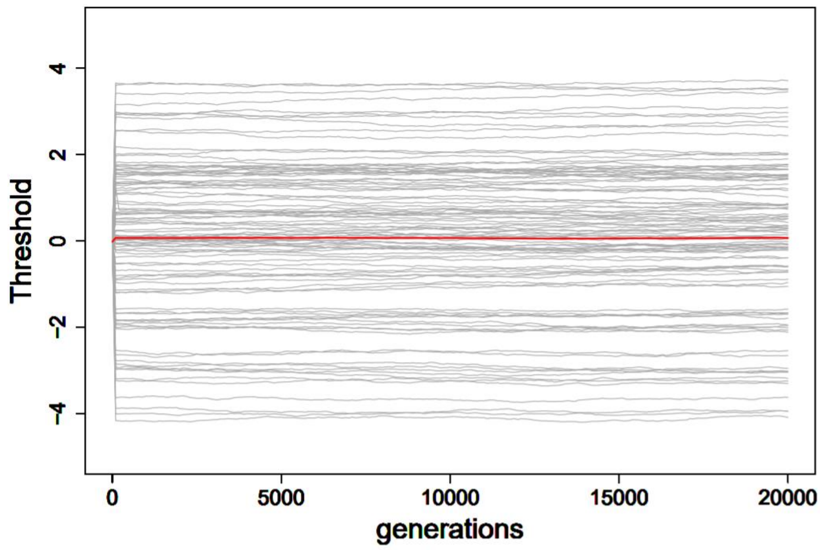


iii) Partly modular agents with modular cognition



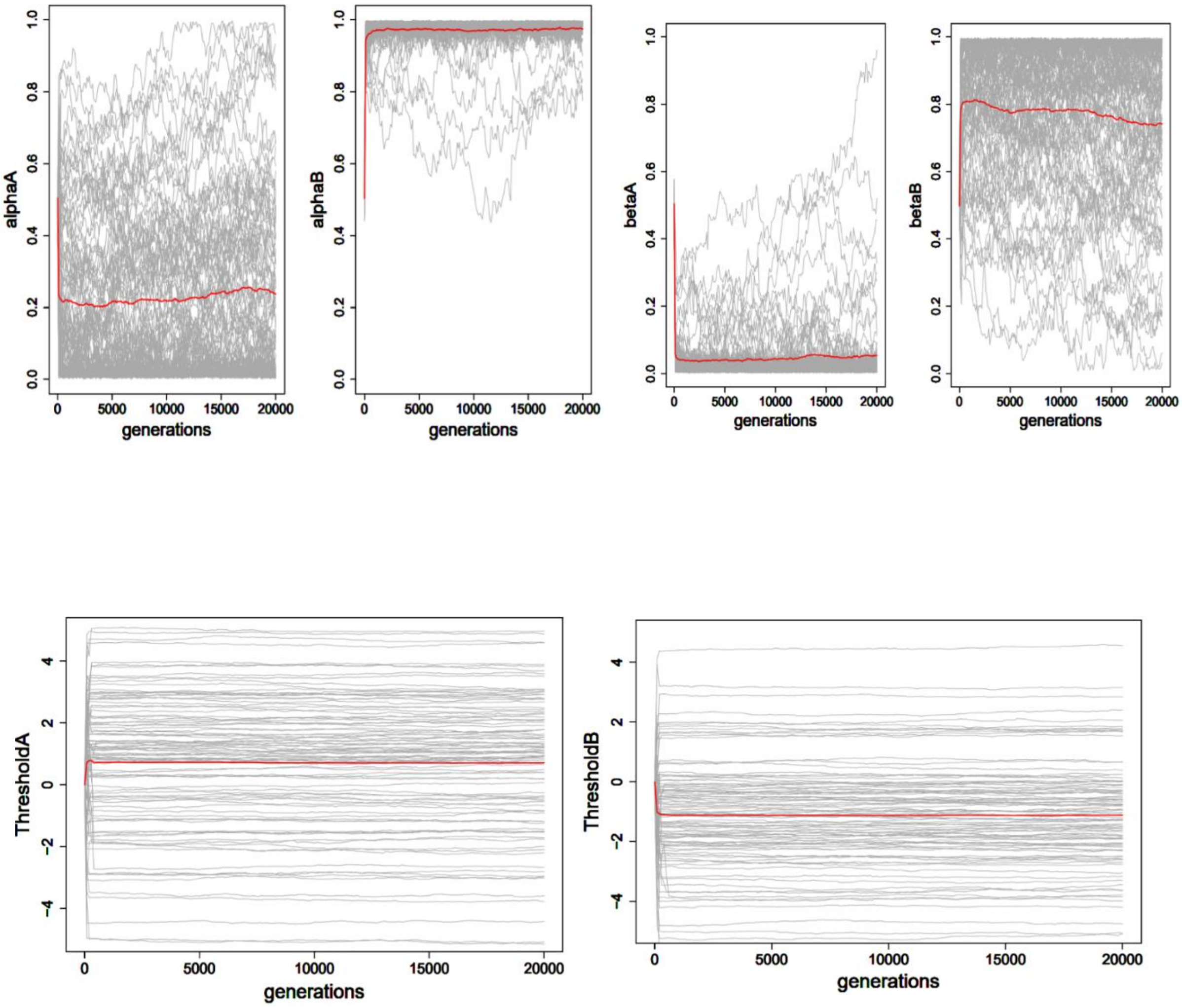
iv) Domain-general agents



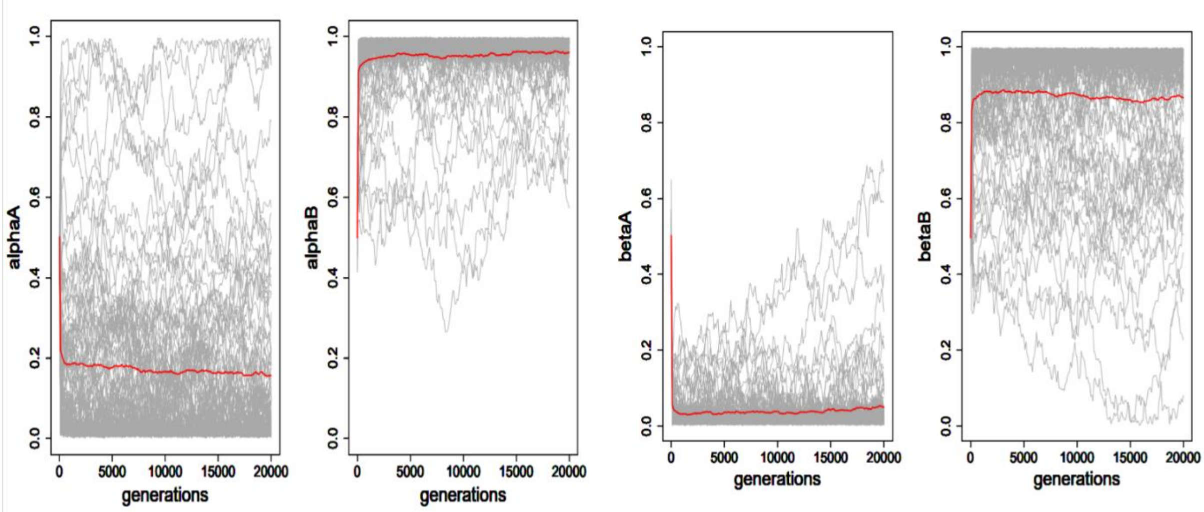


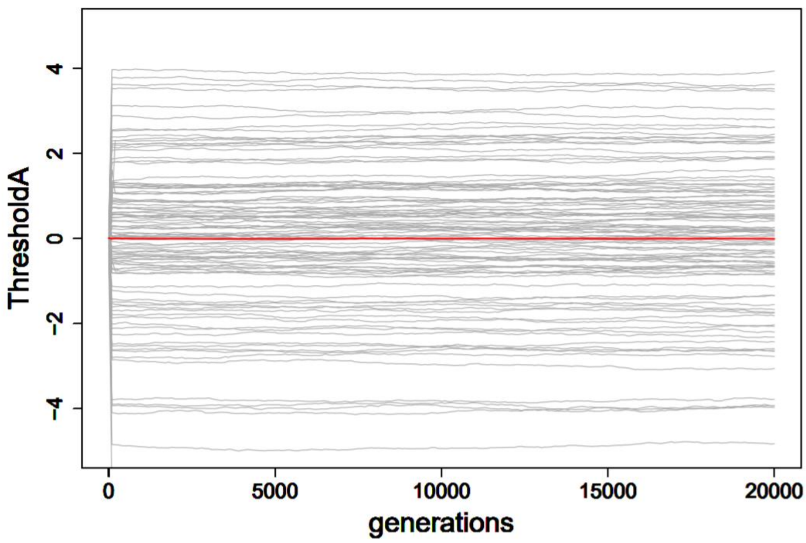
Appendix 2C: The line graphs for runs where the priors and the fitness tied to matching skillsets to the environment favour state 0 in domain A but state 1 in domain B ($p_A = 0.1$, $p_B = 0.9$, $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$)

i) Fully modular

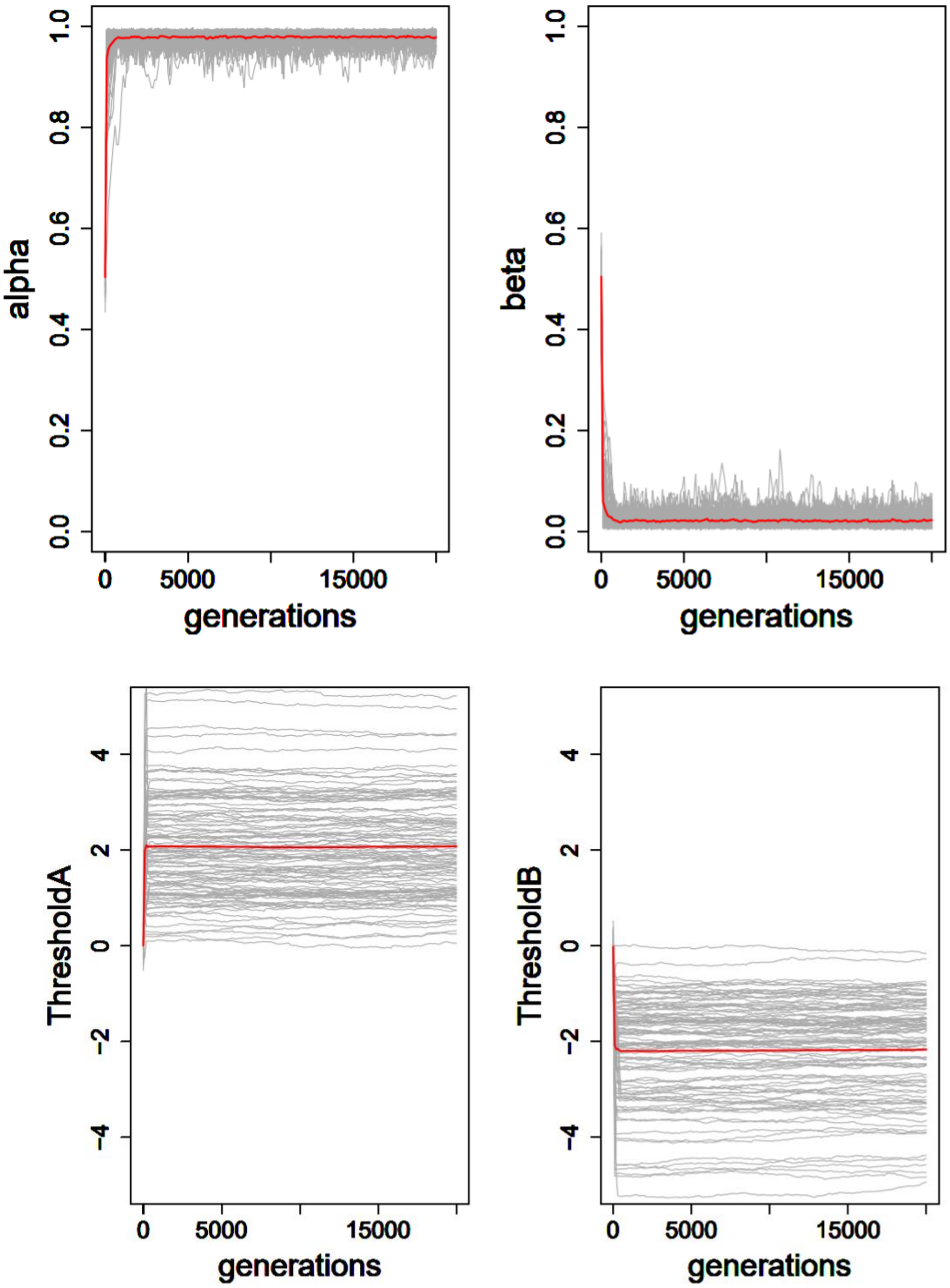


ii) Partly modular agents with modular motivation

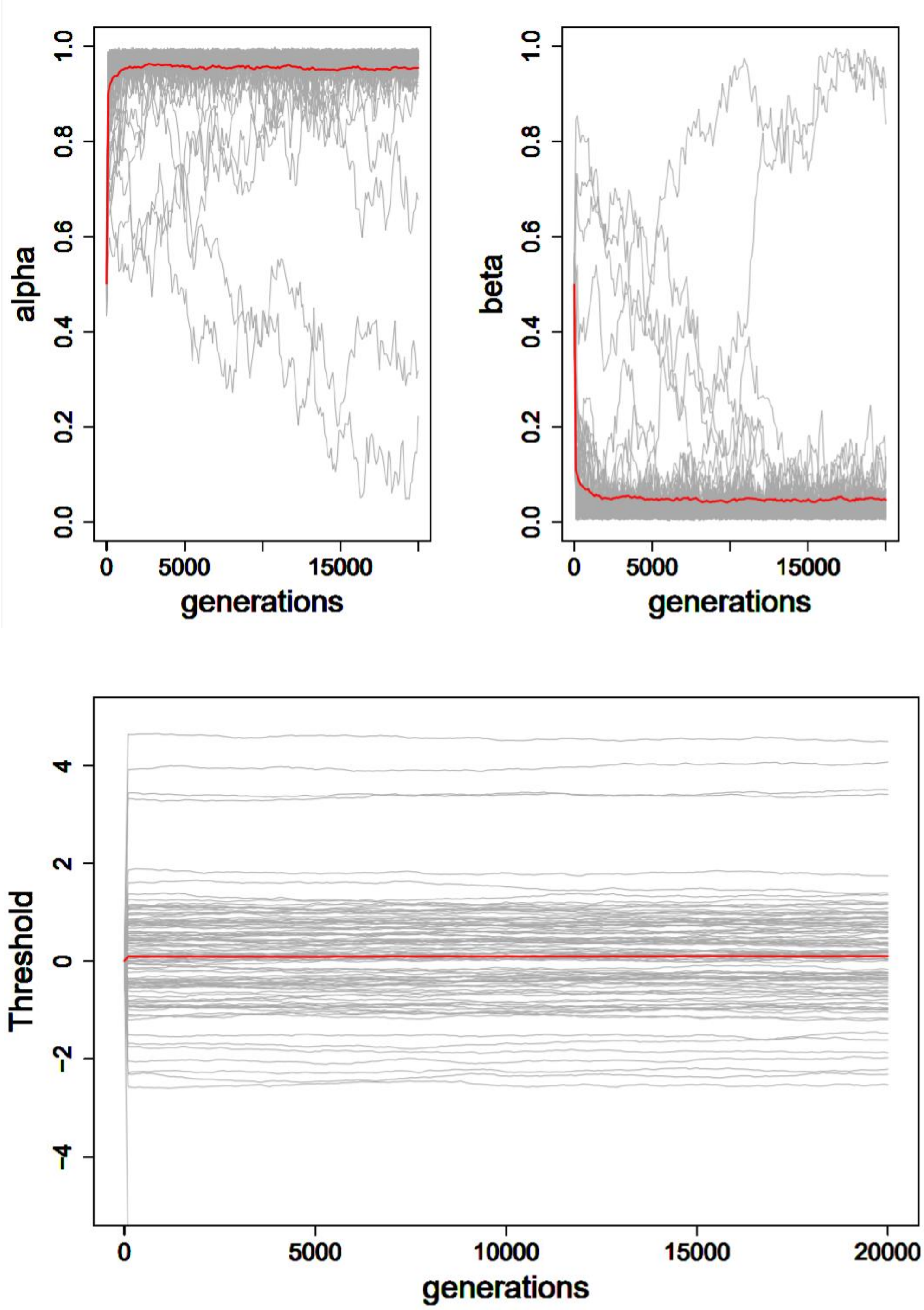




iii) Partly modular agents with modular cognition

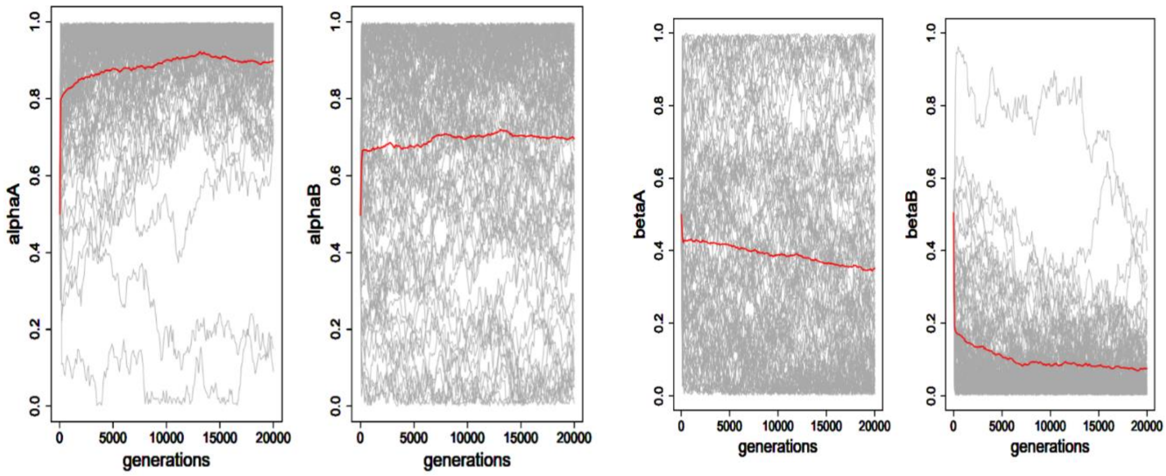


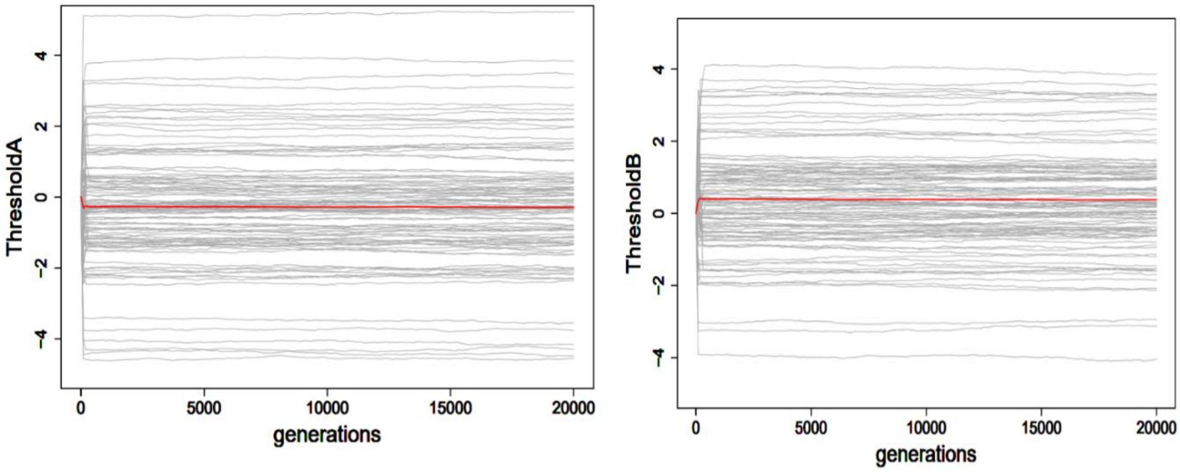
iv) Domain-general agents



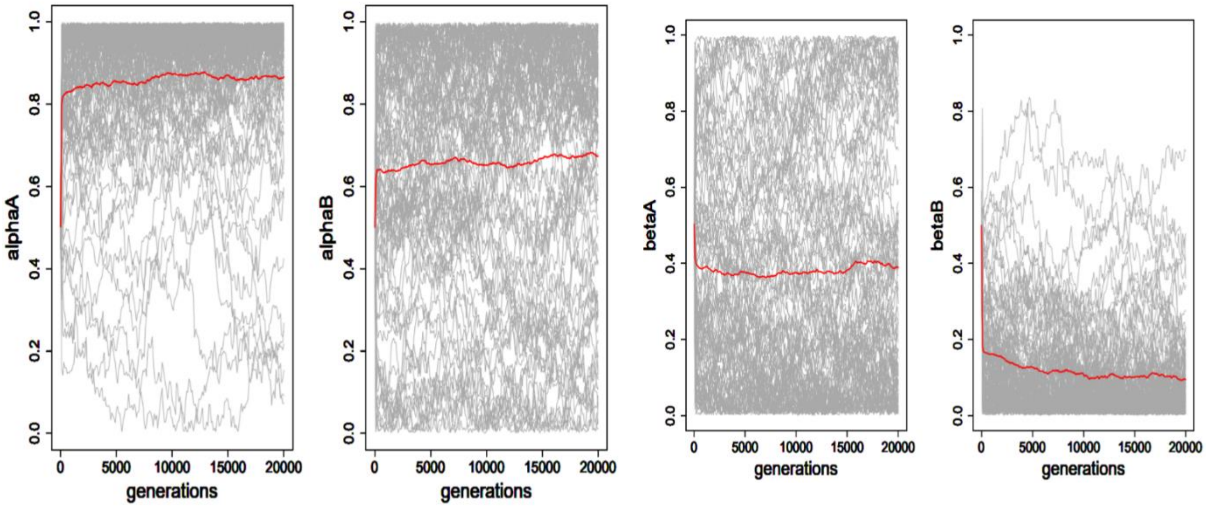
Appendix 2D: The line graphs for runs where the priors and the fitness tied to matching skillsets to the environment clash over both domains. In domain A state 0 is common though state 1 is more fit to match on. In domain B, state 1 is common though state 0 is more fit to match on ($p_A = 0.1$, $p_B = 0.9$, $f_{Aone} = 4$, $f_{Azero} = 0.25$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$).

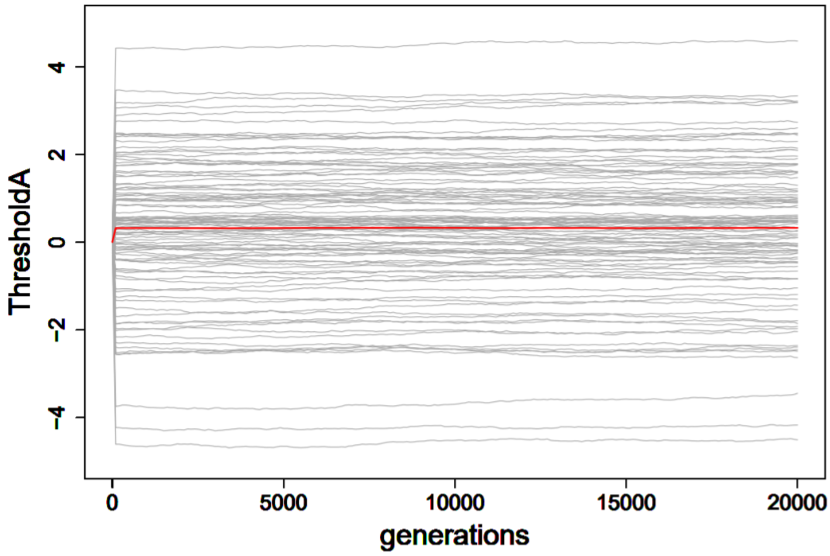
i) Fully modular agents



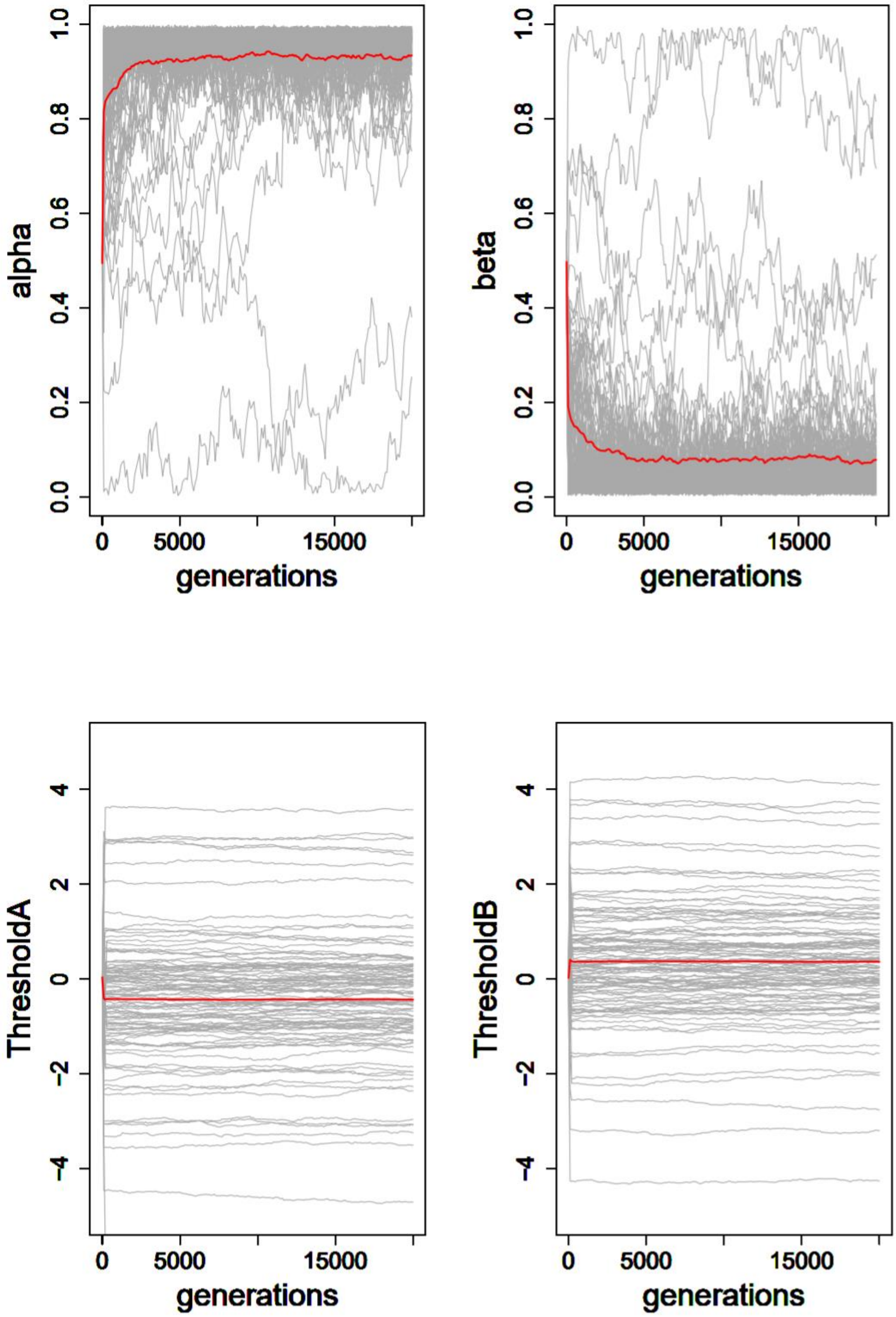


ii) Partly modular agents with modular motivation

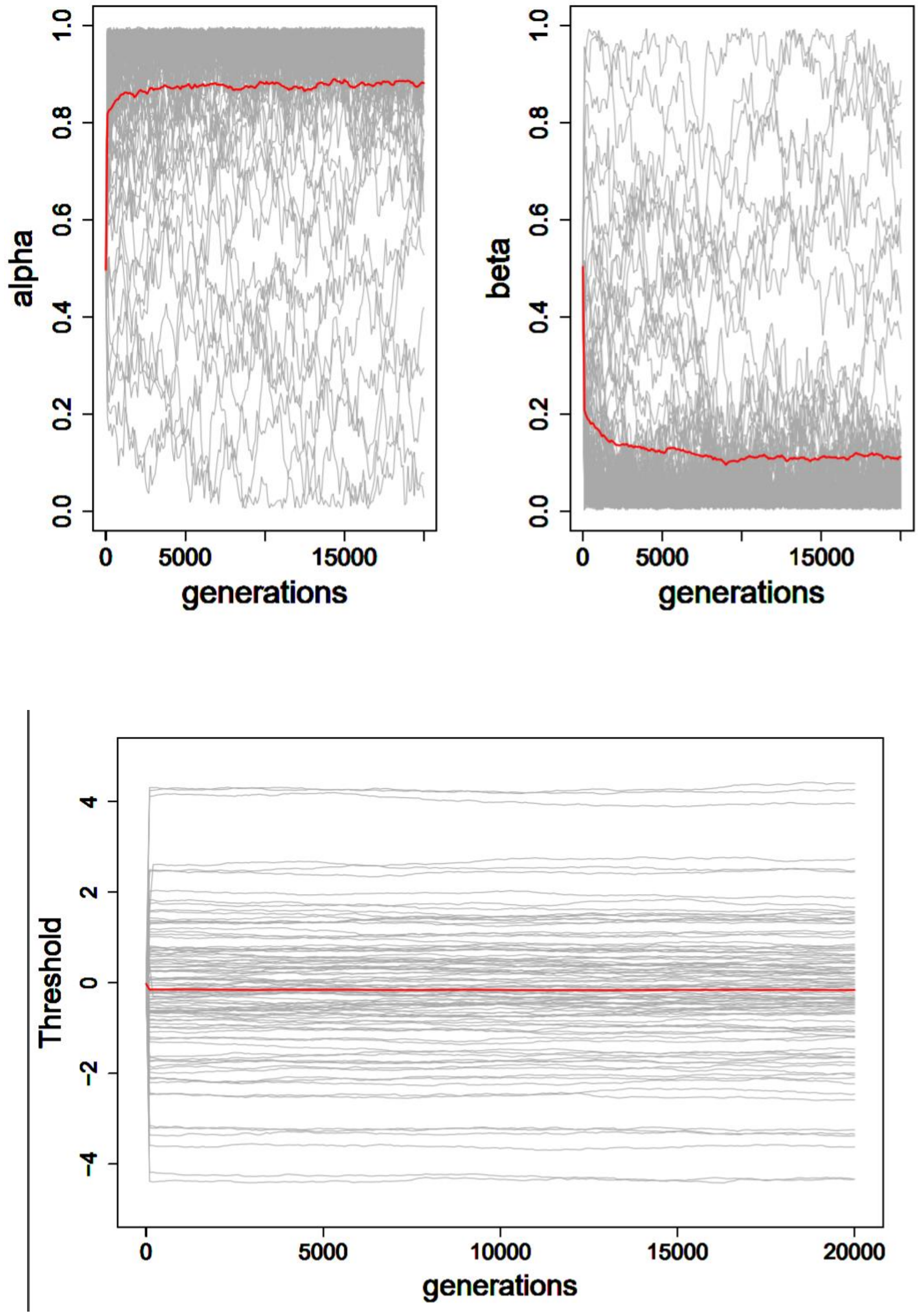




iii) Partly modular agents with modular cognition



iv) Domain-general agents



Appendix 3: The scripts used to run the models and analysis.

As my agent-based models and analysis scripts were divided into one of the four agent types (fully modular, partly modular agents with modular cognition only, partly modular agents with modular motivation only and domain-general agents), I divide my OSF repositories accordingly. Each repository contains the agent-based model (consisting of a h file, an m file and a main file to compile in Objective C), random number generators for Objective C, the bash file to run the simulations for each combination of the parameters of interest. Plus, each individual repository has some R analysis scripts: one for generating the line graphs in appendix 2, one for generating the clustered bar charts seen in section 3.1 and one for generating the heatmaps and regression by psychological components in sections 3.2 and 3.3. These are:

- Full modular agents:
  https://osf.io/xnj9s/?view_only=136842560d9240b48fdd918154375d07
- Partly modular agents with modular cognition only:
  https://osf.io/4xywe/?view_only=730cca4e8c964bf6974e2b1c97b9f08b
- Partly modular agents with modular motivation only:
  https://osf.io/nq3v4/?view_only=0e53cc7f0605464d841c4f9594a62e69
- Domain-general agents:
  https://osf.io/7gje3/?view_only=c97372d5304040c3b49e6ee7f5b9e1f5

Finally, I also include an overall repository where I attach the script for the regression reported in section 3.4 which compared each agent type: https://osf.io/mx8vz/?view_only=40984bf6f40e4654a4b78bbcc8f14fb5.

Appendix 4: The linear combinations performed between all the modular agent types.

These confirm the points made in-text. The fully modular agents and partly modular agents with modular cognition only often accrued more fitness than partly modular agents with modular motivation only. Fully modular agents and agents with modular cognition only were often indistinguishable in terms of fitness, supporting the conclusion that modular cognition is more important to skill learning.

Appendix 4. The linear combinations compare the fitness of all the modular agent types, for the regression reported in Table 3, section 3.4 of the main text. These linear combinations are run for all parameter combinations of interest.

Prior Probabilities	**Linear Combination**		
	Fully modular - modular motivation	**Fully modular - modular cognition**	**Modular motivation - modular cognition**
$p_A = 0.1$, $p_B = 0.1$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=4.19, p=0.04	F(1,39996)=0.16, p=0.688	F(1,39996)=2.30, p=0.13
$p_A = 0.9$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=0.52, p=0.47	F(1,39996)=3.65, p=0.057	F(1,39996)=5.64, p=0.02
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 4$, $f_{Bzero} = 0.25$	F(1,39996)=1.32, p=0.25	F(1,39996)=6.30, p=0.01	F(1,39996)=2.05, p=0.15
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 4$, $f_{Azero} = 0.25$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=4.49, p=0.03	F(1,39996)=2.02, p=0.15	F(1,39996)=8.97, p=0.003

Appendix 5: A regression predicting agent fitness based on whether they had modular or domain-general cognition, and modular versus domain-general motivation

This regression confirms that modular cognition always outperforms domain-general cognition as soon as the agent makes decisions over disparate domains with conflicting priors and fitness pressures. The fitness differences between having modular motivation and domain-general motivation are less clear cut, however. This again confirms my hypothesis that modular cognition may be important to the learning of asocial skills.

Appendix 5. The regression results displaying any differences between domain-general and modular cognition, versus domain-general and modular motivation, for each parameter combination. The domain-general agents were the omitted category of this regression as they were dummy coded as 0.

Prior Probabilities	**Estimates for regression predicting fitness**		
	Intercept	**Cognitive threshold dummy (0 = domain-general, 1 = modular)**	**Motivation threshold dummy (0 = domain-general, 1 = modular)**
$p_A = 0.1$, $p_B = 0.1$, $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	8.05 *** (0.016)	0.020 (0.018)	-0.02 (0.018)
$p_A = 0.9$, $p_B = 0.9$, $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	1.913 *** (0.014)	0.013 (0.016)	-0.038 * (0.016)
$p_A = 0.1$, $p_B = 0.9$, $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 4$, $f_{Bzero} = 0.25$	6.81 *** (0.018)	0.789 *** (0.021)	0.829 *** (0.021)
$p_A = 0.1$, $p_B = 0.9$, $f_{Aone} = 4$, $f_{Azero} = 0.25$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	1.88 *** (0.013)	0.05 *** (0.015)	-0.022 (0.015)

The asterisks denote the significance of our p values with the following key:
** = 0.05*
*** = 0.01*
**** <0.001*
. = trend

Chapter 6
It takes two: Modular psychology allows us to coordinate on various social norms including maladaptive ones

Aysha Bellamy
Charles Efferson
In preparation for publication

Word count: 15,330 excluding references and appendices

Abstract

It is unclear how individuals come to coordinate on social norms over distinct domains. For example, deciding when it is appropriate to dance and which side of the road to drive on. Moreover, the optimal social norm may not necessarily be the norm that most group members choose. I investigated whether agent cognition and/or motivation should be flexible across domains (i.e., domain-general) or should be specialised in each domain (i.e., modular) when acquiring social norms. I ran a model where agents tried to coordinate on two coordination games to mirror this uncertainty that individuals have when choosing their social behaviours over two distinct domains. I found that domain-general psychology could uphold coordination on two social norms with similar fitness landscapes. Modular psychology was instead needed for two distinct domains. When the most common environment is not necessarily optimal to coordinate in, then some simulations ended with agents coordinating on the most fit behaviour and others ended with agents coordinating on the common behaviour. Agents who coordinate on a norm simply because it is common may uphold a maladaptive social norm. To summarise, modular psychology allows us to coordinate on distinct social norms but cannot prevent coordination on maladaptive social norms.

Keywords: social norms, coordination, drift, maladaptive norms, modularity, domain-general

1. Introduction

Picture a conference delegate who is attending a meeting at Oxford. She dresses formally, drives to the conference centre and, upon arrival, accepts a champagne flute ready to toast her fellow speakers. While these discrete series of decisions may seem obvious to those of us that regularly attend conferences, the behaviour itself is highly arbitrary and driven by a series of social norms. Social norms are rules which guide our behaviour and are successful when more members of a social group coordinate on them (Gintis, 2004; Legare, 2017). Our conference delegate clearly has a norm regarding the conventionality of the occasion which influenced her choice of dress. She also has a norm regarding if and when alcohol consumption is appropriate at these events.

There are seemingly infinite social rules that guide behaviour. Beyond our example of a conference, there are norms regarding when to help others (Gintis et al., 2005; 2008; Henrich et al., 2001; Henrich & Muthukrishna, 2021; House, 2018), what behaviours are considered moral (Curtin et al., 2020; Gelfand et al., 2020; McNamara et al., 2019; Mrazek et al., 2013; Price, 2005; Purzycki et al., 2018) and how to signal one's belonging to a group (Smaldino et al., 2018; Sosis et al., 2007). Even deciding which side of the road to drive on is a social norm (Hao et al., 2017; Hindriks, 2019).

This breadth of social norms becomes even more complicated when one considers that societies can uphold arbitrary or maladaptive social norms (Boyd & Richerson, 2007). For example, social norms regarding female genital cutting (Efferson et al., 2020a), foot-binding (Gavrilets, 2020) and suicide trends (Mesoudi, 2009) are individually harmful and yet can be upheld at a group level. When the most fit behaviour is not actually the most common among the group, then the individuals may face a dilemma and perhaps uphold a maladaptive norm (Boyd & Richerson, 2007).

To make coordination even more complex, it is not necessarily clear which behaviour would be optimal to coordinate on for each social norm (Henrich & Muthukrishna, 2021). Consider a social norm regarding dance. Dancing can be ritualistic, such as prior to a ceremony or sporting event (Gelfand et al., 2020; Hartigan, 2011). There are also times when dancing would be inappropriate such as in formal work meetings. Individuals have to track whether dancing (and countless other norms) are beneficial over multiple contexts and so this model investigates how agents come to coordinate in two distinct social domains when uncertain, to reflect this complexity.

It is important to understand the psychological mechanisms that could uphold such a wealth of social norms. There are a multitude of decision-making domains which may be guided by social norms, which has led some researchers to conclude that our decision-making is domain-general (Heyes, 2012; 2018; Simon & Hessen, 2019).Domain-general psychological systems are those that remain flexible to many inputs (Kan et al., 2013; Vergauwe et al., 2010). Consider our conference delegate again. She decides to dress formally, drive to the conference, and propose a toast. All of these behaviours are sufficiently abstract and novel in the ancestral past that they are likely upheld by a flexible cognitive system.

This is in contrast to the predictions of Evolutionary Psychology, which argues that complex human behaviour is driven by massive modularity (Cosmides & Tooby, 1994). This means that human behaviour is driven by a series of modules, or cognitive processes that are each designed to work on one specific input to produce a certain output (Cosmides & Tooby, 1994). These modules may lead to cognitive adaptations that are designed to address recurrent issues that we faced throughout our ancestral past (Price, 2008). Complex social norms can be obtained via higher-order modules (Sperber, 1994; Sperber & Hirschfeld, 1999; 2004). These modules synthesise the input of many separate periphery processing modules and allow us to coordinate on complex social behaviour (Barrett & Kurzban, 2006; Sperber & Hirschfeld, 1999).

Social norms operate across a wealth of distinct domains (Henrich & Muthukrishna, 2021) which may suggest the need for a modular psychological system that can specialise per each domain (Charbonneau, 2016). Modularity need not be upheld by genetically selected cognitive biases alone. Indeed, the dominant concept of modularity in the evo-devo literature sees modules as psychological systems that can emerge throughout development and are shaped by the cultural institutions that one is embedded in (Müller, 2007; Reader, 2006). Researchers also do not need to commit to a viewpoint of full modularity (Cosmides & Tooby, 1994) or full domain-generality (Bolhuis et al., 2011). Human cognition can perhaps best be understood via a mixture of modular processes and domain-general learning rules (Chiappe & McDonald, 2005; Stephen, 2014).

While there has been a keen theoretical debate over this topic, it is necessary for models to investigate the emergence of social norms in agents with domain-general or modular psychology to understand which system is likely to uphold complex human social behaviour (Laker et al., 2021). This current model aims to address this gap in the literature by comparing how a group of agents with domain-general psychology versus modular psychology come to coordinate on social norms in two distinct domains. These domains may or may not provide contrasting selection pressures for coordination.

Defining modularity may be difficult based on previous critiques of this concept. There are many interpretations of a 'module' in literature (Fodor, 1983; Frankenhuis & Ploeger, 2007; Spunt & Adolphs, 2017; Stokes & Bergeron, 2015). To address this debate, I note that the defining feature of a module that most researchers agree on is *functional specialisation.* That is, modules are designed to work on one specific input alone (Pietraszewski & Wertz, 2021). This also corresponds to the notion of ‘independent evolvability’ of modules in evo-devo research (Müller, 2007). Building on these definitions, I model modular agents to have separate psychological systems that can specialise per each distinct domain, For domain-general mechanisms to guide decision-making over multiple domains, then it is necessary to consider the decision to ‘act’ or ‘not act’ as the highest level of abstraction over these domains.

Previous research has focused on the importance of cognitive modularity, though it is my view that motivation is just as important in driving behaviour (Delton et al., 2011; Öhman & Mineka, 2001; Tooby et al., 2006; Williams, 2020). To illustrate with an example, should our motivation to uphold a social norm regarding a painful and humiliating group initiation be the same as our motivation to uphold a social norm regarding which colour clothing to wear?

Previous models have investigated cognition and motivation by fixing one at an arbitrary baseline and allowing the other to evolve (Delton et al., 2011). Modellers likely did this to avoid psychological polymorphism. That is, leaving cognition and motivation to coevolve may affect behaviour in infinite ways. Previous models have had to focus on simplistic representations of human behaviour to avoid polymorphism (Laland, 1993). There is a trade-off between the simplicity of a model and the extent to which it captures complex and meaningful human behaviour, however (Kendal et al., 2018). When ancestral humans began to coordinate on group norms, then their cognitive and motivational systems would have evolved in-tandem. Previous models that fixed one psychological aspect and allowed the other to evolve may therefore make an oversimplification. I thus seek to give equal footing to cognitive and motivational systems by allowing the two to coevolve throughout this model’s run.

Finally, modelling a cognitive and motivational component of agent psychology allows me to test the case of a partly modular agent, between the extremes of full modularity (Cosmides & Tooby, 1994) and full domain-generality (Bolhuis et al., 2011). Partly modular agents may have modular cognitive thresholds (which can evolve separately to the demands of each of the two social norm domains) and domain-general motivation (which must be flexible across the demands of both domains). Alternatively, partly modular agents may have domain-general cognition (which must be flexible across the demands of both social norm domains) and modular motivation (which can evolve separately to the demands of each of the two domains).

To recap, I investigate the likely psychology of agents who aim to coordinate on two distinct social norms. This model can be thought of as having two novel aims:

1. To investigate how cognition and motivation coevolve to influence the agents' ability to coordinate on two distinct social norms.
2. To investigate whether fully domain-general, fully modular, or partly modular psychology (with modular cognition only or modular motivation only) is likely to uphold our decision of whether to coordinate when we are uncertain as to which behaviour is the optimal social norm over various domains.

2: Methods

2.1 Model description

A population of 100 agents are randomly assigned into a dyad. Each dyad plays two coordination games, labelled as domain A and domain B throughout. During these games, the agent chooses between one of two behaviours. Let the agents chosen behaviour be denoted by B_A for domain A and B_B for domain B. Both components could take the realisation $\in\{0,1\}$. These can broadly be thought of as the decision to act (0) or not to act (1).

The agent must try to anticipate what the partner with whom she has been paired will choose. This pressure to coordinate with one's partner makes the coordination game analogous to coordinating on a social norm (Efferson et al., 2008). If the agent does not choose the same option as her partner, then both parties fail to receive a payoff. I refer to this as 'miscoordination'. One option will be worth more points to both parties to coordinate on than the other. This is called the Pareto dominant or optimum (Kets et al., 2021), though in this chapter I refer to this outcome as 'coordinating on an optimal social norm'. If the agents coordinate on the other option, then this would be 'coordinating on a suboptimal social norm'.

The payoff of the agent's behaviour not only depends on her partner's behaviour but also the value of the environmental state. The state of the environment in domain A was represented by a random variable, S_A, and the state of the environment in domain B was represented by a random variable, S_B. Both had support $\in\{0,1\}$. These states affect which behaviour would be optimal to coordinate on. When the state is 1, then coordinating on behaviour 1 would be coordinating on the optimal social norm (i.e., acting is optimal) while coordinating on behaviour 0 would be coordinating on the suboptimal social norm (i.e., not acting is suboptimal). When the state is 0, then coordinating on 0 would be coordinating on the optimal social norm (i.e., not acting is optimal) and coordinating on 1 would be coordinating on the suboptimal social norm (i.e., acting). Miscoordination in either domain produces zero payoff. See table 1 for the parameters of the model that I run.

Table 1: The notations and variables used throughout the model

Symbol	Model description
$S_A, S_B \in \{0,1\}$	The environmental state in domain A or B respectively.
$B_A, B_B \in \{0,1\}$	The behaviour employed by the agent in domains A and B respectively.
$x_A, x_B \in (-\infty, +\infty)$	The cue summary drawn from the environment to help the agent decide which state the environment is likely to be in for both domains A and B respectively.
$s_A, s_B \in \{0,1\}$	The state that the agent believes the environment is in for both domains A and B respectively. These can be wrong.
$T_A, T_B \in (-\infty, +\infty)$	The cognitive threshold of evidence needed to believe that the state is 1 in domains A and B. (Note just *T* for domain-general agents).
$\alpha_A, \alpha_B \in [0,1]$	The agents' motivation to play behaviour 1 when they believe that the state is 1 for domains A and B respectively. (Note just α for domain-general agents).
$\beta_A, \beta_B \in [0,1]$	The agents' motivation to play behaviour 1 when they believe that the state is 0 for domains A and B respectively. (Note just β for domain-general agents).
$p_A, p_B \in \{0.1, 0.5, 0.9\}$	The probability that the state will be 1 in both domains A and B respectively.
$f_{Azero}, f_{Bzero} \in \{0.25, 1, 4\}$	The fitness that the agent gains when she coordinates on the suboptimal social norm with her partner in state 0 of domains A and B. This can take a low (0.25), intermediate (1) or high value (4) and can be assigned independently in both domains A and B.
$f_{Aone}, f_{Bone} \in \{0.25, 1, 4\}$	The fitness that the agent gains when she coordinates on the suboptimal social norm with her partner in state 1 of domains A and B. This can take a low (0.25), intermediate (1) or high value (4) and can be assigned independently in both domains A and B.

To illustrate the logic of this coordination game, I shall walk through an in-depth example. Imagine domain A reflects a social norm regarding dance. Let $B_A = 0$ represent the decision not to dance and $B_A = 1$ represent the decision to dance. That is, $B_A = 1$ always represents the decision to act whilst $B_A = 0$ represents the decision not to act. As this is the first comparison of modular and domain-general decision-making in a population of theoretically-evolving agents, we take the simplest representation of a domain-general system as one that must decide to act or not act over multiple domains (Pietraszewski & Wertz, 2021). This should be seen as the first step towards a full comparison of domain-genneral and modular agents.

Let $S_A = 1$ denote a context where dancing is common, such as during a group ritual (Legare, 2017). Cases where $S_A = 1$ and $B_A = 1$ thus represent the agent acting optimally (i.e., 'dancing as a ritual'). Coordinating on this ritual is likely to increase one's wellbeing and social standing (Monteiro & Wall, 2011). Cases where $S_A = 1$ and $B_A = 0$ instead imply that the agent fails to act optimally (i.e., she 'fails to dance as ritual'). This may occur negative fitness consequences, such as ostracism for failing to uphold the group's norms (Gintis, 2003; van den Berg et al., 2012).

Let $S_A = 0$ represent a case where dancing is inappropriate, such as during a work meeting. Cases where $S_A = 0$ and $B_A = 0$ imply that the agent avoids acting inappropriately (i.e, 'avoids dancing inappropriately'). Cases where $S_A = 0$ and $B_A = 1$ instead imply that the agent acts inappropriately (i.e., 'dances inappropriately'). This may attract negative attention (Mu et al., 2015). The value of the environmental state determines which of these behaviours will be the optimal social norm to coordinate on. Figure 1 gives the formal payoff matrix of a coordination game using this example.

State = 0: Dancing would be inappropriate

		Partner's behaviour	
		0 (dances inappropriately)	**1 (avoids dancing when inappropriate)**
Focal agents' behaviour	**0 (dances inappropriately)**	$f_{Azero} \in \{0.25, 1, 4\}$	0
	1 (avoids dancing when inappropriate)	0	4

State = 1: Dancing would be a ritual

		Partner's behaviour	
		0 (Fails to dance as ritual)	**1 (Dancing as ritual)**
Focal agents' behaviour	**0 (Fails to dance as ritual)**	$f_{Aone} \in \{0.25, 1, 4\}$	0
	1 (Dancing as ritual)	0	4

Figure 1. The payoff matrix for the coordination game in domain A (the dancing domain). Note that the payoff to coordinating on the suboptimal social norm is decided by the f_{Azero} and f_{Aone} parameter.

The payoffs of miscoordination are standardised at 0 and the payoffs of coordinating on the optimal social norm are standardised at 4. Of course, failing to uphold a social norm could also result in negative fitness payoffs (e.g., ostracism, exclusion, attack; van den Berg et al., 2012). In my model, fitness payoffs were always positive as this allowed for reproduction to be tied to cumulative fitness.Thus, the pressure to coordinate in this model is driven entirely by the size of the f_{Azero}, f_{Aone}, f_{Bzero} and f_{Bone} parameters. These represent the payoffs to coordinating on the suboptimal norm and could take realisations $\in\{0.25, 1, 4\}$ respectively. When these are set to 0.25, then there is a large pressure to coordinate on the optimal social norm due to the large payoff difference. Put simply, there is a cost to coordinating on the suboptimal norm. To illustrate with an example, there is a strong selection to coordinate on not dancing when in a formal work meeting, as dancing is likely to be reprimanded.

Alternatively, if the payoff to coordinating on the suboptimal social norm is 4, then there is no difference in the payoffs between coordinating on the ‘suboptimal’ and the ‘optimal’ option. Thus, there would be no particular selection pressure to play behaviour 0 or 1. To illustrate in reference to domain A, imagine a club. The agent may dance on the dance floor or may not dance and join a conversation at the lounge. Either way, the behaviour is likely to be positive. Note that in this case, there are no costs to suboptimal coordination. As behaviour 0 and 1 are expected to give equal payoffs, then the agents do not necessarily need to coordinate on one behaviour over the other.

A similar payoff structure applies to domain B. Domain B may be a similar social norm to domain A, or completely distinct. For a similar example, imagine that domain B represents the decision to chant as a ritual. There is likely to be an overlap between the states where dancing and chanting are ritualistic (Gelfand et al., 2020).

As an example of a distinct domain, imagine that domain B represents the decision of which side of the road to drive on (Hao et al., 2017; Hindriks, 2019). Let $B_B = 0$ represent the decision to drive on the left-hand side of the road and $B_B = 1$ represent the decision to drive on the right-hand side of the road. Let us arbitrarily consider driving on the right-hand side of the road ‘acting’ in this example. When $S_B = 1$ and $B_B = 1$, then the agent successfully acts, or drives on the right-hand side of the road when most other drivers do (e.g., in the US). Let us refer to this as ‘successful right-hand driving’. When $S_B = 1$ and $B_B = 0$, then the agent does not act optimally; they drive on the opposite side of the road to the norm and thus risks a crash. Let us refer to this as ‘risky left-hand driving’. When $S_B = 0$ and $B_B = 0$, then the agent successfully avoids acting; they avoid driving on the right-hand side of the road when most other drivers avoid this (e.g., in the UK). Let us refer to this as ‘successful left-hand driving’. When $S_B = 0$ and $B_B = 1$, then the agent acts unsuccessfully (i.e., they drive on the wrong side of the road and risk a crash). Let us refer to this as ‘risky right-hand driving’.

If me and my partner both drive on the left-hand side of the road when everyone else drives on the right, then we risk a crash. This domain therefore has a high fitness pressure to coordinate on the optimal social norm ($f_{Bzero} = 0.25$ and $f_{Bone} = 0.25$). However, there still may be cases when there is no expected difference in fitness between driving on either side of the road. For example, when there is an obstruction blocking some of the road and so all drivers temporarily shift to driving on the wrong side ($f_{Bzero} = 4$ and $f_{Bone} = 4$; see Figure 2).

State = 0: Driving on the left

		Partner's behaviour	
		0 (Risky right-hand driving)	**1 (Successful left-hand driving)**
Focal agents' behaviour	**0 (Risky right-hand driving)**	$f_{Bzero} \in \{0.25, 1, 4\}$	0
	1 (Successful left-hand driving)	0	4

State = 1: Driving on the right

		Partner's behaviour	
		0 (Risky left-hand driving)	**1 (successful right-hand driving)**
Focal agents' behaviour	**0 (Risky left-hand driving)**	$f_{Bone} \in \{0.25, 1, 4\}$	0
	1 (successful right-hand driving)	0	4

Figure 2. The payoff matrix for the coordination game when domain B is distinct to domain A (the driving domain). Note that the payoff to coordinating on the suboptimal social norm is decided by the f_{Bzero} and f_{Bone} parameter.

In order to decide whether she wishes to dance, or drive on the right-hand side of the road, then the focal agent must first formulate a belief about the state that she thinks the environment is in. Let the agent's belief about the state of the environment in domain A be represented by s_A, and her belief about the state of the environment in domain B be represented by s_B. These could take support $\in\{0,1\}$. The agent either believes that dancing is a ritual ($s_A = 1$), or she believes that dancing would be inappropriate in the current scenario ($s_A = 0$). Likewise, the agent either believes that driving on the right-hand side of the road is common ($s_B = 1$) or that driving on the left-hand side of the road is common ($s_B = 0$). These beliefs about the environmental state can be wrong and so the agents made decisions regarding which social norm to coordinate on under uncertainty.

To formulate her belief regarding the likely state of the environment, each agent draws a cue summary in domains A and B. This cue summary can be thought of as all the evidence that the agent can gather from the environment to calculate whether a behaviour is a social norm. For example, signs that dancing is a social norm would include if all the group members dance together, if the movements are synchronised and if the dance has a symbolic meaning or takes part before an important event (e.g., the haka before playing rugby; Hartigan, 2011).

For simplicity, imagine that all of these cues are represented by one value– the cue summary– in my model. The cue summary is probabilistically– but not deterministically– decided by the value of the environmental state. When $S_A = 1$ or $S_B = 1$, then the agent draws a cue summary from a normal distribution with a Mμ =1, σ =1. When $S_A = 0$ or $S_B = 0$, then the agent draws a cue summary from a normal distribution with a Mμ = -1, σ =1. Let the agent's cue summary in this model be represented by x_A in domain A and x_B in domain B. These could be any random value from the respective distribution curves. The difficulty of distinguishing between these environments was driven by the overlap of the two distribution curves (~31.7%). Whilst the agent could usually form accurate beliefs about the environment, she could occasionally be wrong about the environmental state.

Once the agent has drawn her cue summary (x_A or x_B), she must then translate this into her beliefs about the environmental state in each decision-making domain (s_A or s_B). To do this, the agent compares her cue summary to her cognitive threshold. This is the minimum amount of evidence that an agent needs from the environment to conclude that the state is 1. These thresholds can take any positive or negative value. Let the cognitive threshold in domain A be represented by T_A and the cognitive threshold in domain B be represented by T_B. When $x_A > T_A$, then the agent believes that dancing is a ritual ($s_A = 1$). When $x_A <= T_A$, then the agent does not have enough evidence to believe that dancing is a ritual and thus believes that dancing would be inappropriate ($s_A = 0$). When $x_B > T_B$, then the agent believes that driving on the right-hand side of the road is the norm ($s_B=1$). When $x_B <= T_B$, then the agent does not have enough evidence that driving on the right-hand side of the road is the norm and so they believe that driving on the left side is the norm ($s_B=0$).

The cue summaries were modelled so that 0 would represent unbiased cognitive thresholds. An agent that evolved positive cognitive thresholds thus needed more evidence from the environment before she would believe that the state was 1. Conversely, an agent who evolved negative cognitive thresholds was less discerning and needed less evidence from the environment to believe that the state was 1.

An agent with modular cognition can reach different conclusions about the likelihood of the state being 1 in both domains. This is because T_A and T_B take separate values in both domains and evolve to the unique selection pressures of either domain. Instead, agents with domain-general cognition can only reason generically about the likelihood that the state is 1 over both domains. I achieve this by constraining domain-general cognition so that $T = T_A = T_B$.

Once the agent has formulated her beliefs about the appropriateness of dancing or driving on the right-hand side of the road in each domain, she can then only play a certain behaviour if she is motivated to do so. Each agent drew a random number form the uninform interval [0,1] and compared this with an internal motivational threshold. Imagine that the agent's cue summary exceeds her cognitive threshold ($x_A > T_A$ or $x_B >$

T_B) and she thus believes that the state is 1 ($s_A = 1$ or $s_B = 1$). Whenever the agent believed that the environmental state was 1 ($s_A = 1$ or $s_B = 1$), then they would play behaviour 1 with a motivation of $Probability_{State\ 1}$ (which I label α for future reference). Whenever this random number exceeded the α threshold, then the agent would play behaviour 1. If this random number did not exceed α, then the agent would play behaviour 0. The probability of playing behaviour 0 when the agent believed that the environmental state took the value 1 was therefore given by $1- \alpha$.

To illustrate with reference to domain A, the agent believes that dancing would be ritualistic, and that driving on the right-hand side of the road is common. The agent is then motivated to dance with a threshold of α_A and is motivated to drive on the right-hand side of the road with a threshold of α_B. Alternatively, the agent may play the behaviour which is likely to be suboptimal with a motivation of $1- \alpha_A$ and $1- \alpha_B$ respectively. Thus, α_A and α_B values represent the agent's motivation to play behaviour 1 when she believes that coordinating on behaviour 1 is to coordinate on the optimal social norm. That is, α_A and α_B values represent the agent's desire to act when they believe that coordinating on action is optimal.

Whenever the agent believed that the environmental state was 0 ($s_A = 0$ or $s_B = 0$), then she would act (i.e., play behaviour 1) with a motivation of $Probability_{State\ 0}$ (which I label β for future reference). The probability of not acting (i.e, playing behaviour 0) when the agent believed that the environmental state took the value 0 was therefore given by $1- \beta$. β is therefore the agent's desire to play behaviour 1 when she thought that this was the suboptimal skill to master in her environment, or the suboptimal social norm to coordinate on.

To illustrate with reference to domain B, imagine that the agent's cue summary does not exceed her cognitive threshold ($x_A <= T_A$ or $x_B <= T_B$) and she thus believes that the state is 0 ($s_A=0$ or $s_B=0$). The agent believes that dancing would be inappropriate and that driving on the left-hand side of the road is common. The agent is then motivated to dance with a threshold of β_A and is motivated to drive on the right-hand side of the road with a motivation of β_B. Alternatively, the agent may play the behaviour which is likely to be optimal with a motivation of $1\text{-}\beta_A$ or $1\text{-}\beta_B$ respectively. Thus, β_A and β_B values represent the agent's motivation to play behaviour 1 when she believes that coordinating on behaviour 1 is to coordinate on the suboptimal social norm. That is, β_A and β_B values represent the agent's desire to act when they believe that coordinating on action is suboptimal.

An agent with modular motivation can show a distinct desire to dance (as a ritual or inappropriately) and a distinct desire to drive on the right-hand side of the road (successfully or as a risky decision). This is because the motivational thresholds α_A, α_B, β_A and β_B all take separate values and can evolve to the unique selection pressures of domains A and B. Instead, a domain-general agent can only show a generic desire to play behaviour 1 when she believes that this is optimal in both the dancing and driving domain. I achieve this by constraining domain-general motivation, so that $\alpha = \alpha_A = \alpha_B$. Domain-general agents also have a generic desire to play behaviour 1 whenever they believe that this is suboptimal in both the dancing and driving domain. I achieve this by constraining domain-general motivation, so that $\beta = \beta_A = \beta_B$.

I model modularity separately for the cognitive and motivational thresholds. These possible combinations give four agent types:

1. Fully modular agents can reason distinctly about whether dancing or driving on the right-hand side of the road is likely to be the optimal social norm to coordinate on in both domains (T_A, T_B) and can show a distinct desire towards

dancing or driving on the right-hand side of the road in both domains (α_A, α_B, β_A, β_B).

2. Partly modular agents with modular cognition. These agents can reason distinctly about whether dancing or driving on the right-hand side of the road is likely to be the optimal social norm to coordinate on in both domains (T_A, T_B), though can only show a generic desire towards playing behaviour 1 across both domains (α, β). They can only show a generic desire towards dancing and driving on the right-hand side of the road.
3. Partly modular agents with modular motivation. These agents can only reason generically about the likelihood of the optimal behaviour to coordinate on across both the dancing and driving domains (T), though they can show a distinct desire towards dancing or driving on the right-hand side of the road (α_A, α_B, β_A, β_B).
4. Domain-general agents can only reason generically about the likelihood of the optimal behaviour to coordinate on across both the dancing and driving domains (T) and can only show a generic desire towards playing behaviour 1 across both domains (α, β). They can only show a generic desire towards dancing and driving on the right-hand side of the road.

We ran four models each forcing one of the agent types above. The aim of this model is to investigate whether the agents can coordinate on two distinct social norms in two distinct domains and, if they can coordinate, how the agents' cognition and motivation support this behaviour. Regardless of the four agent types, cognition and motivation are the variables that are left to endogenously evolve throughout my model.

These variables decide the focal agent's behaviour (B_A and B_B). The payoff of the agent's behaviour also depends on her partner's behaviour (B_A and B_B) and the environmental state in both domains (S_A and S_B). The possible outcomes of the agent's behaviour are to coordinate with her partner on the optimal social norm (when her behaviour matches her partner's and matches the environmental state); to coordinate on the suboptimal social norm (when her behaviour matches her partner's but does not match the environmental state) or miscoordination (when her behaviour does not align with her partner's).

This model is unique in considering the payoffs from two different domains with uncertainty in each domain regarding which behaviour is the optimal social norm to coordinate on. Each agent receives an exogenous fitness value of 1 point on top of the fitness that she accrues via her interactions with her partner in the two coordination games. This reflects the fitness from the agent's decisions in areas besides her ability to coordinate on certain social norms. These values were then summed to give a total fitness value per agent. Note that the agents chance of having offspring was directly proportional to their total fitness. For every one of the 100 agents, I translated their total fitness value into a fitness value which was a cumulative proportion of the entire generation. Agents who had more fitness thus had a higher cumulative proportion value. To assign offspring genotype, I simply allowed a random interval between [0,1] which corresponded to this cumulative proportion space. Thus, parental agents with a larger proportional fitness would have a higher chance of having offspring, as the offspring agents were more likely to be sampled from these larger proportional fitness values.

Once the offspring's parental agent was identified, the offspring then inherited her parental agent's psychological variables. Note that I allowed cognition and motivation to be inherited independently from different parental agents. This did not mirror sexual reproduction, but was instead intended to reflect how continuous, complex traits such as psychological phenomena are likely to be coded at multiple allele sites.

Offspring agents inherit the cognitive and motivational thresholds from the parental agents with a small rate of mutation (see SI text for full model details). Agents inherited the cognitive thresholds with a mutation rate of 0.5. As this was a continuous variable with no fixed end points, I assumed mutations to be common, but constrained mutation to only occur in small chunks in each time step ($M\mu = 0$, $\sigma = 0.00125$). Agents inherited their motivational thresholds with a lower mutation rate of 0.05. Mutations occurred in steps of plus or minus 0.01 at each timestep. Note that mutation was constrained so that the motivational thresholds could only take a value between 0 and 1. The agents could never be less than 0%, or more than 100%, motivated to play a certain behaviour. Mutation was modelled independently for the two psychological components. The offspring and parental generation did not co-exist, and reproduction was proportional and affected by mutations, as is typical of a Wright Fisher model (Suchow et al., 2017).

2.2. Parameters

We ran a population of 100 agents for 20,000 generations and repeated this for 100 simulations for each possible combination of parameter space. See appendix 1 for the full parameter space of my model. First, I vary the probability with which the environmental state is 1 in domain A (denoted by p_A) and the probability with which the environmental state is 1 in domain B (denoted by p_B). These are modelled separately and could take realisations $p_A \in \{0.1, 0.5, 0.9\}$ and $p_B \in \{0.1, 0.5, 0.9\}$ respectively.

That is, the environmental state in one coordination domain does not depend on the environmental state in the other domain.

We also vary the f_{Azero}, f_{Bzero}, f_{Aone} and f_{Bone} parameters. This is the fitness given to agents for coordinating on the suboptimal social norm. These are modelled independently for both domains. The payoff for coordinating on the suboptimal social norm is either much lower than coordinating on the optimal social norm (0.25), is intermediate (1), or is the same as coordinating on the optimal social norm (4). This gives a strong, intermediate, or weak selection pressure to coordinate on the optimal behaviour respectively.

The full parameter space of my model is too large to investigate here. For clarity, I focus the analysis on the following combinations of the priors and the fitness tied to suboptimal coordination:

1. Runs where the agents try to coordinate in two similar domains, such as dancing as ritual in domain A and chanting as ritual in domain B. On these runs, the priors make a certain environmental state likely, and the fitness tied to suboptimal coordination favour coordination in this state which is likely to occur. Here, I consider a run where both the priors ($p_A = 0.1$, $p_B = 0.1$) and the fitness tied to suboptimal coordination favour coordination on inaction in state 0 ($f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).
2. Runs where the agents decide whether to coordinate in two similar domains, though the state which is most likely to occur would not necessarily be the most fit state to coordinate in. This run could lead to an outcome where the most fit behaviour to coordinate on is not necessarily the one that is popular to uphold by the group. To illustrate with an example, a group may have a social norm to dance and chant regularly as a ritual, but this will be a costly norm to uphold if

the group lives on a calorific knife-edge. Here, I consider a run where the priors favour state 1 ($p_A = 0.9, p_B = 0.9$) though the fitness tied to suboptimal coordination favour inaction i.e., coordination in state 0 (f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4).

3. Runs where the agents try to coordinate in two distinct domains, such as dancing as ritual in domain A and driving on a certain side of the road in domain B. On these runs, both the prior probabilities and the fitness tied to suboptimal coordination favour coordinating on inaction in state 0 of domain A but coordinating on action in state 1 of domain B ($p_A = 0.1, p_B = 0.9;$ f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 4, f_{Bone} = 0.25).
4. Runs where the agents try to coordinate in two distinct domains, though the environmental state which is likely to occur is not the most fit state to coordinate in. This run could lead to an outcome where the most fit behaviour to coordinate on is not necessarily the one that is most popular in the group. On these runs, state 0 is common in domain A though there is a higher fitness payoff to coordinating on action (state 1); and state 1 is common in domain B though there is a higher fitness payoff to coordinating on inaction (state 0) ($p_A = 0.1, p_B = 0.9$, f_{Azero} = 4, f_{Aone} = 0.25, f_{Bzero} = 0.25, f_{Bone} = 4). This run can potentially shed light on how maladaptive social norms may be upheld in two distinct social domains.

To summarise the model, the agents must decide to ‘act’ or ‘not to act’ in two social coordination domains. To make this decision, each agent must first decide which environmental state the domain is in (does it favour coordination on acting [$s_A = 1, s_B = 1$] or not acting? [$s_A = 0, s_B = 0$]). To decide this, each agent has a private signal (x_A,

x_B) which can be thought of as the cues from the environment. These cues must exceed the cognitive threshold ($x_A > T_A, x_B > T_B$) for the agent to believe that the state is 1 ($s_A = 1, s_B = 1$). This signal can thought of as all the cues that the agent needs to decide that acting is beneficial in each domain (e.g., cues that dancing is a ritual is if all actions are synchronised and if it takes place before a community event). If the private signal does not exceed the threshold ($x_A <= T_A, x_B <= T_B$), then the agent instead believes that the state is 0 and that they should not act ($s_A = 0, s_B = 0$). These beleifs can be wrong. After formulating these beliefs, the agent must be motivated in order to act. The agent is motivated to act with a probability of α when they believe that the state is 1 (and1-α for not acting), and are motivated to act with a probability of β when they believe that the state is 0 (and 1-β for not acting). Once the agent 'acts' ($B_A = 1, B_B = 1$) or 'does not act' ($B_A = 0, B_B = 0$) in each domain, they then receive fitness based on whether this behviour actually matches the environmental state ($S_A = 0, S_A = 1,$ and $S_B = 0, S_B = 1$) and their partner's behaviour ($B_A = 0, B_A = 1,$ and $B_B = 0, B_B = 1$) across both domains. Miscoordiation gives no fitness (when the agent's behviour clashes with both the environment and their partner). The agents get optimal fitness (4) when their behviour matches both the environmental state and their partner's behaviour. Finally, the fitness pressures are driven by the parameters associated to suboptimal coordination: when the agent's behviour matches their partner's, but clashes with the environmental state.

The fitness of these behaviours will be affected by the probability that the state is 1 (p_A and p_B) and the fitness parameters set in each domain (f_{Aone}, f_{Azero}, f_{Bone}, f_{Bzero}). All agents receive an exogenous fitness value of 1 on top of this fitness. Fitness is then changed to a cumulative proportion value, so that more fit agents are more likely to have more offspring. Reproduction is therefore proportional to fitness. The offspring inherit cognitive (T_A, T_B) and motivational thresholds (α_A, α_B, β_A, β_B) from separate agents with a small rate of mutation. The offspring generation then overwrite the parental agent at each time step, as characteristic of a Wright-Fisher model (Suchow et al., 2017).

Our ultimate aim is to compare how modular and domain-general agents come to decide to act or not act in each social coordination domain. To achieve this, fully modular agents were coded to have a separate cognitive and motivtional threshold per domain (T_A, T_B, α_A, α_B, β_A, β_B). Partly modular agents either had modular cognition but domain-general motivation (T_A, T_B, α, β) or they had domain-general cognition and modular motivation (T, α_A, α_B, β_A, β_B). Finally, the fully domain-general agents have constrained cognition so that $T = T_A = T_B$, and constrained motivation so that $\alpha = \alpha_A = \alpha_B$ and $\beta = \beta_A = \beta_B$. This allows domain-general aagents to decide to act or not act over multiple decision-making domains simultaneously.

3. Results

All models converged on a stable psychological architecture across the 20,000 generations over all simulations (see appendix 2). For each case, I explore whether the final generation of agents coordinate (section 3.1), the agents' underlying psychological architecture by the final generation (section 3.2), which aspects of psychology (i.e., cognition and/or motivation) affect agent fitness (section 3.3) and compare the fitness of the four agent types (section 3.4). Finally, Section 3.5 summarises the main findings of this model.

3.1. Did the agents coordinate on a social norm?

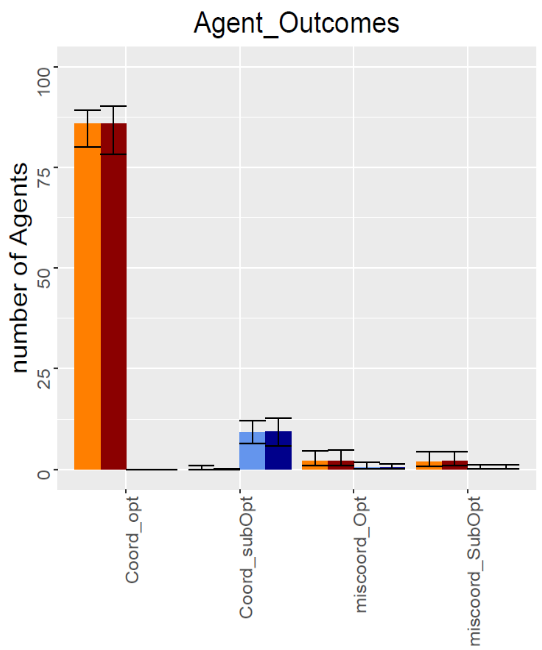
When the payoffs to coordination are sufficient, and the priors are meaningful, then the dominant outcome of a coordination game should be for all parties to coordinate on the Pareto optimum (Kets et al., 2021). Thus, it is expected that agents who were able to identify the unique priors and fitness payoffs in each domain would coordinate on the Pareto optimum. Figure 3 reveals that there were very few cases of miscoordination for any agent type.

In line with these expectations, all agents managed to coordinate on the optimal behaviour of 0 when this was both common and the most fit social norm to coordinate on ($p_A = 0.1, p_B = 0.1;$ $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$; figure 3i). That is, all agent types can avoid dancing or chanting when this would be inappropriate in the current environment.

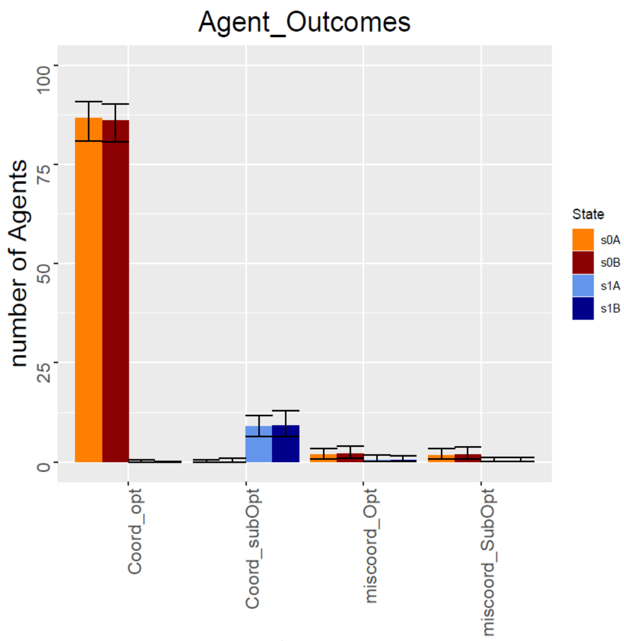
When the agents made decisions in two similar domains where the priors favoured state 1 but the fitness tied to suboptimal coordination favoured coordination in state 0 ($p_A = 0.9, p_B = 0.9;$ $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), there is a two-behaviour split amongst all agent types. Some agents can coordinate on the optimal social norm and others come to coordinate on the suboptimal one. Note that the analysis in appendix 3 confirms that this was not a two-strategy split in the population per se. By the final generation, all agent types come to coordinate on one behaviour entirely. Whether that behaviour was the optimal or suboptimal social norm changed between runs, however.

To illustrate what this means with an example, imagine that dancing as ritual is initially popular, as may be the case to signal group membership in a cooperative group (Legare, 2017; Reddish et al., 2013). However, a sudden shift in the environment means that the group lose access to a food source and the group find themselves living on a calorific knife-edge. In this case, dancing is likely to be costly. My results suggest that most agents continue to uphold the suboptimal social norm after events such as a sudden spatial or temporal shift changes the behaviour which would be optimal to display in the current environment (Deffner et al., 2020).

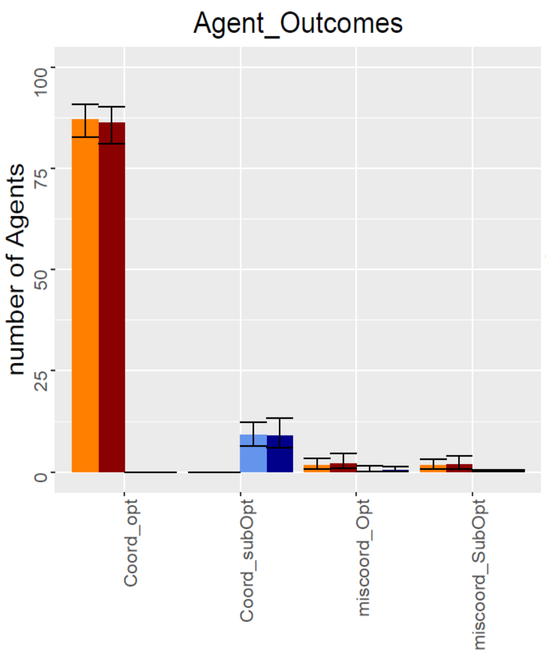
Figure 3i) $p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$



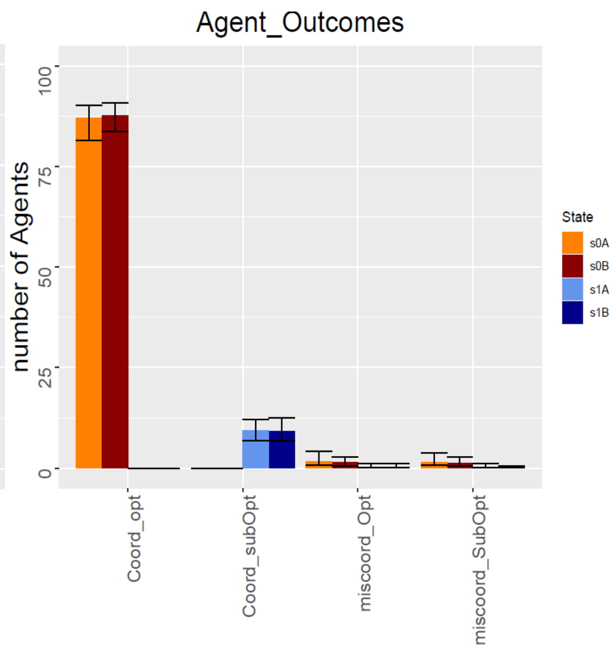
Fully modular



Modular motivation

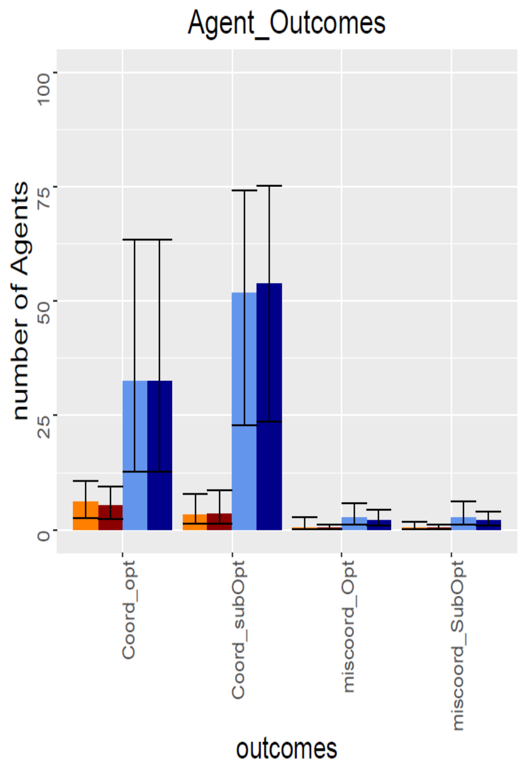


Modular cognition

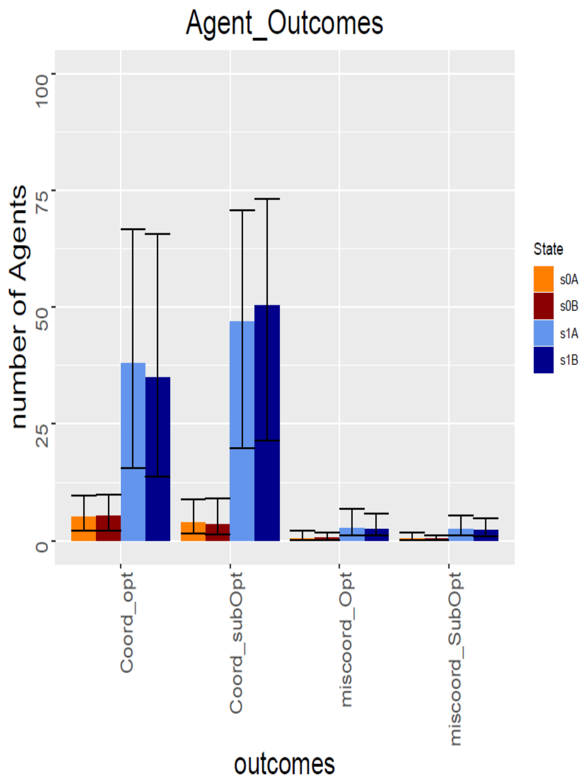


Domain-general

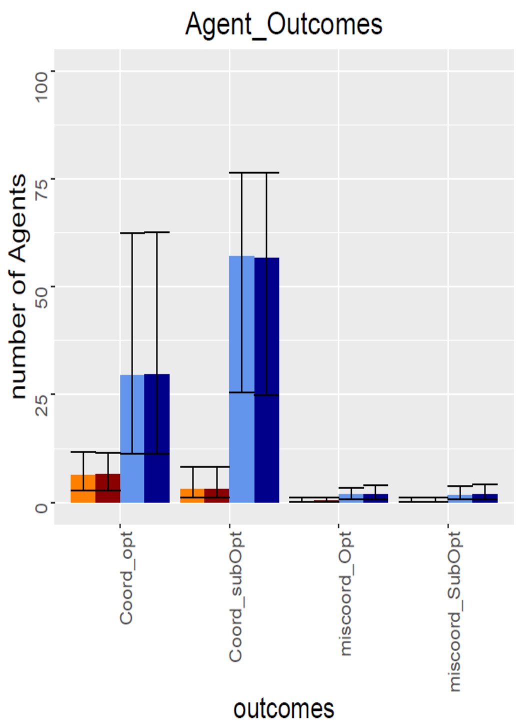
Figure 3ii) $p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$



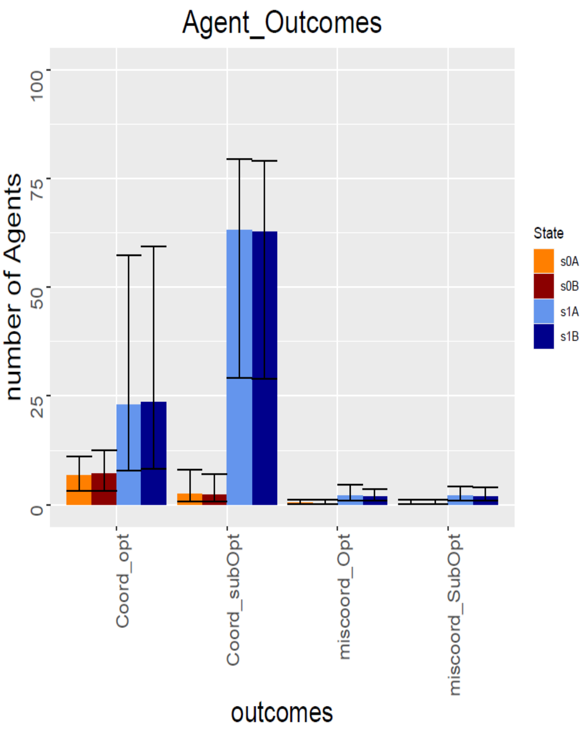
Fully modular



Modular motivation



Modular cognition



Domain-general

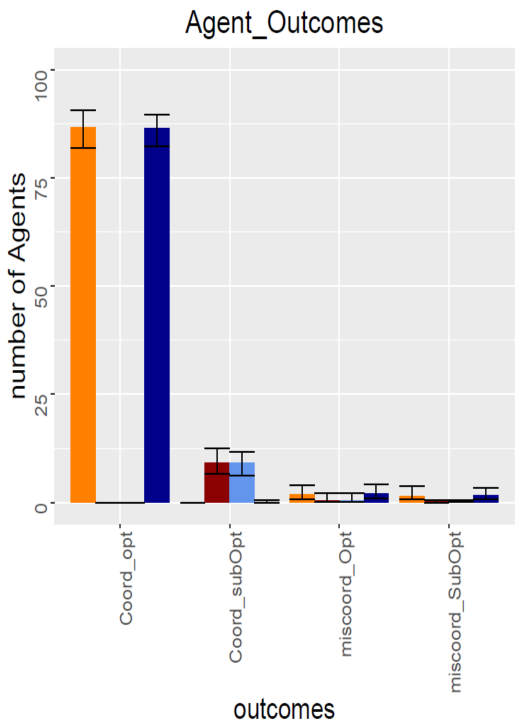
Figure 3. Bar charts showing the agent outcomes on runs where the agents try to coordinate in two similar domains. Note that the agent outcomes on the x axis are dependent on the focal agent's behaviour. Therefore, miscoord_Opt represents cases of miscoordination where the focal agent chose what would have been the Pareto optimum and miscoord_Subopt represents cases of miscoordination where the focal agent chose the option which would have been suboptimal. Coord_subopt represents both agents coordinating on the suboptimal social norm, and Coord_opt represents coordination on the Pareto optimum. These bar charts give the behavioural strategies for fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents in response to runs where i) both domains favour coordination on behaviour 0; $p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$ and ii) coordination over similar domains where the most common environmental state is not the most fit to coordinate on; $p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$. Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

When the agents made decisions in two distinct domains such as dancing and driving, then all three modular agent types could coordinate on the optimal behaviour of 0 in domain A and 1 in domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$; figure 4i). Finally, domain-general agents would typically coordinate on the optimal social norm in both domains, though there was a lot of noise in their behavioural strategies which led to some coordination on suboptimal behaviours and some miscoordination, too. This may be because domain-general agents could not track the contrasting priors across the two distinct domains.

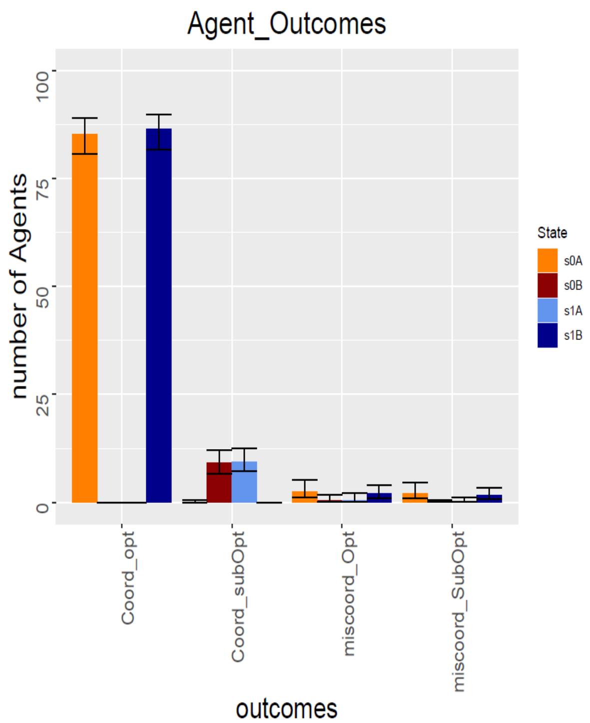
Finally, in cases where the agents tried to coordinate in two distinct domains where the most common state was not necessarily the most fit state to coordinate on ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then the agents with modular motivation (both fully modular and partly modular with modular motivation only) were more likely to coordinate on the suboptimal as opposed to the optimal social norm (see figure 4ii). The agents with domain-general motivation (both domain-general agents and partly modular agents with modular cognition only) seemed to just play behaviour 0 or 1. All bar charts display a two-strategy equilibrium. This is because some runs seemingly ended with the agents coordinating on the behaviour which would be the most fit and others ended with the agents coordinating on the behaviour which would have been common given the priors (see appendix 3).

Taken together, modularity may be important when the agents made decisions in two different domains of social norms. Particularly, modular motivation may help the agents to coordinate in a domain where the most probable environment would not necessarily be the most fit to coordinate in, though they may coordinate on a suboptimal social norm (Figure 4ii). For example, the group may coordinate on dancing as a suboptimal norm if a sudden loss of resources means that the group find themselves living on a calorific knife-edge.

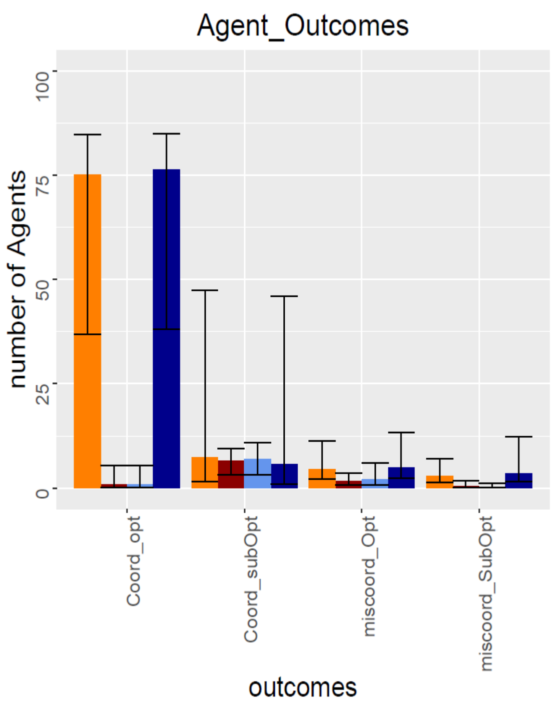
Figure 4i) $p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$



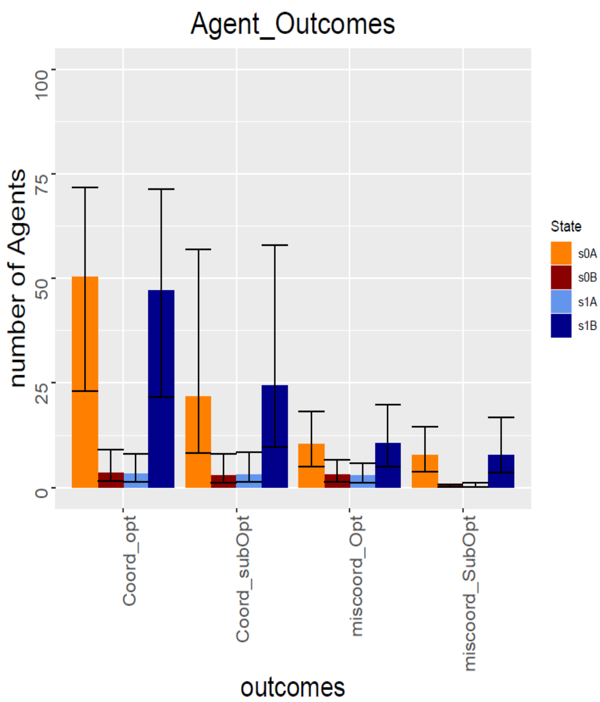
Fully modular



Modular motivation

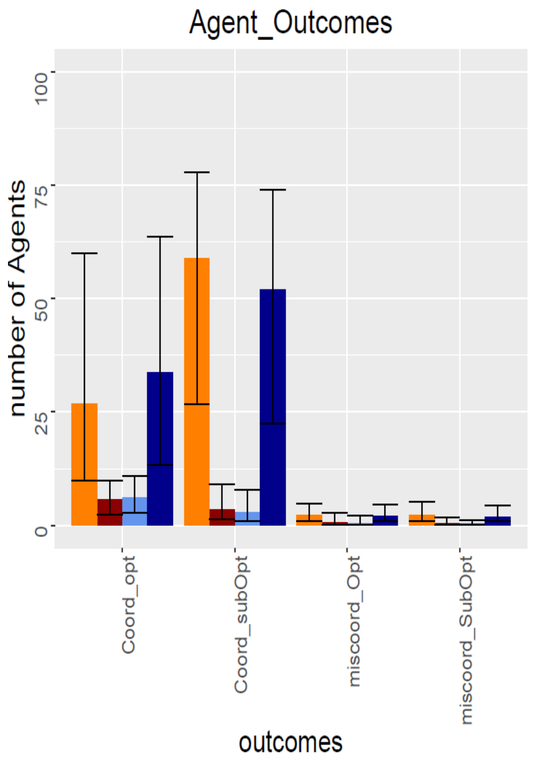


Modular cognition

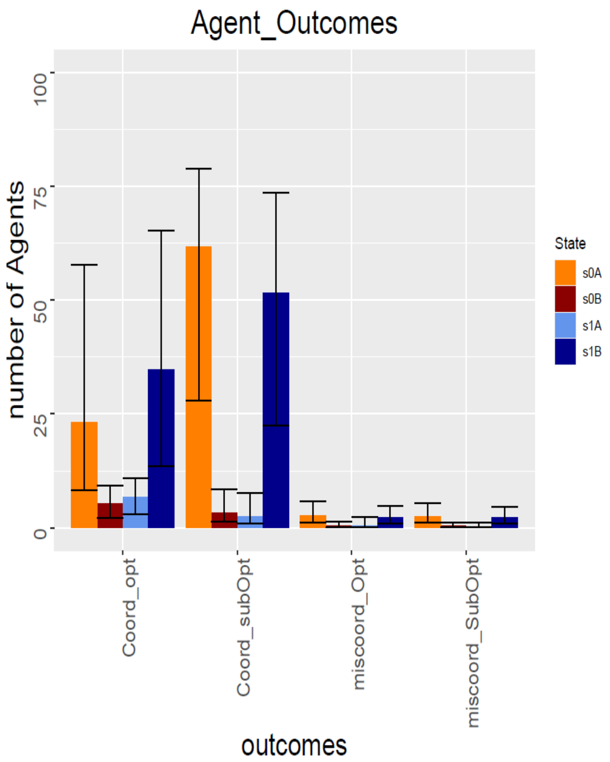


Domain-general

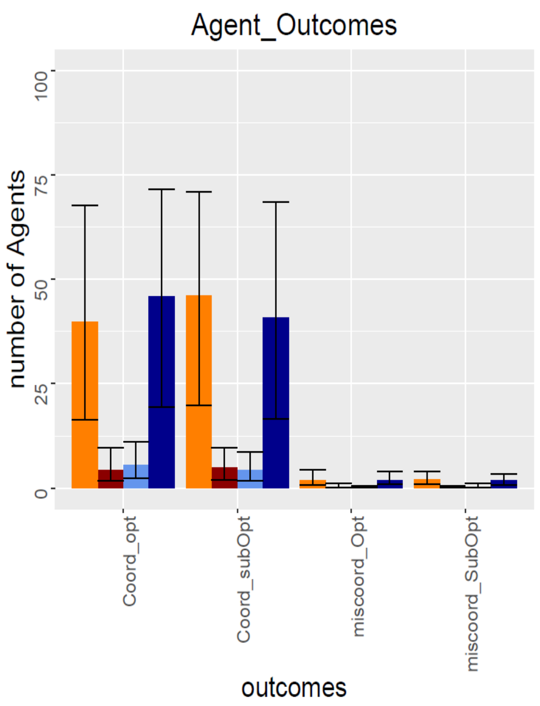
Figure 4ii) $p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$



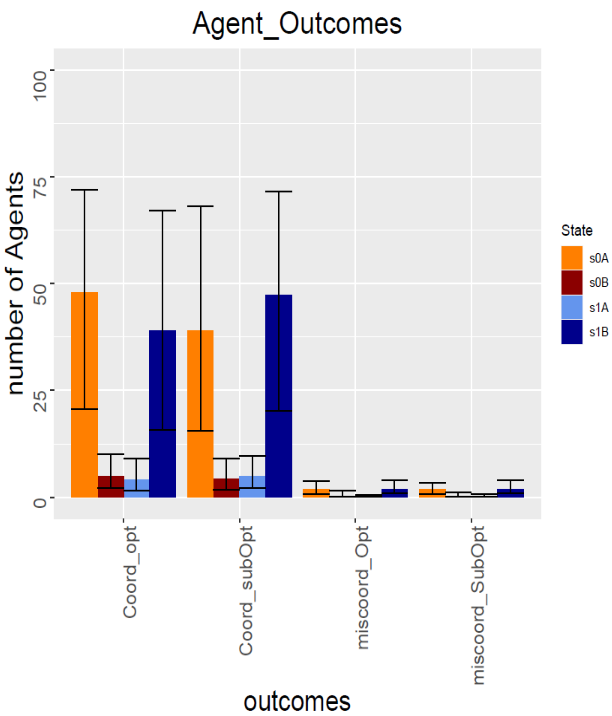
Fully modular



Modular motivation



Modular cognition



Domain-general

Figure 4. Bar charts showing the agent outcomes on runs where the agents try to coordinate in two distinct domains. Note that the agent outcomes on the x axis are dependent on the focal agent's behaviour. Therefore, miscoord_Opt represents cases of miscoordination where the focal agent chose what would have been the Pareto optimum and miscoord_Subopt represents cases of miscoordination where the focal agent chose what would have been the suboptimal option. Coord_subopt represents both agents coordinating on the suboptimal social norm, and Coord_opt represents coordinating on the Pareto optimum. These bar charts give the behavioural strategies for fully modular agents, agents with modular motivation only, agents with modular cognition only and domain-general agents in response to i) distinct domains in which behaviour 0 is favoured in domain A and 1 in B; $p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$ and ii) coordination over two distinct domains where the most common environmental state is not the most fit to coordinate on; $p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$. Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

3.2: The psychological architecture of the four agent types

To visualise how cognition and motivation coevolved in the four agent types, I created binned heatmaps (figures 5-8). Cognitive thresholds could take any value from $-\infty$ to $+\infty$. I used the code in appendix 3 to calculate the smallest and largest threshold values in the final generation of agents, then divide all threshold values into nine equally spaced bins. These correspond to the nine panels in each of the heatmaps in figures 5-8. Then, I partitioned the α and β motivation values into bins. The agent's α and β motivation values could take any form from [0,1], which gave ten bins such as {[0, 0.1], (0.1, 0.2], (0.2, 0.3]…(0.9, 1]}. I calculated two separate bins for the α and β motivation values, giving 100 bins defined jointly over the α and β values.

For each agent type, I calculated the agents' cognitive threshold to decide which of the nine panels of the heatmap to plot to. I then calculated the agents' α and β motivational threshold bins and allow this to track the exact coordinates of this panel to plot to (with α on the x axis and β on the y axis). Finally, I allowed the density of agents with the same cognitive and motivational thresholds to be represented by the heatmap colours. Darker colours mean that more agents share this psychological space, whilst white patches denote that no agents share this exact cognitive and motivational threshold. On top of plotting all agents, I also plotted a black square to each heatmap. This black square denotes the cognitive and motivational threshold of the average agent.

Cognition and motivation clearly coevolved to drive agent behaviour. Figure 5 represents the agents' decision to coordinate in two similar social domains where state 0 is favoured ($p_A = 0.1, p_B = 0.1; f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4$). Agents who aimed to coordinate on behaviour 0 in this run would have accrued the most fitness. The psychological architecture of all four agent types managed to uphold this. Agents with unbiased cognitive thresholds had strong selection on both α and β to be 0, which would motivate them to play behaviour 0 (middle panels of the heatmaps in figure 5). Here, I use the term 'unbiased' cognition to refer to cognitive thresholds close to 0. This is because the summary distributions tended to associate negative cue summary values to state 0 and positive cue summary values to state 1. Thus, a cognitive threshold of around 0 meant that the agents were likely to accurately perceive the environmental state in each domain. A threshold below 0 would bias the agent towards believing that the state was 1, and a threshold above 0 would bias the agent towards believing that the state was 0.

Some agents had positive cognitive thresholds, meaning that their private signal was unlikely to exceed their cognitive threshold ($x_A <= T_A$ or $x_B <= T_B$) and so they were likely to believe that the state was 0 ($s_A = 0$ or $s_B = 0$) (bottom panels of the heatmaps in figure 5). Here, all agents had strong selection acting on β to be 0. The selection acting on α was weaker as the agents were unlikely to believe that the state was 1 and thus were unlikely to use this value.

Finally, some agents had negative cognitive thresholds and thus the agents were likely to have a private signal that exceeded their threshold ($x_A > T_A$ or $x_B > T_B$) and so they were likely to believe that the state was 1 ($s_A = 1$ or $s_B = 1$) (top panels of the heatmaps in figure 5). These agents had a *cognitive bias*. Here, I use the term cognitive bias to refer to cases where the agents' cognitive thresholds shift in the opposite direction to what would be expected based on the priors. Thus, the agent becomes likely to develop an erroneous belief about the value of the environmental state (Efferson et al., 2020b).

As these agents believed that the state was 1, then the agents had strong selection acting on α to be 0. As these agents rarely believed that the state was 0, then β was noisier as this value was less likely to be used. β tended to be lower, however. Regardless of the agents' beliefs, the agents were always motivated to choose behaviour 0 and thus coordinated on the optimal social norm. To illustrate with an example, consider an environment where dancing would be inappropriate such as a work meeting. The agents may be biased to believe that dancing was appropriate in such a scenario, but if they were highly unmotivated to dance then their motivation would compensate for this cognitive bias and so they would still coordinate on not dancing.

There were two dominant strategies that emerged across the simulations where the most common state to coordinate on was not necessarily the most fit across two similar domains. In figure 6, the priors favoured state 1 though the fitness tied to suboptimal coordination favoured coordination on behaviour 0 in state 0 ($p_A = 0.9$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). Half the simulations ended with agents that coordinated on the behaviour that would be the most fit (0) while the other half ended with coordination on the behaviour that was likely to be the most common (1).

When the agents had negative cognitive thresholds– and thus were likely to believe that the state was 1 ($s_A = 1$ or $s_B = 1$)– then there was strong selection acting on α to be either 0 or 1. There was weaker selection acting on β and so this threshold took a range of values (top panels of figure 6). Agents with unbiased cognitive thresholds instead had two clusters, where either α and β values were both low (and so the agents were motivated to play behaviour 0) or α and β values were both high (and so the agents were motivated to play behaviour 1; middle panels of figure 6). When the agents had positive cognitive thresholds– and thus were likely to believe that the state was 0 ($s_A = 0$ or $s_B = 0$)– then strong selection acted on β to be 0 or 1, though α was noisier and took a range of values (bottom panels of figure 6).

All agent types had an average α of 0.4 and β of 0.4 and thus leant towards playing behaviour 0. Note that the analysis in appendix 3 finds no evidence of a split-strategy. The final generations on this run always coordinate, but it changes from simulation to simulation whether they come to coordinate on an optimal, but rare, social norm; or a common, but suboptimal, social norm. This implies that there is selection acting on agent psychology to ensure smooth coordination. Whether this coordination is on the optimal or a suboptimal social norm however seems to be down to drift, or random sampling error over the course of evolution (Rorabaugh, 2014).

To illustrate in the dancing domain, imagine a social group who have converged on dancing as ritual but now this group live on a calorific knife-edge. Dancing would be common but would also be suboptimal as it wastes calories. The agents would be torn between dancing as this was common, or not dancing as this was optimal. Overall, the psychological architecture of the agents suggests that they were likely to try to coordinate on not dancing as this was the most fit.

Figure 5i: Fully modular, domain A. p_A = 0.1, p_B = 0.1. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4.

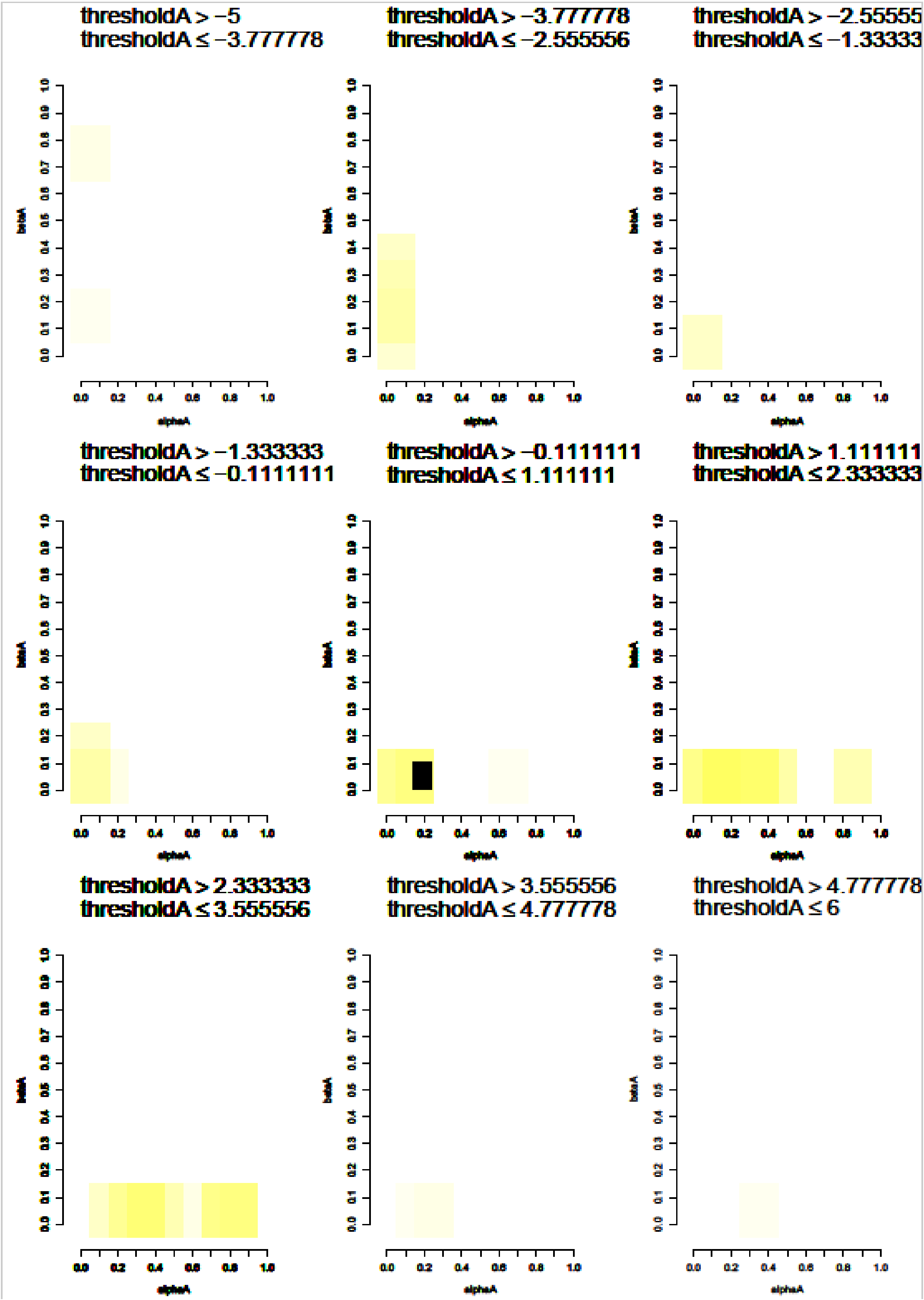


Figure 5i: Fully modular, domain B. $p_A = 0.1$, $p_B = 0.1$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$

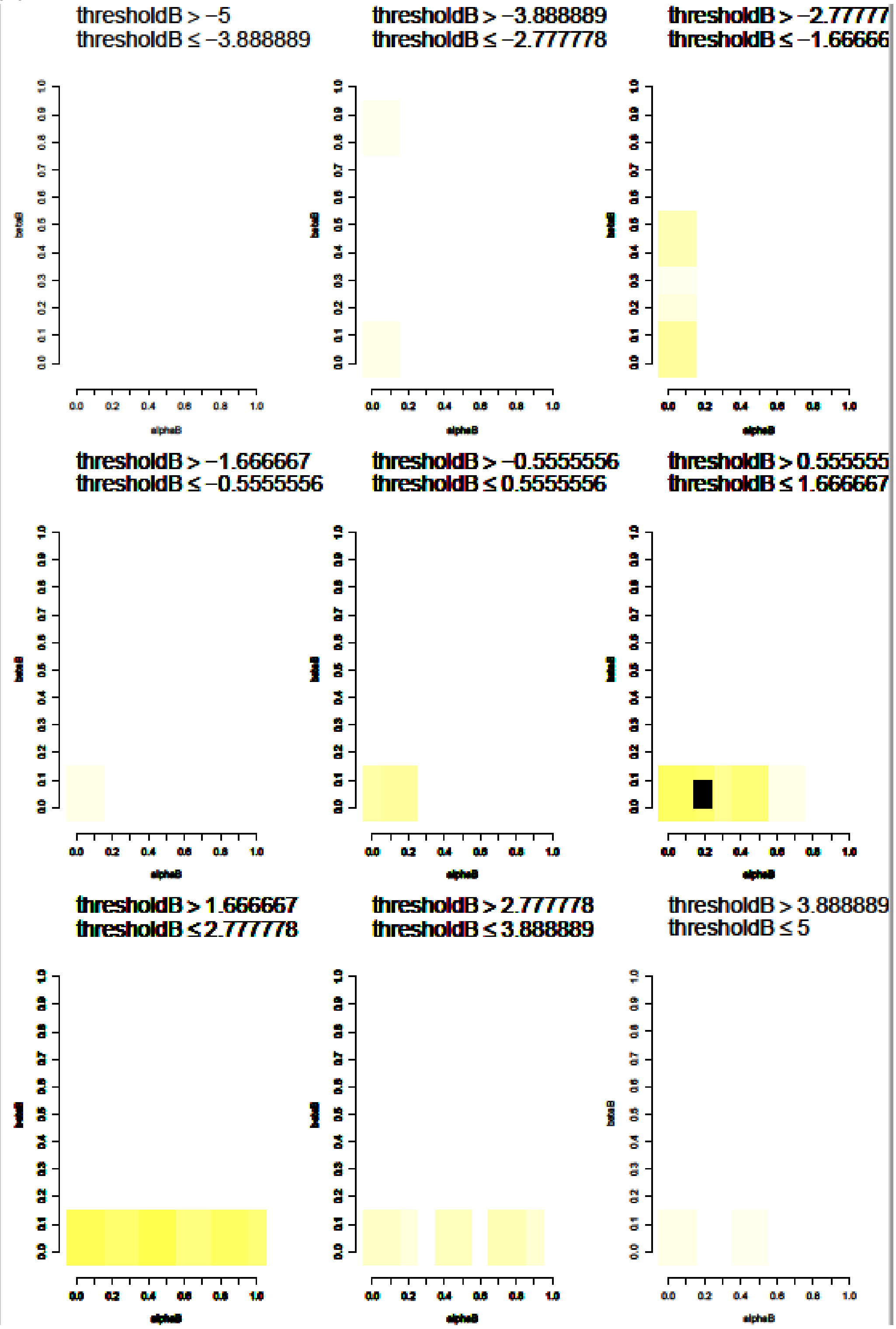


Figure 5ii: Modular motivation, domain A. $p_A = 0.1$, $p_B = 0.1$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$

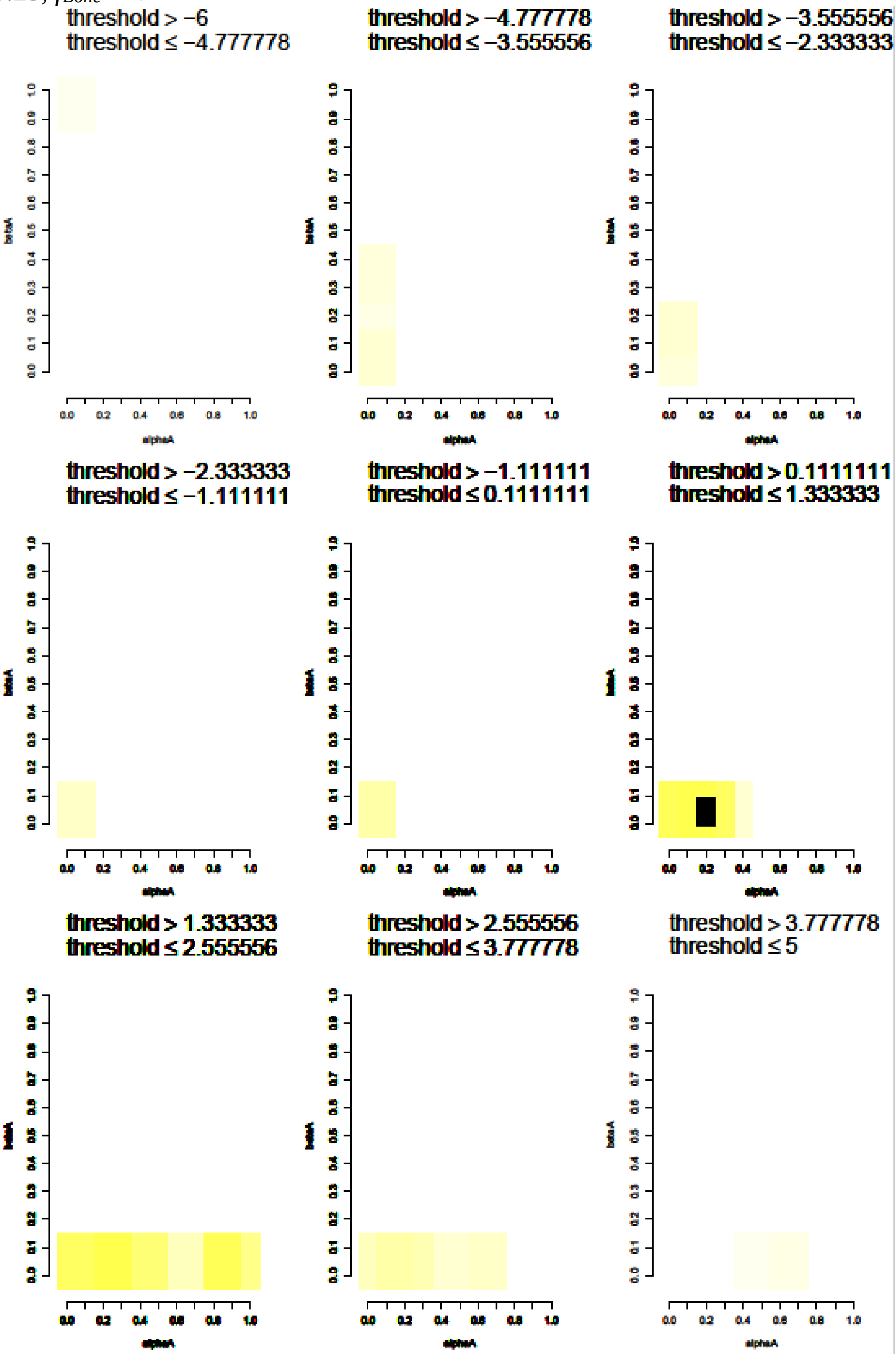


Figure 5ii: Modular motivation, domain B. p_A = 0.1, p_B = 0.1. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4.

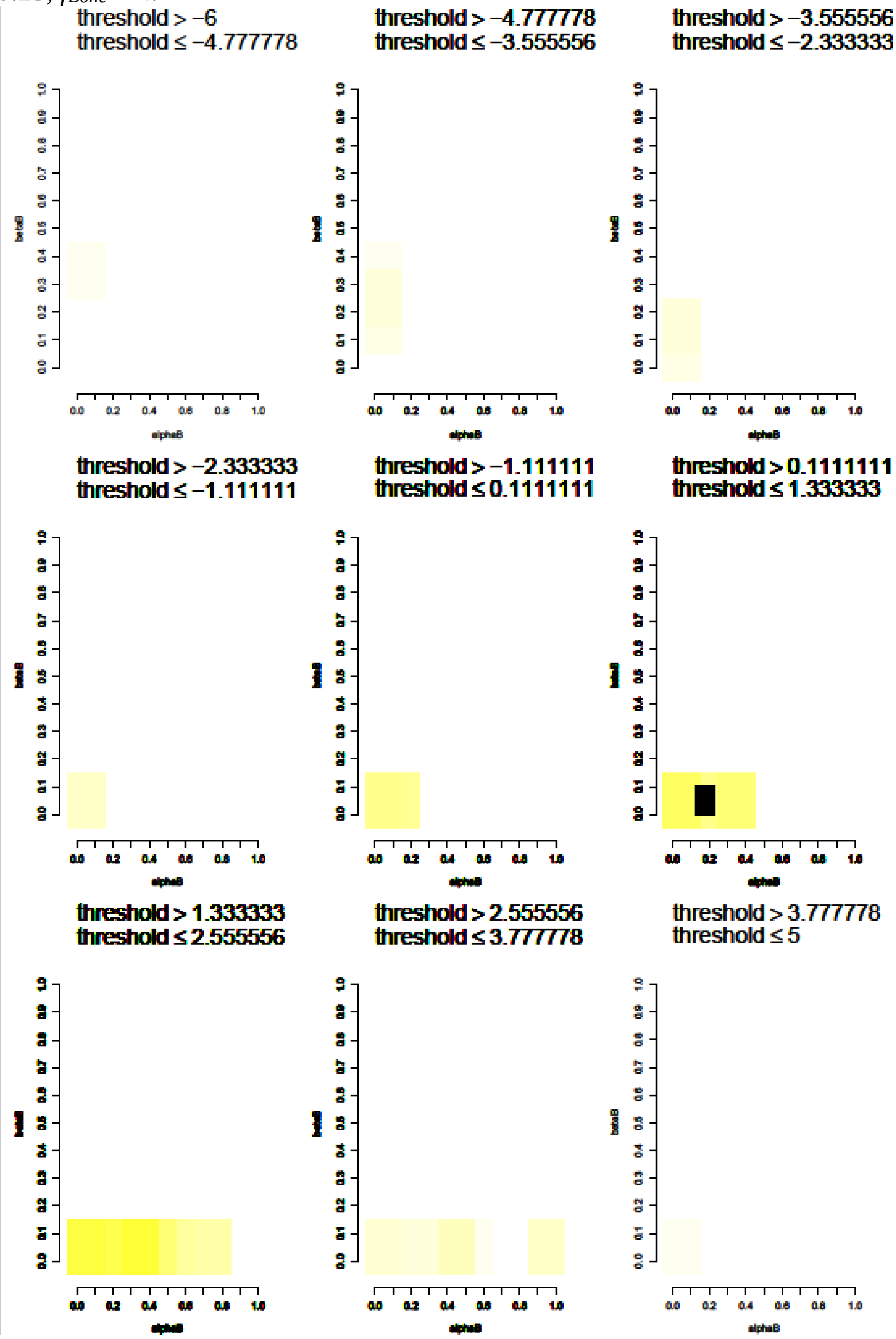


Figure 5iii: Modular cognition, domain A. $p_A = 0.1$, $p_B = 0.1$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$

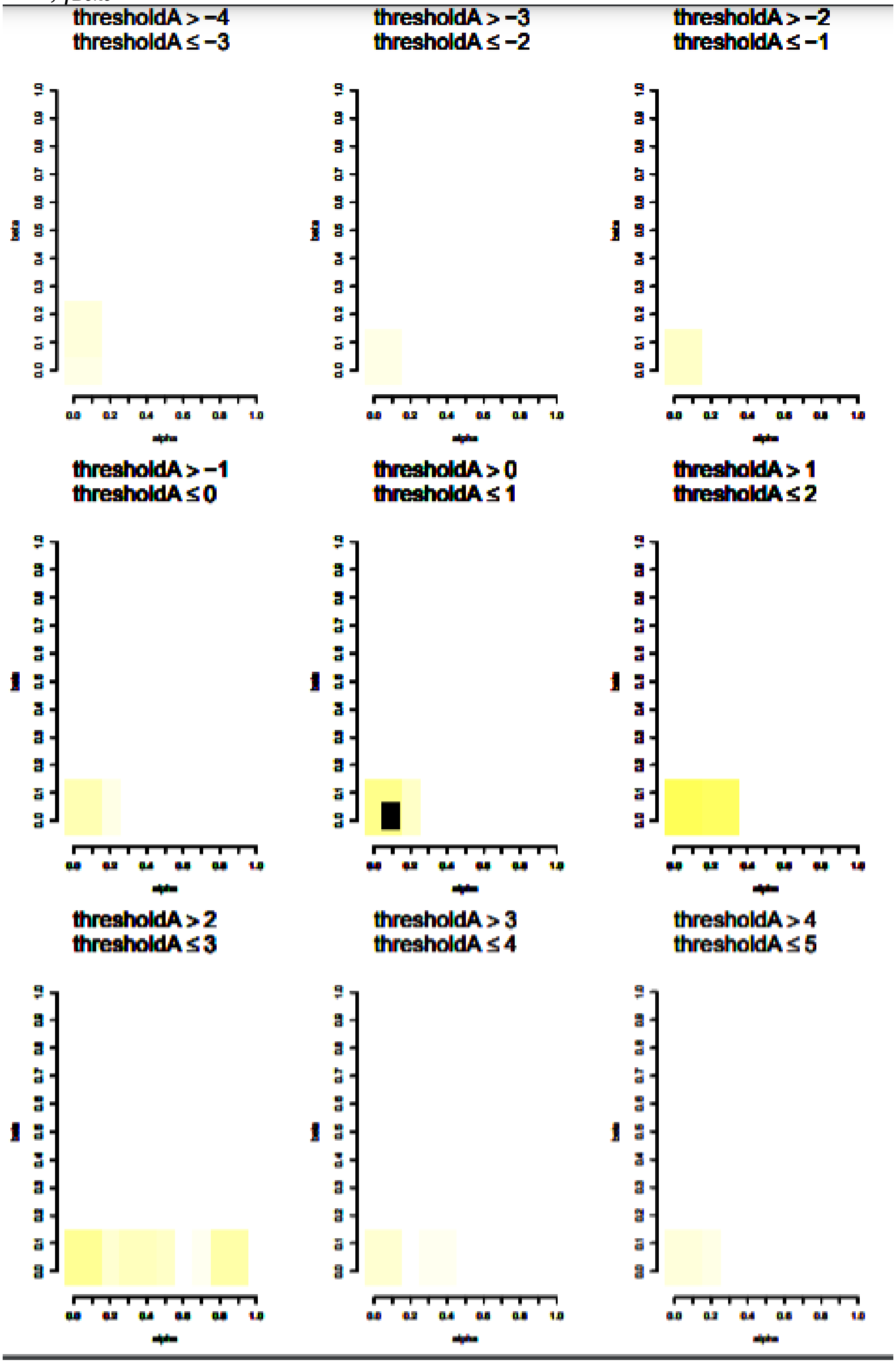


Figure 5iii: Modular cognition, domain B. p_A = 0.1, p_B = 0.1. f_{Azero} = 0.25, *f_{Aone}* = 4, *f_{Bzero}* = 0.25, *f_{Bone}* = 4

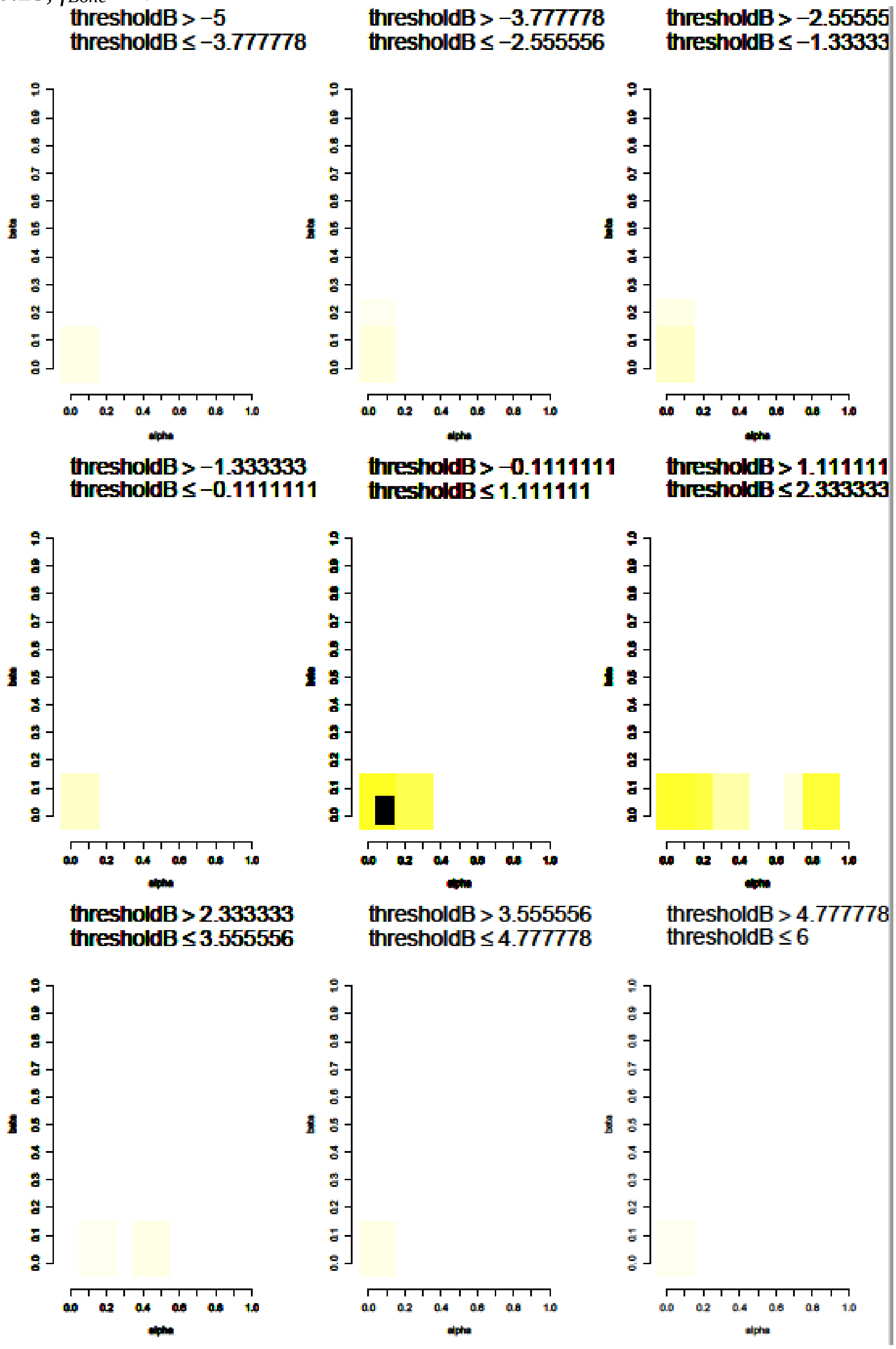


Figure 5iv: Domain-general, $p_A = 0.1$, $p_B = 0.1$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

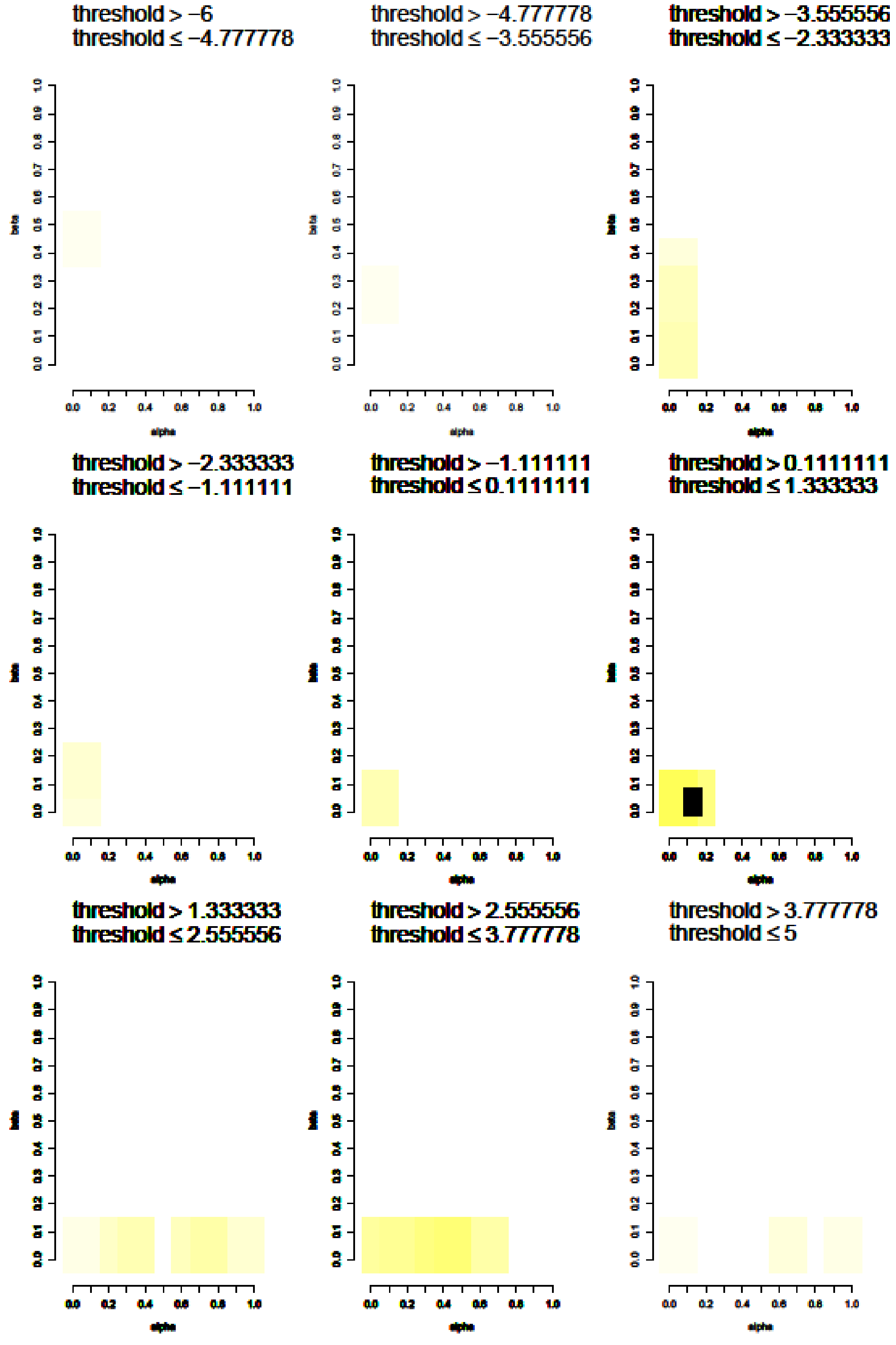


Figure 6i: Fully modular, domain A. p_A = 0.9, p_B = 0.9. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$

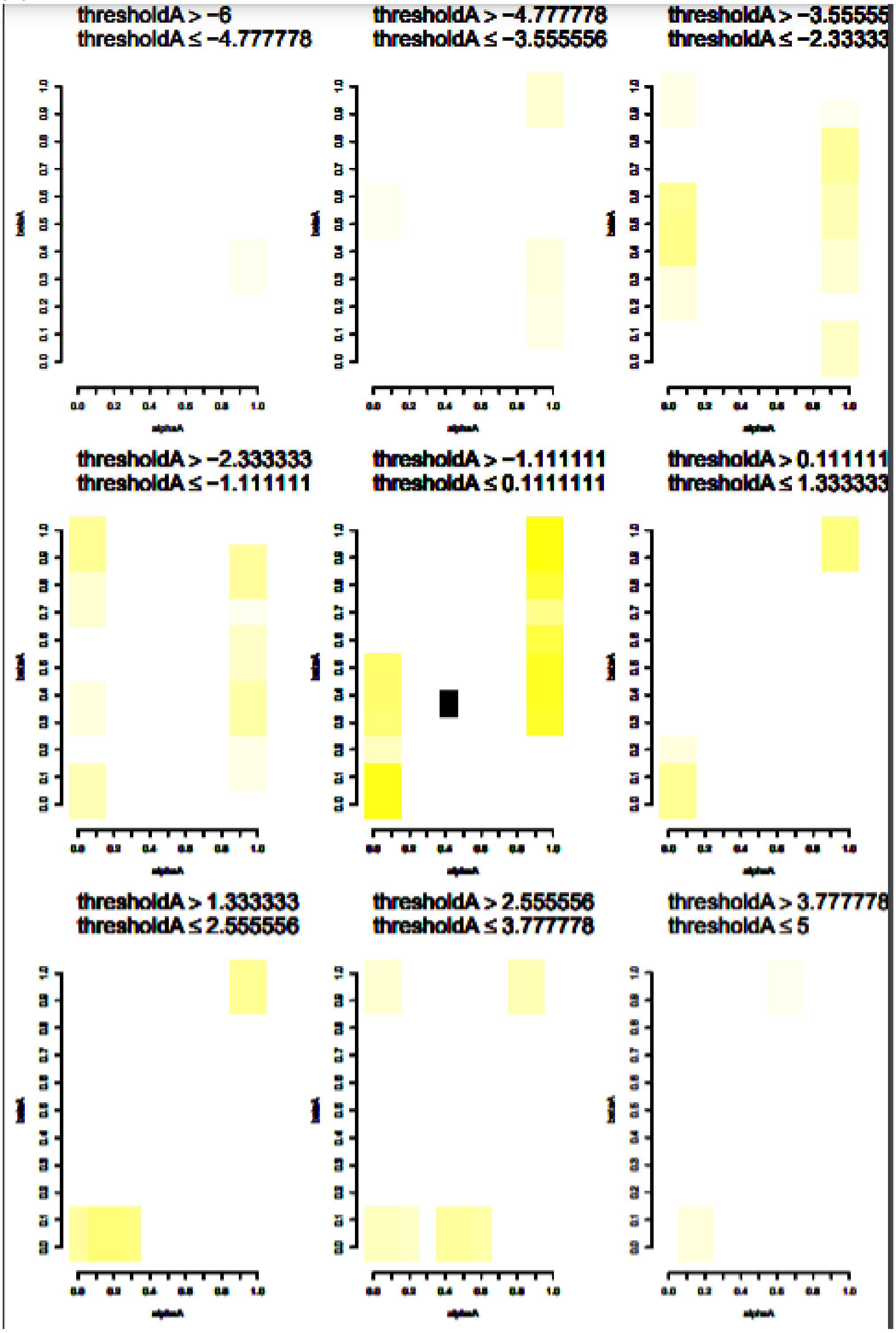


Figure 6i: Fully modular, domain B. $p_A = 0.9$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$

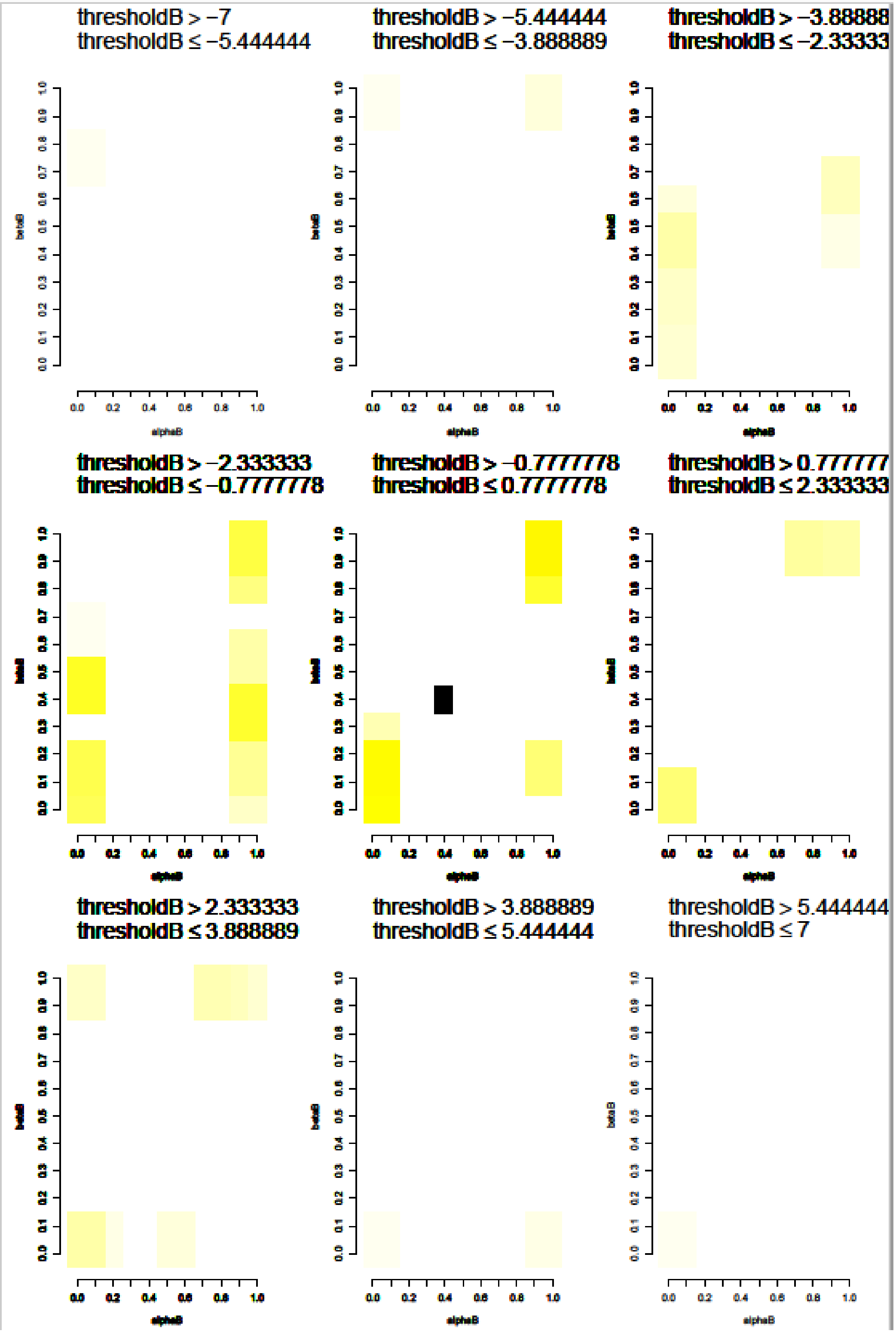


Figure 6ii: Modular motivation, domain A. p_A = 0.9, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4.

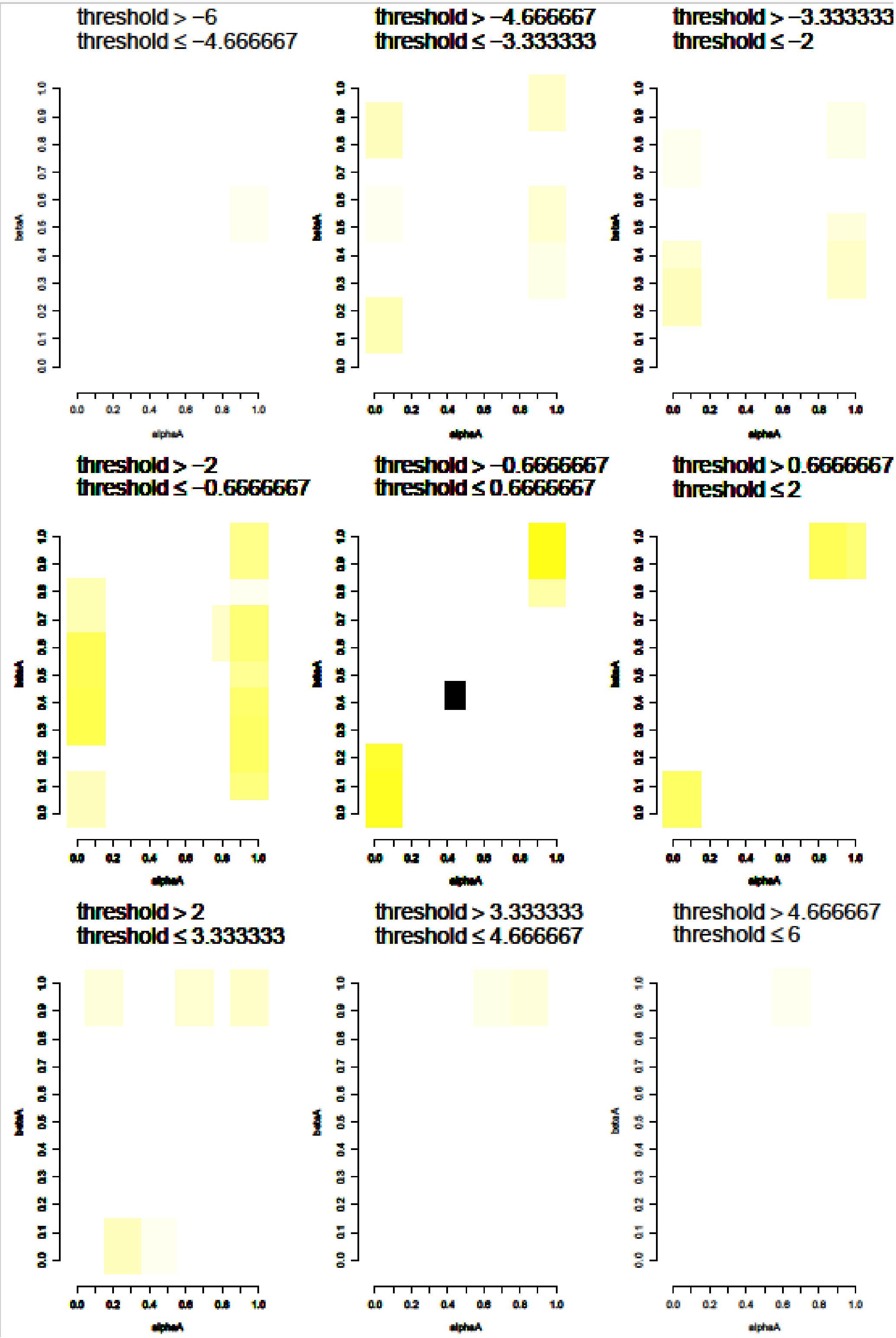


Figure 6ii: Modular motivation, domain B. p_A = 0.9, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4.

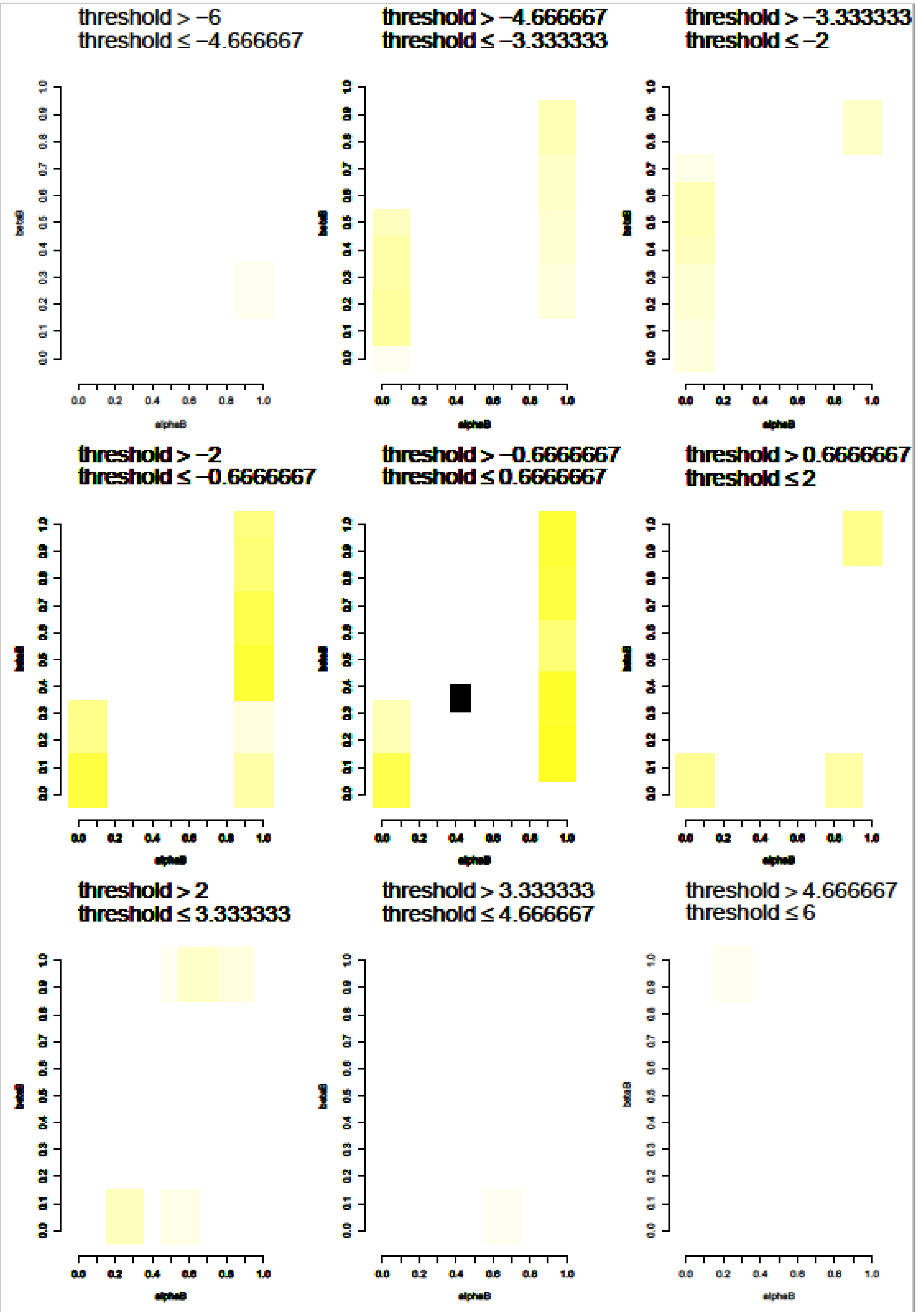


Figure 6iii: Modular cognition, domain A. $p_A = 0.9$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

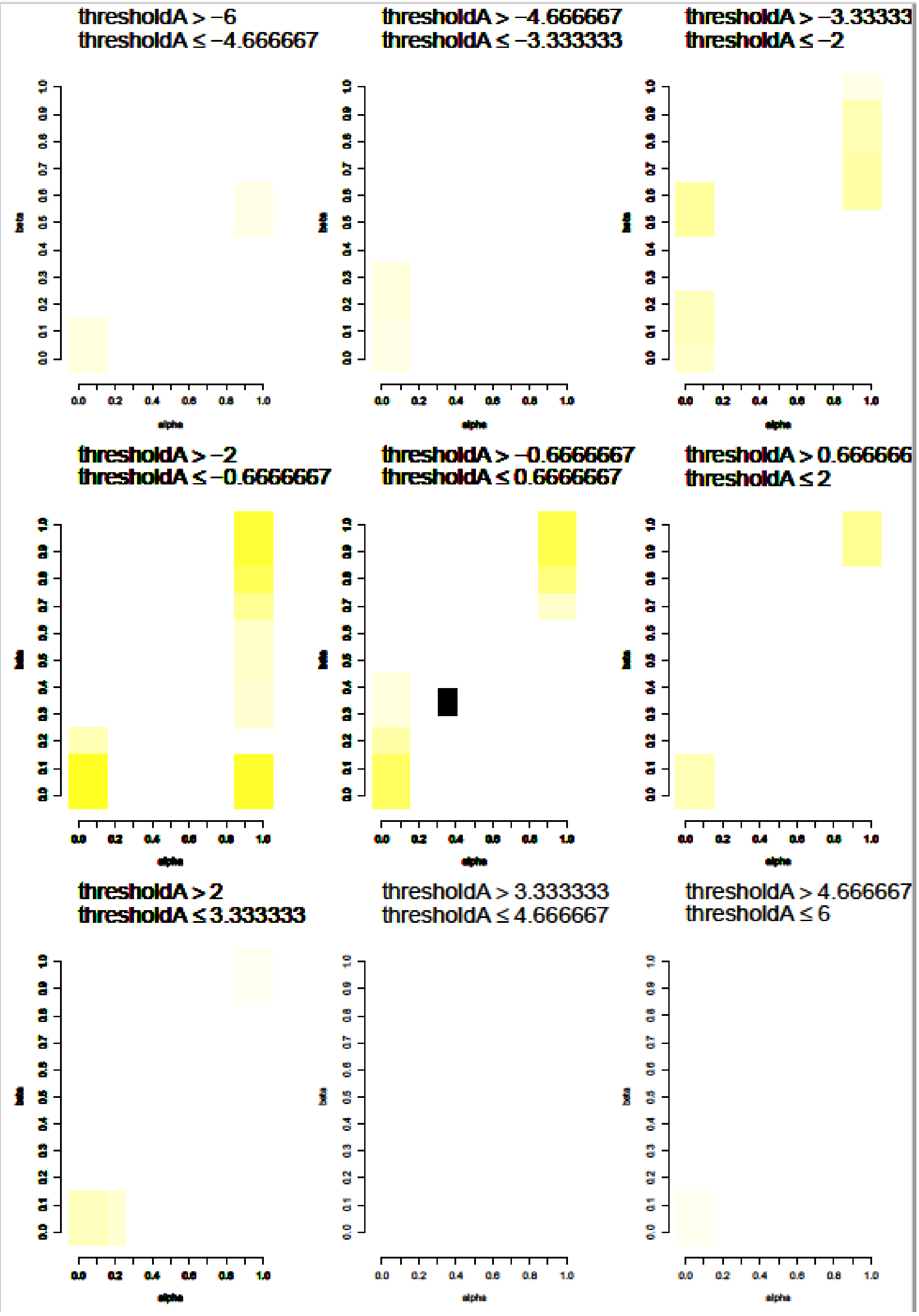


Figure 6iii: Modular cognition, domain B. p_A = 0.9, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4

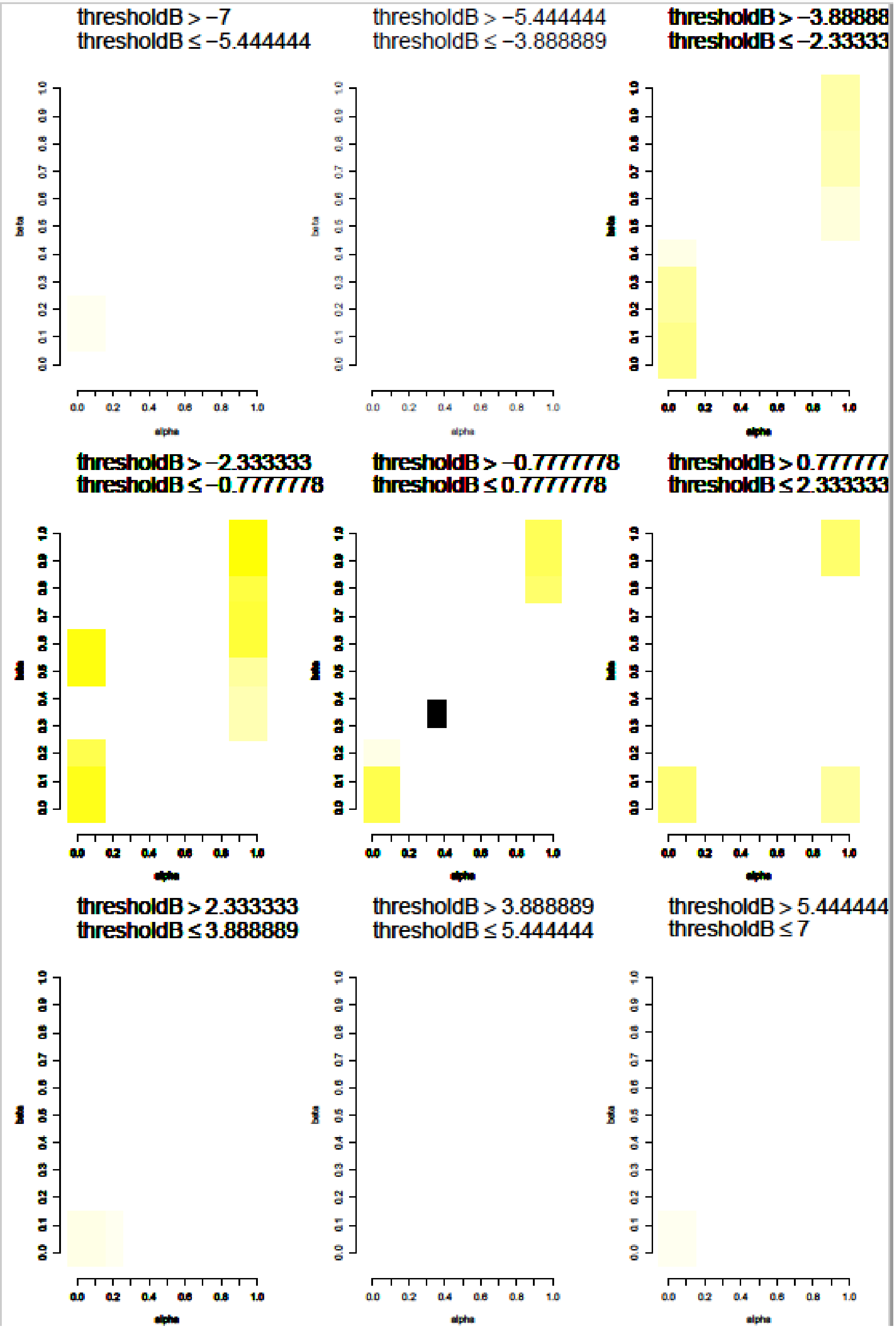


Figure 6iv: Domain-general p_A = 0.9, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4.

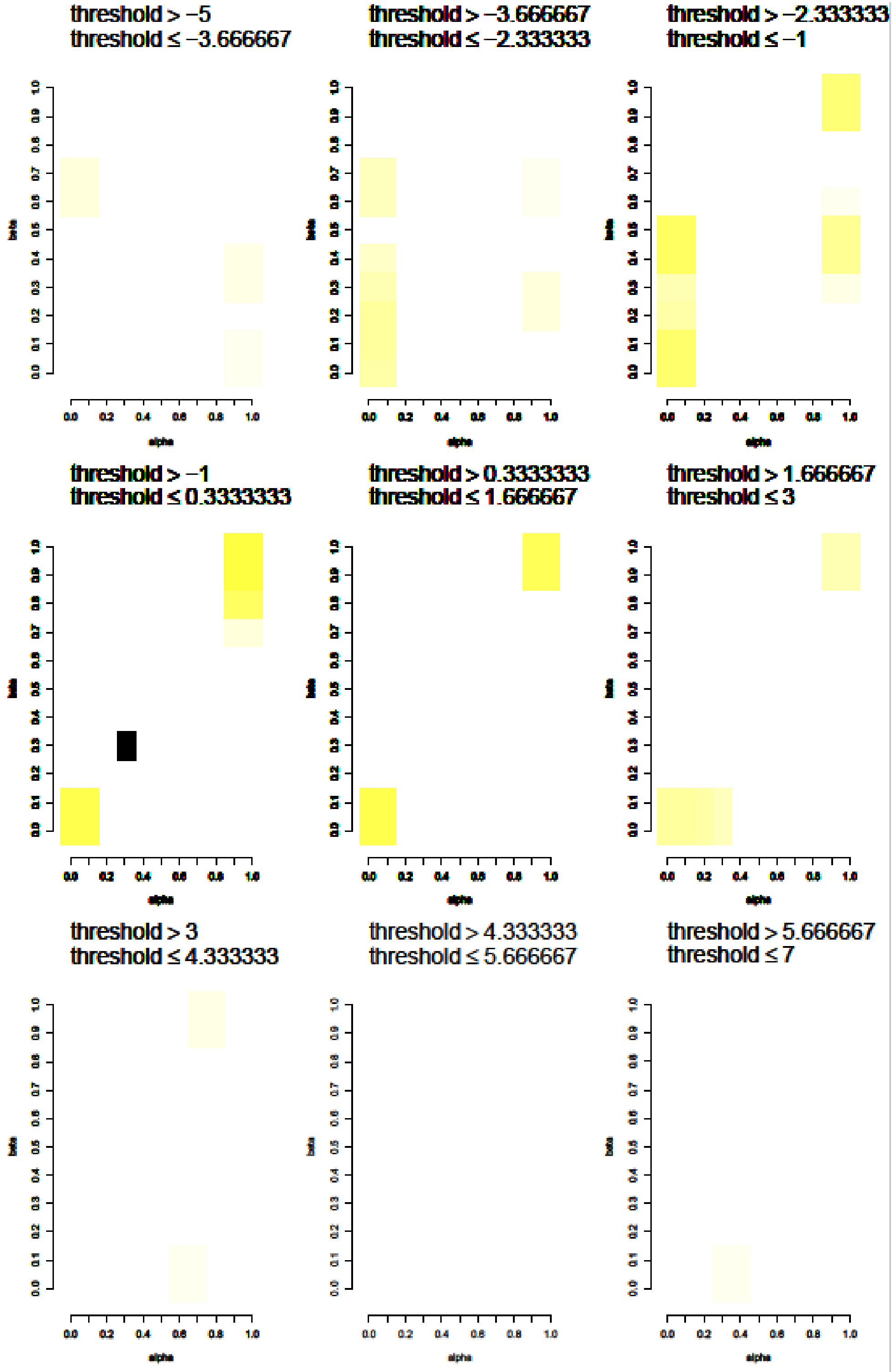


Figure 5 and Figure 6. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) agents with modular motivation only; (iii) agents with modular cognition only and (iv) fully domain-general agents when deciding to coordinate in two similar domains. Figure 5 gives the results for runs where both the priors and the fitness tied to suboptimal coordination favour coordination on behaviour 0 ($p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), while Figure 6 gives the results for runs where the priors favour state 1 but the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 ($p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).

Figure 7 represents the decision to coordinate over two distinct domains, where state 0 is favoured in domain A but state 1 is favoured in domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$). Fully modular agents and agents with modular motivation only had a similar pattern of results. In domain A, agents with negative cognitive thresholds had strong selection on α_A to be 0, though weaker selection acting on β_A which took a range of values (top panels of figures 7i and 7ii A). Agents with unbiased cognitive thresholds had a low α_A and β_A (i.e., to play behaviour 0; middle panels of figures 7i and 7ii A). Agents with positive cognitive thresholds instead had strong selection on β_A to be 0, and a weaker selection on α_A which took a range of values (bottom panels of figures 7i and 7ii A).

This flipped in domain B. Agents with negative or unbiased cognitive thresholds had strong selection acting on α_B to be 1, with weaker selection acting on β_B though this tended to take a higher value (top and middle panels of figures 7i and 7ii B). Agents with positive cognitive thresholds had β_B at 1 with a weaker selection on α_B which took a range of values (bottom panels of figures 7i and 7ii B). Thus, fully modular agents and agents with modular motivation only were always motivated to play behaviour 0 in domain A and play behaviour 1 in domain B, which would mean that they were likely to choose the optimal social norm to coordinate on in each domain. They could avoid dancing when inappropriate and would drive on the right-hand side of the road when this was common.

Note that agents with modular cognition only could not specialise their domain-general motivation to the contrasting demands of the two distinct domains and thus their modular cognition became important. These agents had a larger T_A value (0.66 to 2) and a smaller T_B value (-1.67 to -0.33). They thus needed more evidence to believe in a rare event in domain A and less evidence to believe in a common event in domain B. These agents were less likely to believe that dancing was a ritual when this was rare and were more likely to believe that driving on the right-hand side of the road was common when this was likely. These adjustments in cognitive thresholds helped the agents to make the most of their domain-general motivation, as the average agent had a high α (0.9) and a low β (0.1; see figure 7iii). Whenever the agent believed that the state was 0, then their β value evolved to be low and so they were motivated to play behaviour 0. This was more likely to be the case in domain A. Whenever the agent believed that the state was 1, then their α motivation was high and so they were motivated to play behaviour 1. This was more likely to be the case in domain B. Thus, agents with modular cognition could avoid dancing when they believed that dancing as a ritual was rare and could drive on the right-hand side of the road when they believed that this was common. This highlights that modular cognition was also important when coordinating in two distinct domains. In fact, all three modular types managed to coordinate on the optimal behaviour over the two distinct domains (see figure 4i, section 3.1).

Domain-general agents also had a high α and low β on average (see figure 7iv), though they had an unbiased T as domain-general cognition could not specialise to the contrasting demands of two distinct social domains. There was a lot of noise in the agents' psychological architecture, which may explain the range of behavioural outcomes (see figure 4i; section 3.1). Sometimes, domain-general agents coordinated on the optimal social norm and sometimes they coordinated on the suboptimal social norm, and sometimes they experienced miscoordination.

When the priors made a certain environmental state likely but the opposite environmental state was more fit to coordinate on across two distinct domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$; see figure 8), then half the simulations ended with the agents converging on the behaviour which was likely to be optimal and half of the simulations ended with the agents converging on the behaviour which was most likely to match the environmental state. Fully modular agents and agents with modular motivation only had simulations where β_A evolved to be 0 or 1, with a noisier α_A in domain A. In domain B, fully modular agents and agents with modular motivation only had α_B evolve to be either 0 or 1 with a noisier β_B.

Both these agent types with modular motivation had an average α_A and β_A of approximately 0.6, plus an average α_B and β_B of approximately 0.4 (see figures 8i and 8ii). That is, agents with modular motivation would sometimes coordinate on a behaviour which would be the most likely environmental state though overall they were more likely to coordinate on the behaviour which would have had the highest expected payoff. Returning to the example of an environment where dancing was common but costly as this group existed on a calorific knife-edge, then agents with modular motivation may have been more likely to coordinate on not dancing to conserve calories.

The agents with domain-general motivation (domain-general agents and agents with modular cognition only) had simulations which ended with all agents having either a low α and β (and thus being motivated to play behaviour 0) or having a high α and β

(and thus being motivated to play behaviour 1). These agents' average α and β was 0.5 (see figures 8iii and 8iv). Half the simulations converged on agents who chose the behaviour that they expected to have the highest payoff (and thus these individuals within these populations would coordinate on the optimal social norm). The other half of the simulations would converge on agents who chose the behaviour that was likely to be the most common (and thus these individuals within these populations would have coordinated on the suboptimal social norm). Note that the agents in this latter run may have been at risk of coordinating on a maladaptive social norm. This again highlights that modular motivation may be important to track cases where the most fit behaviour to coordinate on was not necessarily the most likely to occur. Section 3.3 further investigates how the agents' cognitive and motivational thresholds affected their fitness.

Figure 7i: Fully modular, domain A. p_A = 0.1, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 4, f_{Bone} = 0.25.

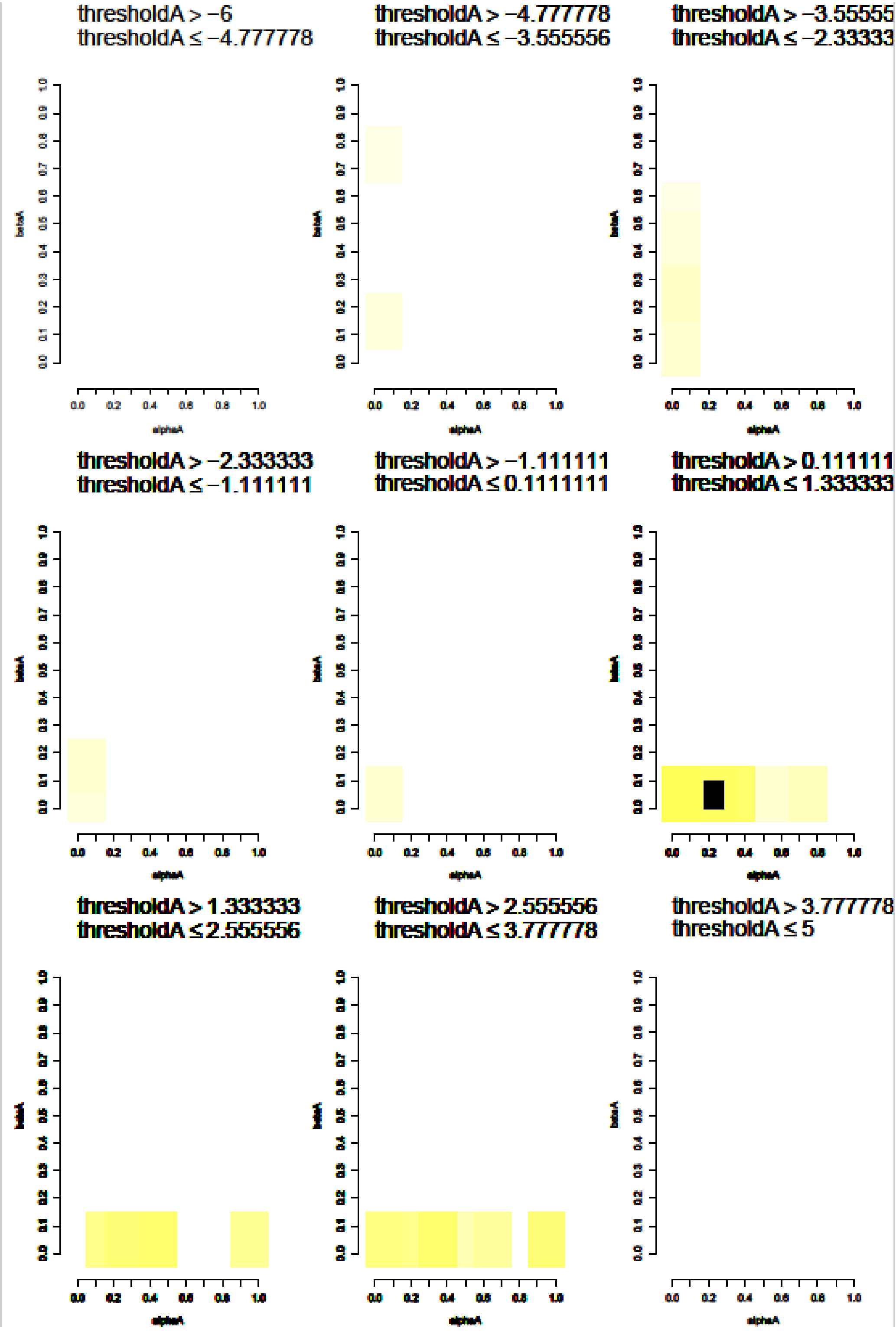


Figure 7i: Fully modular, domain B. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$.

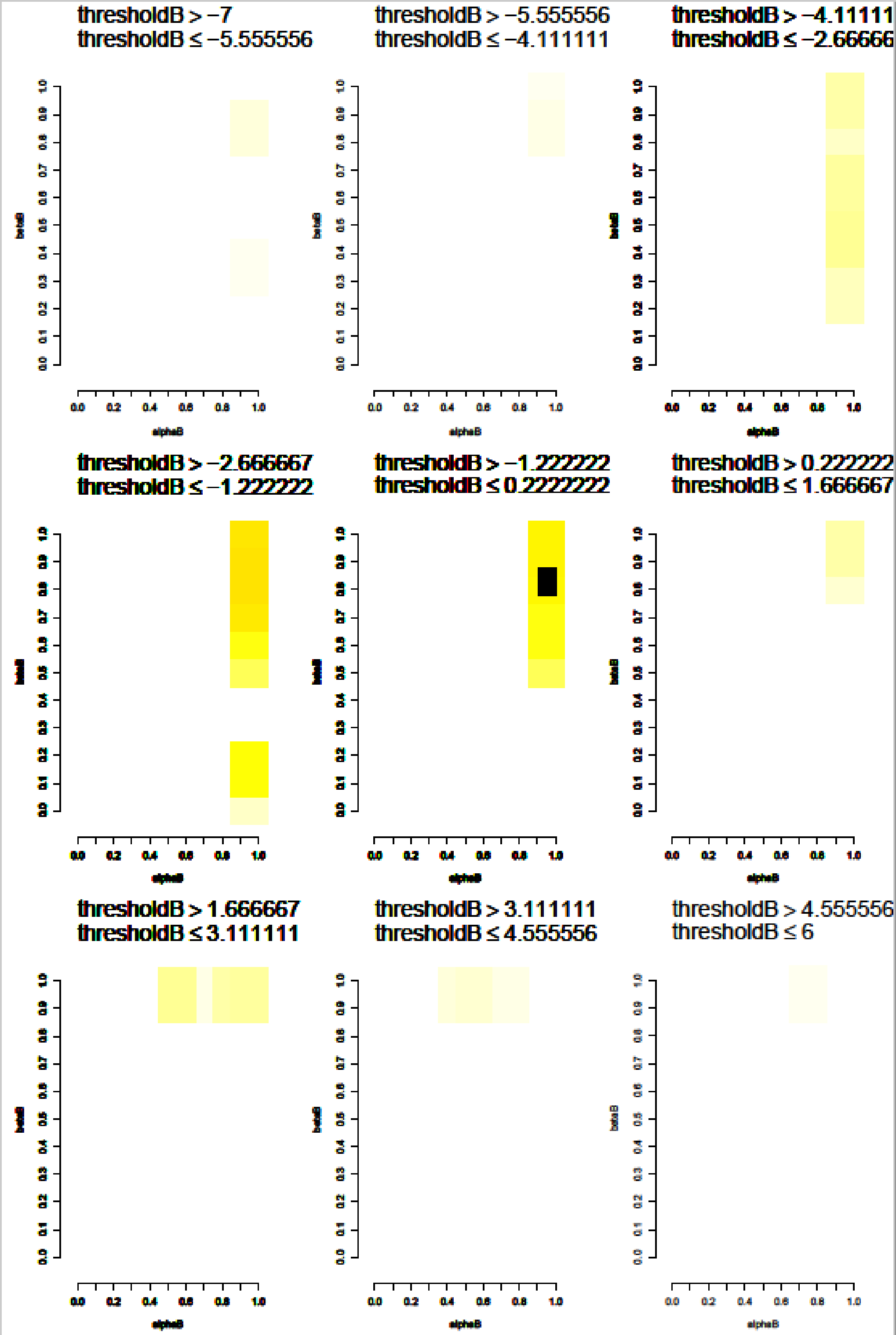


Figure 7ii: Modular motivation, domain A. p_A = 0.1, p_B = 0.9. f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 4, f_{Bone} = 0.25.

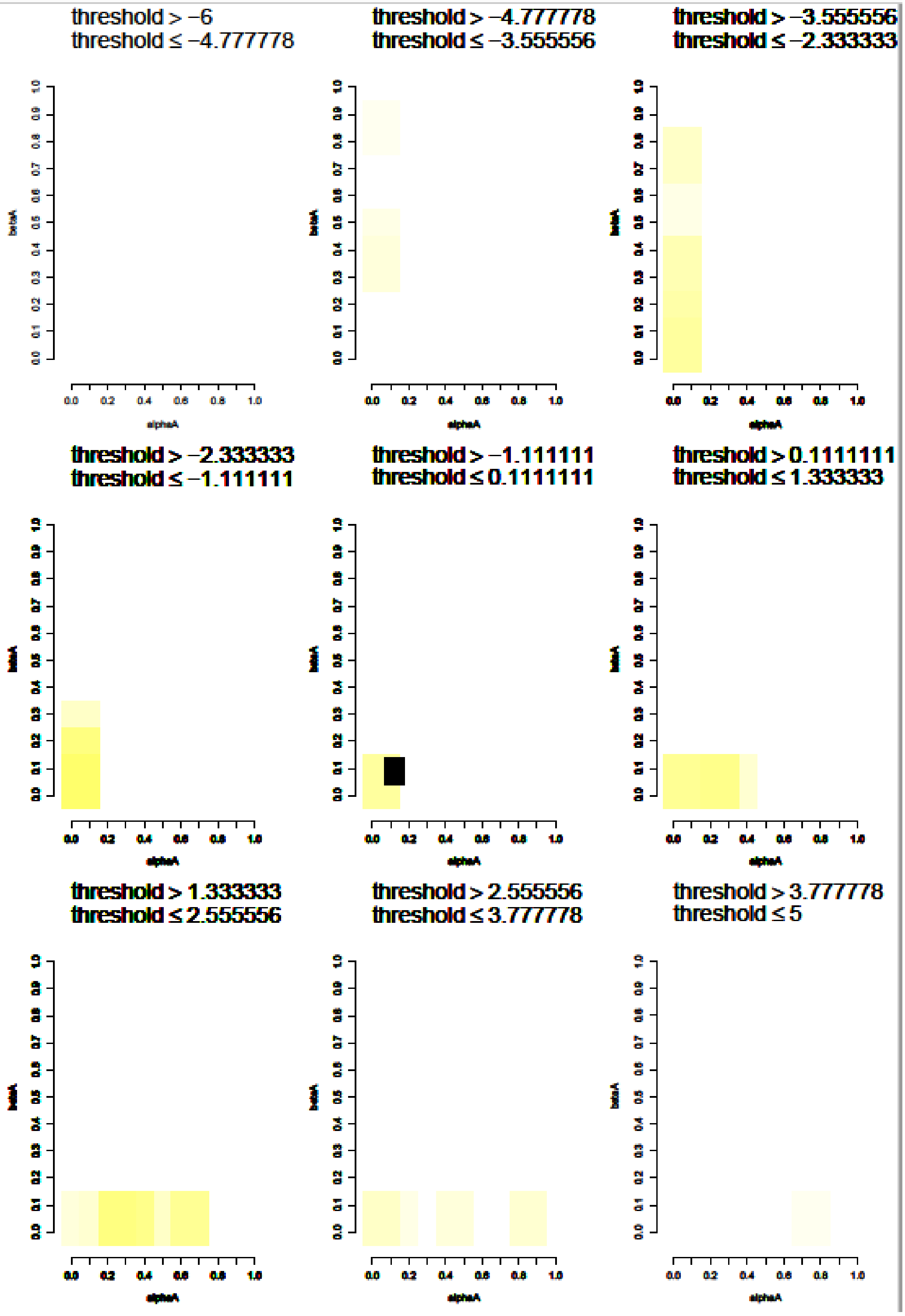


Figure 7ii: Modular motivation, domain B. $p_A = 0.1$, $p_B = 0.9$. $f_{zero0} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$.

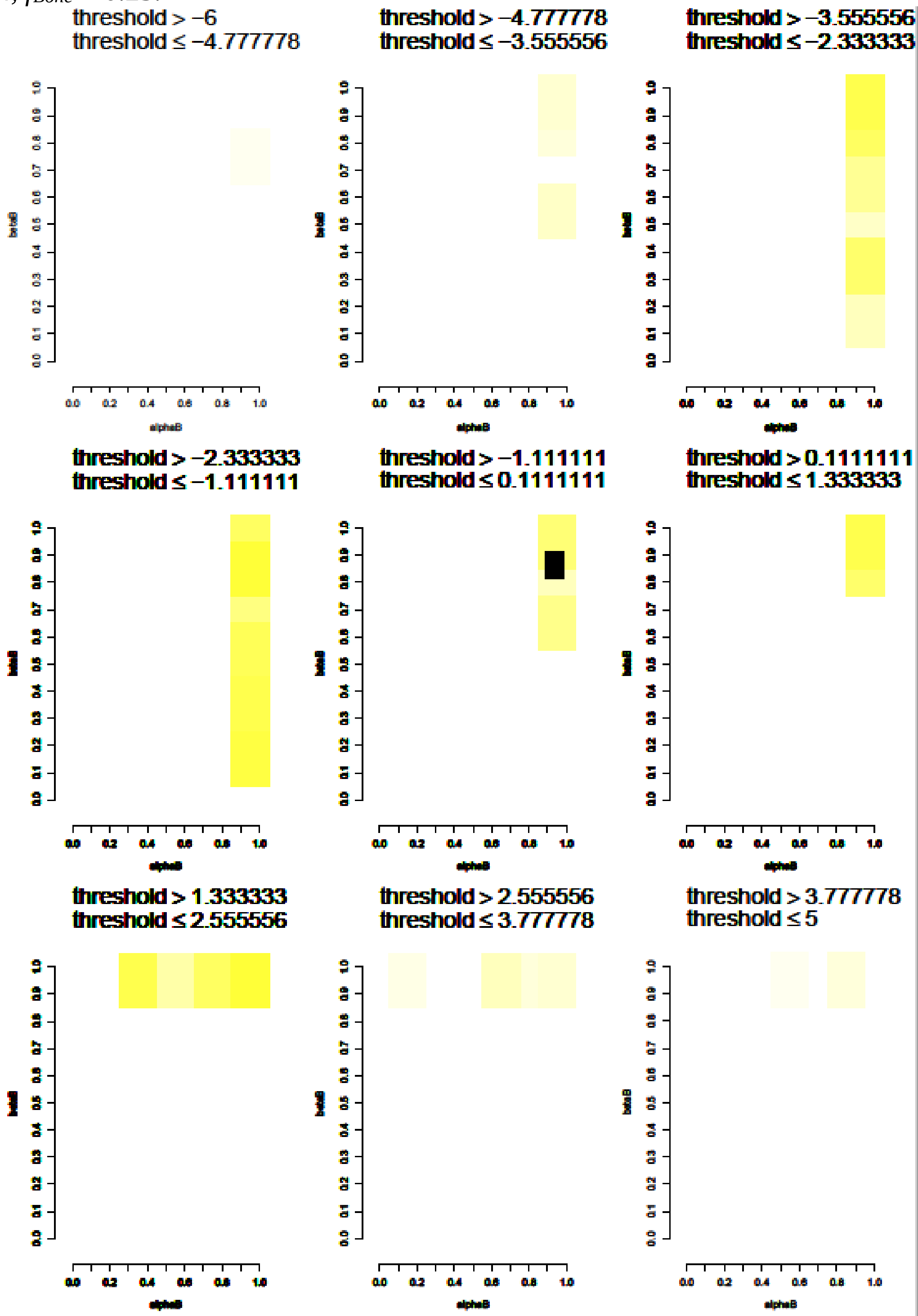


Figure 7iii: Modular cognition, domain A. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$.

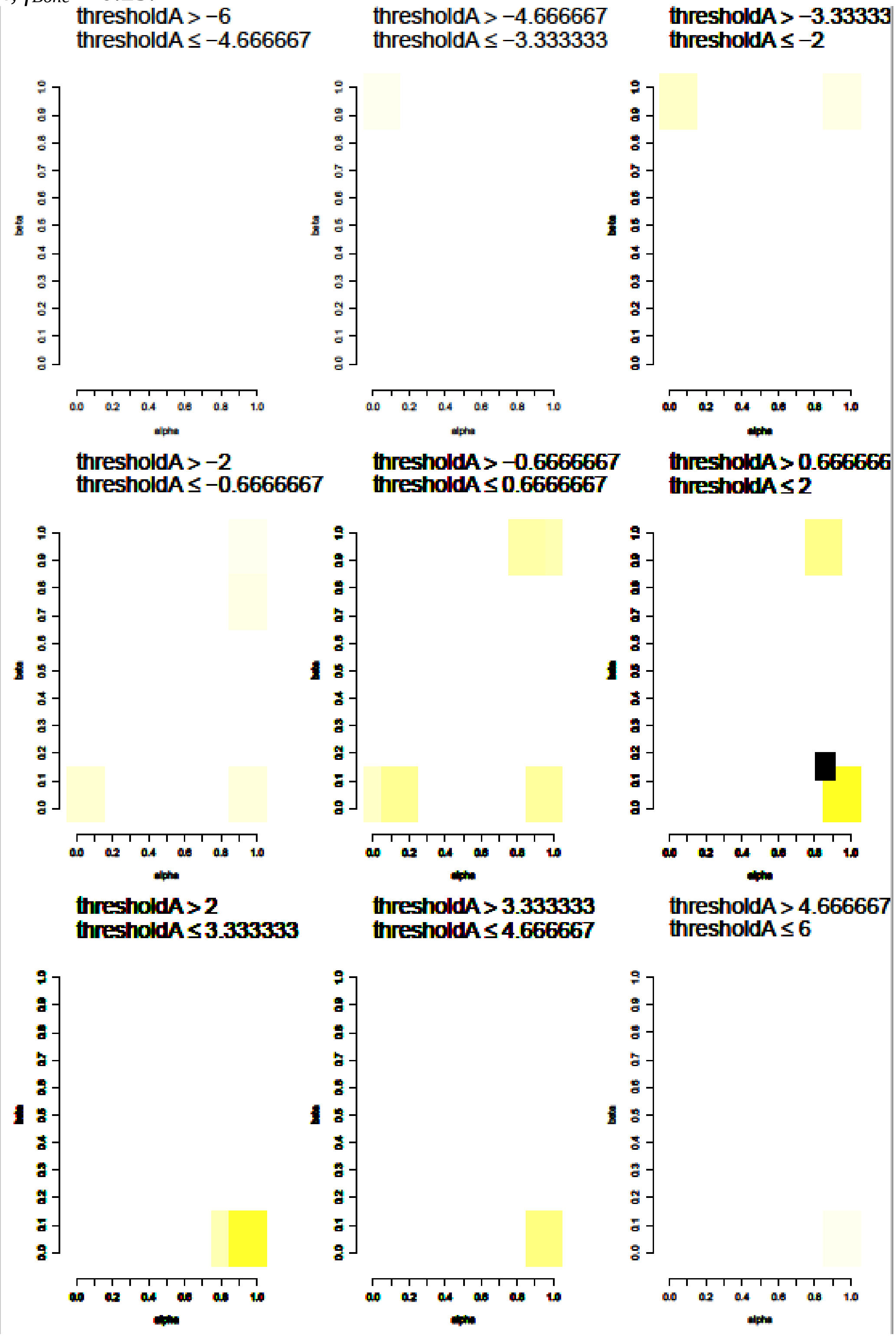


Figure 7iii: Modular cognition, domain B. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$.

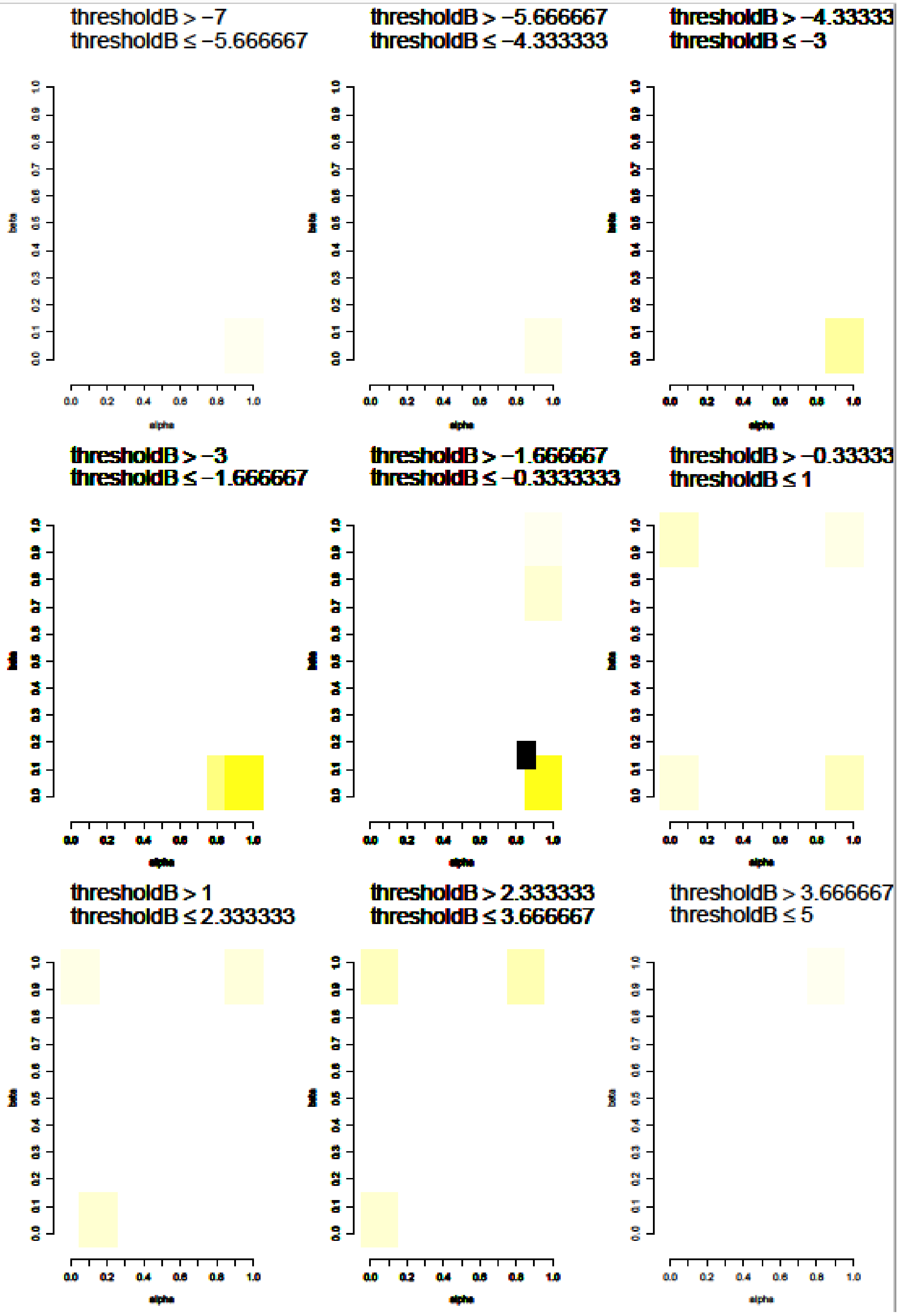


Figure 7iv: Domain-general. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$.

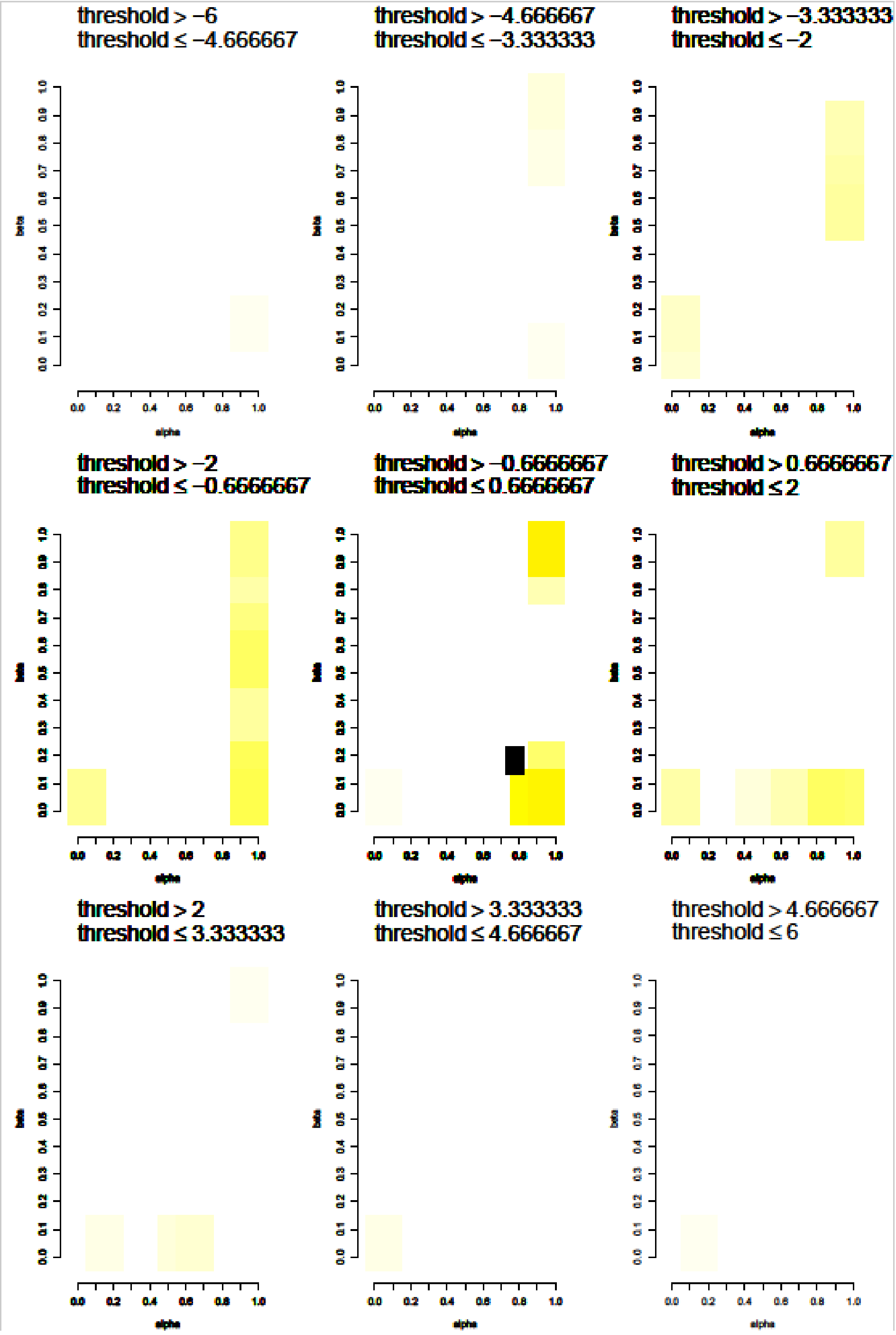


Figure 8i: Fully modular, domain A. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

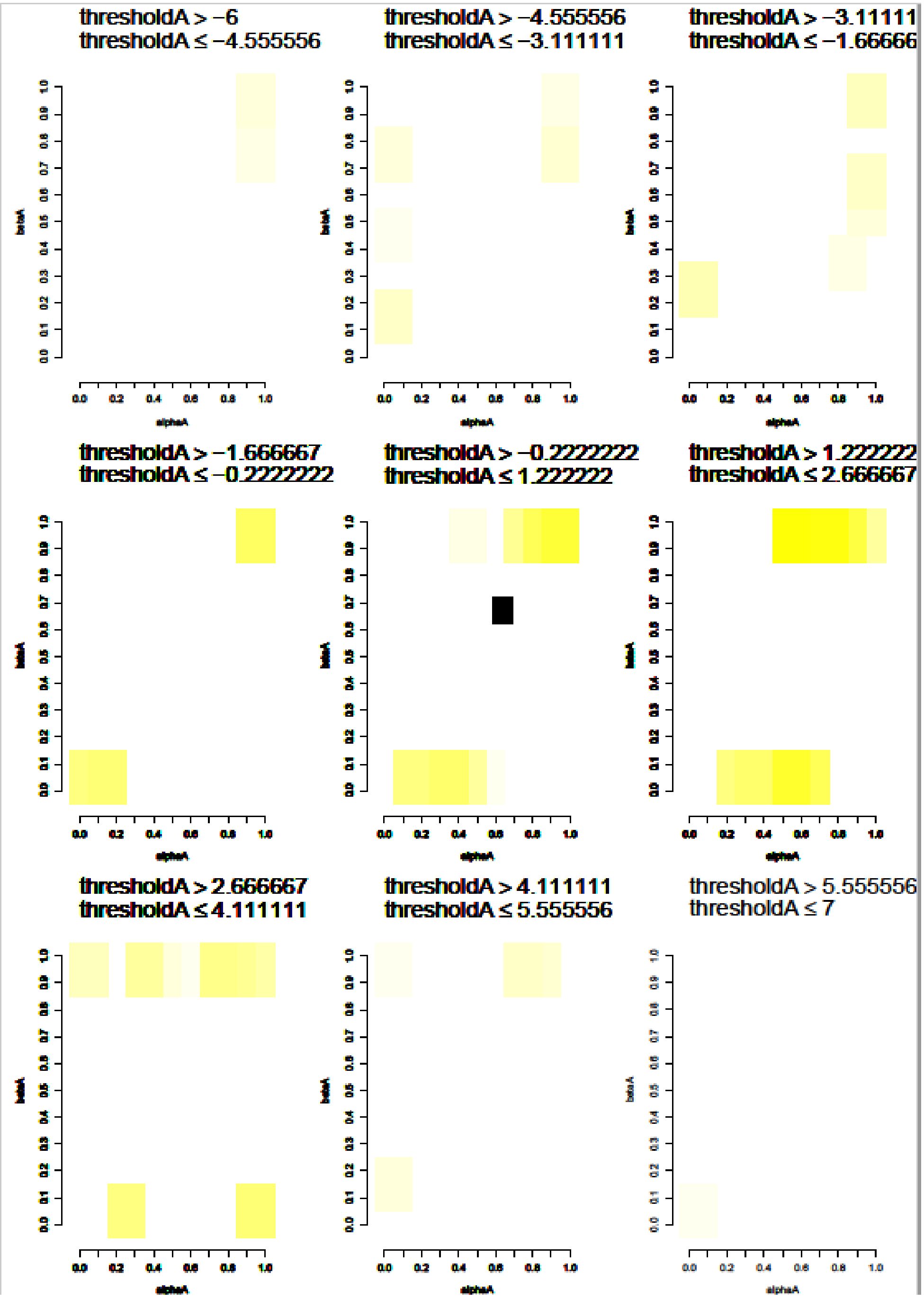


Figure 8i: Fully modular, domain B. p_A = 0.1, p_B = 0.9. f_{Azero} = 4, f_{Aone} = 0.25, f_{Bzero} = 0.25, f_{Bone} = 4.

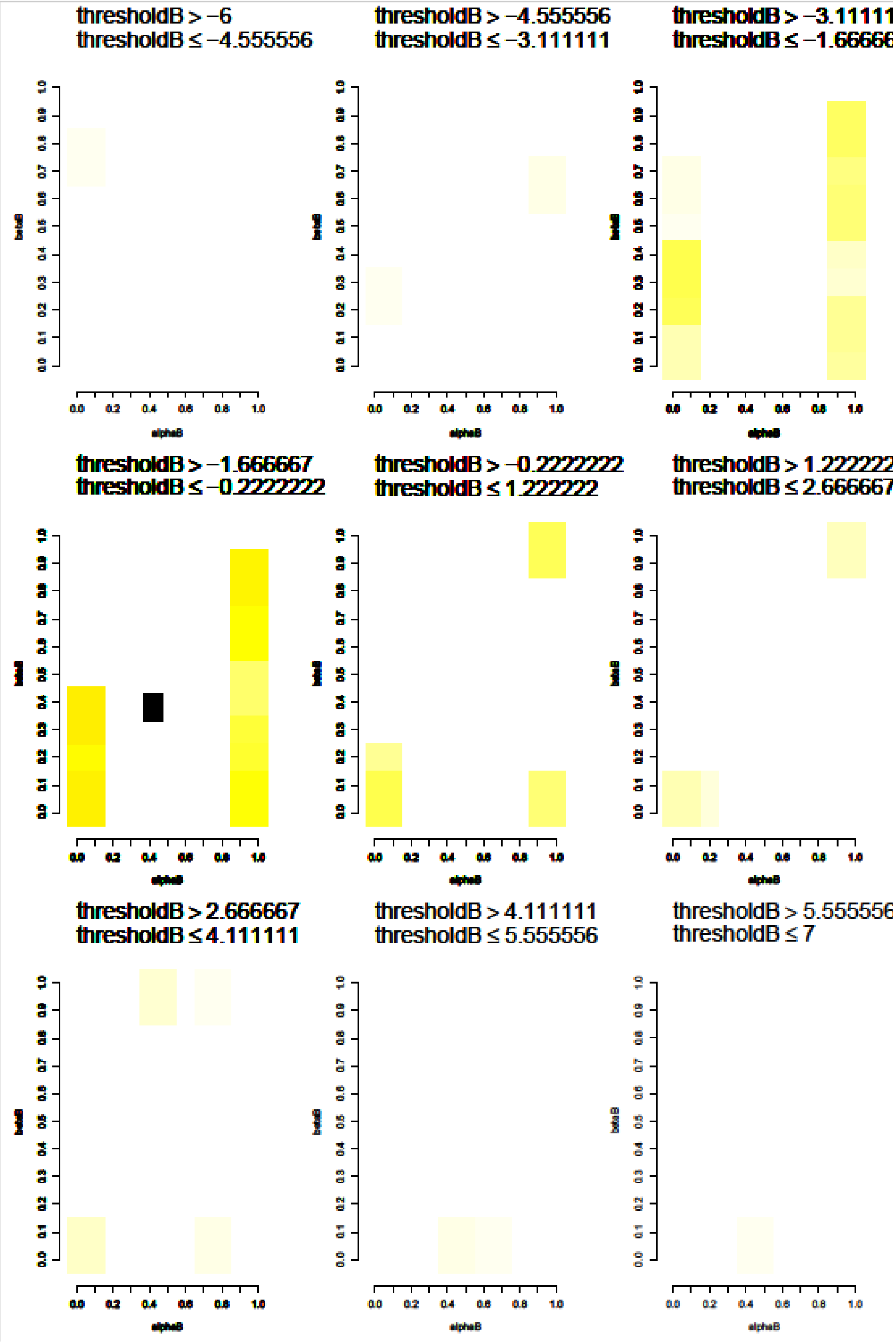


Figure 8ii: Modular motivation, domain A. p_A = 0.1, p_B = 0.9. f_{Azero} = 4, f_{Aone} = 0.25, f_{Bzero} = 0.25, f_{Bone} = 4.

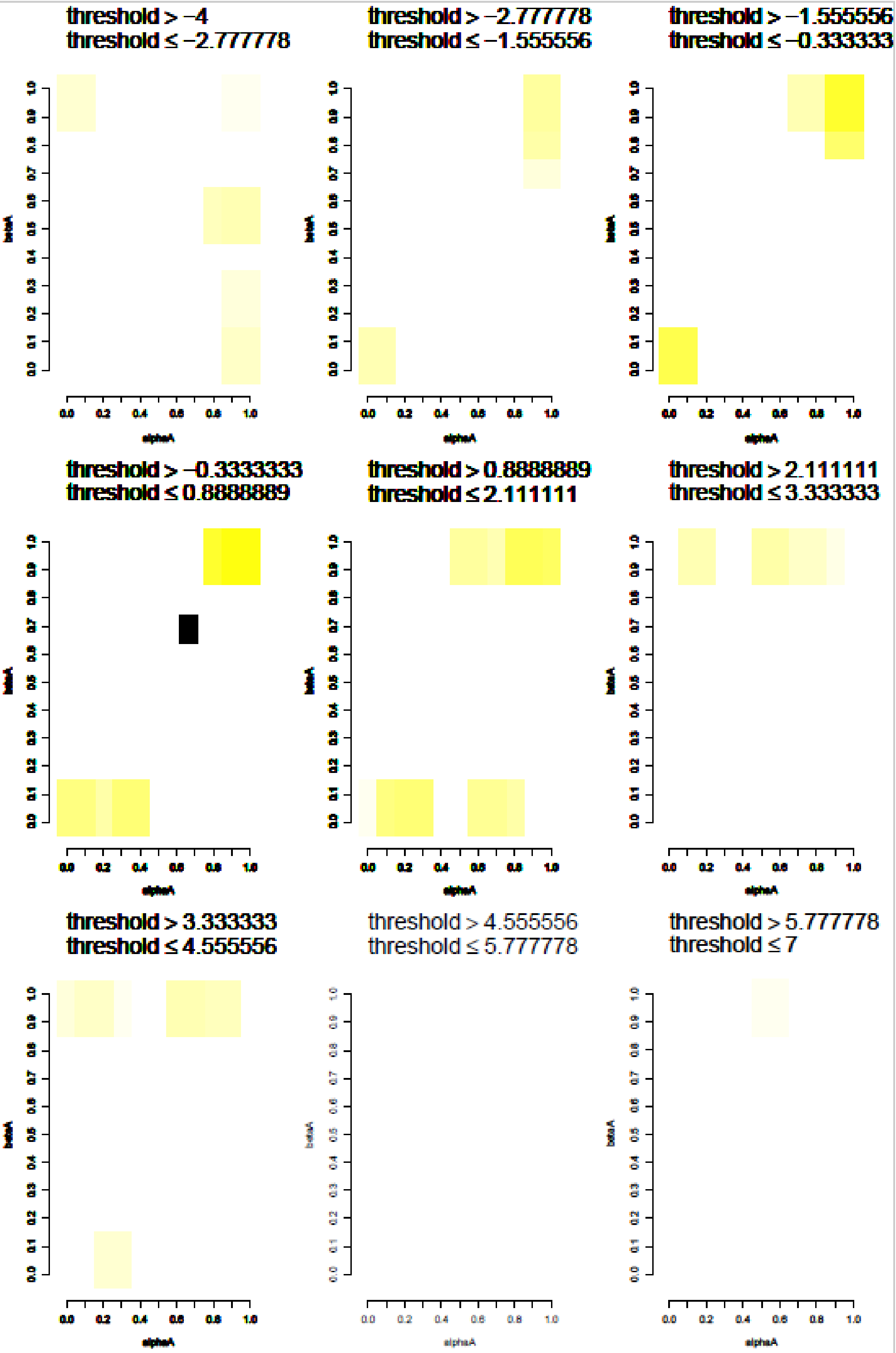


Figure 8ii: Modular motivation, domain B. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

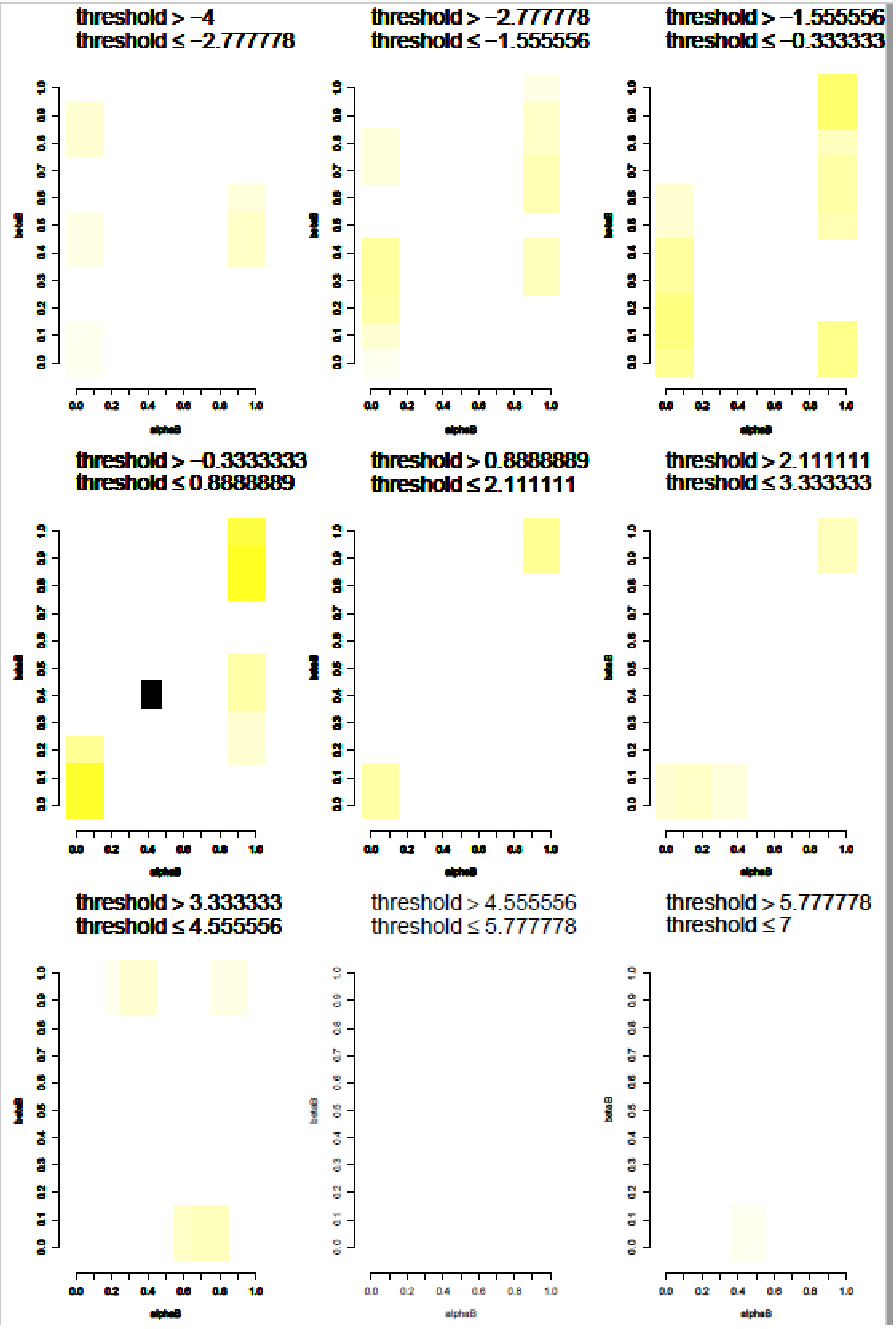


Figure 8iii: Modular cognition, domain A. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

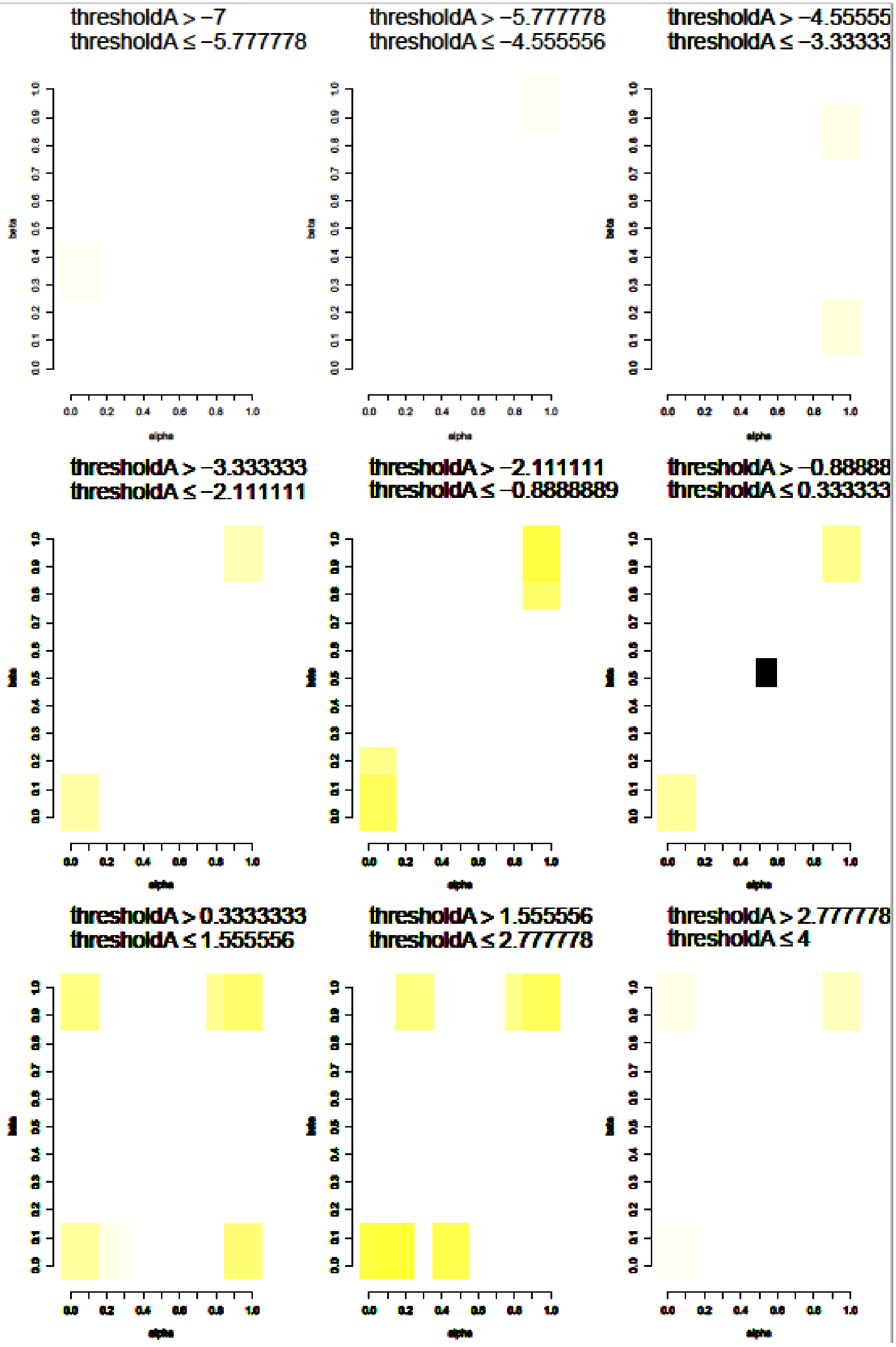


Figure 8iii: Modular cognition, domain B. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

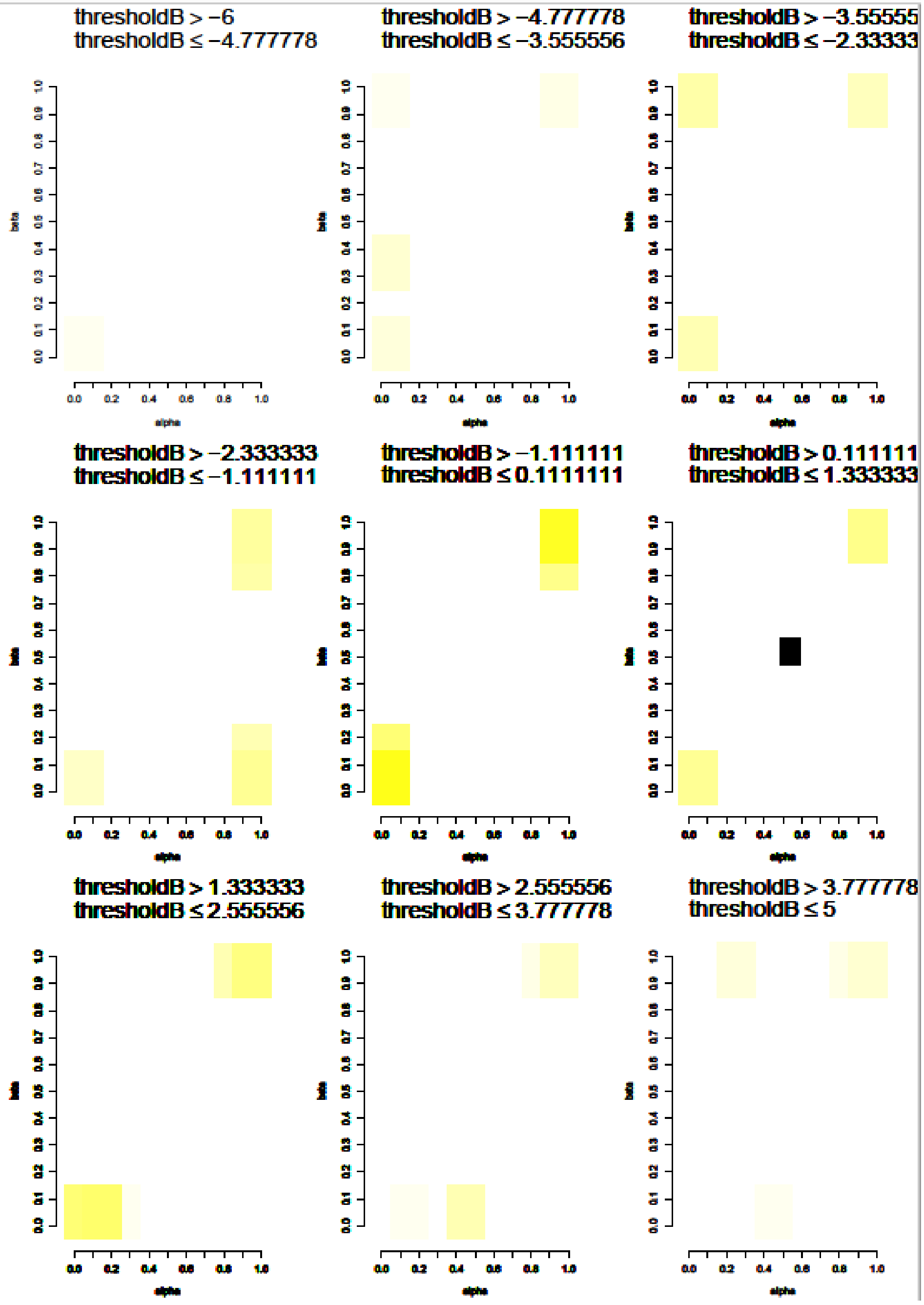


Figure 8iv: Domain-general. $p_A = 0.1$, $p_B = 0.9$. $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$.

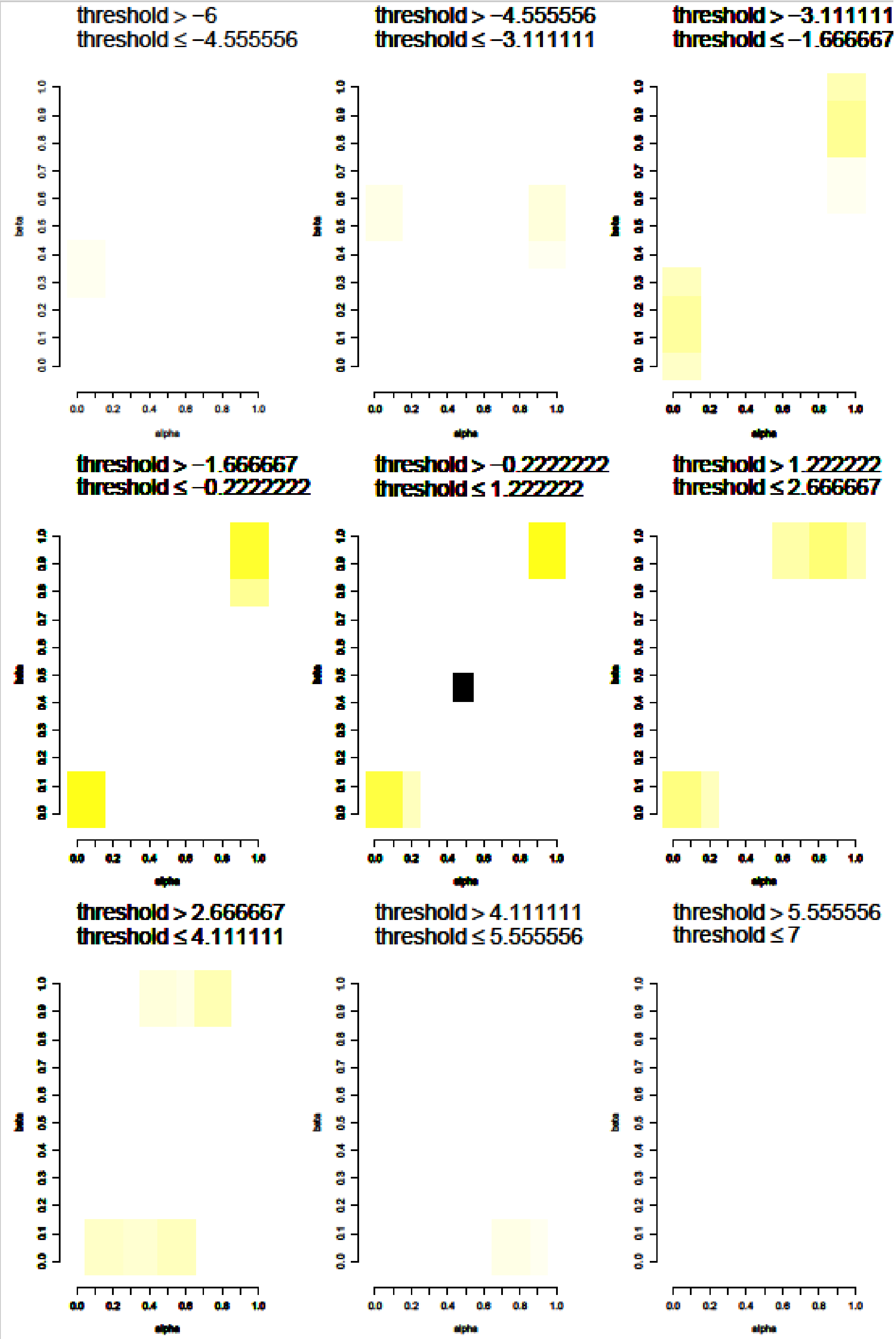


Figure 7 and Figure 8. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) agents with modular motivation only; (iii) agents with modular cognition only and (iv) fully domain-general agents when deciding to coordinate in two distinct domains. Figure 7 gives the results for runs where both the priors and the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 of domain A but behaviour 1 in state 1 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$). Figure 8 gives the results for runs where the priors favour state 0 in domain A and state 1 in domain B, but the fitness tied to suboptimal coordination favour coordination on behaviour 1 in state 1 of domain A and behaviour 0 in state 0 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).

3.3. Did the agents' psychological components affect their fitness?

I ran a regression to see how cognition and motivation coevolved to affect the fitness of all four agent types when coordinating in two social norm domains. I start with cases where the agents coordinate in two similar domains (table 2). Fully modular agents had a larger T_A predicting fitness on both runs. On runs where the priors favoured state 1 but the fitness tied to coordination favoured state 0, then domain-general agents had a large T predict fitness (estimate = 0.02, p = 0.002). This is a cognitive bias as they needed more evidence to believe in a common event. Conversely, agents with modular cognition had a small T_B predicting fitness (estimate = -0.02, p = 0.005) and so they needed less evidence to believe in a common event. Overall, there was an inconsistent effect of cognition on fitness when the agent made decisions across two similar coordination domains.

All agent types had a higher fitness if they had smaller α values (see table 2). This allowed the agents to coordinate on the optimal social norm of behaviour 0. A smaller α makes sense on the first run, where both the priors and fitness tied to suboptimal coordination favoured coordination on behaviour 0 in state 0 ($p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).

Interestingly, the smaller α still predicted fitness even on runs where the priors favoured state 1, but the fitness tied to suboptimal coordination favoured coordination on behaviour 0 in state 0 ($p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). When the likely environment of coordination was not necessarily the most fit environment to coordinate on, then the most fit agents are those that manage to coordinate in the state with the highest expected payoff. In this case, the most fit agents played behaviour 0. For example, if a group typically dances as ritual but live on a calorific knife-edge, then the most fit agents would be those that coordinate on not dancing.

Table 2. The regression results displaying the psychological components that predict fitness for each of the four agent types in runs where the agents coordinate in two similar social domains. The fitness parameters always favour coordination on behaviour 0 in state 0 of both domains A and B ($f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). The results in italic text give the estimates for runs where state 0 is likely ($p_A = 0.1$, $p_B = 0.1$) and the results in non-italic text give the estimates for runs where state 1 is likely ($p_A = 0.9$, $p_B = 0.9$). Note that agents with domain-general cognition have the same threshold in domain B as they do in domain A, and agents with domain-general motivation have the same α and β motivational thresholds in domain B as they do in domain A. For this reason, I arbitrarily block out domain B and focus on domain A for these agent types.

Psychological component left to evolve	**Agent Types**			
	Fully modular agents	**Modular motivation agents**	**Modular cognition agents**	**Domain-general agents**
Intercept	*8.67 **** *(0.02)* 8.66 *** (0.03)	*8.64 **** *(0.02)* 8.55 *** (0.03)	*8.77 **** *(0.017)* 8.69 *** (0.02)	*8.77 **** *(0.01)* 8.70 *** (0.02)
Threshold A	*0.03 *** *(0.008)* 0.005 (0.007)	*0.01* *(0.009)* -0.001 (0.008)	*-0.007* *(0.008)* -0.0002 (0.007)	*0.004* *(0.007)* 0.02 ** (0.007)
Threshold B	*-0.01* *(0.008)* 0.006 (0.007)		*-0.014 .* *(0.007)* -0.02 ** (0.008)	
***α* motivation A**	*-0.31 **** *(0.07)* -0.44 *** (0.05)	*0.07* *(0.06)* -0.41 *** (0.05)	*-0.35 **** *(0.10)* -0.80 *** (0.07)	*-0.14 ** *(0.06)* -0.66 *** (0.06)
***α* motivation B**	*0.08* *(0.06)* -0.23 *** (0.05)	*-0.13 .* *(0.07)* -0.37 *** (0.05)		
***β* motivation A**	*0.002* *(0.13)* 0.007 (0.06)	*0.03* *(0.15)* 0.09 (0.06)	*-3.05 **** *(0.40)* 0.08 (0.08)	*-0.81 **** *(0.17)* -0.10 (0.06)
***β* motivation B**	*-0.41 **** *(0.20)* -0.21 (0.06)	*0.15* *(0.12)* -0.02 (0.07)		

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p < 0.01)*
** (p< 0.05)*
· (trend: p=0.05 – 0.10 significance)

Table 3 represents runs where the agents decided whether to coordinate in two distinct domains, such as dancing and driving. The agents with modular cognition had a larger T_A (fully modular estimate = 0.017, p = 0.008; modular cognition only estimate = 0.22, p<0.001) and a smaller T_B (fully modular estimate = -0.03, p = 0.007; modular cognition only estimate = -0.32, p<0.001) predicting fitness (see table 3). They needed more evidence to believe in a rare event and less evidence to believe in a common event.

Agents with domain-general cognition did not have *T* values predicting fitness. This suggests that modular cognition was necessary when agents decided whether to coordinate across two distinct domains. When the priors and the fitness tied to coordination clashed ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then modular cognition could not predict fitness and domain-general cognition affected fitness inconsistently as it could not specialise to the demands of the discrepant domains (see table 3).

The pattern of results for motivation is striking. Agents with modular motivation only (estimate = -0.25, p = 0.03) and agents with modular cognition only (estimate = -0.65, p<0.001) had a smaller α_A predicting fitness, while domain-general and fully modular agents had a larger α_A predicting fitness (fully modular estimate = 0.11, p = 0.04; domain-general estimate = 0.27, p<0.001). Similarly, a larger β drove fitness for all agent types except the domain-general agent, who were more fit if they had a smaller β (estimate = -0.35, p<0.001). How the motivational thresholds affected fitness seemed to be tied to the size and direction of the cognitive thresholds across the runs. This suggests a path dependence effect where any shifts in the cognitive thresholds resulted in unique changes in the motivational thresholds (Bentley et al., 2004).

Note that there was an inconsistent effect of motivation on fitness across the four agent types on runs where the most common environment to coordinate in was not the most fit ($p_A = 0.1, p_B = 0.9;$ $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). This could highlight an influence of drift. Even random shifts in the cognitive threshold could affect the value of the motivational threshold and vice-versa. Perhaps this has implications for whether a group comes to coordinate on the common environmental state or the one with the highest payoffs.

The modular agents had more components predicting fitness in runs where the agents had to decide whether to coordinate in two distinct domains (table 3). This may suggest that modular psychology is important when deciding to coordinate over two distinct domains. I explore whether agent psychology is likely to be fully modular or domain-general, or in between these extremes, with an analysis comparing the fitness accrued by all four agent types in section 3.4.

Table 3. The regression results displaying the psychological components that predict fitness for each of the four agent types in runs with inconsistent priors favouring state 0 in domain A and state 1 in domain B ($p_A = 0.1$, $p_B = 0.9$). The results in italic text give runs where the fitness tied to suboptimal coordination also favour coordination on behaviour 0 in state 0 of domain A and coordination on behaviour 1 in state 1 of domain B ($f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$). The result in non-italic text gives the estimates for runs where the fitness tied to suboptimal coordination clash with the prior probabilities, to favour coordination on behaviour 1 in state 1 of domain A or coordination on behaviour 0 in state 0 of domain B ($f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). Note that agents with domain-general cognition have the same threshold in domain B as they did in domain A, and agents with domain-general motivation have the same α and β motivational thresholds in domain B as they do in domain A. For this reason, I arbitrarily block out domain B and focus on domain A for these agent types.

Psychological component left to evolve	Agent Types			
	Fully modular agents	**Modular motivation agents**	**Modular cognition agents**	**Domain-general agents**
Intercept	*8.12 *** (0.14)* 8.15 *** (0.04)	*8.85 *** (0.15)* 8.13 *** (0.04)	*7.43 *** (0.06)* 8.29 *** (0.02)	*5.39 *** (0.05)* 8.30 *** (0.02)
Threshold A	*-0.006 (0.008)* 0.017 ** (0.006)	*-0.004 (0.010)* 0.03 ** (0.009)	*0.22 *** (0.015)* 0.02 ** (0.008)	*-0.03 * (0.012)* 0.005 (0.008)
Threshold B	*-0.03 *** (0.007)* 0.006 (0.007)		*-0.32 *** (0.01)* -0.003 (0.008)	
***α* motivation A**	*0.11 * (0.05)* 0.04 (0.06)	*-0.25 * (0.11)* 0.08 (0.06)	*-0.65 *** (0.09)* -0.011 (0.07)	*0.27 *** (0.06)* -0.005 (0.07)
***α* motivation B**	*0.15 (0.14)* -0.49 *** (0.04)	*-0.32 * (0.13)* -0.28 *** (0.05)		
***β* motivation A**	*-0.23 . (0.13)* 0.51 *** (0.05)	*-0.12 (0.10)* 0.40 *** (0.05)	*0.28 ** (0.09)* 0.05 (0.07)	*-0.35 *** (0.07)* 0.03 (0.07)
***β* motivation B**	*0.44 *** (0.07)* 0.09 (0.05)	*0.10 (0.08)* -0.03 (0.06)		

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p < 0.01)*
** (p< 0.05)*
· (trend: p=0.05 – 0.10 significance)

3.4: Is there a difference in the fitness of the four agent types?

Here, I compare the fitness of the four agent types (fully modular, partly modular with modular motivation, partly modular with modular cognition and domain-general; table 4). This can highlight the psychology that likely underlines our ability to coordinate on various social norms.

Table 4. The regression results displaying any differences in fitness of the four agent types in runs where i) both the priors and fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 ($p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$); ii) runs where the priors favour state 1 but the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 ($p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$); iii) runs where both the priors and the fitness tied to suboptimal coordination favour coordination on behaviour 0 in state 0 of domain A but behaviour 1 in state 1 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$) and iv) runs where the priors favour state 0 in A and state 1 in B, though the fitness tied to suboptimal coordination favour coordination on behaviour 1 in state 1 of domain A, but favour coordination on behaviour 0 in state 0 of domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). Note that the domain-general agents are the omitted category to which the other three agent types are compared.

Estimates for regression predicting fitness	**Prior probabilities and fitness tied to suboptimal coordination**			
	$p_A = 0.1$, $p_B = 0.1$ $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$	$p_A = 0.9$, $p_B = 0.9$ $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$	$p_A = 0.1$, $p_B = 0.9$ $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$	$p_A = 0.1$, $p_B = 0.9$ $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$
Intercept	8.73 *** (0.011)	8.46 *** (0.015)	5.53 *** (0.017)	8.31 *** (0.02)
Modular cognition	-0.06 *** (0.002)	-0.019 (0.02)	2.16 *** (0.02)	0.0002 (0.02)
Modular motivation	-0.09 *** (0.016)	-0.22 *** (0.02)	3.07 *** (0.02)	0.02 (0.02)
Fully modular	-0.12 *** (0.02)	-0.16 *** (0.022)	3.13 *** (0.02)	0.03 (0.02)

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*

*** (p < 0.01)*
** (p< 0.05)*
· *(trend: p=0.05 – 0.10 significance)*

We begin with runs where the agents made decisions over two similar domains where both the priors and costs of suboptimal coordination consistently favoured coordination on behaviour 0 in state 0 ($p_A = 0.1$, $p_B = 0.1$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). The modular agents accrued less fitness than domain-general agents (modular cognition estimate = -0.06, p<0.001; modular motivation estimate = -0.09, p<0.001; fully modular estimate = -0.12, p<0.001). This suggests that domain-general psychology was sufficient to allow coordination on the optimal social norm across two domains with similar fitness landscapes, such as dancing and chanting.

Now I turn to runs where the agents made decisions in two similar domains, but the most common environmental state was not necessarily the most fit to coordinate on ($p_A = 0.9$, $p_B = 0.9$; $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$). For example, when dancing and chanting as ritual were encouraged by a cooperative group, but the group then found themselves living on a calorific knife-edge. Agents with modular motivation actually accrued less fitness than domain-general agents (fully modular estimate = -0.16, p<0.001; modular motivation only estimate = -0.22, p<0.001) while there was no difference in fitness between domain-general agents and agents with modular cognition (estimate = -0.019, p = 0.4). Linear combinations confirmed that agents with modular cognition only had a higher fitness than agent with modular motivation only across these runs (see appendix 4). Thus, modular cognition may be more important to track contrasting priors and fitness pressures over both domains.

Now I consider runs over two distinct domains where behaviour 0 is favoured in domain A but behaviour 1 is favoured in domain B ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 0.25$,

f_{Aone} = 4, f_{Bzero} = 4, f_{Bone} = 0.25). This would be the case when domain A represents dancing and B represents driving. Modular agent types accrued more fitness than domain-general agent types (modular cognition estimate = 2.16, $p<0.001$; modular motivation estimate = 3.07, $p<0.001$; fully modular estimate = 3.13, $p<0.001$). Linear combinations confirmed that fully modular agents accrued more fitness than agents with modular cognition only and agents with modular motivation only (see appendix 4). Thus, both modular cognition and modular motivation may be necessary when trying to coordinate over multiple distinct domains.

Finally, I consider two distinct domains where the most common environmental state is not necessarily the most fit to coordinate in ($p_A = 0.1$, $p_B = 0.9$; f_{Azero}= 4, f_{Aone} = 0.25, f_{Bzero} = 0.25, f_{Bone} = 4). There is no difference in fitness between any of the four agent types (see appendix 5). All agent types may be at risk of coordinating on maladaptive social norms when the most common state clashes with the most fit state to coordinate on across distinct domains. This highlights drift as a deciding factor in whether agents coordinate on optimal or suboptimal social norms in runs where the priors and fitness pressures clash over two distinct domains.

3.5. Summary of the key results

For brevity, my discussion of social coordination will focus on three key findings from the results:

1. Domain-general agents could uphold social coordination on two similar social domains.
2. Modular agent types could uphold social coordination on two distinct domains. Modular cognition could track the likelihood of coordination in each

distinct domain, while modular motivation may track cases where the most fit behaviour to coordinate on was not necessarily popular.

3. When the most common environmental state to coordinate on was not necessarily the most fit to coordinate on in two distinct domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then the runs would sometimes converge on agents who coordinated on the rare, but optimal, social norm; and other runs converge on agents who coordinated on the common, but suboptimal, social norm. Perhaps the equilibrium in the final generation of agents is largely affected by drift (Rorabaugh, 2014). This finding could have implications for our understanding of how maladaptive social norms are upheld at the group level.

4. Discussion

We conducted this model to investigate how (i) cognition and motivation coevolved to allow agents to coordinate on two distinct domains of social norms and (ii) compare how social coordination arose in domain-general, fully modular, or partly modular psychology types. I found that social coordination can be upheld by domain-general psychology provided that the two domains are similar in terms of fitness landscapes. Alternatively, modular psychology was likely to be needed when deciding to coordinate in two distinct domains.

To walk through these findings in more detail, consider an example of two social norms that are likely to align. More cooperative groups may require stricter coordination with group rituals to signal one's group membership (Reddish et al., 2013; Smaldino et al., 2018; Sosis et al., 2007). In line with this case, consider domain A to represent ‘dancing as a ritual’ and domain B to represent ‘chanting as a ritual’. There is clearly likely to be an overlap between the scenarios where dancing and chanting would both be ritualistic (Gelfand et al., 2020). Individuals may engage in collective song and dance before an important ceremony (Monteiro & Wall, 2011) such as the haka before rugby (Hartigan, 2011). Thus, chanting is likely to be upheld as a social norm in cases where dancing is a social norm.

Moreover, the costs of miscoordination will be similar in these two domains. It may be equally costly to fail to sing or dance when others do. Any deviation from a signal of group membership may attract negative attention (Legare, 2017), and it may be equally costly to dance or sing when inappropriate such as during a work meeting (Mu et al., 2015). When two such domains are similar in terms of fitness landscapes, then domain-general psychology appears to be sufficient to uphold social coordination in both domains.

Modular psychology may instead be necessary when deciding to coordinate in two distinct domains, such as dancing in domain A and driving on a certain side of the road in domain B. There is unlikely to be a correlation between which societies prefer to dance ritualistically and which societies prefer right-hand driving. This is because dance is likely a ritualistic norm (Legare, 2017) whilst deciding which side of the road to drive on is unlikely to be ritualistic (Hao et al., 2017).

In this scenario, it appears that the agents need to be fully modular. That is, both modular cognition and modular motivation are needed to specialise to the demands of both domains. Agents with modular cognition may track the commonality of a behaviour in two distinct domains. Agents with modular cognition could be aware that dancing as ritual is rare but driving on the right-hand side of the road is common, or vice-versa. Agents with modular motivation can perhaps track cases where the most likely environmental state would not necessarily be the most fit to coordinate in. To illustrate with an example, perhaps these agents avoid coordinating on a dance norm in cases where dancing is common but may waste calories as there is a lack of food.

This suggests that, when priors and costs of suboptimal coordination are not related over multiple domains, then a fully modular psychology may be necessary. There are multiple domains in which one has to coordinate in, and few of these are necessarily related (Henrich & Muthukrishna, 2021). For example, we must coordinate on certain social norms for safety (e.g., deciding which side of the road to drive on; Hao et al., 2017) and other norms we coordinate on to uphold social cohesion, such as which behaviours are considered moral (Curtin et al., 2020; Price, 2005).

While previous researchers have used the breadth of social norms to argue for extremely flexible domain-general psychology (Bolhuis et al., 2011), my model instead suggests that modular psychology is beneficial to ensure coordination in multiple social domains. Future work should investigate whether this arises from an evolved massively modular system (Cosmides & Tooby, 1994; Sperber & Hirschfeld, 1994) or emerges as a result of the environment in which we are raised in during development (Müller, 2007; Reader, 2006), or a mixture of both (Barrett & Kurzban, 2006).

Regardless of whether the agent has domain-general or modular psychology, there is a split occurring on runs where the most common state according to the priors is not necessarily the most fit state to coordinate in. Consider runs where the priors and the costs of suboptimal coordination favoured coordination on a different behaviour in both domains ($p_A = 0.1$, $p_B = 0.9$; $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzro} = 0.25$, $f_{Bone} = 4$). Half the runs found agents that coordinated on the rarer, but optimal, social norm; and the other half of runs found agents that coordinated on the common, but suboptimal, social norm.

To illustrate an example of when the most popular state to coordinate on is not necessarily the most fit, consider the declining fertility rates amongst middle- and upper-class WEIRD populations. The social and economic environments of these WEIRD populations typically favour having less children and investing heavily in them (Morita, 2018). Most individuals conform to this social norm, even though it may be more fit to have more children (Colleran, 2016). Thus, the most popular norm to coordinate on is not always the most fit.

It seems that coordinating on a behaviour because it is fit and coordinating on a behaviour because it is likely to be common are both strategies that can robustly emerge in agents who decide to coordinate under uncertainty. There was clearly a selection pressure acting on my agents to coordinate, as cases of miscoordination were rare by the final generation (see section 3.1). Perhaps this strong selection pressure to coordinate is similar to the runaway selection pressure for a norm psychology found in previous models (Gintis, 2003; 2011). That is, once large brains and other genetic biases have emerged– such as a tendency to feel shame when we break a social norm (Gintis, 2003; 2004)– then this creates a runaway selection pressure for a psychology which can uphold social norms and generate a wealth of new norms to coordinate on (Gintis, 2011; Markov & Markov, 2020; Tomasello & Gonzalez-Corbua, 2017). Although selection acts to create a pure-coordination equilibria by the final generation of agents, the exact norm that the agent comes to coordinate on may be influenced by drift. Drift describes the role of random sampling errors in affecting cultural evolution (Rorabaugh, 2014).

The question then becomes why agent psychology is biased towards coordination, though drift seemingly plays more of a factor in whether these agents coordinated on a rare but optimal social norm or a common but suboptimal one. Perhaps this is because cost asymmetries need only be non-trivial to affect the evolutionary trajectory of behaviour (Efferson et al., 2020b). This creates a scenario where the relatively small difference between zero fitness for miscoordination and some small fitness for suboptimal coordination creates more of a selection pressure on agent psychology than the large fitness differences between suboptimal and optimal social coordination. The greatest selection pressure is seemingly to avoid miscoordination. Drift will then play more of a role in whether we coordinate on an optimal or a suboptimal social norm.

Perhaps human societies have historically solved this issue of drift with social institutions which reward coordination on an arbitrary social norm (Henrich & Muthukrishna, 2021; Lenfesty & Morgan, 2019). This includes institutions such as the legal system (Chudek & Henrich, 2011) and religions (Lenfesty & Morgan, 2019; Gray & Watts, 2017). For example, the agent may be punished for failing to dance as a group norm by being ostracised (Rudert et al., 2020), and there are laws which punish driving on the wrong side of the road (Hao et al., 2017). Coordination may also be increased by introducing punishment to the model. This can reduce the appeal of deviating from certain social norms (Henrich et al., 2006; Mulder, 2008; Rudert et al., 2020) and may even increase the frequency of arbitrary or maladaptive social norms (Abbink et al., 2017; Bhui et al., 2019; Salali et al., 2015).

A wealth of gene-culture coevolutionary literature suggests that norms are socially learned (Boyd & Richerson, 1985 (chapter 7); Henrich & Muthukrishna, 2021; Legare, 2017; Legare & Nielsen, 2015; Wen et al., 2020). Moreover, some researchers argue that our ability to coordinate on a wide range of social norms may be underlined by a mixture of modular psychological processes plus domain-general learning rules (Chiappe & McDonald, 2005). Perhaps these domain-general social learning rules are useful to reach a stable equilibrium in environments where agents must choose between coordinating on a behaviour which is likely to be rarer but give higher payoffs or coordinating on a behaviour which is common but gives suboptimal payoffs.

It is interesting to note that both modular and domain-general agents experienced a split where half of simulations ended with converging on a behaviour with the highest payoff, and half ended with converging on a behaviour that happened to be common but may give suboptimal payoffs. Put simply, even in the absence of institutions that promote coordination, a group may come to coordinate on a maladaptive social norm simply via drift.

Before this equilibrium is reached, there may be cases where a population consists of some agents who wish to coordinate on a behaviour which may be rarer in the environmental state but tied to higher payoffs, whilst other agents wish to coordinate on a behaviour simply as it is common in the environment. This has implications for the study of female genital cutting specifically. This behaviour does not behave like a typical social norm (Efferson et al., 2015). We would intuitively expect all– or at least most– members of a society to coordinate on a behaviour if it is a social norm. However, female genital cutting does not reach full saturation even in pro-cutting societies. Most societies are intermediate between the two extremes of having all women cut or uncut, and only one society is believed to cut 100% of women (Efferson et al., 2015). Though this initially seemed to suggest that this behaviour was not a social norm, my model perhaps suggests that these populations are not yet in equilibrium between coordinating on a behaviour due to expected payoffs or coordinating on a behaviour which is already popular. Those in the intermediate populations who choose to cut may coordinate on this behaviour simply as it happens to be common. The remaining members of the population may instead choose not to cut as they coordinate on the behaviour which is likely to be the most fit for the individual woman.

Cases where the population exists in a two-strategy equilibrium may be particularly effective to target as policy at intervention stage to reduce the spread of harmful social norms (Efferson et al., 2020a). When the members of a social group are yet to converge on the same social norm, then an understanding of how path effects and drift influences our behaviour could allow policy makers to design behavioural 'nudges' towards coordination on less harmful social norms (Berger, 2021). The effectiveness of these 'nudges' is precisely because these populations can still be drastically affected by drift. If a policy intervention is designed which can target behaviour towards a certain end-point, then this could produce a favourable shift in the evolutionary trajectory of this population's coordination behaviours.

Sometimes, these nudges are known as social tipping points. Tipping points occur when there is social information to suggest that other members of the group are willing to coordinate on a different behaviour than the most popular. These tipping points may increase coordination on optimal social norms (Berger, 2021). A tipping point can be designed to influence a population towards coordination on a certain social norm which is likely to be less harmful for the social group involved.

While the findings of this model may have important implications in our understanding of maladaptive social norms, it is worth noting that this model used a simple coordination game in order to investigate how domain-general versus modular cognition and/or motivation affected our ability to coordinate when uncertain in two distinct domains. These simplicities could be addressed by future research.

First, I only considered a limited range of p_A and p_B parameter values which were modelled separately in each domain. This is unlikely to capture the complexity of the wide range of social norms that human societies coordinate on (Henrich & Muthukrishna, 2021). Second, this model considered a coordination game where the agents only had two behavioural strategies to choose from. In reality, we are likely to have to converge on multiple social equilibria. Let us return to the conference delegate from the introduction. How did she know that it was appropriate to toast her colleagues with a champagne flute? After all, alcohol consumption is not considered appropriate in all societies (Mandelbaum, 1965). How did she choose champagne, over the many other different alcoholic beverages? Why did she propose a toast, as opposed to a thousand other things she may have done while drinking? When there is an open-ended number of behavioural strategies at play, then social coordination will perhaps be even more important (Henrich & Muthukrishna, 2021). Finally, this model assigned the agents to play the coordination games in pairs. This is the most simplistic form of this task, though an *n*-person coordination game is perhaps more realistic as this would mirror social coordination at the level of the entire social group (Wilson & Rhodes, 1997). Future extensions to my model may wish to consider these complexities.

In summary, domain-general psychology is sufficient to uphold coordination on two similar social norms, but modular psychology is necessary to uphold coordination on two distinct social norms. When the most common norm is not the fittest norm to coordinate on, then drift has the largest influence on whether my agents coordinated on the optimal, but rare; or the suboptimal, but common, social norm. Agents who coordinate on a social norm simply as it is common may uphold a maladaptive social norm. The influence of drift suggests that, in cases where the group has yet to reach a pure-coordination equilibrium regarding a certain behaviour, then these groups may be particularly effective to target with policy or intervention to reduce the spread of harmful social norms such as female genital cutting. In conclusion, modular psychology may uphold social coordination across a wealth of distinct social domains, but it does not prevent a social group from coordinating on maladaptive social norms.

5. References

Abbink, K., Gangadharan, L., Handfield, T., & Thrasher, J. (2017). Peer punishment promotes enforcement of bad social norms. *Nature communications*, *8*(1), 1-8. https://doi.org/10.1038/s41467-017-00731-0.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review, 113*(3), 628–647. https://doi.org/10.1037/0033-295X.113.3.628.

Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1547), 1443-1450. https://doi.org/10.1098/rspb.2004.2746.

Berger, J. (2021). Social Tipping Interventions Can Promote the Diffusion or Decay of Sustainable Consumption Norms in the Field. Evidence from a Quasi-Experimental Intervention Study. *Sustainability*, *13*(6), 3529. https://doi.org/10.3390/su13063529.

Bhui, R., Chudek, M., & Henrich, J. (2019). How exploitation launched human cooperation. *Behavioral Ecology and Sociobiology*, *73*(78), 1-14. https://doi.org/10.1007/s00265-019-2667-y.

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biol*, *9*(7), e1001109. https://doi.org/10.1371/journal.pbio.1001109.

Boyd, R., & Richerson, P. J. (1985). Frequency-dependent Bias and the Evolution of Cooperation. In R. Boyd & P.J. Richerson (Eds.) *Culture and the Evolutionary Process*, (p 204-241). Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (2007). Cultural adaptation and maladaptation: Of Kayaks and Commissars. In S.W. Gangestad & J.A. Simpson (Eds.) *The evolution* *of the mind: Fundamental questions and controversies* (p327-332). New York: Guilford Press.

Charbonneau, M. (2016). Modularity and recombination in technological evolution. *Philosophy & Technology*, *29*(4), 373-392. https://doi.org/10.1007/s13347-016-0228-0.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *The Journal of general psychology*, *132*(1), 5-40. https://doi.org/10.3200/GENP.132.1.5-40.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*(5), 218-226. DOI: https://doi.org/10.1016/j.tics.2011.03.003.

Colleran, H. (2016). The cultural evolution of fertility decline. *Philisophical Transactions of the Royal Society B,*, *371*(1692), 20150152. http://dx.doi.org/10.1098/rstb.2015.0152.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. Hirschfeld and S.Gelman (Eds.) *Mapping the mind: Domain specificity in cognition and culture* (pp. 84-116). New York: Cambridge University Press.

Curtin, C. M., Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M.T., Fitzpatrick, S., Gurven, M., Kanovsky, M., Kushnick, G., Laurence, S., Pisor, A., Scelza, B., Stich, S., von Rueden, C. & Henrich, J. (2020). Kinship intensity and the use of mental states in moral judgment across societies. *Evolution and Human Behavior*, *41*(5), 415-429. https://doi.org/10.1016/j.evolhumbehav.2020.07.002.

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences,* 108(32), 13335-13340. DOI: https://doi.org/10.1073/pnas.1102131108.

Efferson, C., Lalive, R., & Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. *Science, 321*(5897), 1844-1849. DOI: 10.1126/science.1155805.

Efferson, C., McKay, R., & Fehr, E. (2020b). The evolution of distorted beliefs vs. mistaken choices under asymmetric error costs. *Evolutionary Human Sciences, 2*, E27. doi:10.1017/ehs.2020.25.

Efferson, C., Vogt, S., Elhadi, A., Ahmed, H. E. F., & Fehr, E. (2015). Female genital cutting is not a social coordination norm. *Science, 349*(6255), 1446-1447. DOI: 10.1126/science.aaa7978.

Efferson, C., Vogt, S., & Fehr, E. (2020a). The promise and the peril of using social influence to reverse harmful traditions. *Nature human behaviour, 4*(1), 55-68. DOIhttps://doi.org/10.1038/s41562-019-0768-2.

Fodor, J. A. (1983). *The modularity of mind*: *An essay on faculty psychology.* England, MIT press.

Frankenhuis, W. E., & Ploeger, A. (2007). Evolutionary psychology versus Fodor: Arguments for and against the massive modularity hypothesis. *Philosophical Psychology*, *20*(6), 687-710. https://doi.org/10.1080/09515080701665904.

Gavrilets, S. (2020). The dynamics of injunctive social norms. *Evolutionary Human Sciences*, *2*, 60-80. doi:10.1017/ehs.2020.58.

Gelfand, M. J., Caluori, N., Jackson, J. C., & Taylor, M. K. (2020). The cultural evolutionary trade-off of ritualistic synchrony. *Philosophical Transactions of the Royal Society B*, *375*(1805), 20190432. https://doi.org/10.1098/rstb.2019.0432.

Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of theoretical biology*, *220*(4), 407-418. https://doi.org/10.1006/jtbi.2003.3104.

Gintis, H. (2004). The genetic side of gene-culture coevolution: internalization of norms and prosocial emotions. *Journal of economic behavior & organization*, *53*(1), 57-67.

Gintis, H. (2011). Gene–culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 878-888. https://doi.org/10.1098/rstb.2010.0310.

Gintis, H., Bowles, S., Boyd, R. T., & Fehr, E. (Eds.). (2005). *Moral sentiments and material interests: The foundations of cooperation in economic life* (Vol. 6). US: MIT press.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*(2), 241-253. https://doi.org/10.1007/s11211-008-0067-y.

Gray, R. D., & Watts, J. (2017). Cultural macroevolution matters. *Proceedings of the National Academy of Sciences*, *114*(30), 7846-7852. https://doi.org/10.1073/pnas.1620746114.

Hao, J., Sun, J., Chen, G., Wang, Z., Yu, C., & Ming, Z. (2017). Efficient and robust emergence of norms through heuristic collective learning. *ACM*

Transactions on Autonomous and Adaptive Systems (TAAS), *12*(4), 1-20. https://doi.org/10.1145/3127498.

Hartigan, R. (2011). Embarrassing Time, Performing Disunity Rugby, the haka, and Aotearoa-New Zealand in the United Kingdom. *Performance Research*, *16*(2), 37-43. https://doi.org/10.1080/13528165.2011.578728.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, *91*(2), 73-78. DOI: 10.1257/aer.91.2.73.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D.& Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767-1770. DOI: 10.1126/science.1127333.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes, C. (2012). New thinking: the evolution of human cognition. *Philisophical Transactions of the Royal Society B, 367*(1599), 2091-2096. https://doi.org/10.1098/rstb.2012.0111.

Heyes C. (2018). Enquire within: cultural evolution and cognitive science. *Philosophical Transaction Royal Society. B 373*: 20170051. http://dx.doi.org/10.1098/rstb.2017.0051.

Hindriks, F. (2019). Norms that make a difference: social practices and institutions. *Analyse & Kritik, 41*(1), 125-145. https://doi.org/10.1515/auk-2019-410109.

House, B. R. (2018). How do social norms influence prosocial development?. *Current Opinion in Psychology*, *20*, 87-91. https://doi.org/10.1016/j.copsyc.2017.08.011.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637-651. https://doi.org/10.1016/j.cognition.2013.09.001.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Kets, W., Kager, W., & Sandroni, A. (2021). The value of a coordination game. *Available at SSRN:* http://dx.doi.org/10.2139/ssrn.3796188.

Laker, P. R., Hoppitt, W., Weiss, M., & Madden, J. R. (2021). The modularity of a social group does not affect the transmission speed of a novel, socially learned behaviour, or the formation of local variants. *Proceedings of the Royal Society B*, *288*(1947), 20202614. https://doi.org/10.1098/rspb.2020.2614.

Laland, K. N. (1993). The mathematical modelling of human culture and its implications for psychology and the human sciences. *British Journal of Psychology*, *84*(2), 145-169. DOI: https://doi.org/10.1111/j.2044-8295.1993.tb02471.x

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences*, *114*(30), 7877-7883. https://doi.org/10.1073/pnas.1620743114.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in cognitive sciences*, *19*(11), 688-699. https://doi.org/10.1016/j.tics.2015.08.005.

Lenfesty, H., & Morgan, T. J.H. (2019). By reverence, not fear: prestige, religion, and autonomic regulation in the evolution of cooperation. *Frontiers in Psychology*, *10*, 2750. https://doi.org/10.3389/fpsyg.2019.02750.

Mandelbaum, D. G. (1965). Alcohol and culture. *Current Anthropology*, *6*(3), 281-293. https://doi.org/10.1086/200597.

Markov, A. V., & Markov, M. A. (2020). Runaway brain-culture coevolution as a reason for larger brains: Exploring the "cultural drive" hypothesis by computer modeling. *Ecology and Evolution, 10*(12), 6059-6077. https://doi.org/10.1002/ece3.6350.

McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition*, *182*, 95-108. https://doi.org/10.1016/j.cognition.2018.09.008.

Mesoudi, A. (2009). The cultural dynamics of copycat suicide. *PLoS One*, *4*(9), e7252. https://doi.org/10.1371/journal.pone.0007252.

Monteiro, N. M., & Wall, D. J. (2011). African dance as healing modality throughout the diaspora: The use of ritual and movement to work through trauma. *Journal of Pan African Studies*, *4*(6), 234-252.

Morita, M. (2018). Demographic studies enhance the understanding of evolutionarily (mal) adaptive behaviors and phenomena in humans: a review on fertility decline and an integrated model. *Population Ecology*, *60*(1), 143-154. DOI 10.1007/s10144-017-0597-y.

Mrazek, A. J., Chiao, J. Y., Blizinsky, K. D., Lun, J., & Gelfand, M. J. (2013). The role of culture–gene coevolution in morality judgment: Examining the interplay between tightness–looseness and allelic variation of the serotonin transporter gene. *Culture and brain*, *1*(2-4), 100-117. DOI 10.1007/s40167-013-0009-x.

Mu, Y., Kitayama, S., Han, S., & Gelfand, M. J. (2015). How culture gets embrained: Cultural differences in event-related potentials of social norm violations. *Proceedings of the National Academy of Sciences*, *112*(50), 15348-15353. https://doi.org/10.1073/pnas.1509839112.

Mulder, L. B. (2008). The difference between punishments and rewards in fostering moral concerns in social decision making. *Journal of Experimental Social Psychology*, *44*(6), 1436-1443. https://doi.org/10.1016/j.jesp.2008.06.004.

Müller, G. B. (2007). Evo–devo: extending the evolutionary synthesis. *Nature reviews genetics*, *8*(12), 943-949. https://doi.org/10.1038/nrg2219.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483.

Pietraszewski, D., & A. E., Wertz (2021). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science* DOI: 10.1177/1745691621997113 [preprint].

Price, M. E. (2005). Punitive sentiment among the Shuar and in industrialized societies: Cross-cultural similarities. *Evolution and Human Behavior*, *26*(3), 279-287. https://doi.org/10.1016/j.evolhumbehav.2004.08.009.

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Social Justice Research*, *21*(2), 228-240. DOI: https://doi.org/10.1007/s11211-008-0064-1.

Purzycki, B.G., Pisor, A.C., Apicella, C., Atkinson, Q., Cohen, E., Henrich, J., McElreath, R., McNamara, R.A., Norenzayan, A., Willard, A.K. & Xygalatas, D. (2018). The cognitive and cultural foundations of moral behavior. *Evolution and Human Behavior* 39(5), 490-501. https://doi.org/10.1016/j.evolhumbehav.2018.04.004.

Reader, S. (2006). Evo-devo, modularity, and evolvability: Insights for cultural evolution. *Behavioral and Brain Sciences, 29*(4), 361-362. doi:10.1017/S0140525X06399086.

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: synchrony, shared intentionality and cooperation. *PloS one*, *8*(8), e71182. https://doi.org/10.1371/journal.pone.0071182.

Rorabaugh, A. N. (2014). Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *Journal of Archaeological Science*, *49*, 255-264. https://doi.org/10.1016/j.jas.2014.05.016.

Rudert, S. C., Ruf, S., & Greifeneder, R. (2020). Whom to punish? How observers sanction norm-violating behavior in ostracism situations. *European Journal of Social Psychology*, *50*(2), 376-391. DOI: 10.1002/ejsp.2606.

Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior*, *36*(2), 86-94. https://doi.org/10.1016/j.evolhumbehav.2014.09.004.

Simon, C., & Hessen, D. O. (2019). Selection as a domain-general evolutionary process. *Behavioural processes*, *161*, 3-16. https://doi.org/10.1016/j.beproc.2017.12.020

Smaldino, P. E., Flamson, T. J., & McElreath, R. (2018). The evolution of covert signaling. *Scientific reports*, *8*(4905), 1-10. DOI: https://doi.org/10.1038/s41598-018-22926-1.

Sosis, R., Kress, H. C., & Boster, J. S. (2007). Scars for war: Evaluating alternative signaling explanations for cross-cultural variance in ritual costs. *Evolution and Human Behavior*, *28*(4), 234-247. https://doi.org/10.1016/j.evolhumbehav.2007.02.007.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L.A. Hirschfeld & S.A. Gelman (Eds.) *Mapping the mind: Domain Specificity in Cognition and Culture* (pp. 39-87). Australia: Cambridge University Press.

Sperber, D., & Hirschfeld, L. (1999). Culture, cognition, and evolution. In R. Wilson & F.Keil (Eds.), *MIT encyclopedia of the cognitive sciences*, (pp. 111-132). Cambridge, MA: MIT Press.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in cognitive sciences*, *8*(1), 40-46. https://doi.org/10.1016/j.tics.2003.11.002.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nature Reviews Neuroscience*, *18*(9), 559-567. DOI: https://doi.org/10.1038/nrn.2017.76.

Stephen, I. D. (2014). Putting the theory before the data: is “massive modularity” a necessary foundation of evolutionary psychology?. *Frontiers in psychology*, *5*, 1158. https://doi.org/10.3389/fpsyg.2014.01158.

Stokes, D., & Bergeron, V. (2015). Modular architectures and informational encapsulation: A dilemma. *European Journal for Philosophy of Science*, *5*(3), 315-338. https://doi.org/10.1007/s13194-015-0107-z.

Suchow, J. W., Bourgin, D. D., & Griffiths, T. L. (2017). Evolution in mind: Evolutionary dynamics, cognitive processes, and bayesian inference. *Trends in cognitive sciences*, *21*(7), 522-530. https://doi.org/10.1016/j.tics.2017.04.005.

Tomasello, M., & Gonzalez-Cabrera, I. (2017). The role of ontogeny in the evolution of human cooperation. *Human Nature*, *28*(3), 274-288. DOI: https://doi.org/10.1007/s12110-017-9291-1.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for n-person exchange: the evolutionary roots of organizational behavior. *Managerial and Decision Economics*, *27*(2-3), 103-129. https://doi.org/10.1002/mde.1287.

van den Berg, P., Molleman, L., & Weissing, F. J. (2012). The social costs of punishment. *Behavioral and Brain Sciences, 35*(1), 42-43. DOI:10.1017/S0140525X11001348.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource?. *Psychological science*, *21*(3), 384-390. https://doi.org/10.1177/0956797610361340.

Wen, N. J., Willard, A. K., Caughy, M., & Legare, C. H. (2020). Watch me, watch you: ritual participation increases in-group displays and out-group monitoring in children. *Philosophical Transactions of the Royal Society B, 375*(1805), 20190437. https://doi.org/10.1098/rstb.2019.0437.

Williams, D. (2020). Epistemic Irrationality in the Bayesian Brain. *The British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axz044.

Wilson, R.K. & Rhodes, C.M. (1997). Leadership and Credibility in N-Person Coordination Games. *Journal of Conflict Resolution, 41*(6), 767-791. DOI: https://doi.org/10.1177/0022002797041006003.

6. Appendices

Appendix 1: The full parameter space for the social norm model

The bash file (*runSimulations*.txt)* found at OSF link (see appendix 3) details the full parameter space controlled by the researcher. Most of these parameters were fixed. The parameters that I varied included:

- probStateAOne and probStateBOne variables (given by p_A and p_B in-text). These took one of 3 values: 0.1, 0.5 and 0.9.

- The fitness tied to coordinating with one's partner in the suboptimal state (given by f_{Azero}, f_{Aone}, f_{Bzero}, and f_{Bone} in-text). These took one of 3 values: 0.25, 1 and 4.

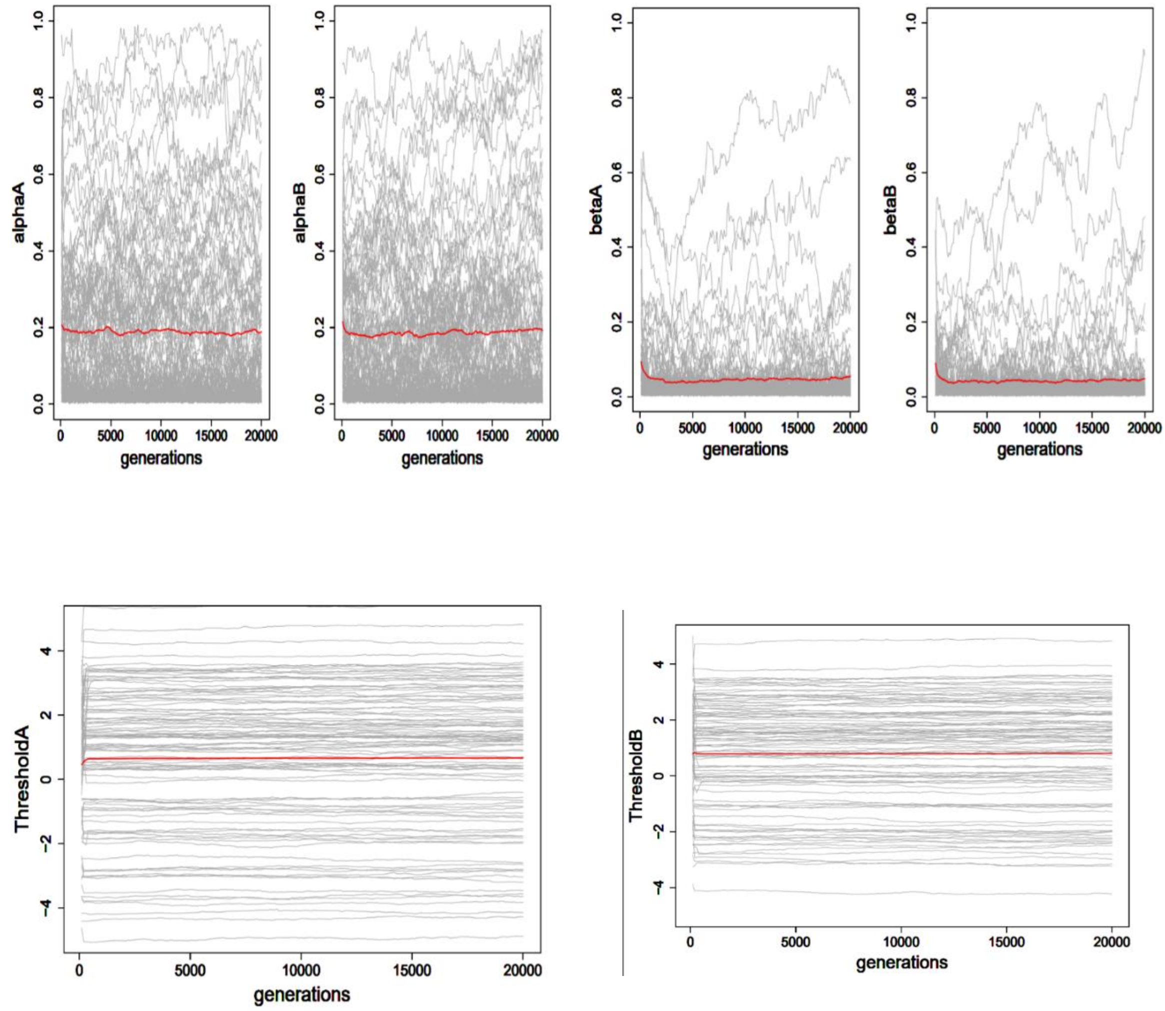
- The modularity of the agent type. I controlled these with the probFullMod, probTwoCog (for partly modular agents with modular cognition) and probTwoMotiv (for partly modular agents with modular motivation) respectively. These dummies took a 0 or 1 value. For example, if probFullMod was 1, then this run would be for all fully modular agents. If probFullMod was 0, then none of the agents would be fully modular and instead one of the other modularity types would be explored. Note that there is no variable modelled for domain-general agents. We modelled this so that, if probFullMod, probTwoCog and probTwoMotiv were all at 0, then all agents in this run would be domain-general by default. Whilst we consider runs to be unanimous across agent types, an interested researcher could force a split between all 4 agent types (for example, by coding probFullMod, probTwoCog and probTwoMotiv values to take 0.25 value each for an equal split). Then the agent type that rises to fixation may be evolutionary dominant over the others.

So, for each agent type we varied both the probability of the state being 1 and the fitness tied to suboptimal coordination. Each of these variables took one of 3 values, and we varied 6 such variables. This gave 3^6 variables which gave 729 runs. I ran these for each of the 4 agent types, giving 2916 runs overall. These runs are too extensive to investigate in the appendices and so I merely note that I upload the data to a public repository should they be of interest, but the analysis was only ran for the cases of interest for brevity: https://www.dropbox.com/sh/9bvr3svmnf21q8r/AADCXuhi4jr329aqsUgJ8EUaa?dl=0. There are multiple files uploaded, one for each of the parameter combinations. The files called *disagg** contain the data for the final generation of agents. These values were used to create the heatmaps in section 3.2, and the regressions in section 3.3. The file called *popDataAgentTypes** represented the cognitive and motivational values averaged over the four agent types of interest over the generations. This was used to make the line graph in appendix 2. The files called *popDataAgentTypesCorrExt** give the agent strategies for the agents at the extreme end of the modularity scale (e.g., fully modular, and domain-general agents). The files called *popDataAgentTypesCorrMod** give the agent strategies for the partly-modular agents (e.g., agents with modular motivation only and agents with modular cognition only). These datasets were used to make the clustered barchart in section 3.1. Finally, the *popDataMot** and *popDataThresh** files give data aggregated over the agent types (for the motivational thresholds and the cognitive thresholds respectively) but was unused in the analysis.

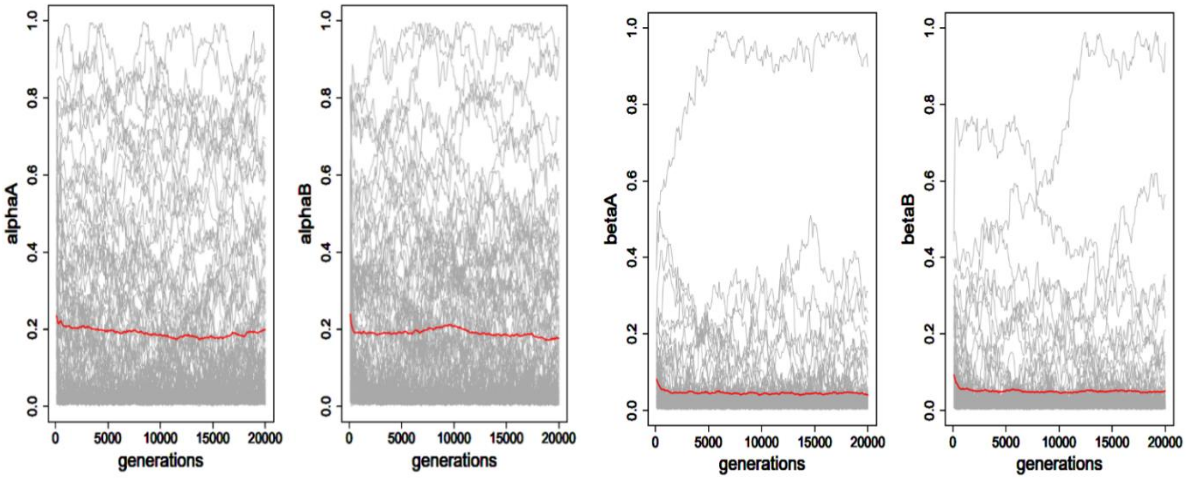
Appendix 2: The line graphs confirming that these runs converged on a stable psychological architecture.

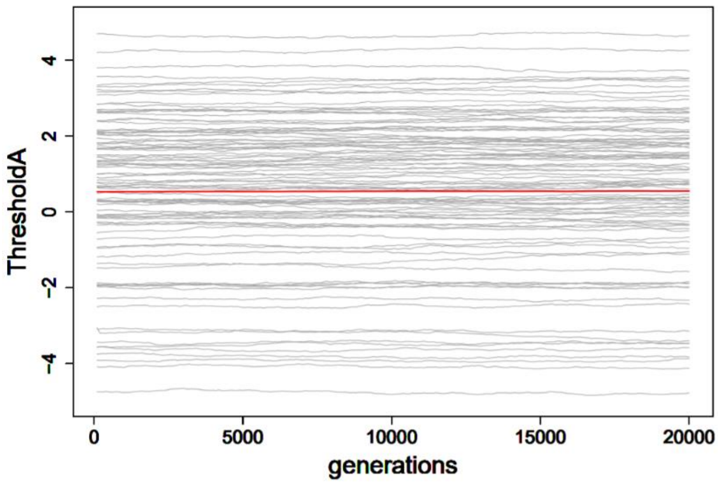
Appendix 2A: The line graphs for runs where the priors and fitness tied to coordination favoured coordinating on state 0 (p_A =0.1, p_B =0.1, f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4).

i) Fully modular

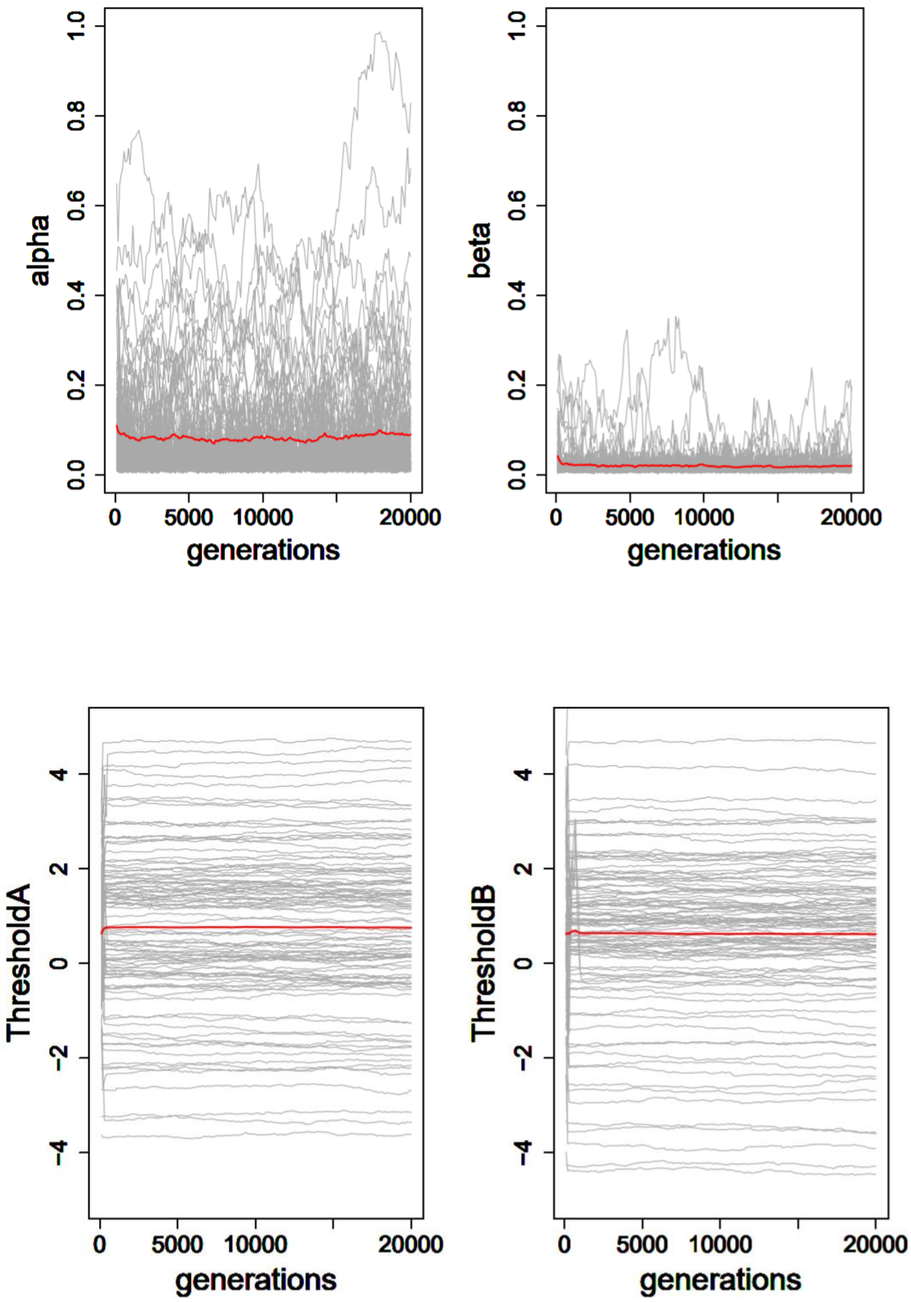


ii) Modular motivation

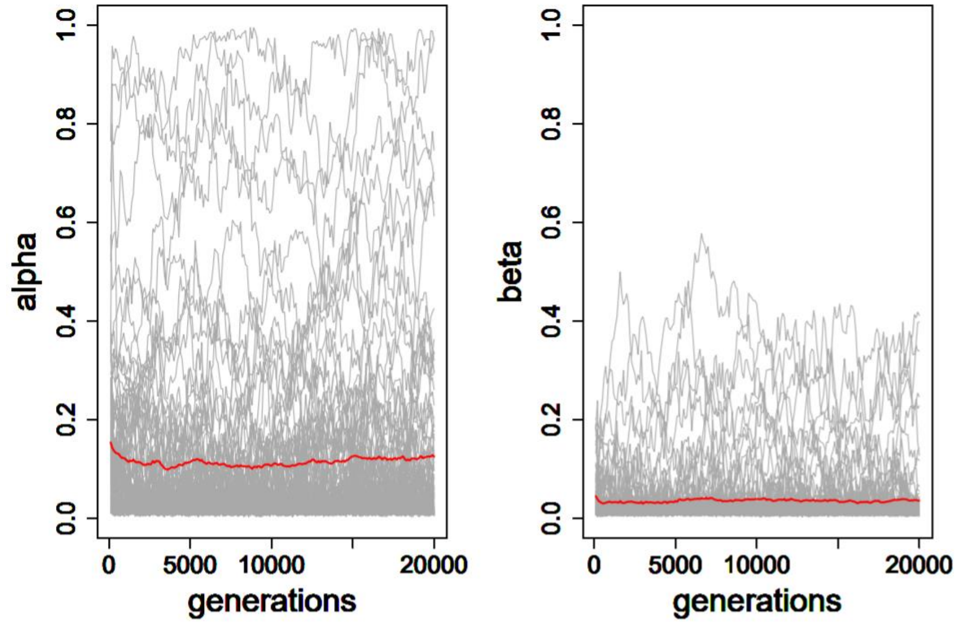


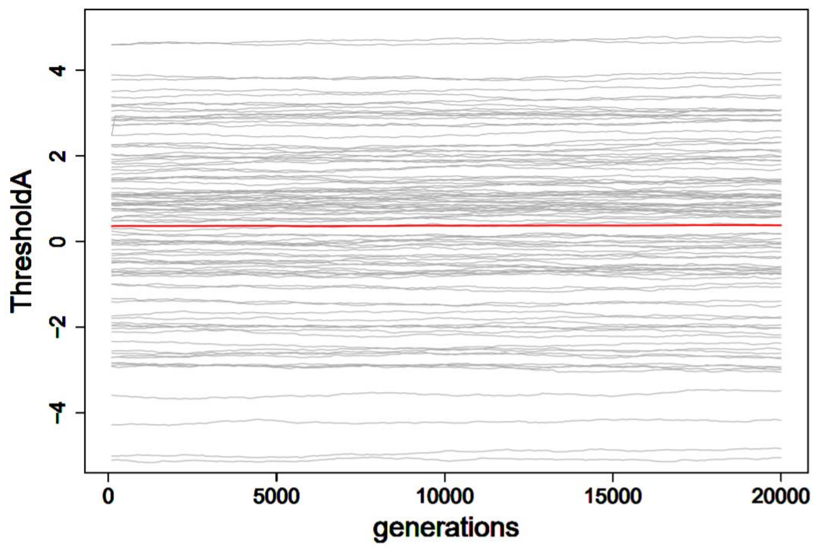


iii) Modular cognition



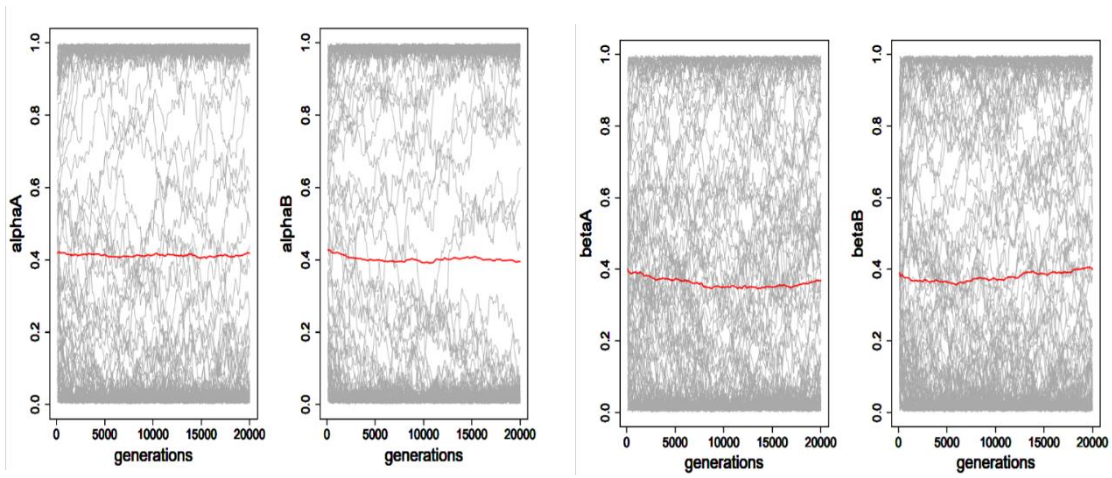
iv) Domain-general

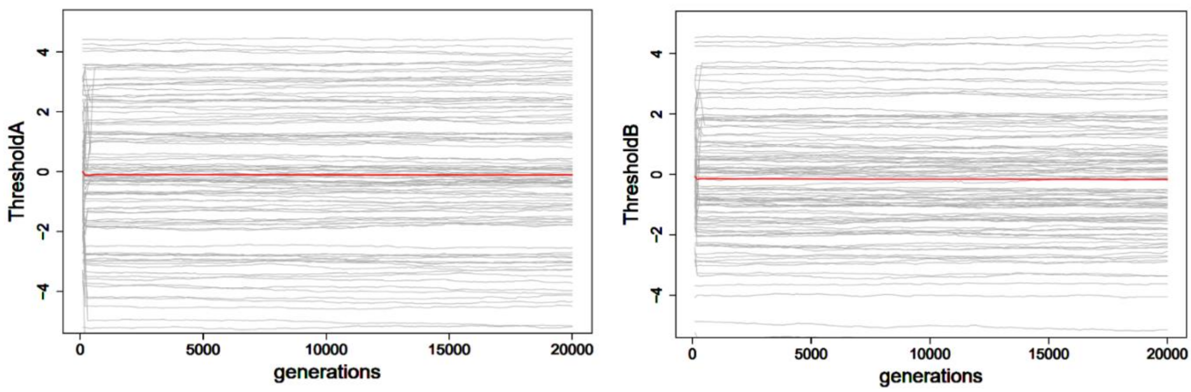




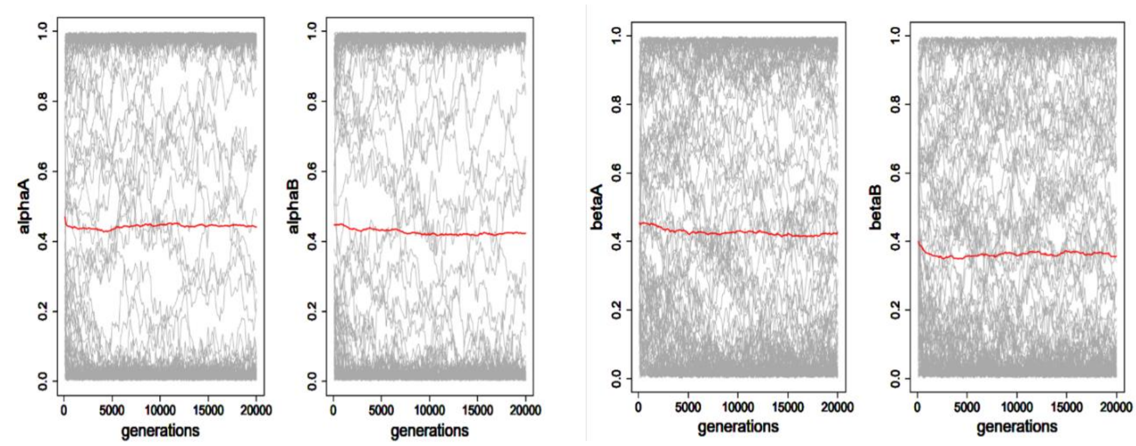
Appendix 2B: The line graphs for runs where the priors favoured state 1 though the fitness tied to coordination favoured coordinating on state 0 (p_A =0.9, p_B =0.9, f_{Azero} = 0.25, f_{Aone} = 4, f_{Bzero} = 0.25, f_{Bone} = 4).

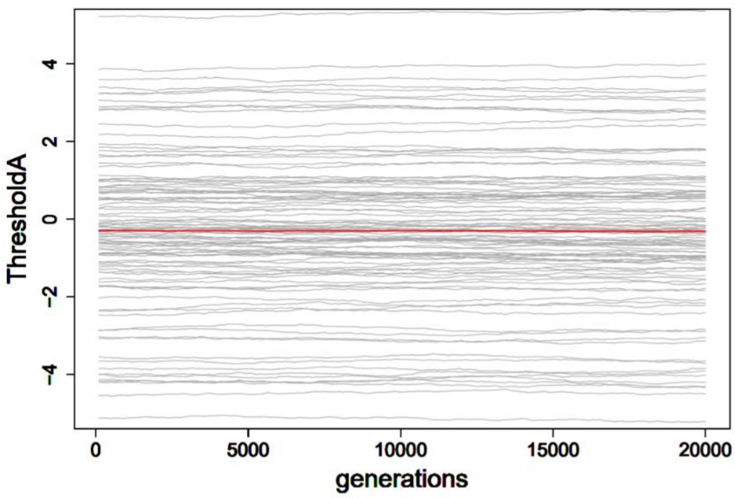
i) Fully modular



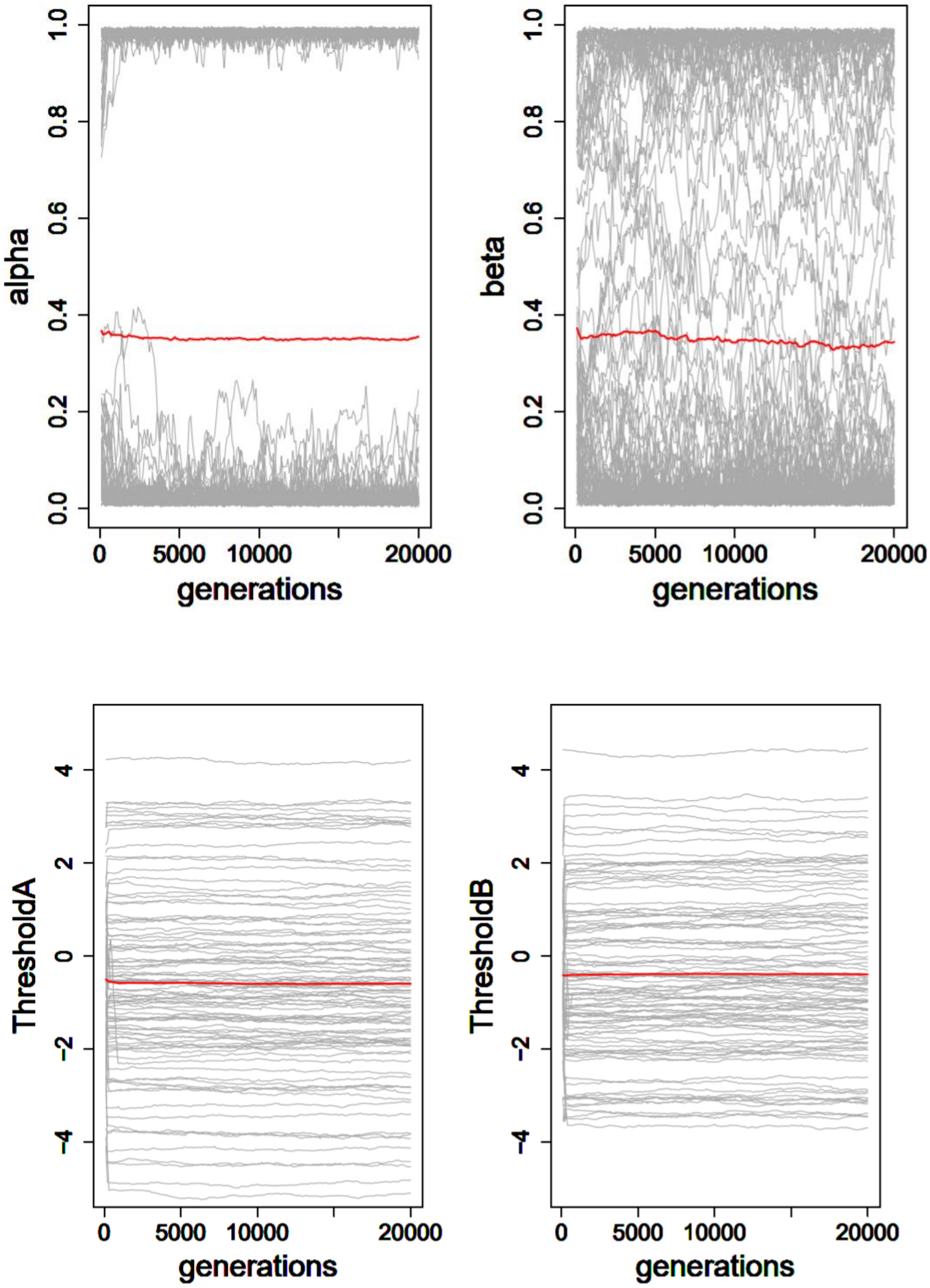


ii) Modular motivation

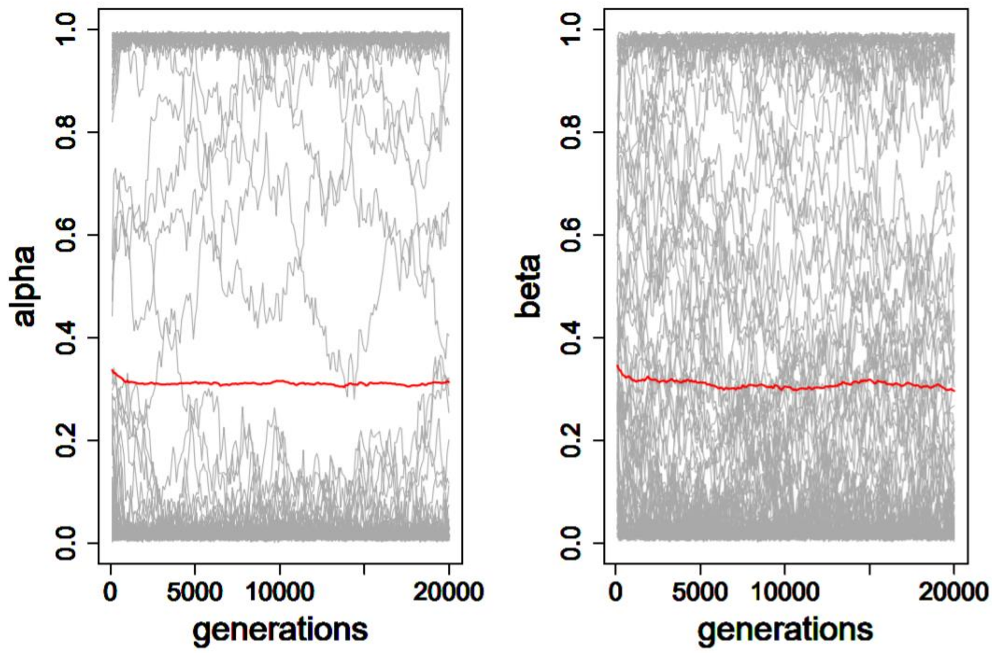


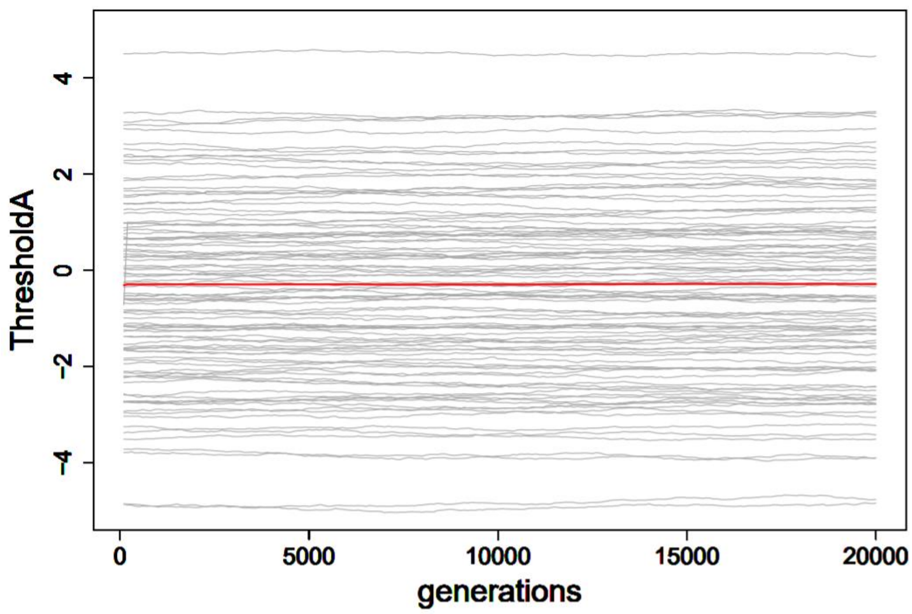


iii) Modular cognition



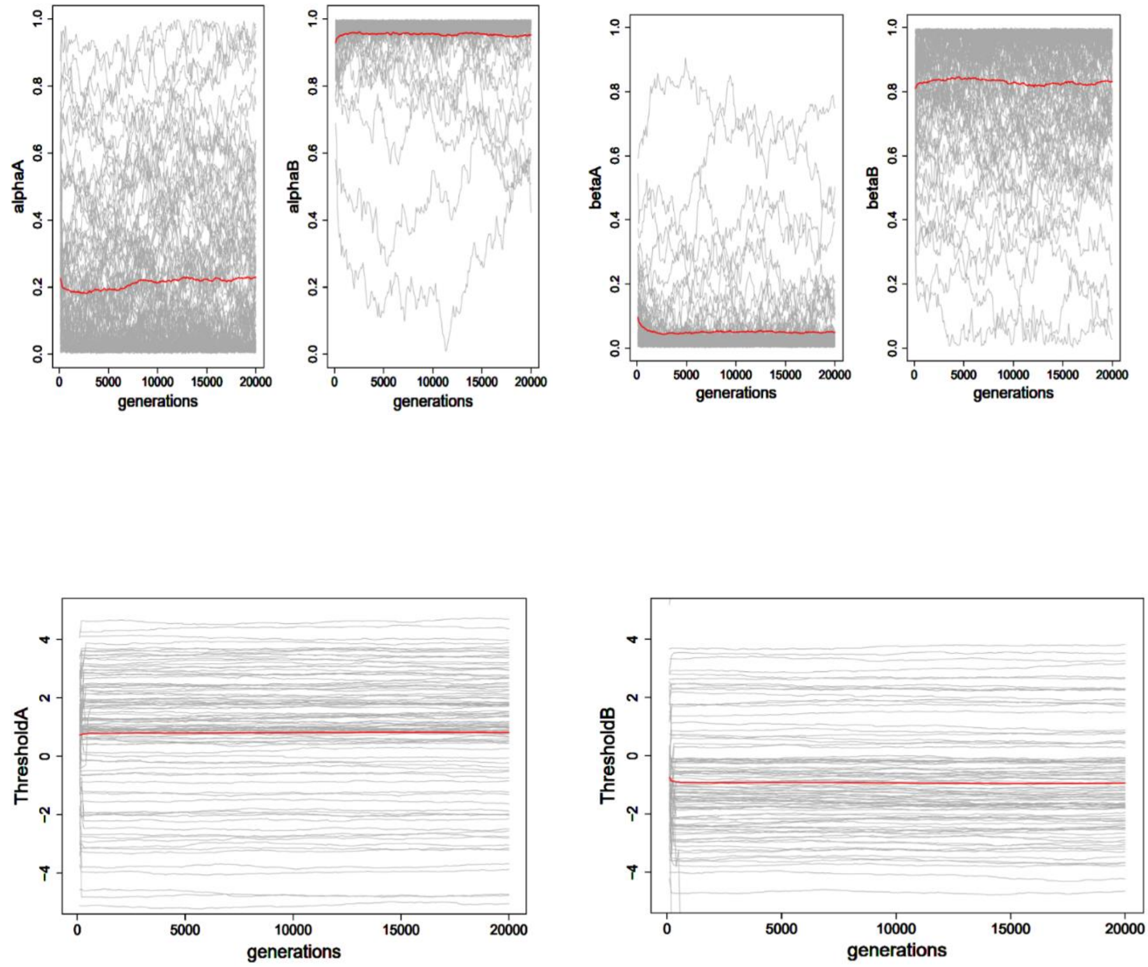
iv) Domain-general



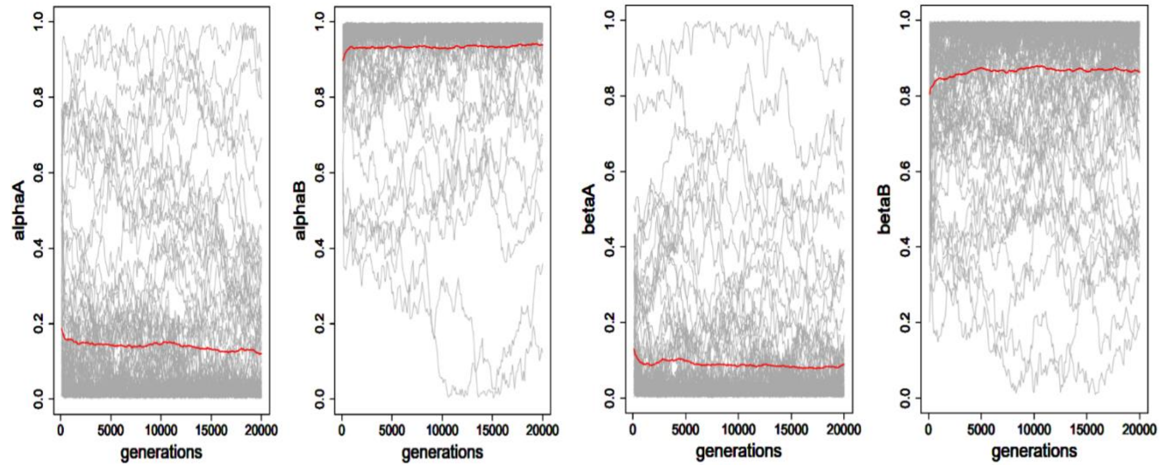


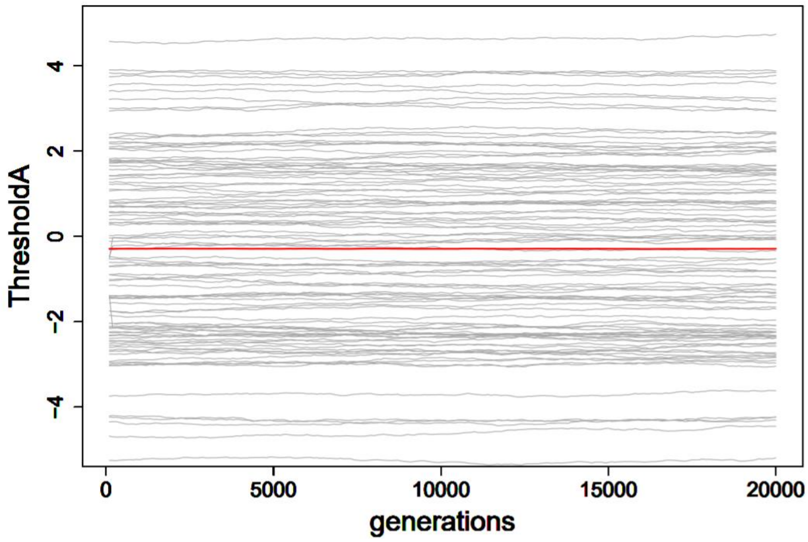
Appendix 2C: The line graphs for runs where the priors and fitness tied to coordination favoured state 0 in domain A but state 1 in domain B ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$).

i) Fully modular

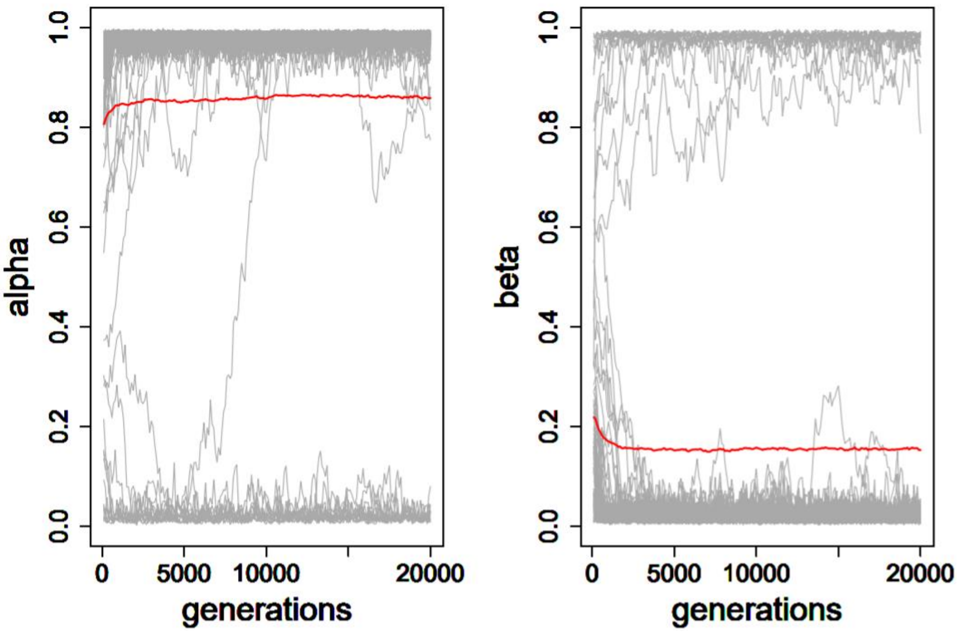


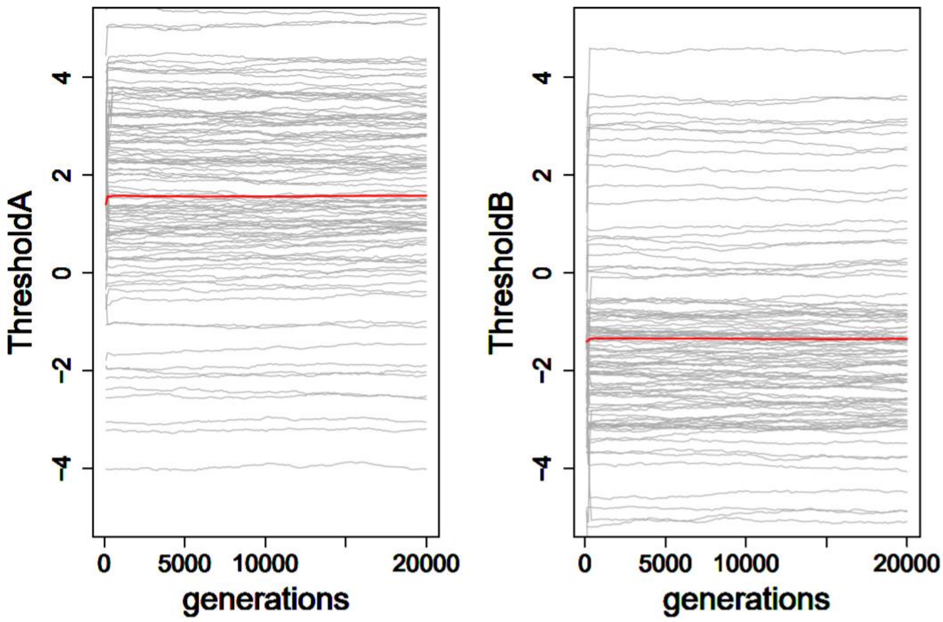
ii) Modular motivation



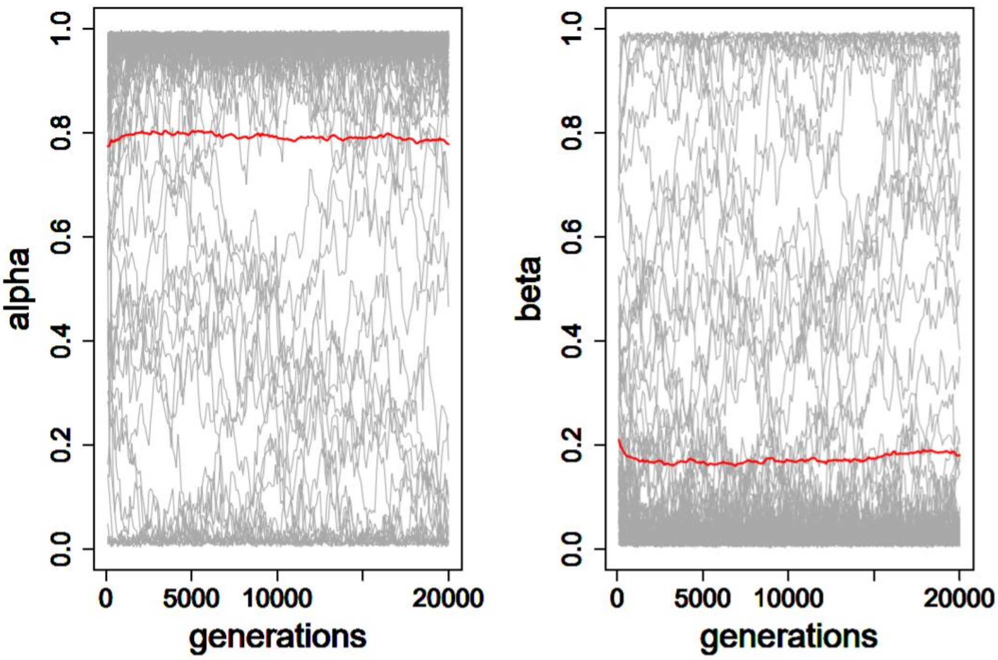


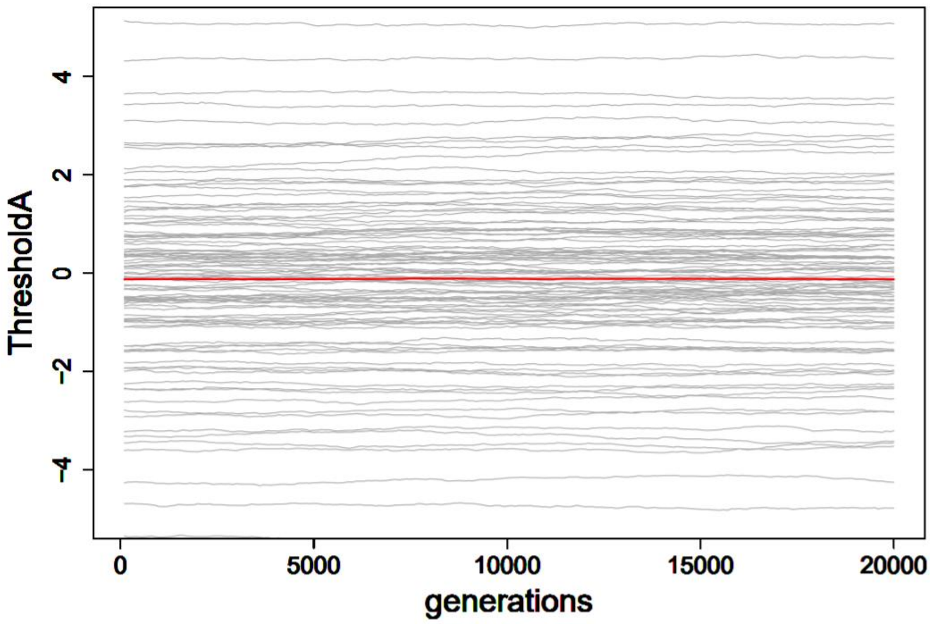
iii) Modular cognition





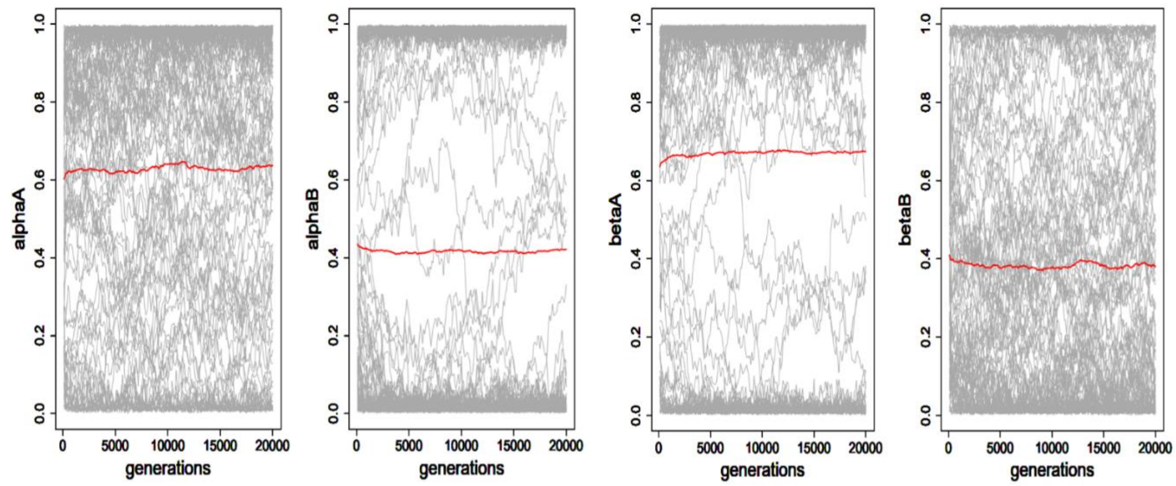
iv) Domain-general

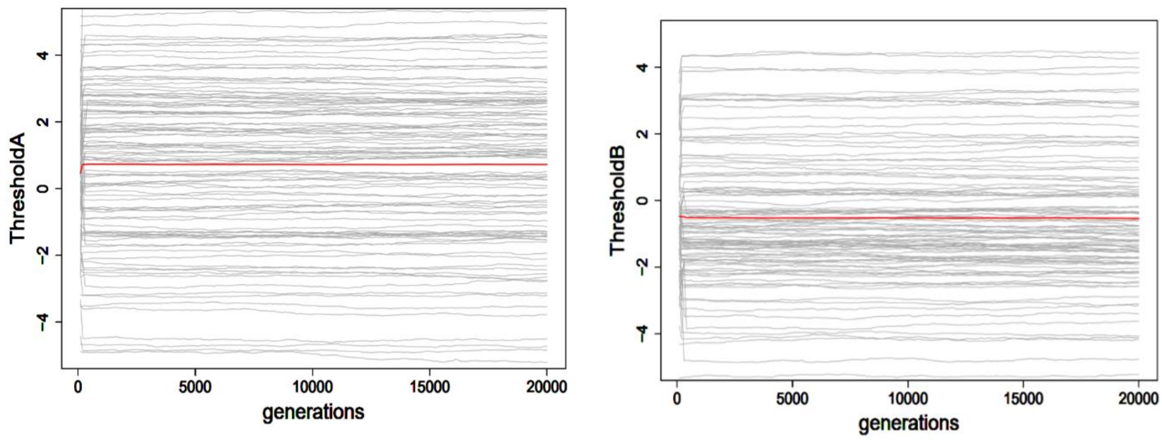




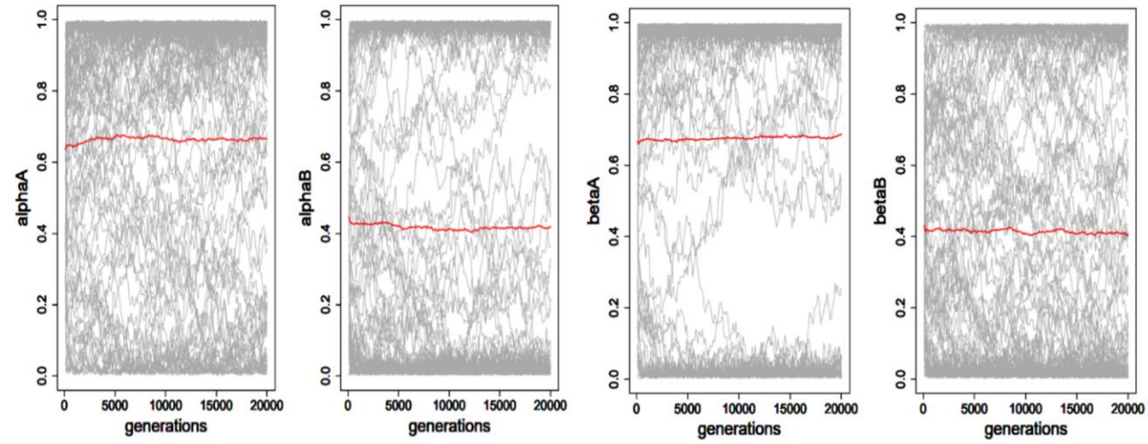
Appendix 2D: The line graphs for runs where the priors favour state 0 in domain A and state 1 in domain B, but the fitness pressures tied to suboptimal coordination favour coordination in state 1 of domain A but state 0 of domain B ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$).

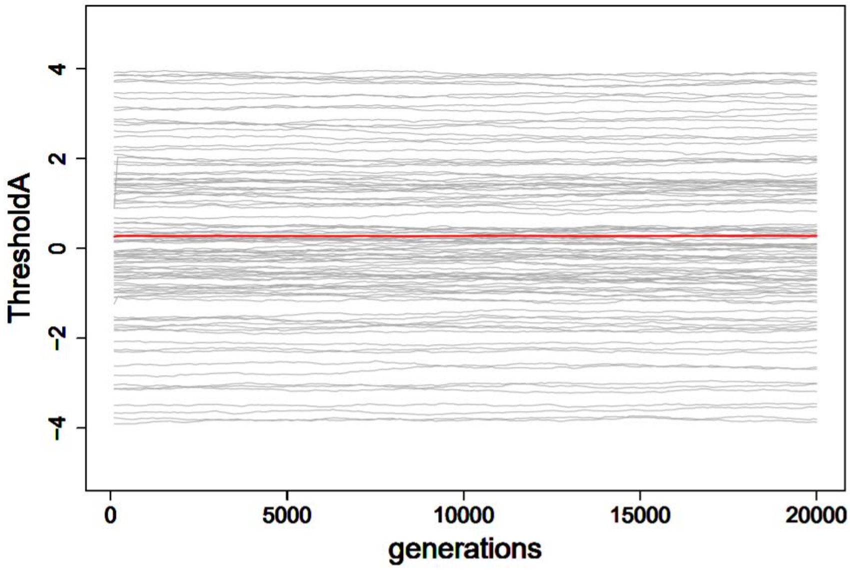
i) Fully modular



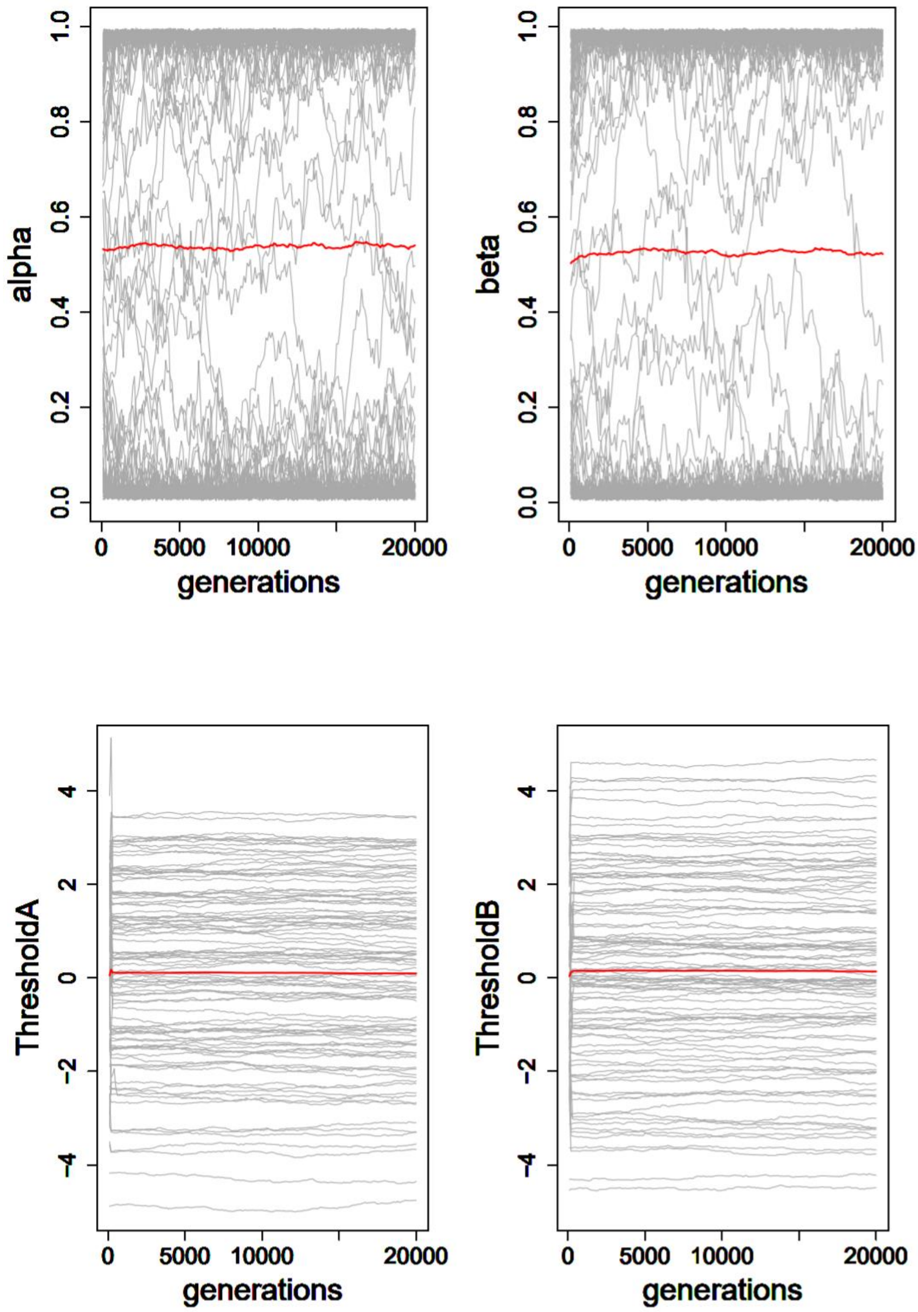


ii) Modular motivation

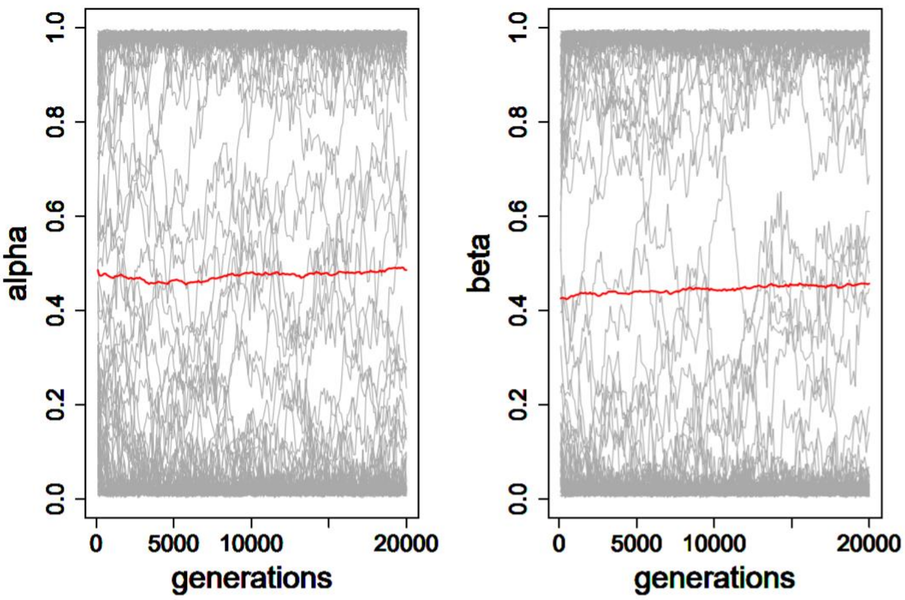


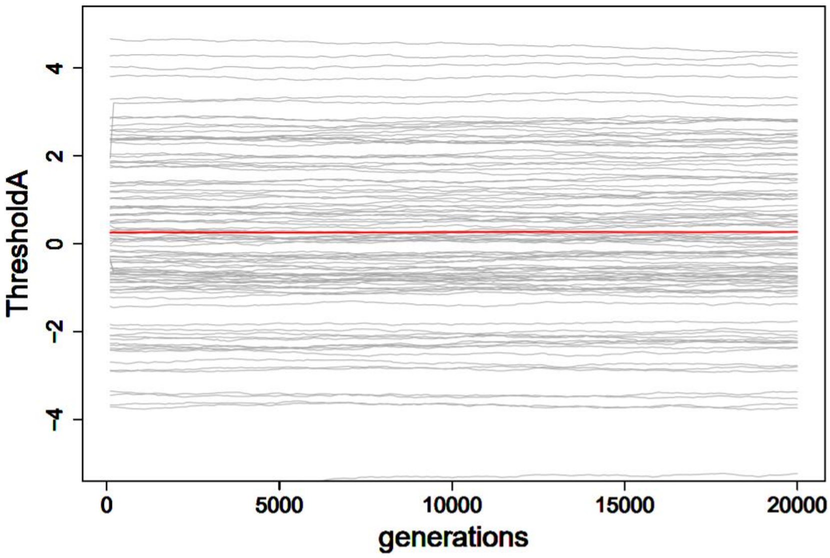


iii) Modular cognition



iv) Domain-general





Appendix 3: The scripts used to run the models and analysis.

As my agent-based models and analysis scripts were divided into one of the four agent types (fully modular, partly modular agents with modular cognition only, partly modular agents with modular motivation only and domain-general agents), I divide my OSF repositories accordingly. Each repository contains the agent-based model (consisting of a h file, an m file and a main file to compile in Objective C), random number generators for Objective C, the bash file to run the simulations for each combination of the parameters of interest. Plus, each individual repository has some R analysis scripts: one for generating the line graphs in appendix 2, one for generating the clustered bar charts seen in section 3.1 and one for generating the heatmaps and regression by psychological components in sections 3.2 and 3.3. These are:

- Full modular agents: https://osf.io/h63aw/?view_only=b4fbb2bf69584b098fafc02ad7edbd4f
- Partly modular agents with modular cognition only: https://osf.io/b4pyq/?view_only=9b9ba391891f4674a475ed80ff0ea9e0
- Partly modular agents with modular motivation only: https://osf.io/wn75b/?view_only=22471d7634b14b27912003ab9885203f
- Domain-general agents: https://osf.io/s8nk6/?view_only=7ace173ef889486c9c426f759f2e38b3

Finally, I also include an overall repository where I attach the script for the regression reported in section 3.4 which compared each agent type: https://osf.io/p7qmw/?view_only=78b60935e26c40ed92a515e6c800805d.

Appendix 4: The linear combinations performed between all the modular agent types.

The linear combinations compare the fitness of all the modular agent types, for the regression reported in Table 3, section 3.4 of the main text. These linear combinations are run for all parameter combinations of interest.

On runs where both the priors and fitness pressures tied to suboptimal coordination favour state 0 ($p_A = 0.1$, $p_B = 0.1$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then agents with modular cognition only had worse fitness than fully modular agents. If anything, this suggests that modular psychology is a hinderance when making decisions in two similar domains as I argued in the main text.

When the priors and fitness pressures clashed over two similar domain ($p_A = 0.9$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then agents with modular cognition (fully modular or partly modular only) consistently did better than agents with partly modular motivation. Thus, modular motivation may be a hinderance over two similar domains. This supports my interpretation in-text that modular motivation may be necessary to track cases where the most probable social norm is not the one that gives the highest payoff to coordinate on.

When the agents decide over two distinct domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$), then fully modular agents always accrue more fitness than the other agent types. Finally, when the priors and fitness tied to suboptimal coordination clash over two distinct domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then no one agent type does better than any other. Drift may have more of an influence in the strategies that the agents come to coordinate on.

Prior Probabilities	Linear Combination		
	Fully modular - modular motivation	**Fully modular - modular cognition**	**Modular motivation - modular cognition**
$p_A = 0.1$, $p_B = 0.1$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=2.70, p=0.10	F(1,39996)=9.43, p=0.002	F(1,39996)=1.97, p=0.16
$p_A = 0.9$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=6.57, p=0.01	F(1,39996)=46.47, p<0.001	F(1,39996)=106.94, p<0.001
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 4$, $f_{Bzero} = 0.25$	F(1,39996)=9.80, p=0.002	F(1,39996)=2610.70, p<0.001	F(1,39996)=1918.00, p<0.001
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 4$, $f_{Azero} = 0.25$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	F(1,39996)=0.33, p=0.56	F(1,39996)=2.12, p=0.15	F(1,39996)=0.75, p=0.39

Appendix 5: A regression predicting agent fitness based on whether they had modular or domain-general cognition, and modular versus domain-general motivation

The regression results displaying any differences between domain-general and modular cognition, versus domain-general and modular motivation, for each parameter combination. The domain-general agents were the omitted category of this regression as they were dummy coded as 0.

For environments where the agents decided over two similar domains and behaviour 0 was optimal to coordinate on ($p_A = 0.1$, $p_B = 0.1$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), then agents with modular cognition and motivation accrued less fitness than their domain-general counterparts. This supports my in-text argument that domain-general psychology is sufficient over two similar domains and, if anything, modular psychology is a hinderance.

When the agents made decisions over two similar domains where the priors favour state 1 but the fitness tied to suboptimal coordination favoured coordinating in state 0 ($p_A = 0.9$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$),, then agents with modular motivation only had less fitness than domain-general agents.

When the agents made decisions over two distinct domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 0.25$, $f_{Aone} = 4$, $f_{Bzero} = 4$, $f_{Bone} = 0.25$), then agents with modular cognition and modular motivation always accrued more fitness than their domain-general counterparts. This supports the notion that fully modular psychology is necessary when coordinating on social norms over two distinct domains.

Finally, in environments where the priors and fitness favoured two different states over two different domains ($p_A = 0.1$, $p_B = 0.9$, $f_{Azero} = 4$, $f_{Aone} = 0.25$, $f_{Bzero} = 0.25$, $f_{Bone} = 4$), there was no difference in fitness between agents with modular or domain-general psychology. This may suggest that drift has more of an outcome on agent behaviour in these runs.

Prior Probabilities	**Estimates for regression predicting fitness**		
	Intercept	**Cognitive threshold dummy (0 = domain-general, 1 = modular)**	**Motivation threshold dummy (0 = domain-general, 1 = modular)**
$p_A = 0.1$, $p_B = 0.1$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	8.72 *** (0.010)	-0.05 *** (0.011)	-0.07 *** (0.011)
$p_A = 0.9$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	8.44 *** (0.013)	0.018 (0.015)	-0.183 *** (0.015)
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 0.25$, $f_{Azero} = 4$, $f_{Bone} = 4$, $f_{Bzero} = 0.25$	6.06 *** (0.015)	1.11 *** (0.018)	2.02 *** (0.018)
$p_A = 0.1$, $p_B = 0.9$ $f_{Aone} = 4$, $f_{Azero} = 0.25$, $f_{Bone} = 0.25$, $f_{Bzero} = 4$	8.31 *** (0.013)	0.006 (0.015)	0.025 . (0.015)

The asterisks denote the significance of our p values with the following key:
** = 0.05*
*** = 0.01*
**** <0.001*
. = trend

<u>Chapter 7</u>
<u>I only want to help: Disentangling the influence of cognition and motivation on the emergence of cooperation</u>

Aysha Bellamy
Charles Efferson
In preparation for publication

Word count: 13,574 excluding references and appendices

Abstract

It is unclear how human societies are cooperative across multiple domains, as behaviour does not clearly align to ‘cooperation’ or ‘defection’. For example, it is sometimes cooperative to hunt (and share resources with the group) and sometimes cooperative to avoid hunting (to avoid depleting resources). I investigated how agents decided to cooperate (or not) when there was uncertainty over which behaviour was considered cooperative across two prisoner’s dilemmas. This mimicked the uncertainty regarding the cooperativeness of an agent’s behaviour in two distinct domains, such as hunting or building shelter. I allowed (i) agent cognition and motivation to coevolve and (ii) forced cognition and/or motivation to be flexible across both domains (domain-general) or allowed cognition and/or motivation to specialise to each domain (modular). I found that cooperation emerged as a ‘mistake’ in domain-general agents, who could not track the contrasting pressures across two distinct domains. Moreover, motivation was more important than cognition in driving (un)cooperative behaviour. These findings have implications for previous evolutionary accounts of cooperation which may have focused on cognitive biases at the expense of motivation. When deciding to cooperate, what we *want to do* was as important as what we *think we ought to do.*

Keywords: Cooperation, Prisoner’s Dilemma, modularity, domain-general, motivation, cognition

1: Introduction

The costly levels of human cooperation seen across most societies may pose a challenge to traditional evolutionary theories. Cooperation involves accepting a cost in order to provide a benefit to another (Chudek et al., 2013). It is difficult to see how cooperation may have emerged due to individual level selection, if individuals act to maximise their own fitness only (Boyd et al., 2011; Colman, 2006; Dawkins, 1976). Despite this, it has consistently been found that individuals will donate large amounts of money to anonymous strangers even in one-shot interactions (Fehr et al., 2002; Fehr & Fischbacher, 2005; Sally, 1995; Vogt et al., 2015). This behaviour is one-shot cooperation. Some evolutionary researchers have argued that one-shot cooperation may arise via self-interest.

Error management theory (EMT) posits that when making decisions under uncertainty, one should aim to make the least costly of two errors (Haselton et al., 2015). Whenever individuals met a new person throughout the ancestral past, there would have been some uncertainty over the longevity of their interactions. The individual may have made one of two mistakes here. She may have incorrectly assumed that she would meet this other person again, and so offered a one-off donation of her resources. If this donation was not reciprocated, then this is an error called one-shot cooperation. Alternatively, the individual may have incorrectly assumed that she would never meet the other individual again and so would free ride or refuse to cooperate with the other individual. Of course, defecting makes sense if this other individual was also a defector (Molleman et al., 2019; Zimmerman & Efferson, 2017). However, if the other individual was a conditionally cooperative partner, then they would have reciprocated any cooperation. The focal individual's decision to free ride would therefore anger this other individual (Molleman et al., 2019). This latter error would be erroneous defection (Krasnow & Delton, 2016) and could remove the benefit of a long lasting and mutually beneficial cooperative relationship. Erroneous defection is therefore more costly than one-shot cooperation. EMT predicts that our psychology should lean towards one-shot cooperation as the least costly error whenever we are uncertain about the longevity of future interactions (Delton et al., 2013; Krasnow & Delton, 2016).

Strong support for EMT would show that cooperation can emerge in conditions of uncertainty. Delton et al. (2011) have shown this using an agent-based model. In their model, agents played a prisoner's dilemma game. This tracks the agents' decision to cooperate or defect with a partner. Defection always gives the highest payoff at the individual level, though mutual cooperation gives a higher payoff than mutual defection and so the game represents a dilemma (Holt & Roth, 2004). Delton et al. assigned agents to play either a one-shot prisoner's dilemma, where mutual defection would be the most beneficial strategy or a repeated prisoner's dilemma. A repeated dilemma essentially becomes a coordination game, where it is better for the agent to choose the same option as her partner (Bear et al., 2017).

Crucially, the agents were unsure as to whether they played the one-shot or repeated dilemma. This meant that the agents were making decisions under uncertainty, as per the predictions of EMT (Haselton et al., 2015). This allowed Delton et al. to see exactly when one-shot cooperation would become a less costly error than erroneous defection, and so psychology would evolve towards making the former mistake. As few as five potential rounds of interaction were needed to bias agents towards one-shot cooperation on the one-shot prisoner's dilemma. When cooperative interactions are likely to last for five or more exchanges, then it becomes more costly to erroneously defect and miss out on this cooperative relationship than it does to offer one-shot cooperation on a one-shot prisoner's dilemma. Thus, our psychology evolves to be biased towards the least costly error and one-shot cooperation will be common. This model supports the predictions of an EMT approach to cooperation.

Delton et al's. model and EMT models focus on the uncertainty regarding the length of cooperative interactions. However, there are other realms of uncertainty when it comes to cooperation. Societies commonly differ in terms of which domain they cooperate in (Chudek & Henrich, 2011). For example, some societies will cooperate when they hunt but not when they build shelter, or vice versa. Moreover, the behaviours in a given domain that are considered cooperative in one society may be considered uncooperative in another. For example, sometimes it is cooperative to hunt (and share the resources with your group) and sometimes it is cooperative to avoid hunting (so as to not over deplete resources) (Safin et al., 2015). This variation may be difficult for group members to learn, and for evolutionary researchers to explain (Chudek & Henrich, 2011). My model thus has the novel aim of investigating how agents come to display certain behaviours when they are unsure as to which behaviours are cooperative or uncooperative in two distinct domains.

This model will also investigate the importance of cognition and motivation in upholding cooperative behaviour. The decision to cooperate in Delton et al's. (2011) model was divided into a cognitive component (i.e., do I believe that the interaction is one-shot or long lasting?) and a motivational component (i.e., given what I believe, do I want to cooperate?). Delton et al. then ran the model twice. In the first run, cognition was fixed at an arbitrary baseline and only the motivation to cooperate could evolve. In the second run, motivation was fixed so that the individual always wanted to defect when they thought that the prisoner's dilemma was one-shot, and always wanted to cooperate when they thought that the dilemma was repeated. Cognition was then left to evolve.

Delton et al. likely made this decision to avoid potential polymorphism. This means that, if cognition and motivation were left to coevolve, then they could potentially affect behaviour in a myriad of ways that are hard to disentangle (Laland, 1993). There is a trade-off between a model's simplicity and the extent to which it captures meaningful human behaviour, however (Heyes, 2016; Kendal et al., 2018). It is my view that fixing either cognition or motivation, and allowing the other to evolve, is an oversimplification. As ancestral humans shifted towards cooperative group living, then both their cognitive and motivational systems would have evolved in tandem. This current model allows both cognition and motivation to coevolve to capture this complexity.

As ancestral humans shifted towards group living, then cooperation would have made sense. Cooperative groups could bring down better game (Price, 2006), build better shelter (Chudek & Henrich, 2011) and could band together to out-compete others in times of warfare (Ihara, 2011). Some researchers argue that individuals engage in costly cooperation today due to the influence of these selection pressures from our ancestral past (Price, 2008).

In fact, Evolutionary Psychology argues that individuals have numerous evolved modules each designed to deal with a recurrent issue faced by humans throughout the ancestral past (Cosmides & Tooby, 1994). A module is a cognitive process which has been designed to work on a certain input in order to produce a specific behavioural output. Some modules have become cognitive biases, as they lead to behaviour which is suboptimal in the current environment. For example, a preference for high-fat and high-sugar foods would have been adaptive in the ancestral past when our ancestors lived on a calorific knife-edge but is maladaptive now that processed food is readily available and we lead increasingly sedentary lifestyles (Power & Schulkin, 2013). In a similar vein, a module which favours cooperation due to the conditions of the ancestral past may produce a maladaptive cognitive bias nowadays. This is because industrialised societies often have large scale or anonymous interactions that are unlikely to lead to long lasting, mutual relationships (Price, 2008).

Some researchers disagree with the claim that human cognition is modular, however. After all, our decision to cooperate must remain flexible to multiple scenarios. Some of these scenarios are novel and did not occur in the ancestral past. For example, the decision of whether or not to pay taxes or donate funds to start-up projects. Due to these novel scenarios, some researchers instead argue that our cognition to cooperate should be thought of as domain-general (Bolhuis et al., 2011). Domain-general cognition consists of a central processor that flexibly oversees decision-making across a variety of environmental inputs. Whilst there has been theoretical debate over whether cognition is modular or domain-general (Burke, 2014; Fodor, 2001; Spunt & Adolphs, 2017), this model will be the first to the author's knowledge to investigate how domain-general versus modular cognition and/or motivation may differentially affect the evolutionary trajectory of cooperative behaviour.

Previous research in Evolutionary Psychology have focused on cognitive modules (Cosmides & Tooby, 1994), though it is my intuition that motivational processes are likely to be just as important. Should our motivation to cooperate be the same when deciding whether to join in a collaborative group hunting effort to bring down a mammoth, as compared to the decision of joining in a communal building project? There are likely to be different costs and benefits associated to each cooperative domain that our cognition and motivation should weigh.

Agents with modular cognition may be able to reach specific conclusions about the likelihood of a behaviour being cooperative in different domains, whilst agents with domain-general cognition may reason generically about the likelihood of the behaviour that may be cooperative over multiple domains. An agent with modular motivation can show a different desire to adapt a certain behaviour in each domain, based on how likely it is that this behaviour is thought to be cooperative. Agents with domain-general motivation can only show a generic desire to choose a certain behaviour across multiple domains, and thus only show a generic drive towards being helpful or unhelpful. Whilst this model is not a full comparison of these architecture, it represents a promising first step to a full comparison. To achieve this, I consider a decision-making domain to represent ‘act’ or ‘not act’, as this level of abstraction will allow me to compare domain-general agents performance.

To summarise, I investigate the emergence of cooperation in a theoretically evolving population. This model has two novel aims:

1. To investigate how cognition and motivation coevolve to influence the emergence of cooperation when the agents are uncertain about which behaviour is ‘cooperative’ in each decision-making domain.
2. To investigate whether cognition and motivation are likely to be domain-general or modular when uncertain as to which behaviour is cooperative over two distinct domains.

2. Methods

2.1. Model description

A population of 100 agents are randomly assigned into dyads. Each dyad plays two prisoner’s dilemmas. In these games, the agent has to choose one of two options:

to cooperate or defect. Cooperation always produces a benefit, *b*, to one's partner at a cost, *c* to oneself. A prisoner's dilemma has a payoff matrix where $b>c>0$. That is, the benefit given to the other individual must outweigh the initial cost of cooperation. In the current dilemmas, I standardise c=1. As the costs of cooperation are standardised, then the appeal of cooperation in these prisoner's dilemmas are driven entirely by the size of the *b* parameter (see Section 2.2 for further details). See Table 1 for the parameters of this model.

Table 1. The notations and variables used throughout the model

Symbol	**Model description**
$b \in \{2,4\}$	Benefits of cooperation.
$c = 1$	Cost of cooperation.
$S_A,S_B \in\{0,1\}$	The environmental state in domain A or B respectively.
$B_A,B_B \in\{0,1\}$	The behaviour employed by the agent in domains A and B respectively.
$x_A,x_B \in (-\infty, +\infty)$	The cue summary drawn from the environment to help the agent decide which state the environment is likely to be in for both domains A and B respectively.
$s_A,s_B \in\{0,1\}$	The state that the agent beliefs the environment is in for both domains A and B respectively. These can be wrong.
$T_A,T_B \in (-\infty, +\infty)$	The cognitive threshold of evidence needed to believe that the state is 1 in domains A and B (note just *T* for domain-general agents).
$\alpha_A,\alpha_B \in[0,1]$	The agents' motivation to play behaviour 1 (i.e., hunt or build more) if they believe that the state is 1 for domains A and B. (Note just α for domain-general agents).
$\beta_A,\beta_B \in[0,1]$	The agents' motivation to play behaviour 1 (i.e., hunt or build more) if they believe that the state is 0 for domains A and B. (Note just β for domain-general agents).
$p_A,p_B \in \{0.1,0.5,0.9\}$	The probability that the state will be 1 in both domains A and B respectively.

The agents play two one-shot prisoners dilemmas, designed to mirror the uncertainty that agents have over which behaviour is cooperative in two distinct domains (Chudek & Henrich, 2011). I call these dilemmas A and B. These represent two distinct domains that a group may cooperate in. For example, imagine that dilemma A represents hunting and dilemma B represents building shelter. Let the state of the environment in dilemma A be represented by a random variable, S_A, and the state of the environment in dilemma B be represented by a random variable, S_B. Both had support $\in\{0,1\}$.. Note that which behaviour is cooperative depends on the value of the environmental state. When the state is 0, then playing behaviour 0 would be cooperating and playing behaviour 1 would be defecting. When the state is 1, then playing behaviour 1 would be cooperating and playing behaviour 0 would be defecting.

To understand how this decision maps onto cooperation, I will walk through an in-depth example of the decision to hunt in dilemma A. The agent has to choose between two behaviours: to act (1) or not to act (0). These can be thought of as the decision to spend more time hunting or the decision to spend less time hunting respectively. The payoff of the agent's decision – and her partner's decision – can be translated to the payoff matrices of a prisoner's dilemma (see Figure 1). In this model, there is no uncertainty over the longevity of interactions (all interactions are one-shot prisoner's dilemmas). Instead, the agent is uncertain about which environmental state she is in and thus is uncertain over which behaviour represents cooperation or defection. To illustrate in reference to hunting, it is sometimes cooperative to hunt more and share resources with one's group (Hill, 2002). Agents who hunt more ($B_A = 1$) are 'hunting more as cooperating', provided that $S_A = 1$. If the agent refuses to hunt but still takes some food ($B_A = 0$), then she is 'hunting less as defecting'.

Cases where $S_A = 0$ instead represent conditions where to go hunting would be to defect. For example, if there is a scarcity of game and to overhunt would be to take resources from one's neighbours (Lu & Wirth, 2011). If the agent decides to hunt less

($B_A = 0$) then she is ‘hunting less as cooperation’. She avoids the temptation to free ride and extract resources via selfish means. If the agent instead decides to hunt more ($B_A = 1$), then she is ‘hunting more as defection’.

Thus, $B_A = 0$ always represents the decision to not act (i.e., hunt less) and $B_A = 1$ always represents the decision to act (i.e., hunt more). Which of these behaviours is considered cooperative depends on the environmental state ($S_A = 0$ or $S_A = 1$). As this is the first comparison of modular and domain-general decision-making in a population of theoretically-evolving agents, we take the simplest representation of a domain-general system as one that must decide to act or not act over multiple domains (Pietraszewski & Wertz, 2021). This is the first step towards a full and comprehensive comparison of modular or domain-general architecture. To illustrate how these behaviours affect fitness in line with the payoff matrix of a traditional prisoner’s dilemma game, see Figure 1.

State = 0: Hunting more as defection

		Partner's behaviour	
		0 (hunting less as cooperation)	**1 (hunting more as defection)**
Focal agents' behaviour	**0 (hunting less as cooperation)**	1 + (4 - 1)	1 -1
	1 (hunting more as defection)	1 + 4	1

State = 1: Hunting more as cooperation

		Partner's behaviour	
		0 (hunting less as defection)	**1 (hunting more as cooperation)**
Focal agents' behaviour	**0 (hunting less as defection)**	1	1 + 4
	1 (hunting more as cooperation)	1 - 1	1 + (4 - 1)

Figure 1. The payoff matrix for the prisoner's dilemma in domain A (the hunting domain). Here, I focus on cases where c = 1 and b = 4 (strong selection pressure to cooperate), though there are cases where b = 2 (weak selection pressure to cooperate).

Note that the costs of cooperation are standardised at $c=1$. Thus, the payoffs to cooperation is driven entirely by the size of the b parameter in this model. This benefit to cost ratio is designed to never be negative, though there may be cases where to fail to uphold the correct local (un)cooperative norms could led to punishment or ostracism (Gintis et al., 2003; Henrich et al., 2006; van den Berg et al., 2012). In my model, fitness payoffs were always positive as this allowed for reproduction to be tied to cumulative fitness.

A similar logic extends to dilemma B (building shelter). Cases where $S_B = 1$ and $B_B = 1$ imply that the agent joins a communal building project (i.e., building more as cooperation) whilst cases where $S_B = 1$ and $B_B = 0$ imply that the agent does not join the communal build but tries to use the communal space afterwards (i.e., building less as defection). Cases where $S_B = 0$ and $B_B = 0$ implies that the agent avoids the temptation to build when it would only benefit herself (i.e., building less as cooperation), whilst cases where $S_B = 0$ and $B_B = 1$ implies a selfish decision to build (i.e., building more as defection). For example, the agent may extend her private property onto communal land.

In order for the agent to choose behaviour 0 or 1, then the agent must first formulate a belief about the likely state of the environment in each domain. The agent's belieff is represented by s_A and s_B respectively, and these both take the support $\in\{0,1\}$. They either believe that hunting or building more is cooperative ($s_A=1$ and $s_B=1$) or they believe that hunting or building more is defection ($s_A=0$ and $s_B=0$). To formulate this belief, all agents draw a cue summary that probabilistically yet imperfectly captures the state of the environment. This cue summary can be thought of as a combination of all the factors that one weighs up when deciding whether a behaviour is helpful for the group. For example, some cues that could indicate a cooperative hunting endeavour include if the hunting party is large, consists of volunteers and is being undertaken before a communal feast. There are other factors about the situation (e.g., our finances) and the other person (e.g., whether we met them far away from our home, if they are anonymous) that must be weighed up before one decides whether to cooperate (Delton et al., 2011). For simplicity, imagine that all of these cues are collapsed into one cue – the cue summary – during this model.

Let the agent's private signal of cue summary in dilemma A be represented by x_A and let the agent's private signal in dilemma B represented by x_B. These can be any random value from the respective distribution curves. The agent draws a cue summary from a distribution of cues separately for both dilemmas A and B. These cue summaries depend on the actual state of the environment. When $S_A = 1$ or $S_B = 1$, then the agent draws a cue summary from a normal distribution with a Mμ =1, σ =1. When $S_A = 0$ or $S_B = 0$, then the agent draws a cue summary from a normal distribution with a Mμ = -1, σ =1. This signal probabilistically but imperfectly represents the environment. The difficulty of distinguishing between these environments was driven by the overlap of the two distribution curves (~31.7%). This overlap means that the agent can be wrong about the state of the environment and so they make decisions under uncertainty.

Once the agents have drawn a cue summary from the environment (x_A or x_B), they can translate this into their beliefs about the environmental state in each decision-making domain (s_A or s_B). They do this using their cognitive threshold, or minimum amount of evidence that they need to conclude that the state is 1 in each domain. This threshold could be thought of as the absolute minimum amount of evidence that an agent needed to believe that the state was 1 in each dilemma. These thresholds could take any positive or negative value. Let the cognitive threshold in dilemma A be represented by T_A and the cognitive threshold in dilemma B be represented by T_B. The cue summaries were modelled so that 0 would represent unbiased cognitive thresholds. An agent that evolved positive cognitive thresholds thus needed more evidence from the environment before she would believe that the state was 1 (i.e., they needed more evidence to believe that action was cooperative). Conversely, an agent who evolved negative cognitive thresholds was less discerning and needed less evidence from the environment to believe that the state was 1 (i.e., they needed less evidence to believe that action was cooperative).

If the agents' cue summary exceeds their cognitive threshold for dilemma A ($x_A > T_A$), then they believe that the state is 1 ($s_A = 1$). They believe that hunting more is cooperative. Likewise, when the agents' cue summary exceeds their cognitive threshold in dilemma B ($x_B > T_B$), then they believe that the state is 1 ($s_B=1$). They believe that building more is cooperative.

In cases where the agents' cue summary is less than or equal to their cognitive thresholds ($x_A <= T_A$ or $x_B <= T_B$), then the agents believe that the environmental state is 0 ($s_A=0$ or $s_B=0$). The agent believes that hunting or building less is cooperation. Agents with modular cognition can formulate different beliefs as to the likelihood of hunting or building more as cooperation across the two decision-making domains. This is because T_A and T_B values can evolve separately. Agents with domain-general cognition can only reason generically about the likelihood of hunting and building more as cooperation across both dilemmas. This is because their cognitive thresholds are constrained so that $T = T_A = T_B$.

Once the agents have formulated their beliefs about the likelihood of the state being 1 then they can only hunt or build if they are motivated to do so. Each agent drew a random number from the uninform interval [0,1] and compared this with an internal motivational threshold. Whenever the agent believed that the environmental state was 1 ($s_A = 1$ or $s_B = 1$), then she would hunt or build more with a motivation of $Probability\ More_{State\ 1}$ (which we label α for future reference). Whenever this random number exceeded the α threshold, then the agent would hunt or build more. If this random number did not exceed the α threshold, then the agent would hunt or build less. The probability of hunting or building less when the agent believed that the environmental state was 1 was therefore given by $1- \alpha$.

Imagine a case where the agents' cue summary exceeds their cognitive threshold ($x_A > T_A$ or $x_B > T_B$). In this case, the agents believe that the environmental state is 1 ($s_A=1$or $s_B=1$). When the agents believe that hunting or building more is cooperation, then the agents will hunt or build more with a motivation of α_A or α_B respectively. Alternatively, the agents may hunt or build less with a motivation of $1- \alpha_A$ or $1- \alpha_B$ respectively. Thus, α_A and α_B values drive the agents' motivation to hunt or build more when they believe that hunting or building more would be cooperative. Thus, α_A and α_B values represent the agent's motivation to act when she believes that this action is coopertive. Note that the agents' beliefs about the environmental state are not always accurate and so there is no guarantee that hunting or building more would in fact be cooperative.

Whenever the agent believed that the environmental state is 0 ($s_A = 0$ or $s_B = 0$), then they would hunt or build more with a motivation of *Probability More*$_{State\ 0}$ (which we label β for future reference). Whenever this random number exceeded her β threshold, then the agent would hunt or build more. If this random number did not exceed her β threshold, then the agent would hunt or build less. The probability of hunting or building less when the agent believed that the environmental state took the value 0 was therefore given by $1- \beta$.

Now imagine a case where the agents' cue summary is less than or equal to their cognitive threshold ($x_A <= T_A$ or $x_B <= T_B$). In this case, the agents believe that the environmental state is 0 ($s_A=0$ or $s_B=0$). They believe that hunting or building more is defection. Agents will then hunt or build more with a motivation of β_A and β_B respectively. Alternatively, the agents may hunt or build less with a motivation of $1-\beta_A$ and $1-\beta_B$ respectively. Thus, β_A and β_B values drives the agents' motivation to hunt or build more when they believe that hunting or building more would be defection. Thus, β_A and β_B values represent the agent's motivation to act when she believes that this action is defection. Note that the agents' beliefs about the environmental state are not always accurate and so there is no guarantee that hunting or building more would in fact be defection.

The agents with modular motivation can change their preferences to hunt and build more or less across the two dilemmas. This is because agents with modular motivation can have α_A, α_B, β_A and β_B values evolve separately in each domain. Domain-general motivation is constrained to be the same across both dilemmas A and B ($\alpha = \alpha_A = \alpha_B$ and $\beta = \beta_A = \beta_B$). This means that they can only show a generic drive towards hunting and building more, or hunting and building less, across both domains.

Note that the implementation of the α and β values in this study do not align with Delton et al's. (2011) (though the idea behind them is similar). In Delton et al's. paper, β always represented the desire to cooperate once the agent's private signal had exceeded her cognitive threshold and she thus believed that the state was 1 (i.e., she believed that the interaction would repeat). α always represented the desire to cooperate when the agent's private signal had not exceeded her cognitive threshold and thus, she did not believe that the state was 1 and instead believed that the state was 0 (i.e., she did not believe that the interaction would repeat and thus believed that the interaction was one-shot).

Instead, my implementation ties α to represent the agent's decisions to play behaviour 1 whenever her private signal exceeds her cognitive threshold, and she thus believes that the state is 1 (i.e., this is the agent's desire to hunt or build more when she believes that hunting or building more is cooperation). β is instead tied to the agent's decisions to play behaviour 1 whenever her private signal does not exceed her cognitive threshold and she thus believes that the state is 0 (i.e., this is the agent's desire to hunt or build more when she believes that hunting or building more is defection). Thus, I tie the motivational signal to a *choice* (i.e., to hunt or build more) rather than to a *cooperative strategy.* This aligned with my aim of investigating how agents choose behaviour when uncertain as to which behaviour would be cooperative or uncooperative.

This model allows the agents' cognitive and motivational thresholds to coevolve. This addresses my first research aim regarding the complexity of how cooperative behaviour emerged throughout the ancestral past. The agents' modularity for cognitive thresholds, and their modularity for motivational thresholds, are modelled separately. This gives four possible combinations:

1. Fully modular agents can reason about the likelihood of hunting more being cooperative, and building more being cooperative, distinctly in each decision-making domain (T_A, T_B) and can show a different desire to hunt or build more in each decision-making domain (α_A, α_B, β_A, β_B).
2. Agents with modular cognition only. These agents can reason about the likelihood of hunting more being cooperative, and building more being cooperative, distinctly in each decision-making domain (T_A, T_B) but can only show a generic desire towards playing behaviour 1 versus behaviour 0 across both domains (α, β) (i.e., they can only show a generic desire towards hunting and building more across both domains).
3. Agents with modular motivation only. These agents can only reason generically about the likelihood for hunting or building more as cooperation across both domains (T), though they can show a different desire to hunt or build more in each decision-making domain (α_A, α_B, β_A, β_B).
4. Domain-general agents can only reason generically about the likelihood for hunting or building more as cooperation across both domains (T) and can only

show a generic desire to hunt and build more, or to hunt and build less, across both domains *(α, β).*

Whatever the psychological architecture implemented from the four possibilities, cognitive thresholds and motivational thresholds are left to endogenously evolve throughout the current model. The agent receives fitness based on her behaviour (B_A and B_B) her partner's behaviour (B_A and B_B) and the environmental state (S_A and S_B) in the two dilemmas. Each agent receives an exogenous fitness value of 1 point on top of the fitness that she accrues via her interactions with her partner. This exogenous fitness captures the agents' payoffs from behaviour in domains besides cooperation. These values were then summed to give a total fitness value per agent. Note that the agents' chance of having offspring was directly proportional to their total fitness. For every one of the 100 agents, we translated their total fitness value into a fitness value which was a cumulative proportion of the entire generation. Agents who had more fitness thus had a higher cumulative proportion value. To assign offspring genotype, we simply drew a random interval between [0,1] which corresponded to this cumulative proportion space. Thus, parental agents with a larger proportional fitness would have a higher chance of having offspring, as the offspring agents were more likely to be sampled from these larger proportional fitness values.

Once the offspring's parental agent was identified, the offspring then inherited her parental agent's psychological variables. Note that we allowed cognition and motivation to be inherited independently from different parental agents. This did not mirror sexual reproduction, but was instead intended to reflect how continuous, complex traits such as psychological phenomena are likely to be coded at multiple allele sites.

Agents inherited the cognitive thresholds with a mutation rate of 0.5. As this was a continuous variable with unspecified end-points, we assumed mutations to be common, but constrained mutation to only occur in small chunks in each time step ($M\mu = 0$, $\sigma = 0.00125$). Agents inherited their motivational thresholds with a lower mutation rate of 0.05. Mutations occurred in steps of plus or minus 0.01 at each timestep. Note that mutation was constrained so that the motivational thresholds could only take a value between 0 and 1. The agents could never be less than 0%, or more than 100%, motivated to hunt or build. Mutation was modelled independently for the two psychological components. The offspring and parental generation did not co-exist. Offspring agents replaced parental agents entirely once they had inherited their traits, and reproduction was proportional to fitness and affected by a small mutation rate, as is typical of a Wright Fisher model (Suchow et al., 2017).

2.2. Parameters

See SI text for the full parameter space of my model. There are two parameters of interest that I vary. First, I vary the probability with which the environmental state could be 1 in domain A (denoted by p_A) and the probability with which the environmental state could be 1 in domain B (denoted by p_B). These are modelled separately and could take realisations $p_A \in \{0.1, 0.5, 0.9\}$ and $p_B \in \{0.1, 0.5, 0.9\}$ respectively. The environmental state in one decision-making domain does not depend on the environmental state in the other domain. For the sake of clarity, this analysis will focus on three parameter combinations of interest:

1. Environments with skewed but equivalent distributions over both domains, favouring state 0 ($p_A = 0.1$, $p_B = 0.1$). These parameters mean that hunting or building more was likely to be defection.

2. Environments with skewed but equivalent distributions over both domains, favouring state 1 ($p_A = 0.9$, $p_B = 0.9$). These parameters mean that hunting or building more was likely to be cooperation.
3. Environments with skewed but disparate distributions of environmental states ($p_A = 0.1$, $p_B = 0.9$). This creates an environment where hunting more is likely to be defection though building more is likely to be cooperation. These environments are realistic. Some groups cooperate to build shelter, but not to hunt, and vice versa (Chudek & Henrich, 2011).

Modularity may be more important for environments with inconsistent prior probabilities of state 1 ($p_A = 0.1$, $p_B = 0.9$). This is because a modular agent can respond to disparate decision-making domains by formulating different beliefs and/or desires across both domains. A domain-general agent may be at a disadvantage when tailoring their beliefs and motivation across two different domains with inconsistent priors.

Second, I vary the fitness associated with mutual cooperation by varying the b parameter. I include a 'weak' environment ($b=2$). This creates an environment where mutual cooperation does not give a substantially higher payoff than mutual defection and so cooperation may be less likely to emerge ($b = 2 > c=1 > 0$). I also consider a 'strong' environment ($b=4$). This creates an environment where mutual cooperation gives a higher payoff than mutual defection and so cooperation may be more likely to emerge ($b = 4 > c=1 > 0$). I run a population of 100 agents for 20,000 generations and repeat this for 100 simulations for each possible combination of parameter space.

To summarise the model, the agents must decide to 'act' or 'not to act' in two prisoner's dilemmas, when they are uncertain as to which behviour is cooperative over both domains. To make this decision, each agent must first decide which environmental state the domain is in (would acting more be coopertion [$s_A = 1, s_B = 1$] or would acting less be cooperation? [$s_A = 0, s_B = 0$]). To decide this, each agent has a private signal (x_A, x_B) which can be thought of as the cues from the environment. These cues must exceed the cognitive threshold ($x_A > T_A, x_B > T_B$) for the agent to believe that the state is 1and that it is cooperative to act more ($s_A = 1, s_B = 1$). This signal can thought of as all the cues that the agent needs to decide that an action is coopertve (e.g., cues that suggest hunting is cooperative include if the hunting party is large and takes place before a communal feast). If the private signal does not exceed the threshold ($x_A <= T_A$, $x_B <= T_B$), then the agent instead believes that the state is 0 and that it is cooperative to act less ($s_A = 0, s_B = 0$). These beleifs can be wrong. After formulating these beliefs, the agent must be motivated in order to act. The agent is motivated to act with a probability of α when they believe that the state is 1 (and1-α for not acting), and are motivated to act with a probability of β when they believe that the state is 0 (and 1-β for not acting).

Once the agent 'acts' ($B_A = 1, B_B = 1$) or 'does not act' ($B_A = 0, B_B = 0$) in each domain, they then receive fitness based on whether this behviour actually matches the environmental state ($S_A = 0, S_A = 1$, and $S_B = 0, S_B = 1$) across both domains, and whether their behaviour matches their partner's behaviour over both dilemmas ($B_A = 0$, $B_A = 1$, and $B_B = 0, B_B = 1$). This will be affected by the probability that the state is 1 (p_A and p_B) in each domain. The fitness assigned to the agent also depends on the size of the cooperative benefits to cost ratio, where $c=1$ throughout and $b\epsilon(2,4)$. All agents receive an exogenous fitness value of 1 on top of this fitness. Fitness is then changed to a cumulative proportion value, so that more fit agents are more likely to have more offspring. Reproduction is therefore proportional to fitness. The offspring inherit cognitive (T_A, T_B) and motivational thresholds ($\alpha_A, \alpha_B, \beta_A, \beta_B$) from separate agents with a small rate of mutation. The offspring generation then overwrite the parental agent at each time step, as characteristic of a Wright-Fisher model (Suchow et al., 2017).

Our ultimate aim is to compare how modular and domain-general agents come to decide to act or not act in each coopertive domain. To achieve this, fully modular agents were coded to have a separate cognitive and motivtional threshold per domain (T_A, $T_{B,}$ α_A, $\alpha_{B,}$ β_A, β_B). Partly modular agents either had modular cognition but domain-general motivation (T_A, $T_{B,}$ $\alpha,$ β) or they had domain-general cognition and modular motivation ($T_,$ α_A, $\alpha_{B,}$ β_A, β_B). Finally, the fully domain-general agents have constrained cognition so that $T= T_A = T_B$, and constrained motivation so that $\alpha = \alpha_A = \alpha_B$ and $\beta= \beta_A = \beta_{B.}$ This matches the definition of domin-general systems as those that must operate to decide to act or not act over multiple decision-making domains simultaneously.

3. Results

The runs converged on stable psychological architecture over all simulations of 20,000 generations (see appendix 1). Here, I discuss the stable psychological architecture across the runs of interest: (i) runs with skewed but consistent prior probabilities such as $p_A = 0.1, p_B = 0.1$ and (ii) $p_A = 0.9, p_B = 0.9$ and (iii) for runs with inconsistent prior probabilities of state 1 in both domains ($p_A = 0.1, p_B = 0.9$). For each of these runs, I explore: the outcomes present in the final generation (section 3.1), the psychological architecture of the final generation of agents (section 3.2) plus how each element of the agents' psychological architecture contributed to the fitness of the agents when playing the prisoner's dilemma games (section 3.3). I then compare the fitness across all four agent types (section 3.4), to see which psychology is likely to underline our decision to cooperate (or not) in two distinct domains.

3.1: Did the agents cooperate?

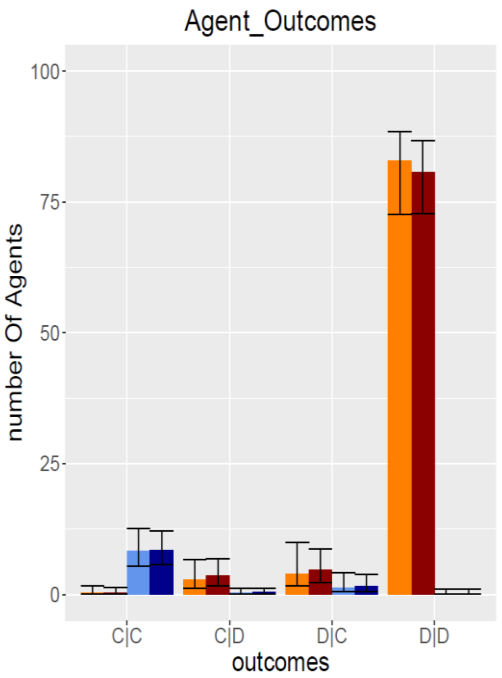
Mutual defection was the dominant outcome across the runs. Mutual defection was expected in environments with weak selection pressures to cooperate *($b=2$)*, as cooperation did not provide a substantially greater benefit than mutual defection. Accordingly, this analysis focuses on environments with a strong pressure to cooperate, where $b=4$ (though see appendix 2 for cases where $b=2$). Despite the large benefits to cooperation, mutual defection was still the dominant outcome (see Figure 2). This may be because I used one-shot Prisoner's Dilemmas. Mutual cooperation may only become a stable strategy when there is an opportunity for the dilemmas to repeat (Delton et al., 2011). Thus, mutual defection may be expected as an outcome in agents who can track the contrasting pressures of which behaviour is likely to be defection in each of the two domains.

The agents clearly adopted a mutual defection outcome on runs where there was a consistent skew in the prior probabilities of the state being 1 (Figure 2i: $p_A = 0.1, p_B = 0.1$; Figure 2ii: $p_A = 0.9, p_B = 0.9$). This makes sense as, when the prior probabilities are highly but consistently skewed in both domains, then the uncertainty over which behaviour is cooperation or defection is reduced. This is because the agent's cue summary is being drawn on top of priors that are already quite conclusive. The agents typically chose not to hunt or build more as cooperation in state 1, and typically chose to hunt or build more as defection in state 0.

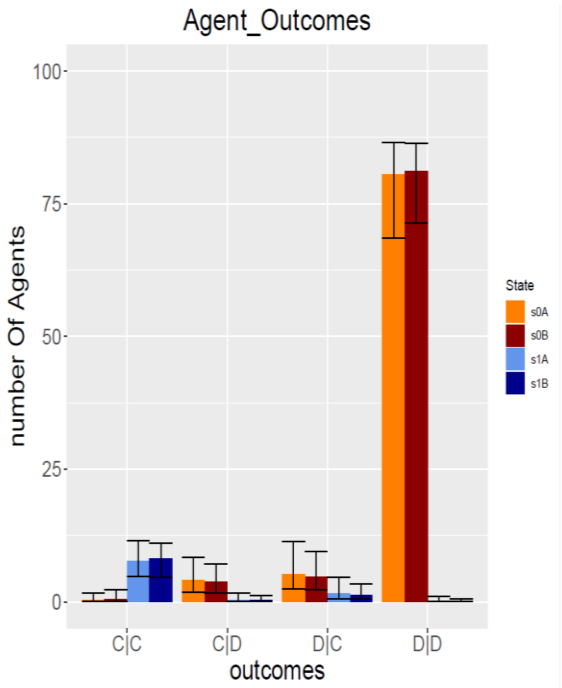
In comparison, on runs where there was an inconsistency in the prior probabilities in the two domains (Figure 2iii; $p_A = 0.1, p_B = 0.9$), then the fully modular and partly modular agents defected in state 0 and state 1. The agent hunted more as defection and built less as defection in the two dilemmas. This may be because the modular agents have cognition and/or motivation systems that were capable of discriminating between the priors of two decision-making domains.

The results for the domain-general agents are interesting, however. These agents experienced some mismatches (i.e., some freeride, and some receive the sucker payoff) and some agents engaged in mutual cooperation across the prisoner's dilemmas with inconsistent priors. As the domain-general agents have psychological architecture which is not equipped to discriminate between the two decision-making domains, then there was more variation in their choice to cooperate or defect at the individual level. That is, domain-general agents may start to cooperate as they could not track the different environmental states across multiple domains.

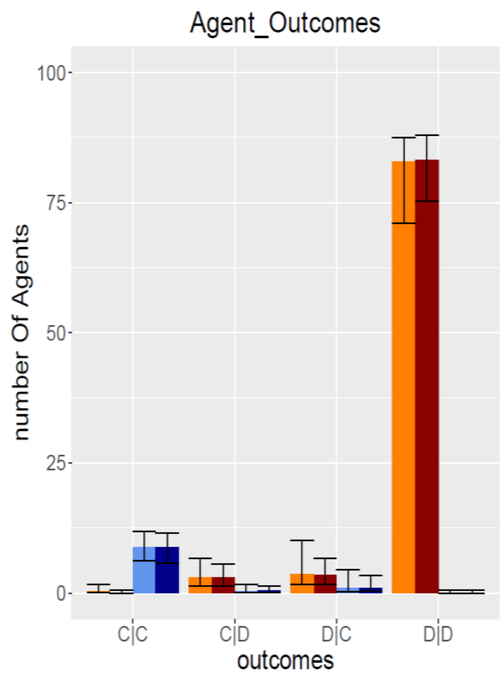
Figure 2(i): $p_A = 0.1$, $p_B = 0.1$.



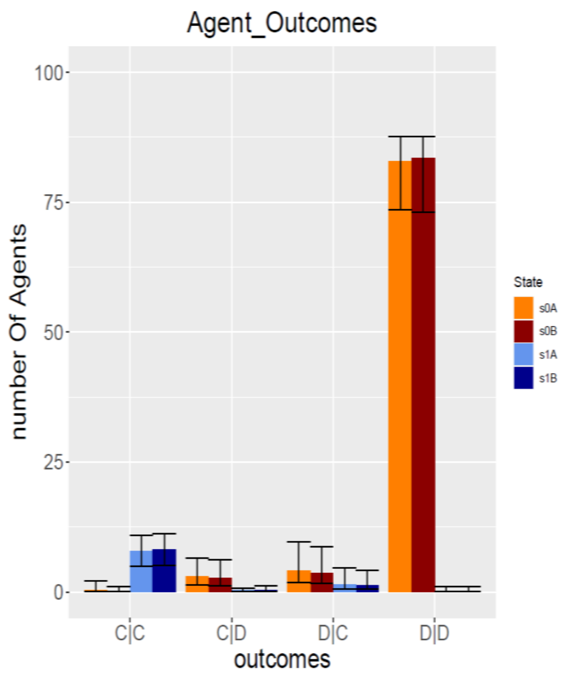
Fully modular



Modular motivation

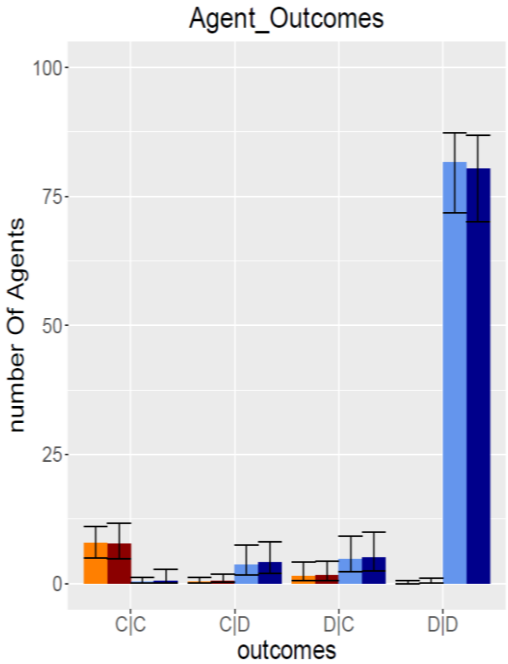


Modular cognition

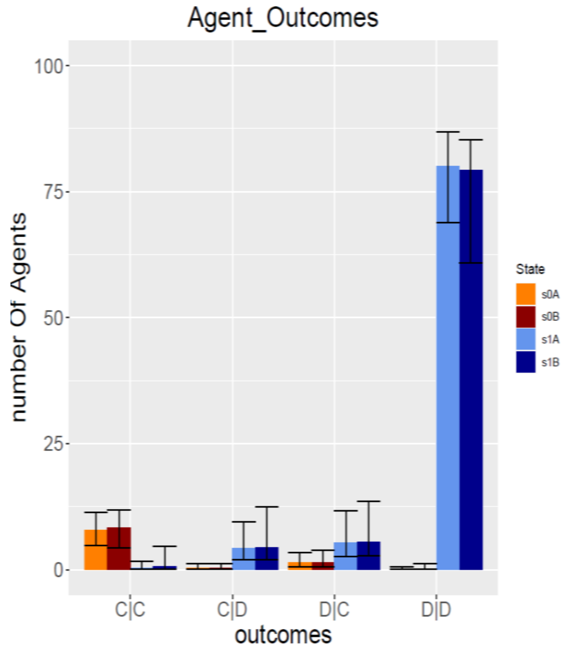


Domain-general

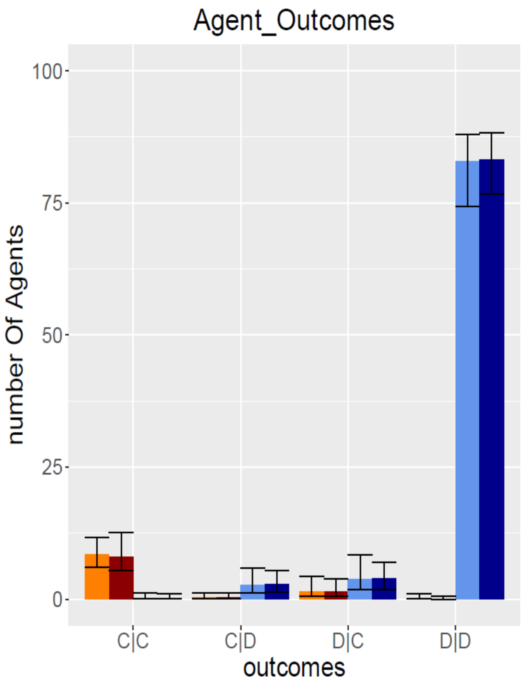
Figure 2(ii): $p_A = 0.9$, $p_B = 0.9$.



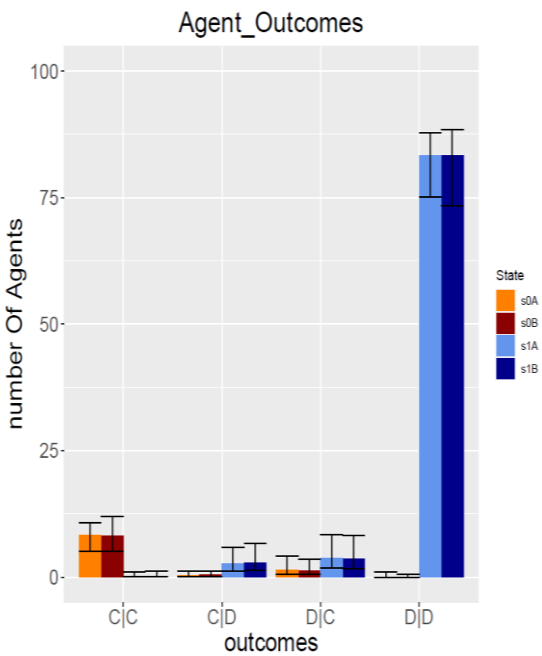
Fully modular



Modular motivation

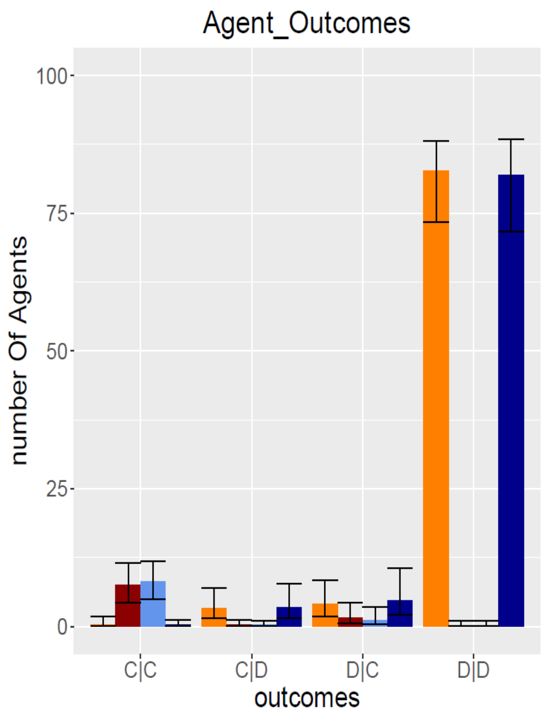


Modular cognition

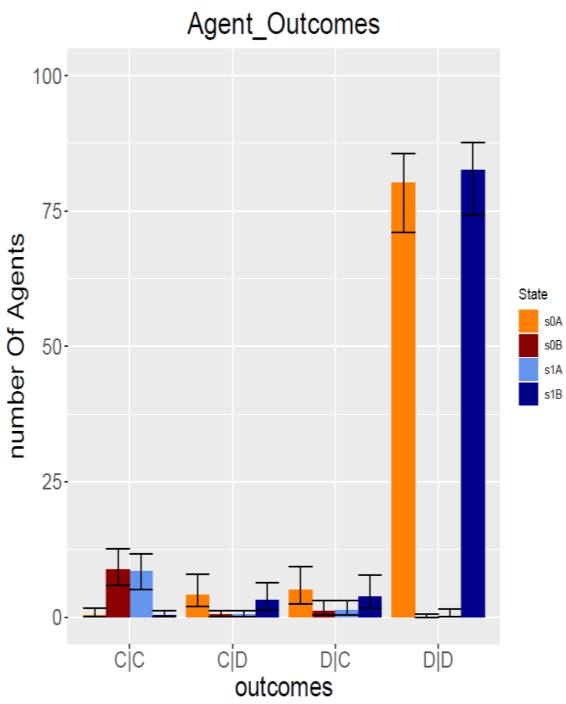


Domain-general

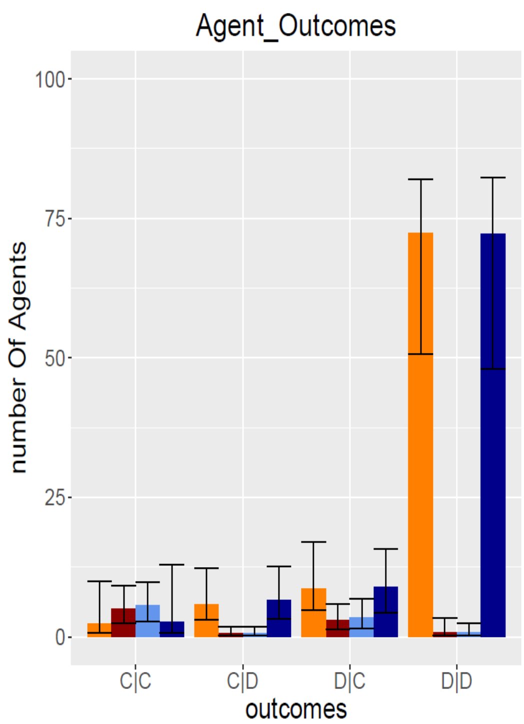
Figure 2(iii) $p_A = 0.1$, $p_B = 0.9$.



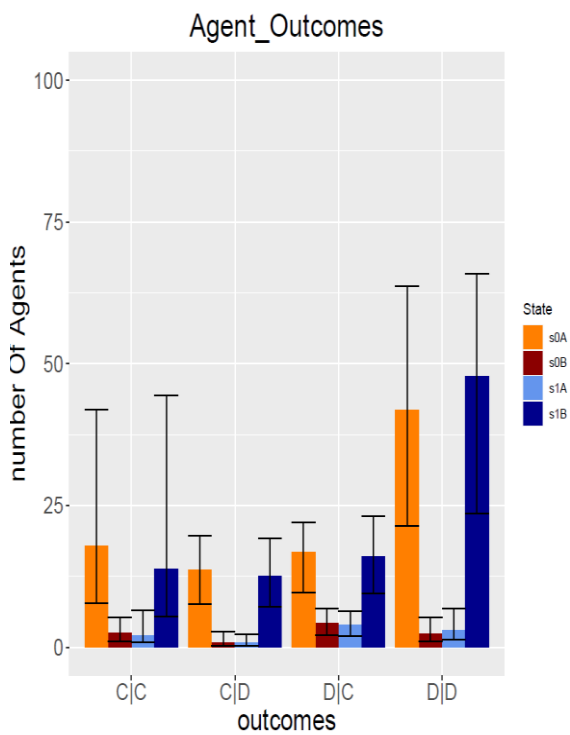
Fully modular



Modular motivation



Modular cognition



Domain-general

Figure 2. Bar charts showing the distribution of cooperation and defection. This is the outcomes for each of the four agent types on runs with (2i) skewed but consistent priors of state 1, where $p_A = 0.1$, $p_B = 0.1$ and (2ii) when $p_A = 0.9$, $p_B=0.9$ and (2iii) for runs with inconsistent priors of state 1 in both domains ($p_A = 0.1$, $p_B=0.9$). Note the x axis represents the four possible behaviours from the agent when playing a prisoner's dilemma: C|C is cooperation conditional on partner cooperation , C|D is cooperation conditional on partner defection, D|C is defection conditional on partner cooperation and D|D is defection conditional on partner defection. The y axis gives the number of agents who adopt each outcome. These bar charts are for environments with strong cooperative benefits (b=4) (though see appendix 2 for the analysis of runs with a weak pressure to cooperate: [b=2]). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

3.2: The psychological architecture of the four agent types

We create a series of binned heatmaps to visualise how the agents' cognition and motivation coevolved to drive behaviour (see Figures 5-8). I used the code in appendix 3 to calculate the smallest and largest cognitive threshold amongst the final generation of agents, and I divided this difference into nine bins. These are the nine sections of each heatmap reported in Figures 5-8. The motivational thresholds could take any value from [0,1], which I divided into ten equally spaced bins {[0, 0.1], (0.1, 0.2], (0.2, 0.3]…(0.9, 1]}. This gave 100 bins defined jointly over the α and β values. For each agent, I calculated their cognitive threshold to decide which section of the heatmap to plot to and I then calculated their α and β motivational threshold bins. These corresponded to the exact coordinates of each graph to plot to, with α on the x axis and β on the y axis. The heatmap colours tracked the density of agents with the same cognitive and motivational thresholds. White spaces denote that zero agents have this exact cognitive and motivational threshold combination, while deeper orange reflects denser patches. Finally, I added a black square to each graph to denote the average agents' cognitive and motivational threshold values.

Figure 3 shows the psychological architecture of the four agent types in response to runs with consistent prior probabilities where state 1 was unlikely ($p_A = 0.1$, $p_B = 0.1$). The four agent types had cognitive thresholds between -7 and +7, and most had unbiased cognition. Here, I use the term 'unbiased cognition' to refer to agents who had cognitive thresholds around 0. As the summary distributions tended to associate negative cue summaries with state 0, and positive cue summaries with state 1, then agents with cognitive thresholds around 0 were likely to perceive the environmental state in each domain without bias. Alternatively, thresholds below 0 made the agent more likely to believe that the state was 1 and thresholds above 0 made the agent more likely to believe that the state was 0.

The majority of agents had cognitive thresholds that were unbiased or slightly positive (see Figure 3). The priors were skewed towards state 0 ($p_A = 0.1$, $p_B = 0.1$). This meant that the agents' private signal was likely to be below their cognitive threshold ($x_A <= T_A$ or $x_B <= T_B$). Due to this, the agents were likely to believe that the state was 0 ($s_A = 0$ or $s_B = 0$). As they were more likely to believe that the state was 0 – and β motivational thresholds were utilised whenever the agent believed that the state was 0 – then there was stronger selection acting upon this value as it was more commonly used. β evolved to be close to 1 for agents with unbiased and positive cognitive thresholds (see figure 3). The agents were motivated to hunt or build more when hunting and building more was defection. To illustrate with the hunting example; the game was scarce, and the unbiased agent was aware that the game was scarce. The agent was still motivated to over-hunt to maximise her own resources at the expense of her neighbours.

Some agents had negative cognitive threshold values and needed *less* evidence from the environment to believe that the state was 1. As the prior probabilities of the run meant that state 1 was very unlikely ($p_A = 0.1$, $p_B = 0.1$), then these agents may have a *cognitive bias.* I use this term to refer to agents who were likely to develop a false belief about the value of the environmental state (Efferson et al., 2020). Here, the agents with negative cognitive thresholds had a cognitive bias that made them more likely to believe in a rare event. To illustrate in reference to dilemma A, this would be the equivalent of the resources being scarce and yet the agent believed that hunting more was cooperation.

Whenever the agents had negative cognitive thresholds, then even a small private signal was likely to exceed their threshold ($x_A > T_A$ or $x_B > T_B$). Thus, the agents were likely to (erroneously) believe that the state was 1 ($s_A = 1$ or $s_B = 1$). As the agents were likely to believe that the state was 1 – and as α motivational thresholds were utilised whenever the agent believed that the state was 1 – then there was more selection acting upon this value. Interestingly, there was strong selection acting on α to be close to 1. The agent was motivated to hunt or build more.

Interestingly, the agents with negative cognitive thresholds *believed* that the state was 1, and thus *believed* that hunting or building more was cooperative. Playing behaviour 1 may appear to be cooperative for this agent, but when one considers that the priors favoured state 0 ($p_A = 0.1$, $p_B = 0.1$) then it can be seen that playing behaviour 1 would actually uphold defection, regardless of the agents' beliefs. To illustrate, this would be the equivalent of the agent believing that game was abundant and that hunting more would be cooperative. However, the game was actually scarce in this environment and so hunting more would be to take resources away from others and so would be defection. Despite the agents beliefs regarding the environment, they were motivated to go hunting and so were motivated to choose hunting as defection. This is an interesting finding as it highlights that motivation can compensate for cognitive biases.

It is also interesting to point out any differences between the four agent types. All four agent types had very similar psychological architecture on runs with consistent prior probabilities ($p_A = 0.1, p_B = 0.1$), with the exception of partly modular agents with modular motivation. These agents had an average cognitive threshold of +3. Despite the agents' positive cognitive thresholds, partly modular agents with modular motivation would still defect, comparable to the other three agent types (Section 3.1, Figure 2i). All four agent types may have shown similar outcomes, but these could be upheld by a wide range of underlying psychological architecture. The fact that agents with modular motivation but domain-general cognition upheld similar outcomes to the other three agent types despite their different cognitive thresholds again highlights the powerful role that motivation played in compensating for cognition when the two structures coevolved.

Now, consider runs where the prior probabilities favour environmental state 1 ($p_A = 0.9$, $p_B = 0.9$; Figure 4). Note that this condition is analogous to $p_A = 0.1$, $p_B = 0.1$. The priors were skewed but consistent in both runs, meaning that the environmental states were likely to take similar values in both domains A and B. Of course, this run in Figure 4 flips the logic as now state 1 was likely to occur ($S_A = 1$ and $S_B = 1$). So, the agents' cognitive and motivation thresholds would expectantly evolve towards the eventuality that the state was 1.

The agents mostly had unbiased cognition though again, partly modular agents with modular motivation developed positive cognitive thresholds. This was a cognitive bias, as it meant that these agents needed more evidence to believe that the environmental state was 1, even though this event was likely. This is interesting, as cognitive biases seem most likely to emerge in agents with domain-general cognition but modular motivation. Despite this adjustment in cognition, partly modular agents with modular motivation still upheld mutual defection as the common outcome (see Section 3.1, Figure 2). This again reflects that motivation compensated for cognitive biases.

Figure 4 reveals that the agents with unbiased cognition or negative cognitive thresholds have strong selection acting upon their α value to be low. This meant that the agent was unmotivated to cooperate whenever they believed that hunting or building more was cooperative. The agents instead hunted or built less as defection. To illustrate with the hunting example, the agent was aware that game was abundant and yet did not go hunting.

Finally, motivation again compensated for cognitive biases. There were some runs where agents have positive cognitive thresholds (see Figure 4). This was unexpected, as agents with positive cognitive thresholds needed *more* evidence from the environment to believe that the state was 1, although the priors made state 1 very likely to occur ($p_A = 0.9$, $p_B = 0.9$). These agents had a cognitive bias to (erroneously) believe that the state was 0. As the agent was likely to believe that the state was 0, then there was greater selection acting on β which evolved to be close to 0. Whenever the agent believed that hunting or building more was defection, then they chose to hunt or build less. This may seem to be a cooperative decision, but again it is imperative to remember that the priors in this run favoured state 1 ($p_A = 0.9, p_B = 0.9$). Regardless of the agents' beliefs, hunting and building more was likely to be cooperative. By hunting and building less, the agent was therefore defecting. To illustrate this choice with the hunting example, the agent may have falsely believed that the game was scarce and so hunting would be defection. She thus avoided hunting. However, in this particular environment the game was actually abundant and, by refusing to hunt, she was not helping to provide any meat to the community. Motivation once again compensated for cognition.

Figure 3(i) $p_A = 0.1$, $p_B = 0.1$. *Fully modular, Domain A.*

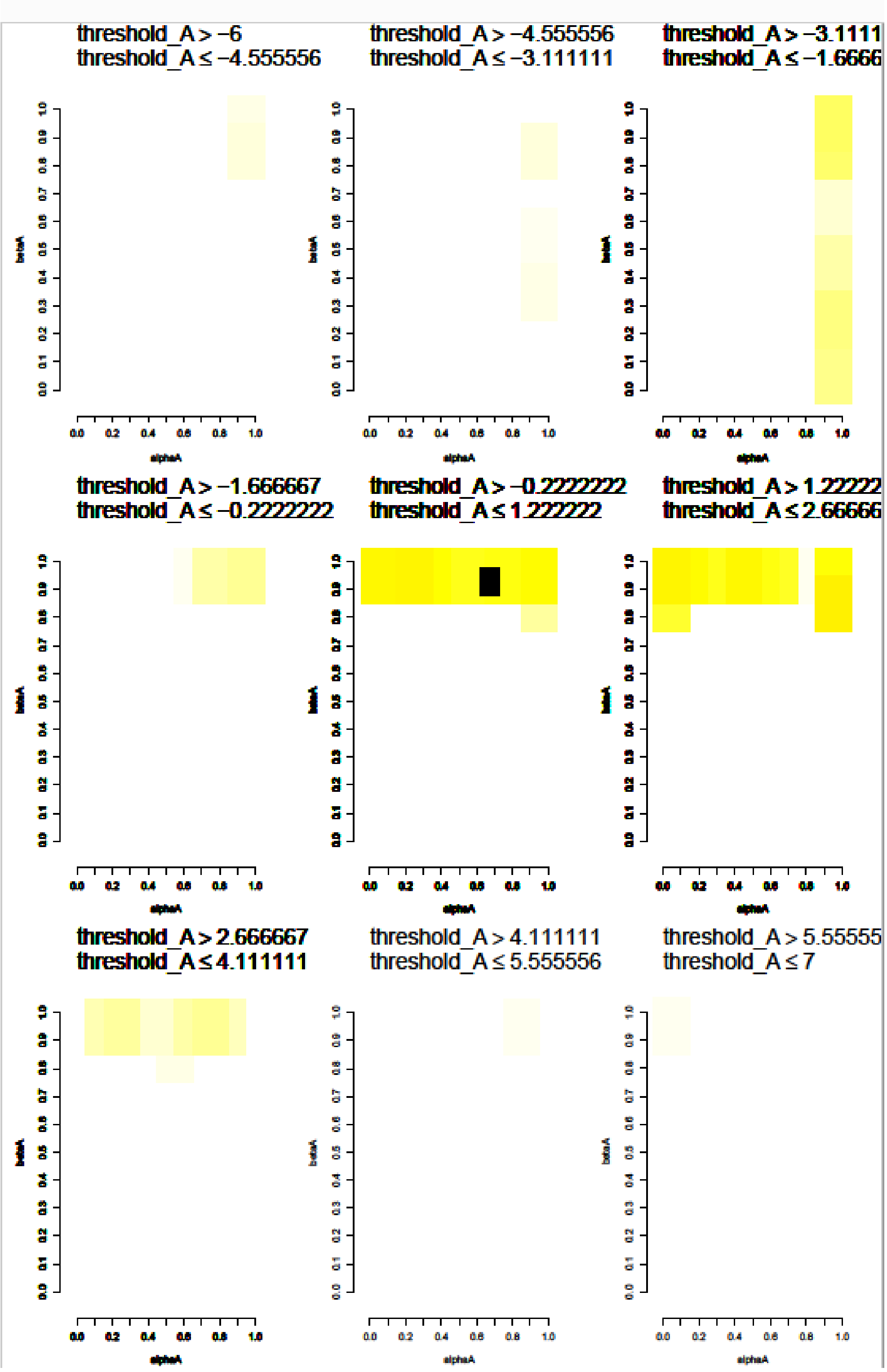


Figure 3(i) $p_A = 0.1$, $p_B = 0.1$. Fully modular, Domain B.

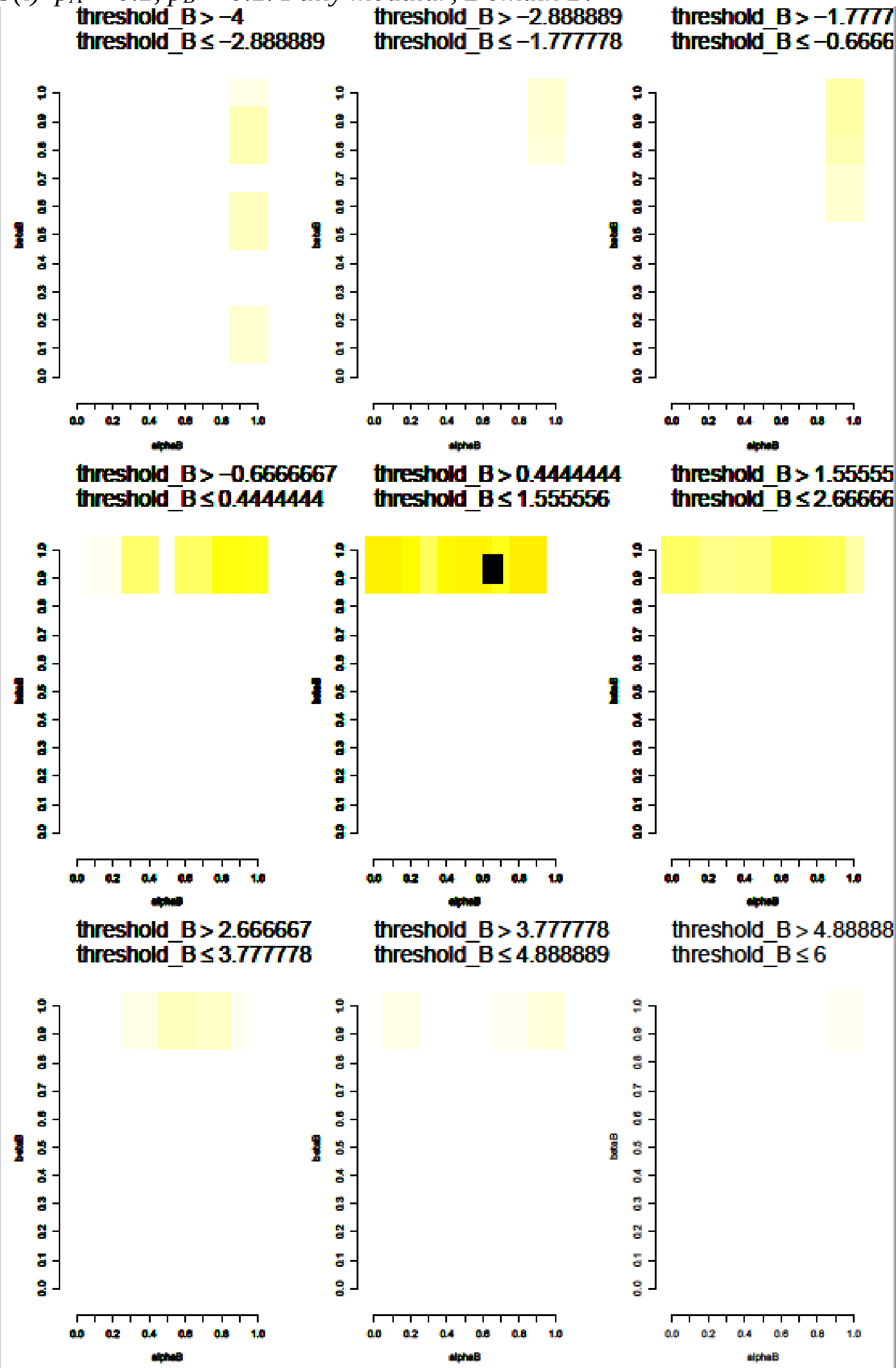


Figure 3(ii) $p_A = 0.1$, $p_B = 0.1$. *Modular motivation, Domain A.*

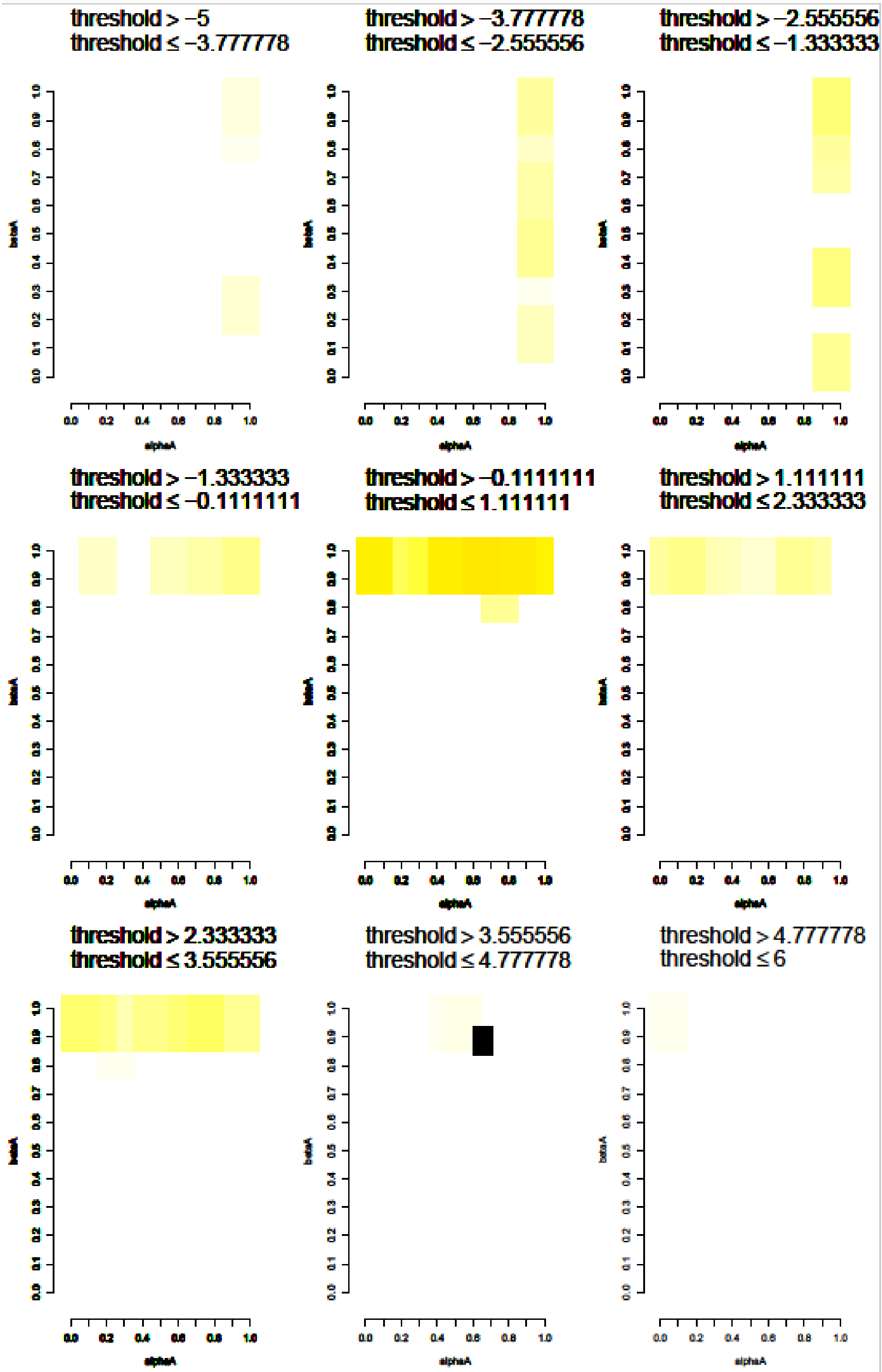


Figure 3(ii) $p_A = 0.1$, $p_B = 0.1$. Modular motivation, Domain B.

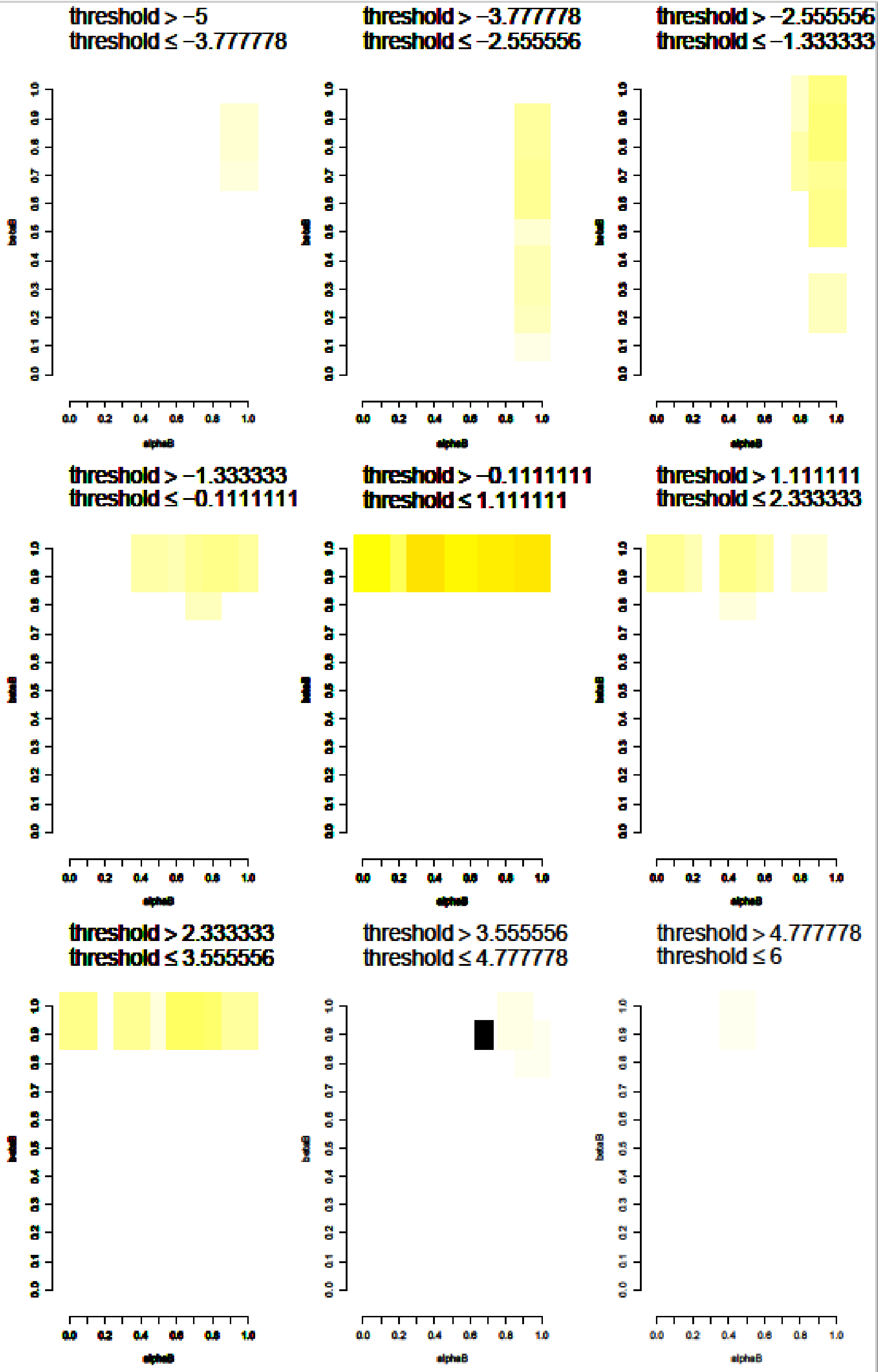


Figure 3(iii) $p_A = 0.1$, $p_B = 0.1$. Modular cognition, Domain A.

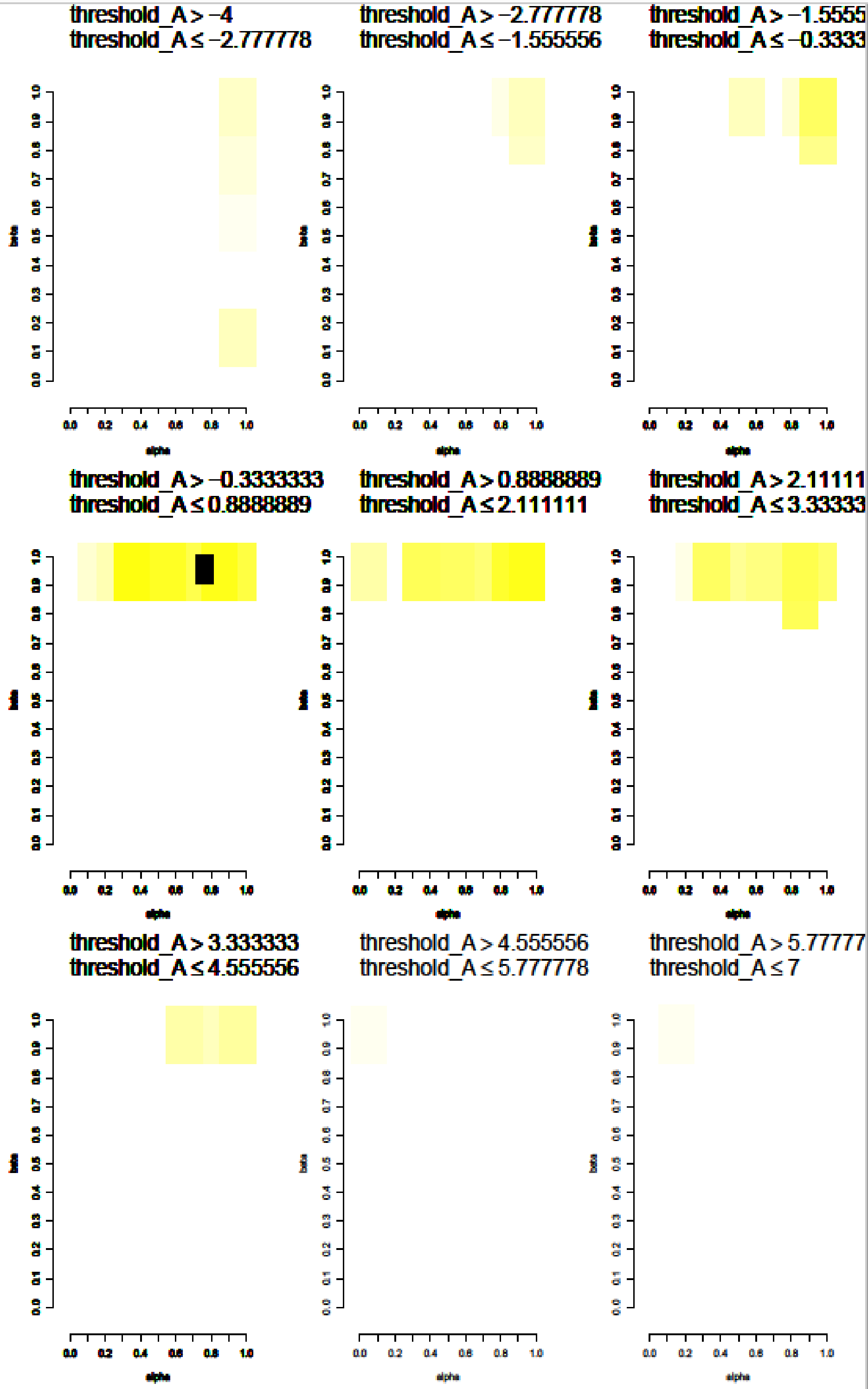


Figure 3(iii) $p_A = 0.1$, $p_B = 0.1$. Modular cognition, Domain B.

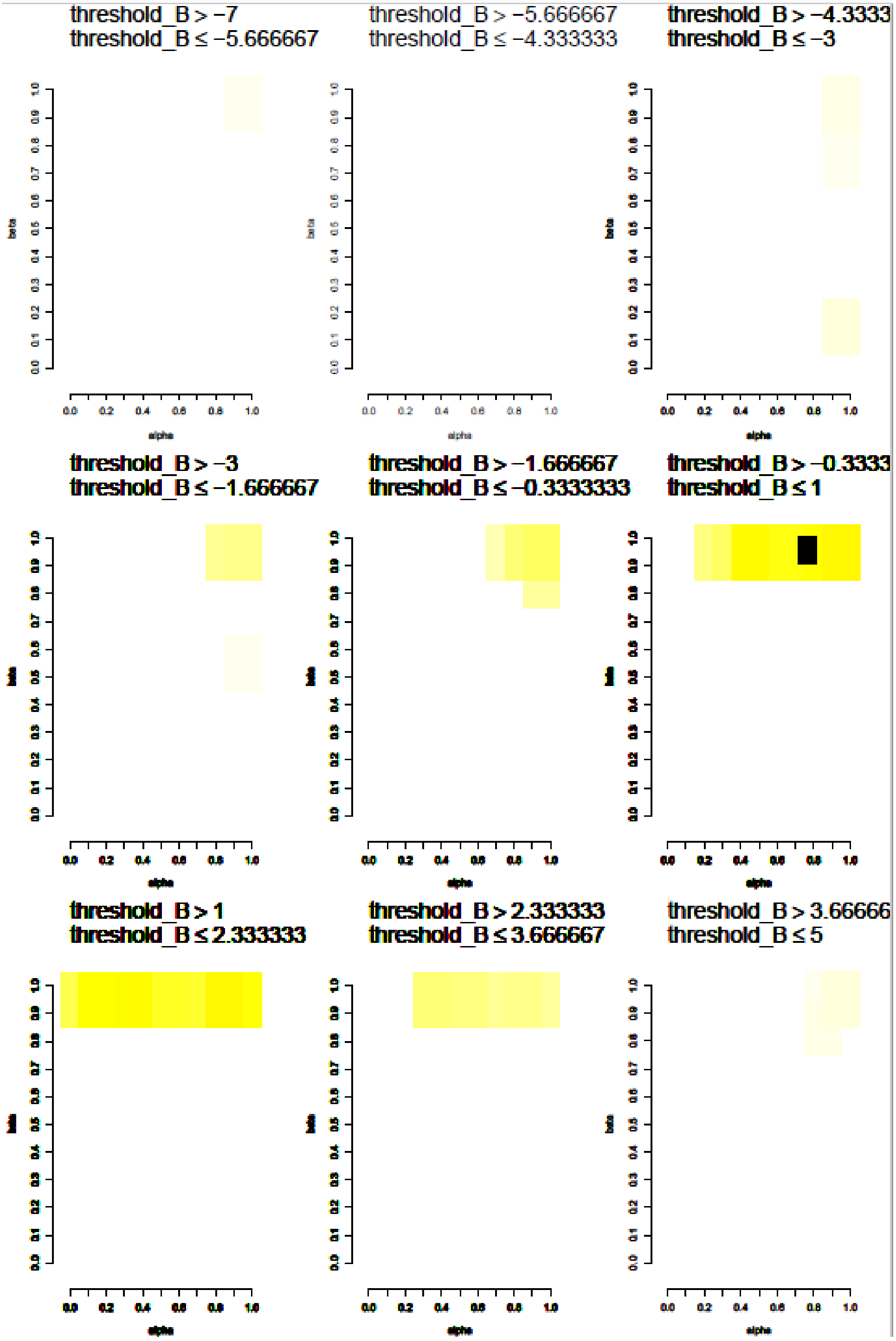


Figure 3(iv) p_A = 0.1, p_B = 0.1. Domain-general

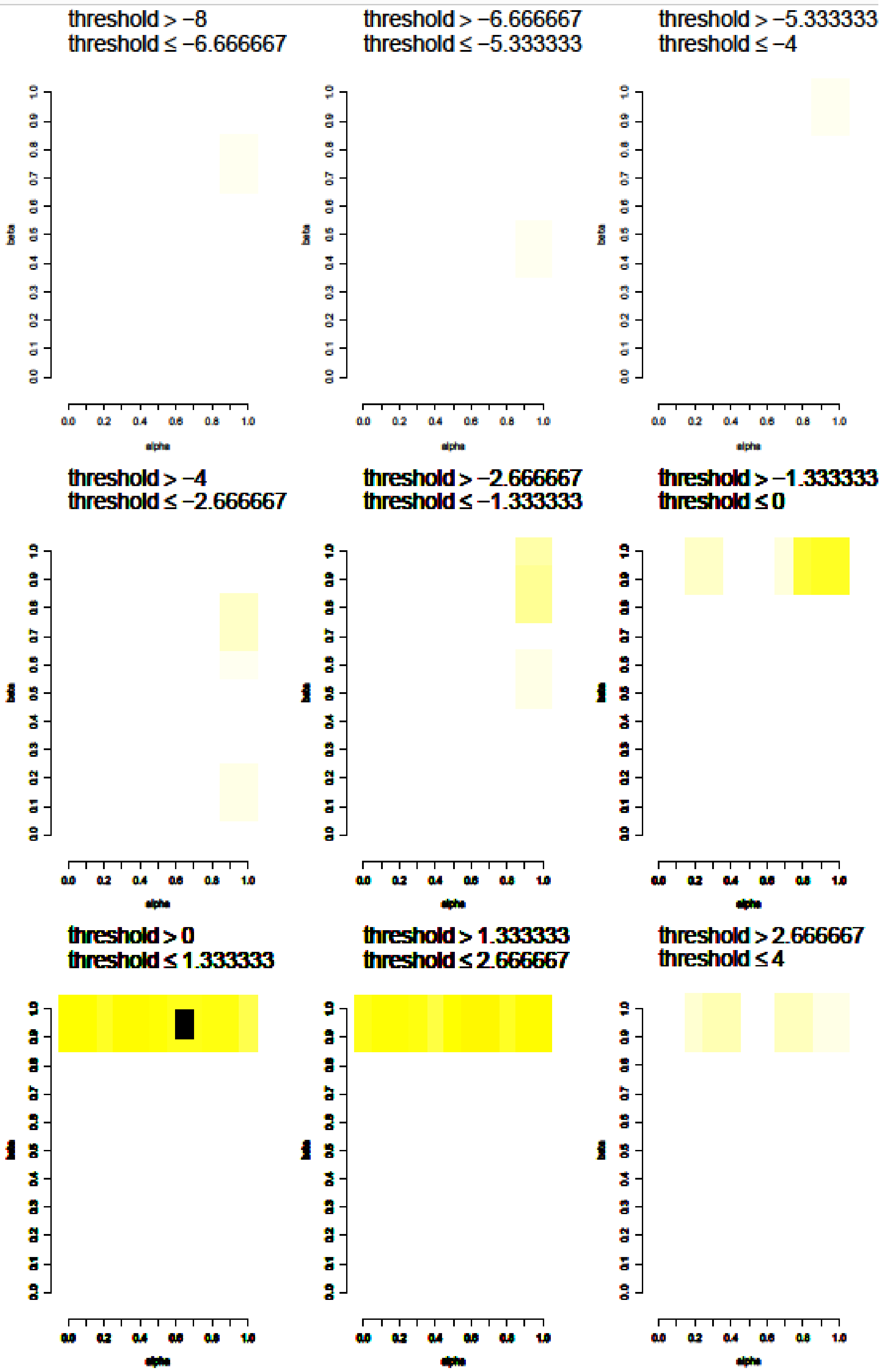


Figure 4(i) p_A = 0.9, p_B = 0.9. Fully modular, domain A

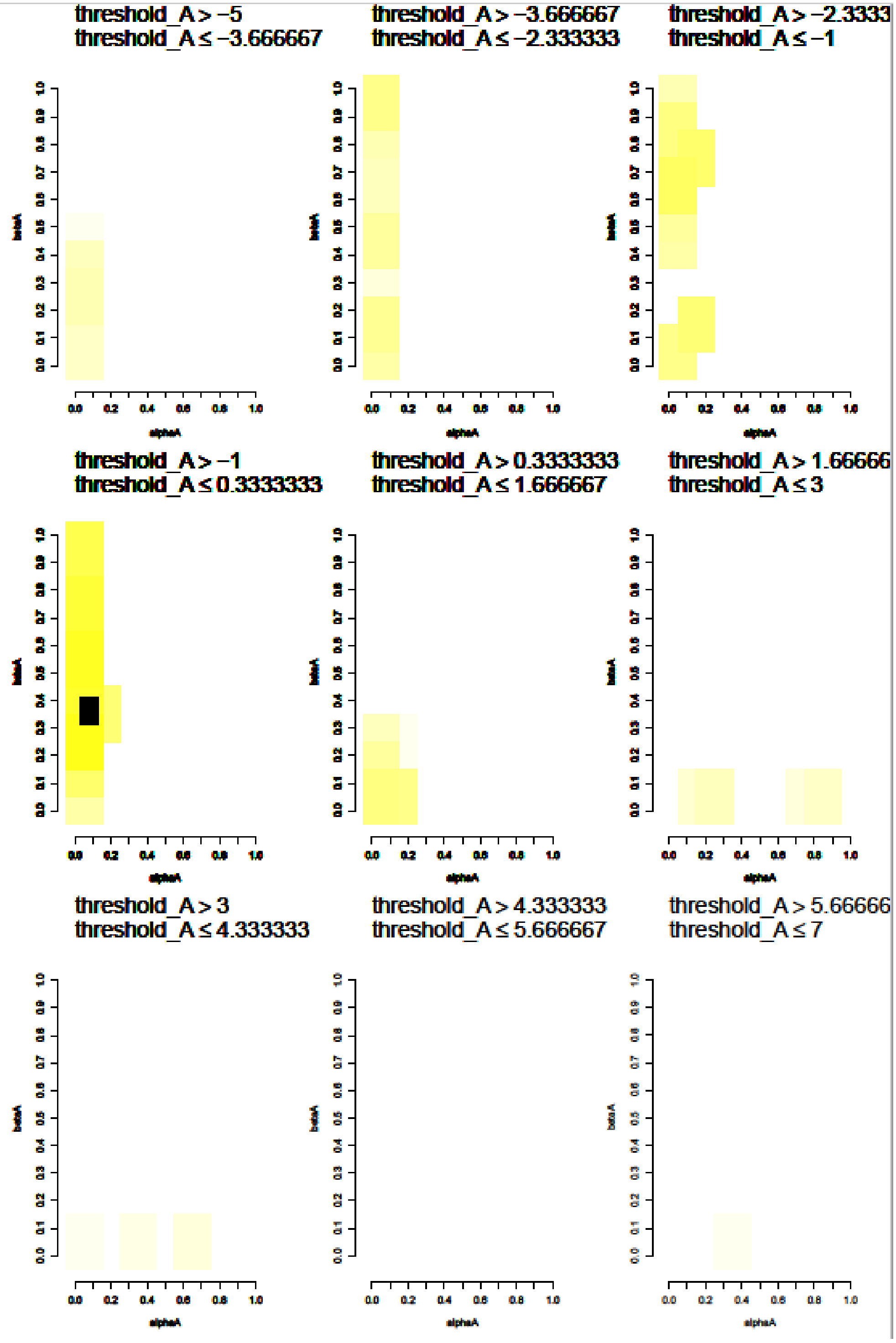


Figure 4(i) $p_A = 0.9$, $p_B = 0.9$. Fully modular, domain B

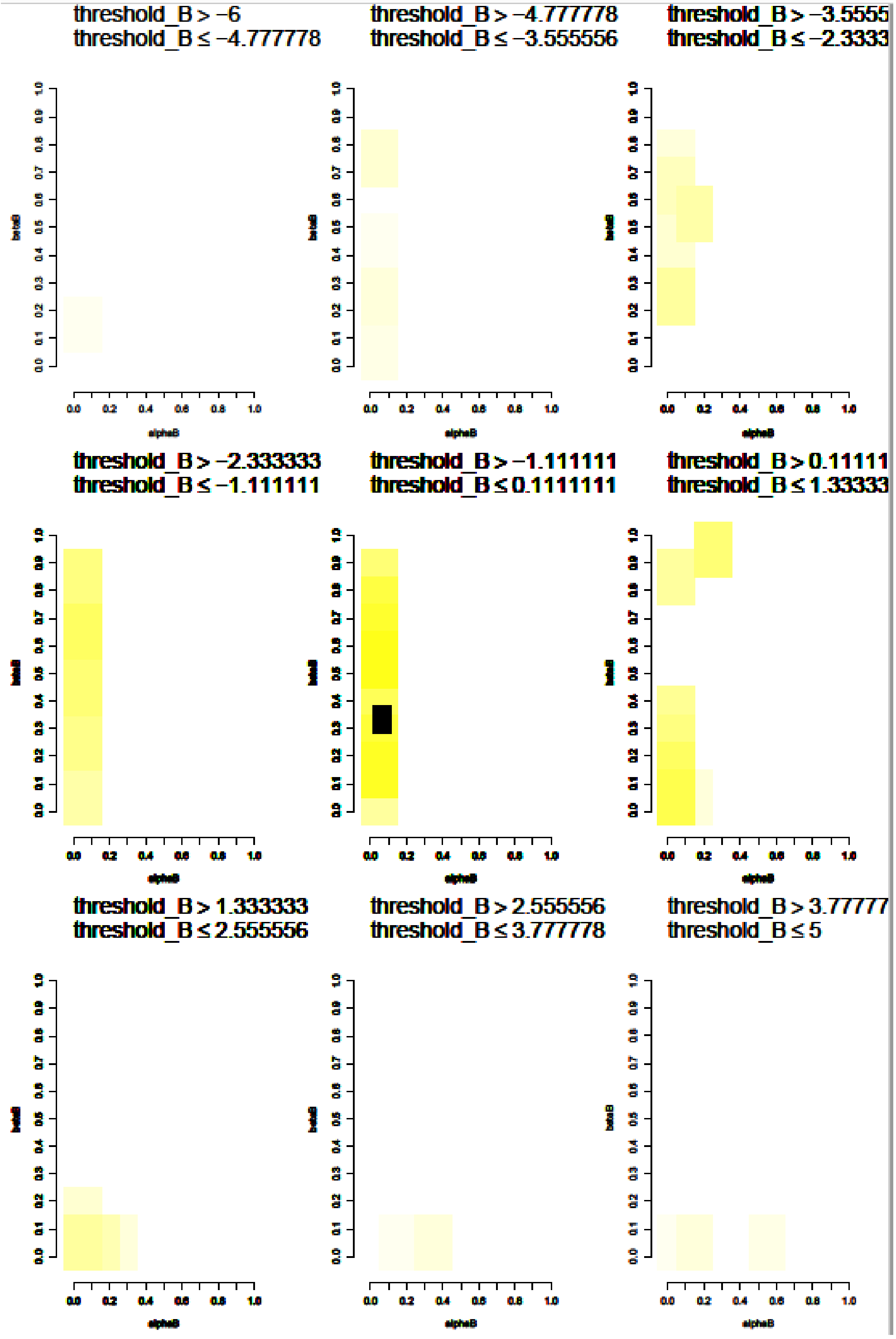


Figure 4(ii) p_A = 0.9, p_B = 0.9. Modular motivation, domain A

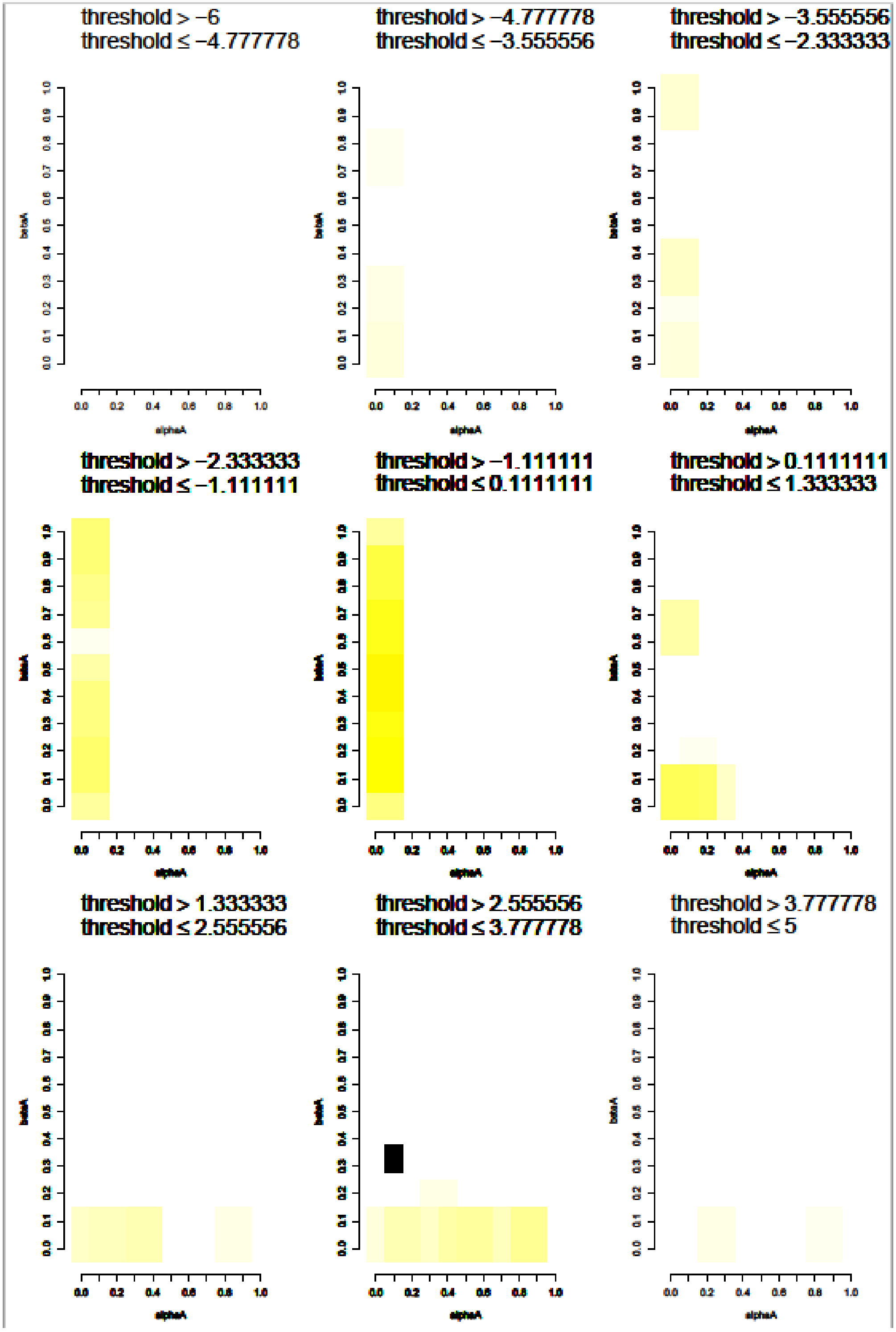


Figure 4(ii) $p_A = 0.9$, $p_B = 0.9$. Modular motivation, domain B

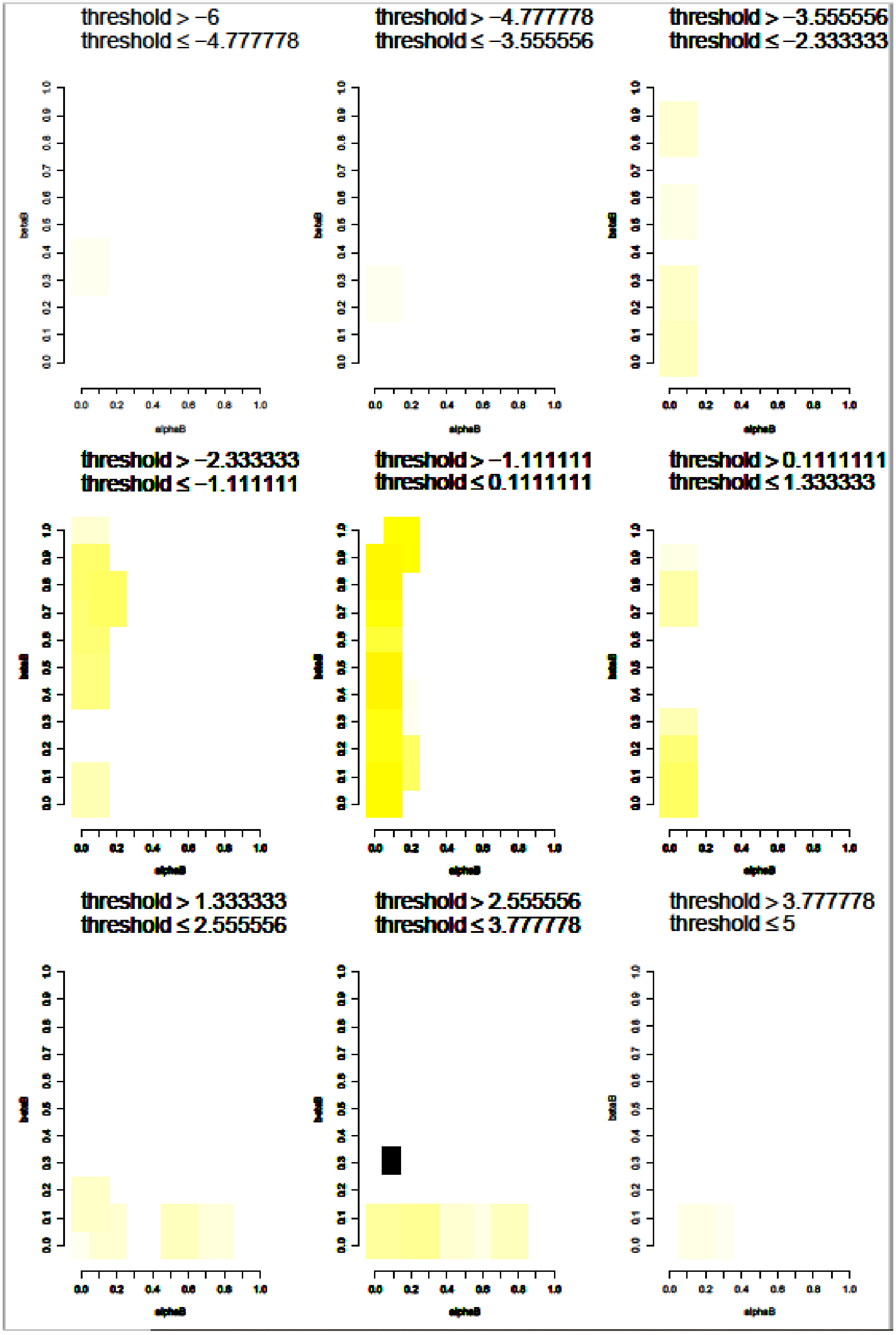


Figure 4(iii) $p_A = 0.9$, $p_B = 0.9$. Modular cognition, domain A

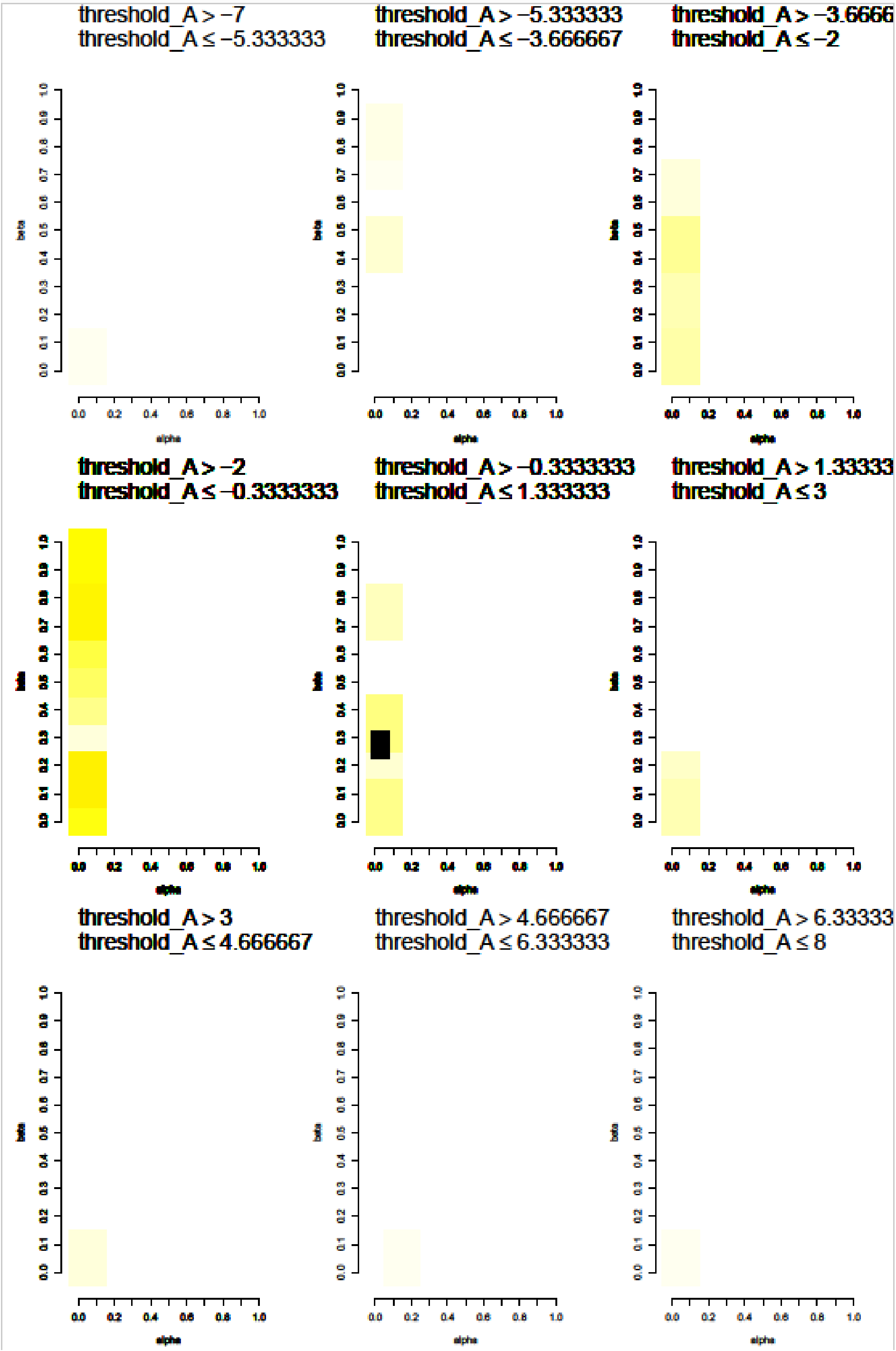


Figure 4(iii) $p_A = 0.9$, $p_B = 0.9$. Modular cognition, domain B

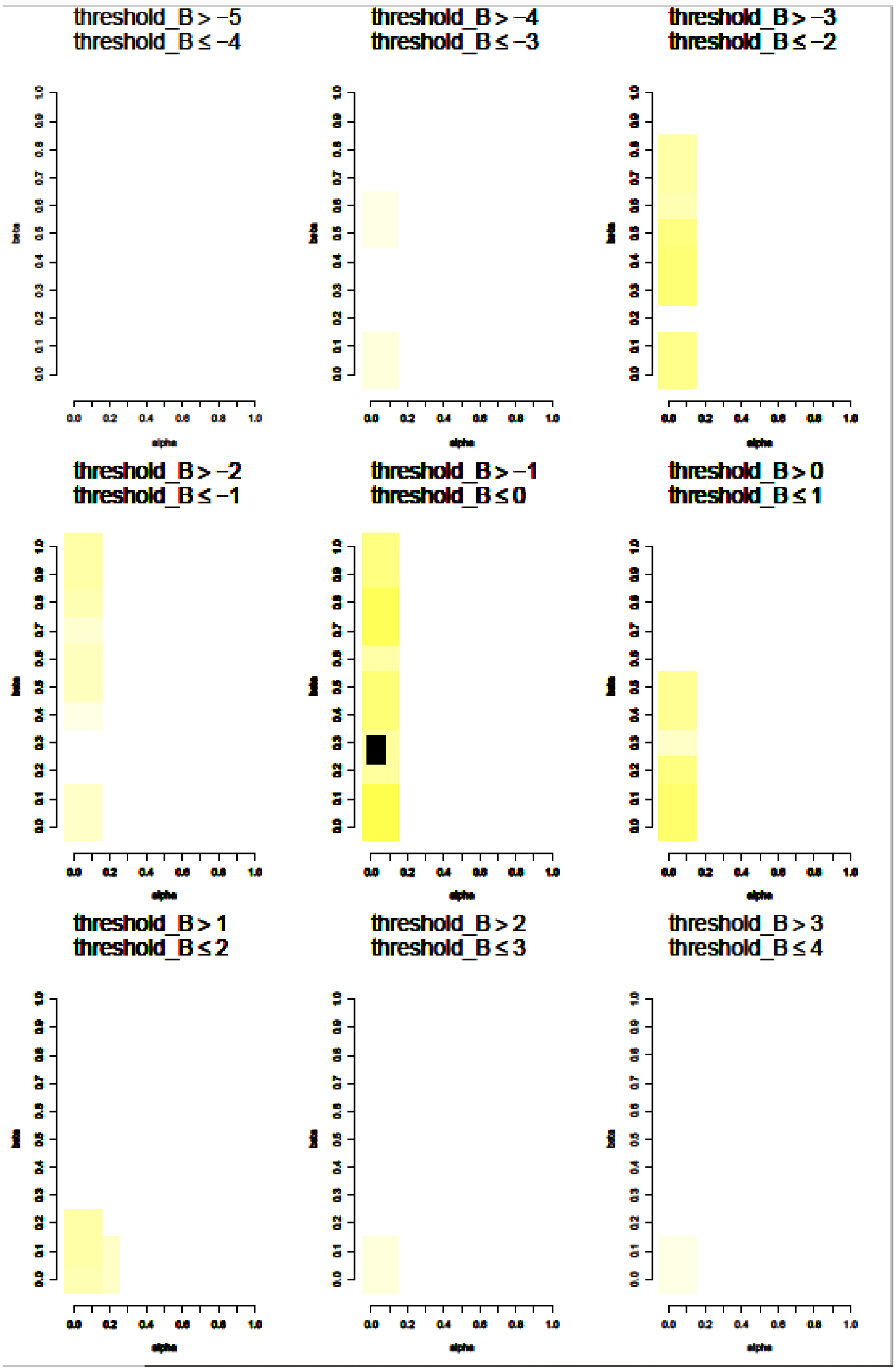


Figure 4(iv) $p_A = 0.9$, $p_B = 0.9$. Domain-general

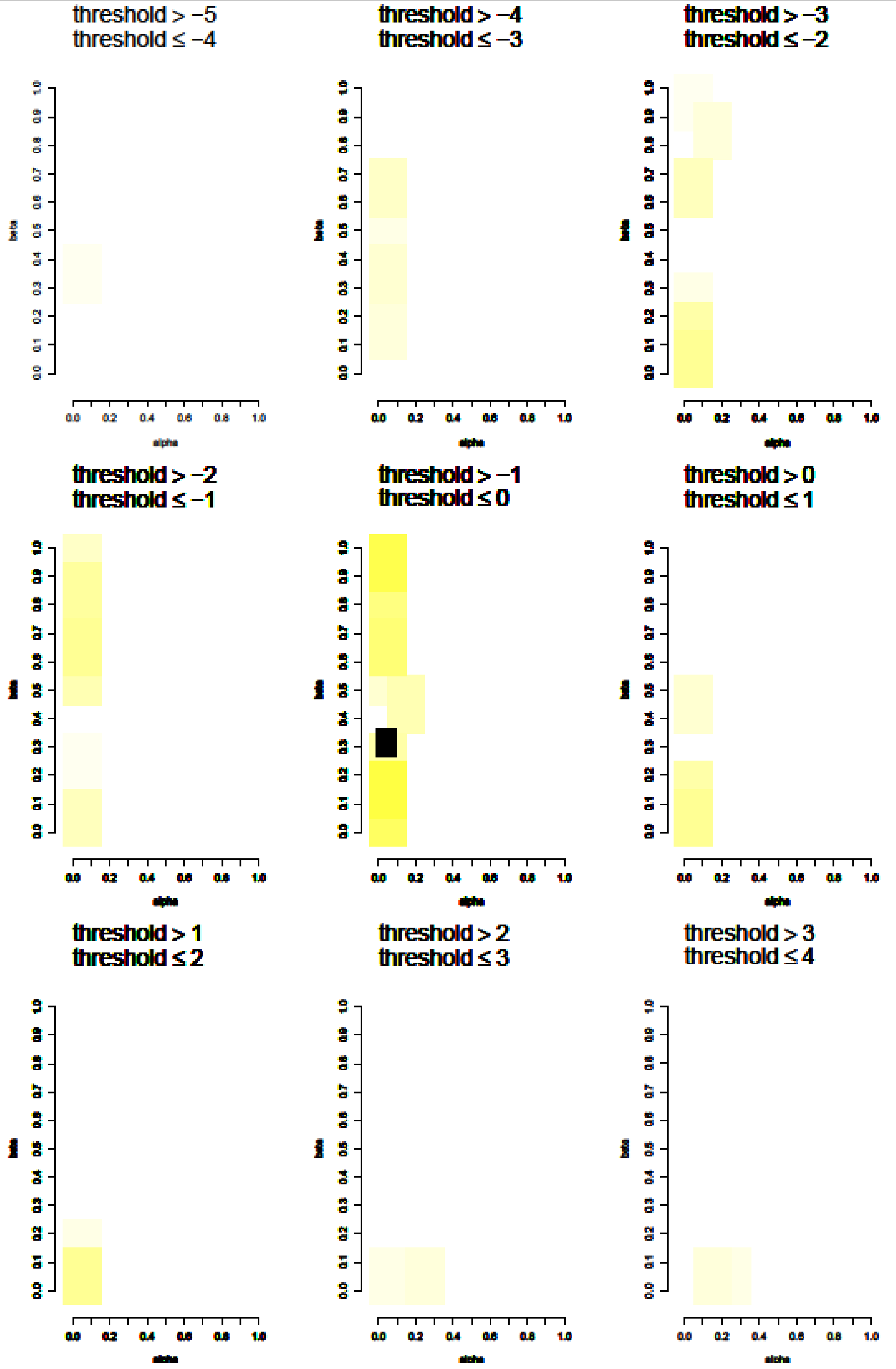


Figure 3 and Figure 4. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) partly modular agents with modular motivation; (iii) partly modular agents with modular cognition and (iv) fully domain-general agents. These figures give the results for runs with skewed but consistent prior probabilities. Figure 3 gives the results for runs with $p_A = 0.1$, $p_B = 0.1$, and Figure 4 gives the results for runs with $p_A = 0.9$, $p_B = 0.9$. Note that these heatmaps represent cases with strong benefits to cooperation (b=4), though see appendix 4 for runs where b=2.

Thus far, all four agent types had similar psychological architecture by the final generation. Of course, the benefits of modularity may become more apparent on runs with inconsistent prior probabilities ($p_A = 0.1$, $p_B = 0.9$; see Figure 5*)*. First, I consider the evolution of modular cognitive thresholds in response to these inconsistent priors. Partly modular agents with modular cognition had a slight adjustment where their T_A values became positive, but T_B became negative. These agents were less likely to believe that the environmental state was 1 when this was rare in domain A but were more likely to believe that the environmental state was 1 when this was common in domain B (this is even easier to see in the line graphs in appendix 1). Thus, these adjustments in the cognitive thresholds were expected. Fully modular agents did not have equivalent adjustments in their T_A or T_B values, however. Fully modular agents had unbiased cognition clustered around 0. This perhaps shows that modular cognition may have only evolved to compensate for the inconsistent priors across two domains when motivation could not specialise to these demands.

Domain-general agents did not experience adjustments in their cognitive thresholds (Figure 5iv) and partly modular agents with modular motivation, but a domain-general cognitive threshold, actually evolved a positive cognitive threshold (Figure 5ii). This positive threshold was likely to lead to an accurate belief regarding the environmental state in domain A ($p_A = 0.1$) but was likely to lead to a cognitive bias in domain B ($p_B = 0.9$). Thus, domain-general cognition either remained unbiased in the face of domains with contrasting selection pressures or adjusted in favour of one domain only as, by definition, a domain-general cognitive system could not adjust to the contrasting demands of multiple domains.

Agents with modular motivation (fully modular and agents with modular motivation but domain-general cognition) had a striking pattern of results. In domain A, when the environmental state was likely to be 0, then strong selection acted on β_A to be high but weak selection acted on α_A to take any value. This was true for agents with unbiased cognition or positive thresholds (who were likely to believe that the state was 0). Thus, modular motivation for domain A matches cases where $p_A = 0.1, p_B = 0.1$. In domain B, when the environmental state was likely to be 1, then strong selection acted on α_B to be low though β_B could take any value. This is true for agents with unbiased cognition or negative thresholds (who were likely to believe that the state was 1), as they were more likely to be using their α_B value. Thus, modular motivation for domain B matches cases where $p_A = 0.9, p_B = 0.9$. Thus, agents with modular motivation can mutually defect across the two distinct domains.

Domain-general motivation could not discriminate well between two decision-making domains when the priors were skewed but inconsistent ($p_A = 0.1, p_B = 0.9$). Agents with domain-general motivation and modular cognition evolved a higher T_A and a lower T_B. This meant that the agents were unlikely to draw a private signal that exceeded their threshold in domain A ($x_A <= T_A$) and thus were likely to believe that the state was 0 ($s_A = 0$). In this case, strong selection acted on β to be high (i.e., to drive the mutual defection outcome seen in Section 3.1, Figure 2iii). Alternatively, the agents were likely to draw a private signal that exceeded their cognitive threshold in domain B ($x_B > T_B$) and thus were likely to believe that the state was 1 ($s_B = 1$). In this case, strong selection acted on α to be low (again this drives the mutual defection outcome seen in Section 3.1, Figure 2iii). This meant that modular cognition evolved in a way which maximised the use of the domain-general motivation systems.

To walk through an in-depth example of what this meant for the agents with modular cognition but domain-general motivation, these agents needed more evidence to believe that hunting was cooperative (perhaps as game was scarce) but needed less evidence to believe that building was cooperative (perhaps as this group typically cooperated to build shelter but not to hunt; Chudek & Henrich, 2011). Thus, the agent typically believed that hunting more was defection and that building more was cooperation. The agent was then motivated to hunt more and build less (i.e., they defected).

Domain-general agents had much more noise in their motivational thresholds, however (Figure 5iv). This noise may explain why domain-general agents display a wider range of choices than any of the other agent types, including some outcomes of mutual cooperation (see Section 3.1, Figure 2iv). When the payoffs of cooperation were sufficient *($b = 4$)*, then domain-general agents seem less motivated to avoid this. Cooperation may have arisen in a domain-general agent who must balance their psychology over two domains with inconsistent priors ($p_A = 0.1$, $p_B = 0.9$).

Figure 5(i) $p_A = 0.1$, $p_B = 0.9$. Fully modular, domain A

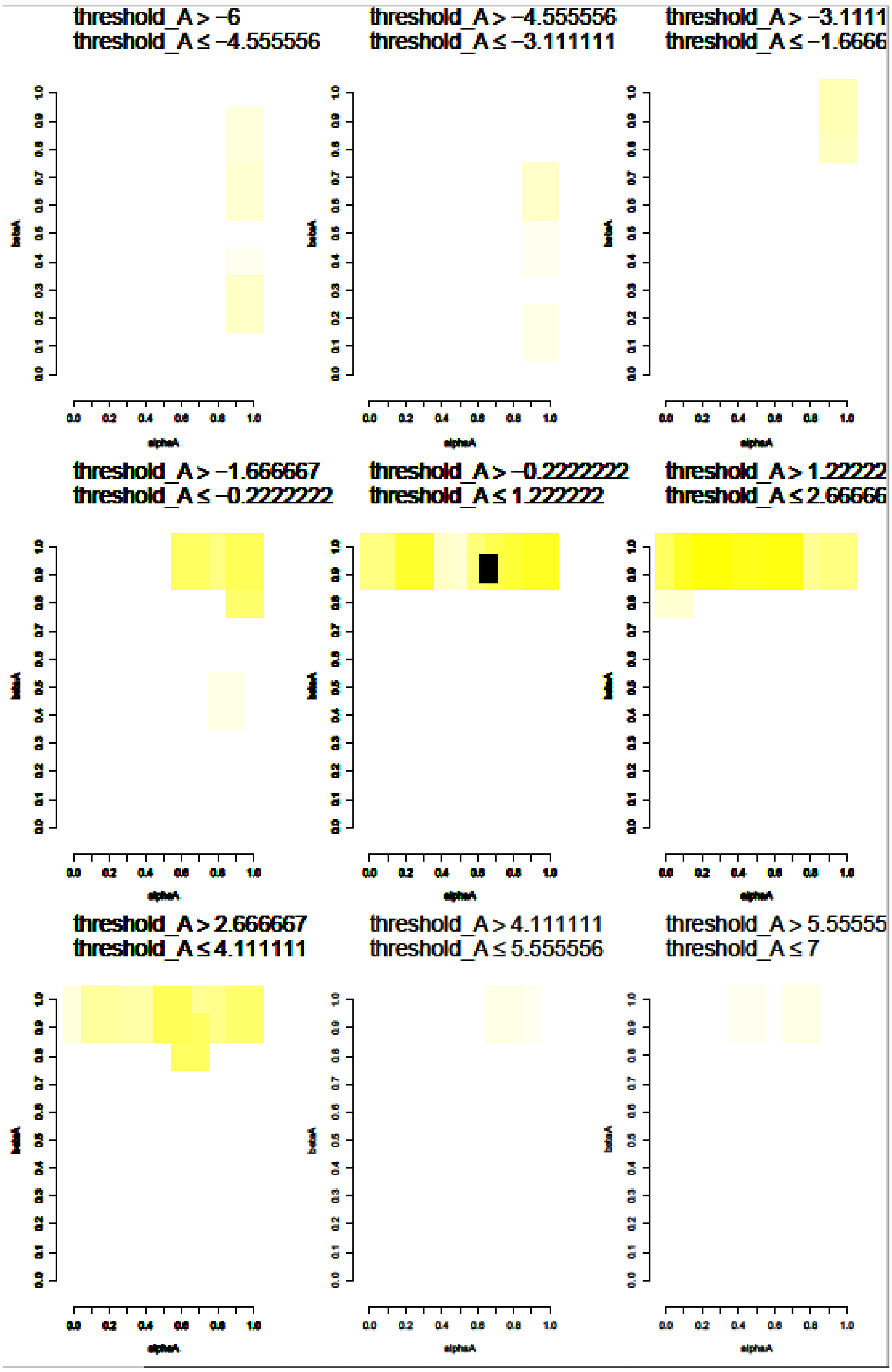


Figure 5(i) $p_A = 0.1$, $p_B = 0.9$. Fully modular, domain B

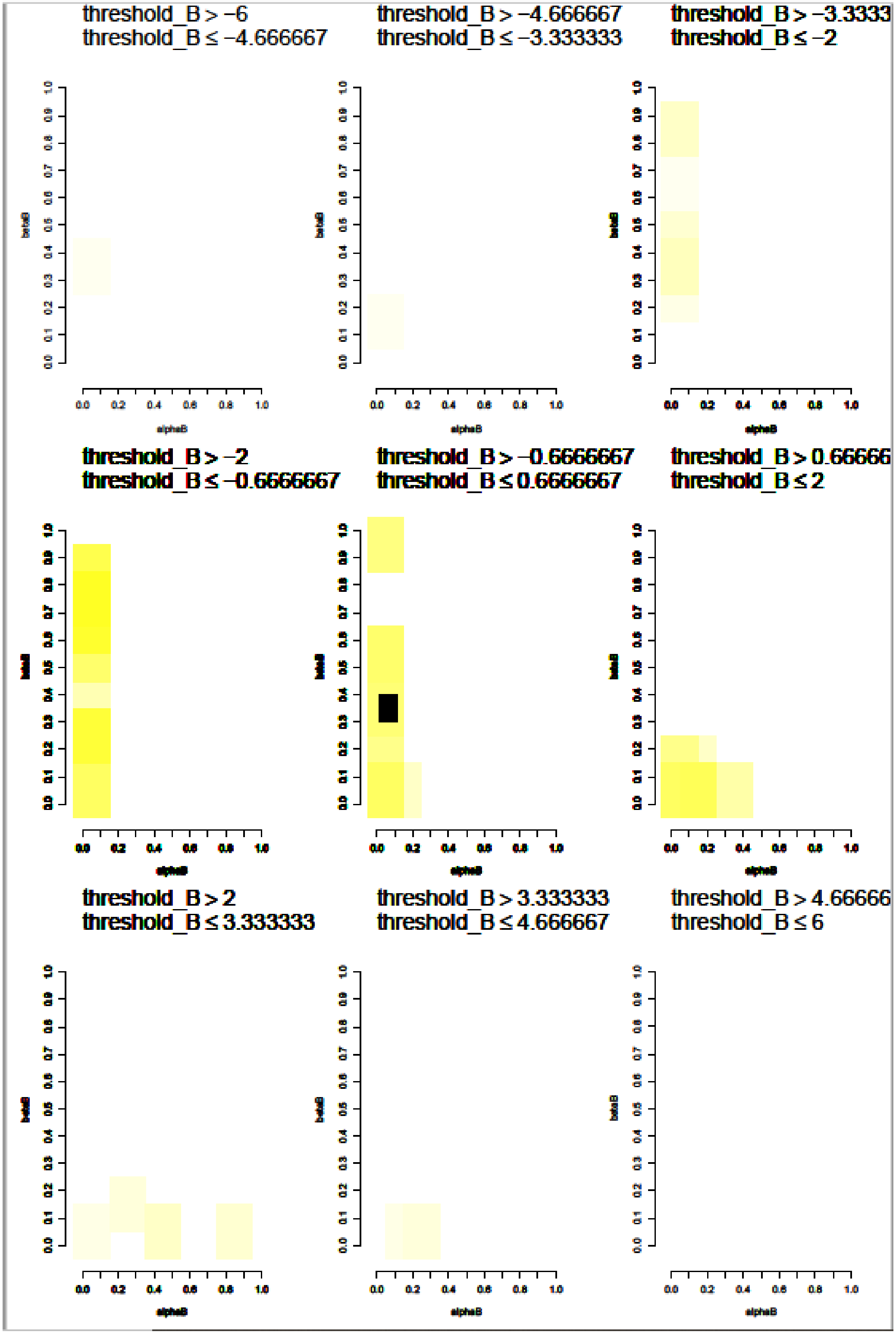


Figure 5(ii) p_A = 0.1, p_B = 0.9. Modular motivation, domain A

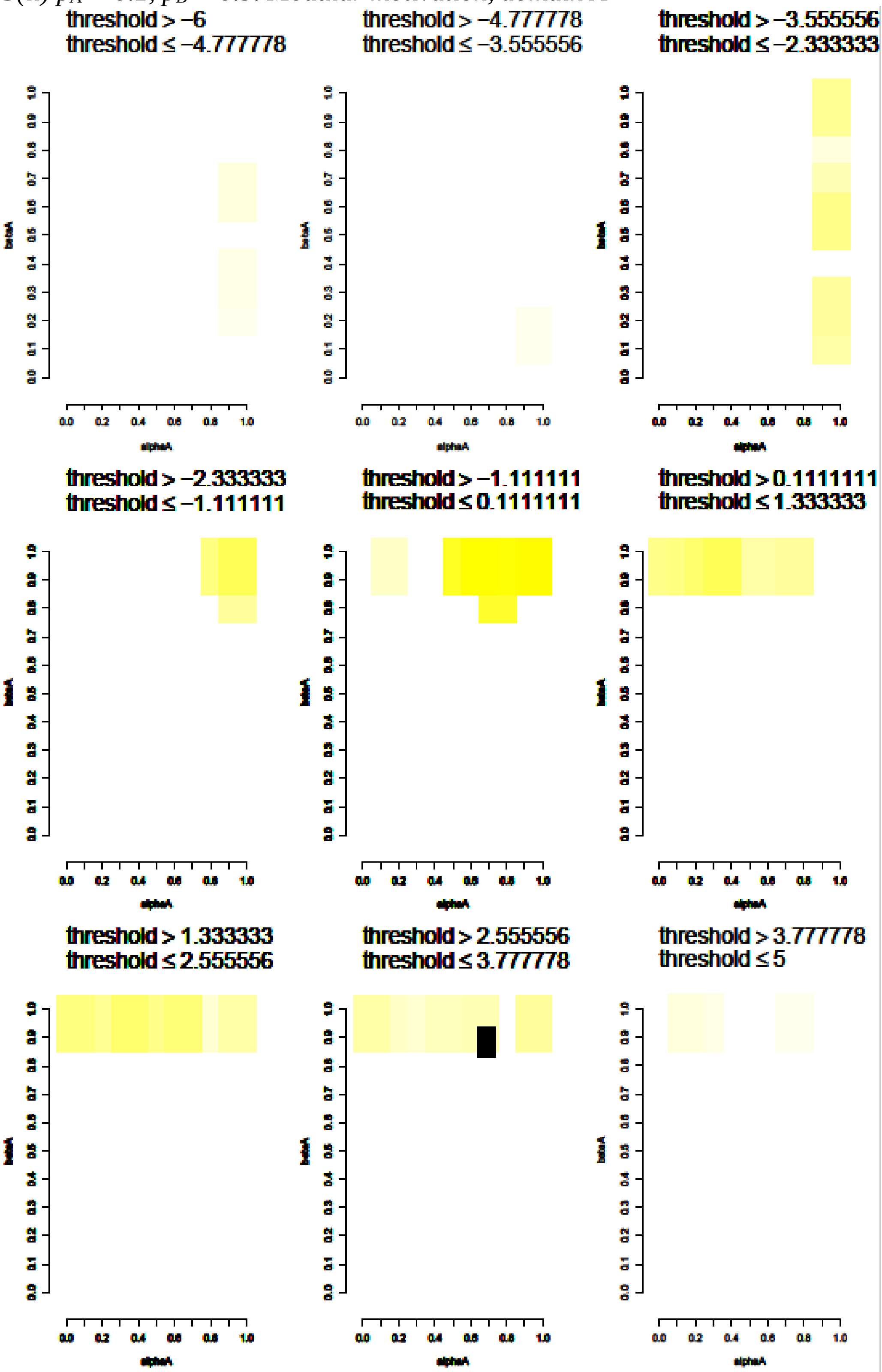


Figure 5(ii) $p_A = 0.1$, $p_B = 0.9$. Modular motivation, domain B

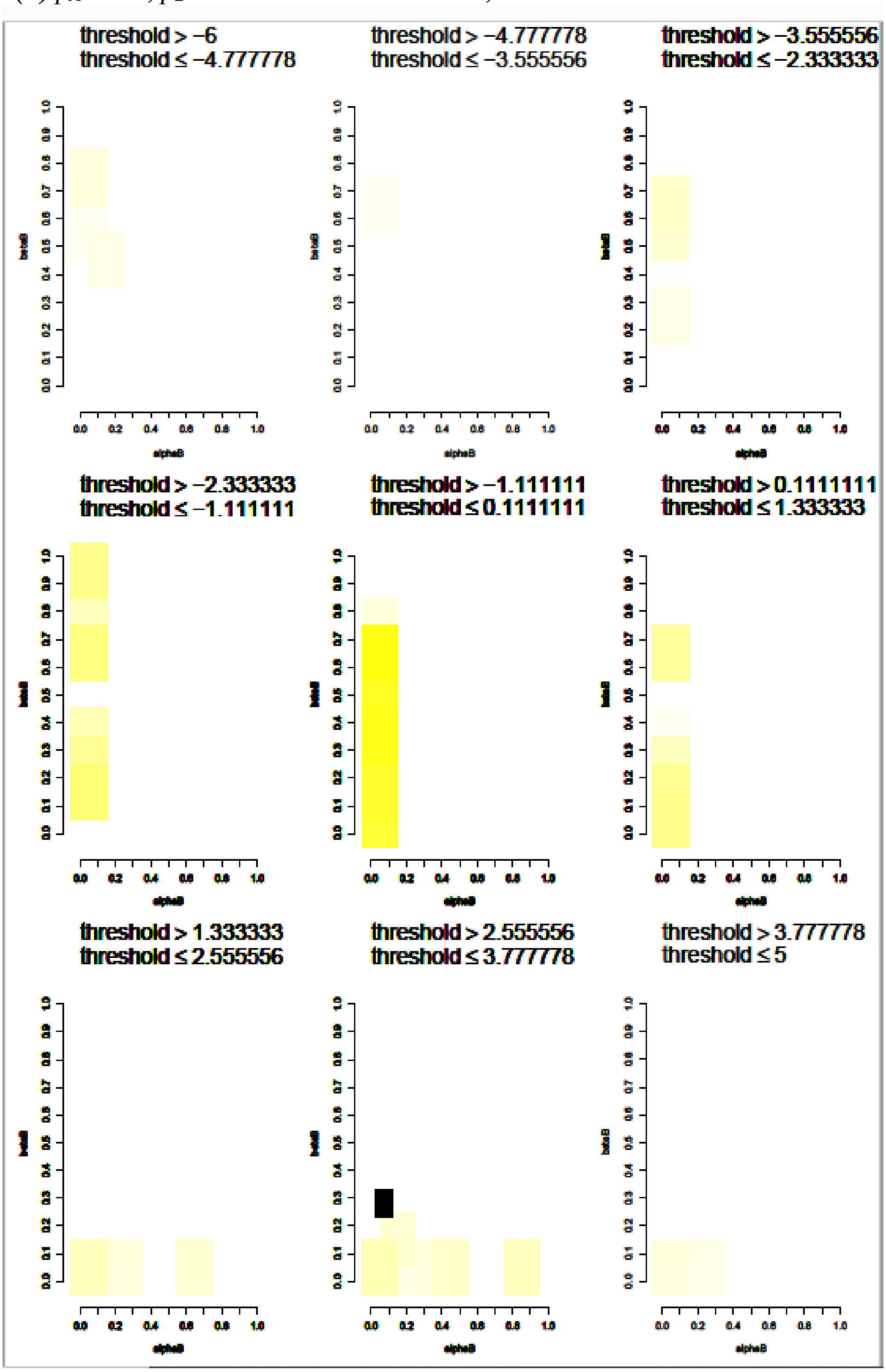


Figure 5(iii) $p_A = 0.1$, $p_B = 0.9$. Modular cognition, domain A

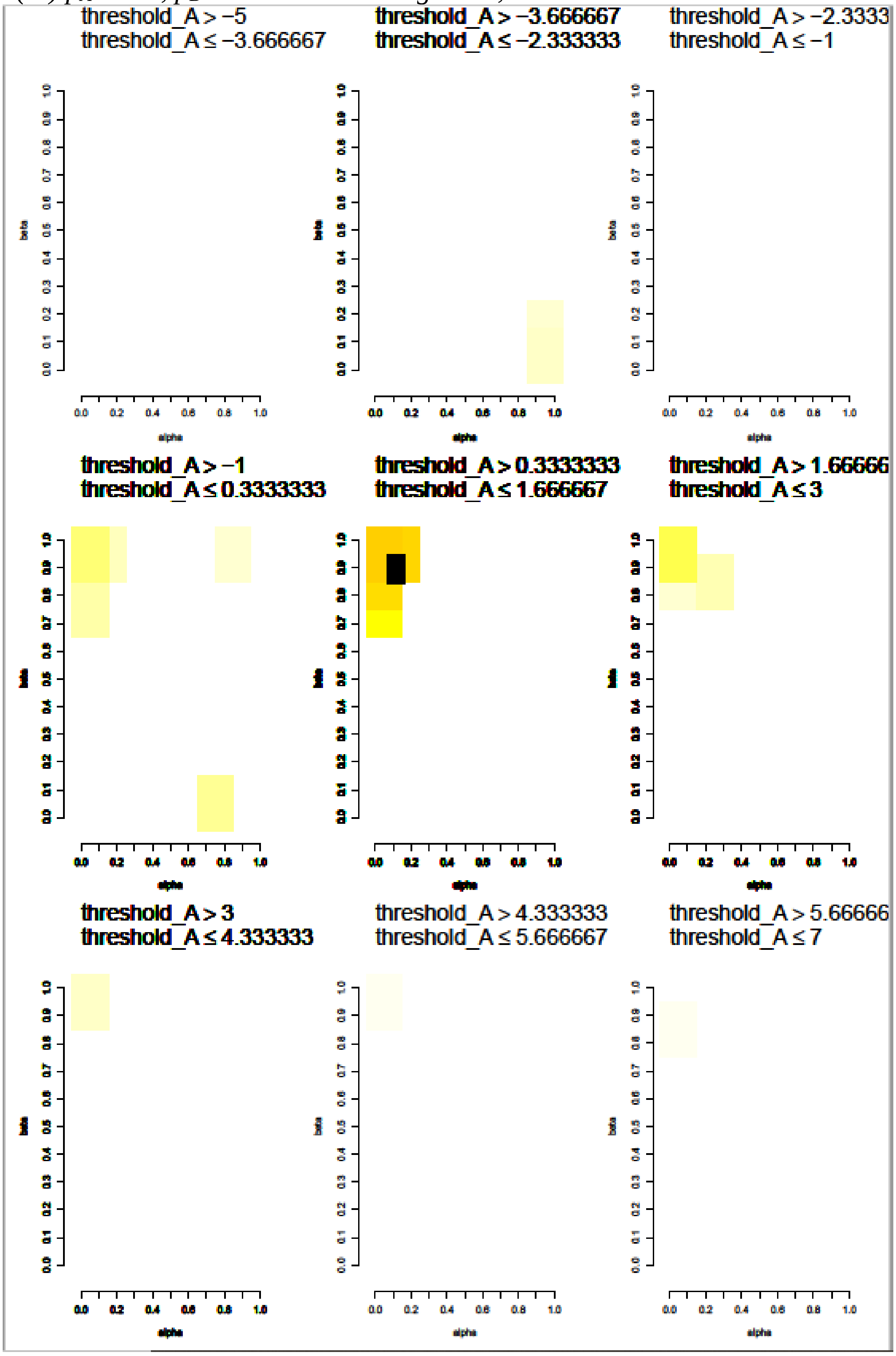


Figure 5(iii) $p_A = 0.1$, $p_B = 0.9$. Modular cognition, domain B

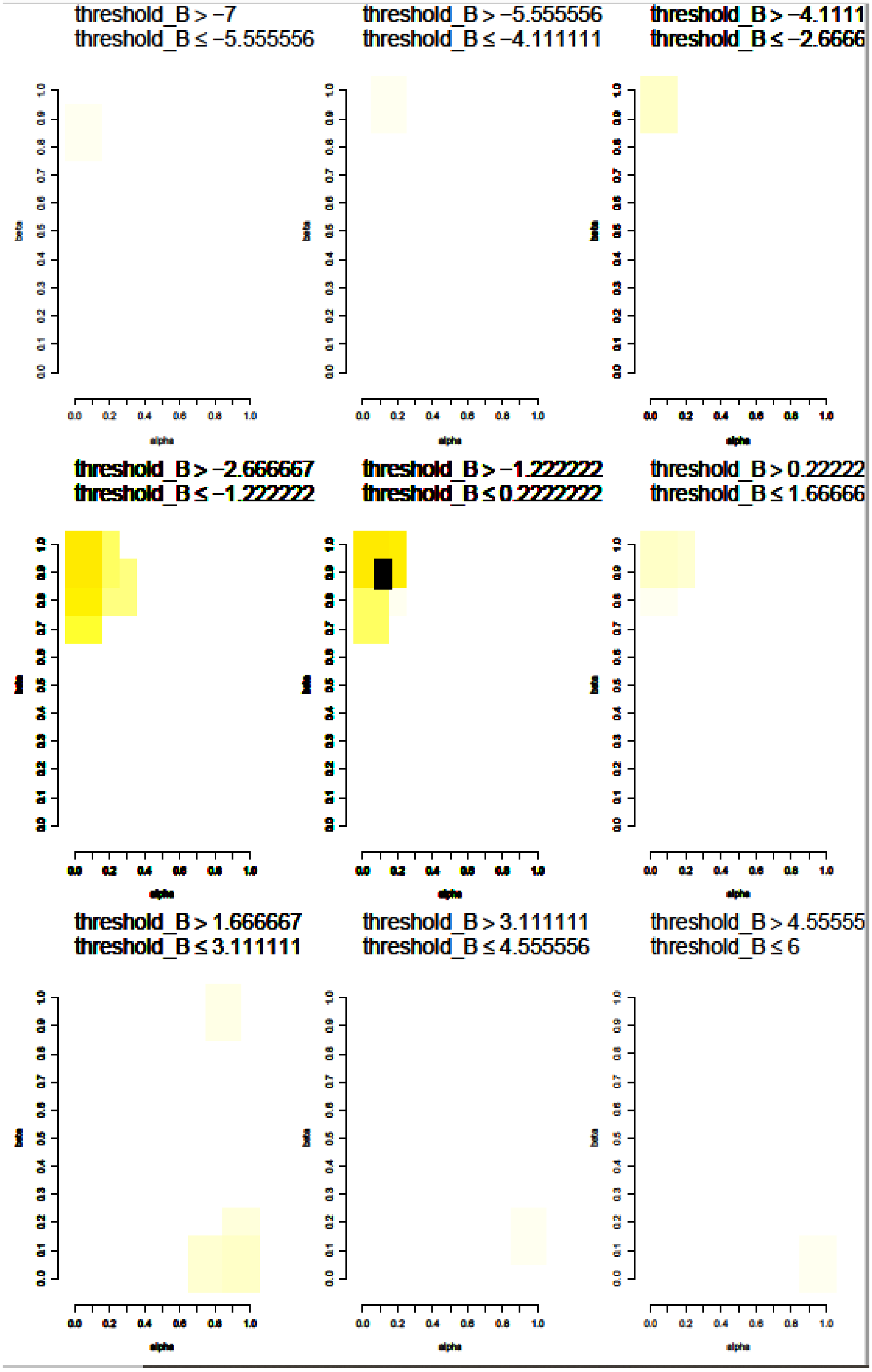


Figure 5(iv) $p_A = 0.1$, $p_B = 0.9$. Domain-general cognition

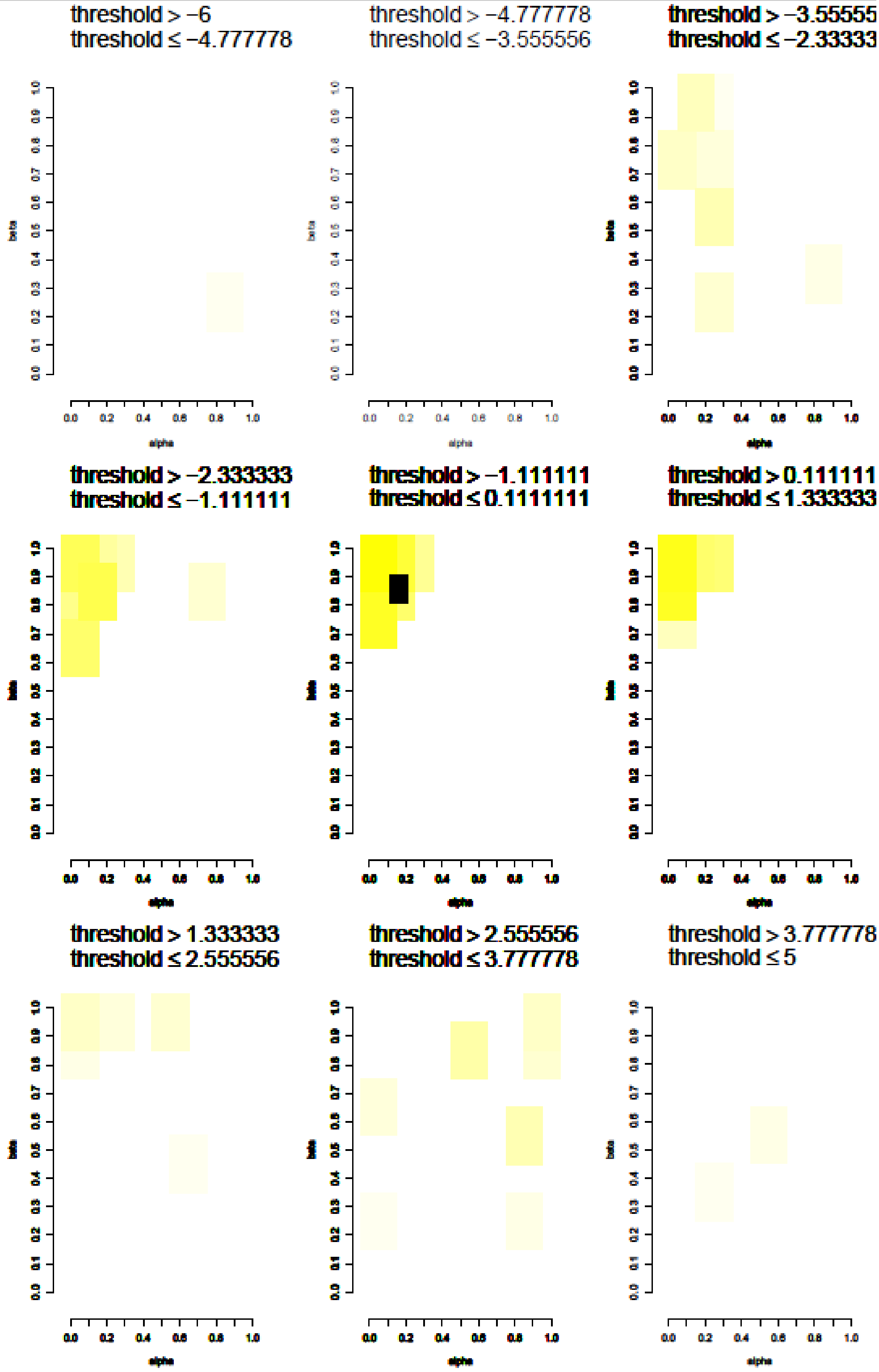


Figure 5. The binned distribution heatmaps displaying the psychological architecture of the final generation's phenotype for (i) fully modular agents; (ii) partly modular agents with modular motivation; (iii) partly modular agents with modular cognition

and (iv) domain-general agents. These figures give the result for runs with inconsistent prior probabilities of state 1 ($p_A = 0.1$, $p_B = 0.9$). Note that these heatmaps represent cases with strong benefits to cooperation (b=4), though see appendix 5 for runs where b=2.

3.3. Did the agents' psychological components affect their fitness?

It is my intuition that cognition and motivation coevolved to drive one's decisions to cooperate or defect, and so both components would affect the agent's fitness. I run a regression to reveal which psychological components predict fitness for the four agent types across the runs of interest (see Table 2).

The domain-general agents' cognitive and motivational thresholds affected their fitness across all runs (see Table 2). When psychological systems had to stay flexible to a range of domains, then any adjustments in these domain-general systems were likely to have fitness consequences in both domains. The modular agents did not have all of their psychological components evolve to drive their fitness and this in itself could be a design feature of modularity. Take agents with modular motivation on runs where $p_A = 0.1$ and $p_B = 0.1$, for example (Table 2i). Here, only β_A and β_B evolved to drive fitness. As the agent was unlikely to believe that the state was 1 ($s_A = 1$ or $s_B = 1$), then they were unlikely to use α_A and α_B, and so these values are unlikely to affect agent fitness. Likewise, take runs where $p_A = 0.9$ and $p_B = 0.9$ (Table 2ii). Here, only α_A and α_B evolved to drive fitness. As the agent was unlikely to believe that the state was 0 ($s_A = 0$ or $s_B = 0$), then they were unlikely to use β_A and β_B and so these were unlikely to affect fitness. In fact, this may imply that modularity was redundant in cases where the prior probabilities of environmental states were consistent in both domains.

Perhaps the real test of modularity comes from runs with skewed but inconsistent prior probabilities of state 1 in the two domains ($p_A = 0.1$ and $p_B = 0.9;$

Table 2iii*)*. This is because modular agents have cognitive and/or motivational systems that could track differing priors.

On these runs, partly modular agents had their modular psychological architecture evolve to compensate for any domain-general psychological architecture. For example, agents with modular motivation but domain-general cognition could not balance their cognitive thresholds to the inconsistent priors of both domains, and so most of their modular motivational thresholds (α_A, α_B, β_A, β_B) evolved to drive fitness (α_A estimate = -0.26, $p<0.001$; α_B estimate = -0.30, p = 0.01; β_A estimate = 0.21, p=0.005 and β_B estimate =0.11, p = 0.11). Agents with modular cognition but domain-general motivation could not balance their motivational thresholds to the inconsistent priors in the two domains, and so both their cognitive thresholds evolved to drive fitness. Interestingly, a smaller T_A (estimate = -0.07, $p<0.001$) and larger T_B (0.09, $p<0.001$) predict fitness. This is perhaps opposite to what one would intuitively expect, as this meant that the agents needed less evidence to believe in a rare event and more evidence to believe in a common event. However, the domain-general motivation still evolved to compensate for these cognitive biases and favour defection (Figure 5) and thus mutual defection is upheld as an outcome in these partly modular agents.

Interestingly, fully modular agents never had all six of their psychological architecture evolve to drive fitness (T_A, T_B, α_A, α_B, β_A, β_B), even on runs with inconsistent prior probabilities of environmental states in the two domains ($p_A = 0.1$, $p_B = 0.9$; Table 2iii). Taken together, these findings highlight that there was a larger difference between a domain-general agent and a partly-modular agent, than there was between a partly-modular agent and a fully-modular agent. That is, once the organism has modularity in either their cognitive or motivational system, there is perhaps little benefit to acquiring modularity in the other system. It seems that allowing both *modular* cognition and *modular* motivation to coevolve was redundant.

Finally, the domain-general agents have interesting results on runs with skewed but disparate environmental state distributions (Table 2iii; $p_A = 0.1, p_B = 0.9$). In this case, larger α (estimate = 0.49, p<0.001) and smaller β values (estimate = -0.48, p<0.001) predicted fitness. Thus, agents had a desire to hunt or build more when they believed that hunting or building more was cooperative. They also had a desire to hunt or build less when they believed that hunting or building less was cooperative. Figure 5iv confirms that fully domain-general agents do not have biased cognition. It thus seems that they were evolving towards preferring mutual cooperation.

This point becomes clearer when I consider the behavioural outcomes of each agent type in Figure 2iii (section 3.1). Domain-general agents partake in mutual cooperation, and no other agents do. These findings show that mutual cooperation could invade mutual defection as a stable behavioural strategy given that (i) the benefits of cooperation were sufficient *(b=4)* and (ii) the agents were domain-general, meaning that they had to balance their cognition and motivation to the demands of multiple environments which gave inconsistent prior probabilities. Cooperation may arise as a trade-off from domain-general systems that cannot specialise to multiple domains.

Table 2. The regression results displaying the psychological components that predict fitness for each of the four agent types in (i) environments with skewed but consistent prior probabilities such as ($p_A = 0.1$, $p_B = 0.1$) and (ii) ($p_A = 0.9$, $p_B = 0.9$) and (iii) environments with inconsistent prior probabilities of state 1 in the two domains ($p_A = 0.1$, $p_B = 0.9$). The results in italic text depict runs where the benefits of cooperation were weak (b=2), whilst the results in non-italic text depict runs where the benefits of cooperation were strong (b=4). Agents with domain-general cognition have the same

threshold in domain B as they did in domain A, and agents with domain-general motivation have the same α and β motivational thresholds in domain B as they do in domain A. For this reason, I arbitrarily block out domain B and focus on domain A for these agent types.

Table 2i) environments with skewed but consistent prior probabilities (p_A =0.1, p_B =0.1).

Psychological component left to evolve	**Agent Types**			
	Fully modular Agents	**Modular motivation agents**	**Modular Cognition agents**	**Domain-general agents**
Intercept	*2.92 **** *(0.10)* 3.81 *** (0.15)	*2.99 **** *(0.04)* 3.85 *** (0.10)	*2.93 **** *(0.07)* 3.69 *** (0.13)	*2.91 **** *(0.07)* 3.60 *** (0.12)
Threshold A	*0.0009* *(0.003)* -0.025 ** (0.008)	*-0.02 **** *(0.003)* -0.05 *** (0.09)	*-0.003* *(0.003)* -0.02 ** (0.007)	*-0.01 **** *(0.003)* -0.05 *** (0.009)
Threshold B	*-0.007 ** *(0.003)* -0.018 * (0.009)		*-0.005 ** *(0.002)* 0.0002 (0.007)	
***α* motivation A**	*-0.02* *(0.02)* -0.25 *** (0.05)	*-0.07 **** *(0.02)* -0.27 *** (0.05)	*-0.14 **** *(0.01)* -0.40 *** (0.05)	*-0.17 **** *(0.013)* -0.38 *** (0.42)
***α* motivation B**	*-0.10 **** *(0.01)* -0.18 *** (0.05)	*-0.08 **** *(0.02)* -0.20 *** (0.05)		
***β* motivation A**	*0.05* *(0.05)* 0.10 (0.09)	*0.09 ** *(0.04)* 0.06 (0.08)	*0.22 *** *(0.07)* 0.20 (0.14)	*0.26 **** *(0.08)* 0.25 . (0.13)
***β* motivation B**	*0.16 .* *(0.10)* 0.022 (0.11)	*0.08 ** (0.04) 0.05 (0.10)		

The asterisks denote the significance of our p values, with the following keys:
**** (<0.001)*
*** (0.01)*
** (0.05)*
· (trend: 0.05 – 0.10 significance)

Table 2ii) environments with skewed but consistent prior probabilities in both domains (p_A =0.9, p_B =0.9).

Psychological component left to evolve	**Agent Types**			
	Fully modular Agents	**Modular motivation agents**	**Modular Cognition agents**	**Domain-general agents**
Intercept	*3.01 **** *(0.01)* 3.43 *** (0.03)	*3.00 **** *(0.01)* 3.55 *** (0.03)	*3.02 **** *(0.006)* 3.48 *** (0.02)	*3.00 **** *(0.006)* 3.48 *** (0.02)
Threshold A	*0.004* *(0.003)* -0.004 (0.08)	*0.01 **** *(0.003)* 0.05 *** (0.01)	*0.006 ** *(0.002)* 0.01 * (0.01)	*0.009 **** *(0.002)* 0.04 *** (0.001)
Threshold B	*0.003* *(0.003)* 0.04 *** (0.008)		*0.004 .* *(0.002)* 0.01 (0.01)	
***α* motivation A**	*-0.08 ** *(0.04)* 0.16 (0.10)	*-0.07 .* *(0.04)* -0.30 ** (0.11)	*-0.35 **** *(0.09)* 0.16 (0.47)	*-0.14 *** *(0.05)* -0.21 (0.18)
***α* motivation B**	*-0.06* *(0.04)* -0.18 (0.14)	*-0.04* *(0.04)* 0.09 (0.11)		
***β* motivation A**	*0.03 .* *(0.02)* 0.21 *** (0.05)	*0.09 **** *(002)* 0.12 * (0.05)	*0.13 **** *(0.01)* 0.42 *** (0.05)	*0.12 **** *(0.01)* 0.41 *** (0.04)
***β* motivation B**	*0.07 **** *(0.02)* 0.36 *** (0.05)	*0.06 **** *(0.02)* 0.33 *** (0.05)		

The asterisks denote the significance of our p values, with the following keys:
**** (<0.001)*
*** (0.01)*
** (0.05)*
· (trend: 0.05 – 0.10 significance)

Table 2iii) environments with skewed but inconsistent prior probabilities of state 1 in both domains (p_A =0.1, p_B =0.9).

Psychological component left to evolve	**Agent Types**			
	Fully modular Agents	**Modular motivation agents**	**Modular Cognition agents**	**Domain-general agents**
Intercept	*3.02 *** (0.04)* 3.54 *** (0.10)	*3.02 *** (0.03)* 3.59 *** (0.08)	*3.09 *** (0.12)* 3.41 *** (0.17)	*2.90 *** (0.10)* 5.00 *** (0.11)
Threshold A	*-0.002 (0.003)* -0.010 (0.007)	*0.0003 (0.004)* -0.02 . (0.01)	*-0.03 *** (0.005)* -0.07 *** (0.01)	*-0.01 (0.005)* -0.02 (0.01)
Threshold B	*0.008 ** (0.003)* 0.04 *** (0.009)		*0.03 *** (0.006)* 0.09 *** (0.012)	
***α* motivation A**	*-0.06 *** (0.01)* -0.09 . (0.04)	*-0.08 *** (0.02)* -0.26 *** (0.06)	*-0.28 * (0.12)* -0.17 (0.17)	*-0.10 (0.09)* 0.49 *** (0.09)
***α* motivation B**	*-0.07 (0.48)* 0.08 (0.12)	*-0.14 *** (0.04)* -0.30 * (0.12)		
***β* motivation A**	*0.05 (0.04)* -0.02 (0.10)	*0.08 * (0.03)* 0.21 ** (0.08)	*0.13 (0.12)* 0.77 *** (0.18)	*0.28 ** (0.11)* -0.477 *** (0.11)
***β* motivation B**	*0.05 ** (0.02)* 0.39 *** (0.05)	*0.06 *** (0.02)* 0.11 (0.07)		

The asterisks denote the significance of our p values, with the following keys:
**** (<0.001)*
*** (0.01)*
** (0.05)*
· (trend: 0.05 – 0.10 significance)

3.4: Is there a difference in the fitness of the four agent types?

Section 3.3 addresses my first research aim regarding how cognition and motivation coevolve to affect agent fitness. I now address my second research aim by comparing the fitness of fully modular, partly modular, and domain-general agents (see table 3).

Table 3. The regression results displaying any differences in fitness of the four agent types in environments with skewed but consistent prior probabilities such as (p_A = 0.1, p_B = 0.1) and (p_A = 0.9, p_B = 0.9) and environments with inconsistent prior probabilities of state 1 in the two domains (p_A = 0.1, p_B = 0.9). The results in italic text depict runs where the benefits of cooperation were weak (b=2), whilst the results in non-italic text depict runs where the benefits of cooperation were strong (b=4). The domain-general agents were the omitted category of this regression.

Estimates for regression predicting fitness	**Prior Probabilities**		
	p_A = 0.1, p_B = 0.1	**p_A = 0.9, p_B = 0.9**	**p_A = 0.1, p_B = 0.9**
Intercept	*3.050 **** *(0.004)* 3.582 *** (0.013)	*3.037 **** *(0.004)* 3.575 *** (0.013)	*3.136 **** *(0.006)* 4.673 *** (0.015)
Modular cognition	*-0.007* *(0.006)* -0.020 (0.018)	*-0.002* *(0.006)* 0.018 (0.018)	*-0.013* *(0.008)* -0.768 *** (0.022)
Modular motivation	*-0.012 ** *(0.006)* 0.051 ** (0.018)	*-0.004* *(0.006)* 0.089 *** (0.018)	*-0.102 **** *(0.008)* -1.064 *** (0.022)
Fully modular	*-0.006* *(0.006)* 0.038 * (0.018)	*-0.002* *(0.006)* 0.048 ** (0.018)	*-0.103 **** *(0.008)* -1.077 *** (0.022)

The asterisks denote the significance of our p values, with the following keys:
**** (p<0.001)*
*** (p<0.01)*
** (p<0.05)*
· (trend: p=0.05 – 0.10 significance)

First, consider runs with skewed but consistent priors in the two domains (p_A = 0.1, p_B = 0.1 and p_A =0.9, p_B =0.9). Provided that the benefits of cooperation were strong (*b=4*), it can be seen that that fully modular agents and agents with modular motivation but domain-general cognition have significantly higher fitness than domain-general agents, both when the runs favoured state 0 ($p_A = 0.1$, $p_B = 0.1$; fully modular estimate = 0.04, p=0.03; modular motivation estimate = 0.05, p=0.003) and when the runs favoured state 1 ($p_A = 0.9$, $p_B = 0.9$; fully modular estimate =0.05, p=0.008; modular motivation estimate = 0.09, p<0.001). When the domains had consistently skewed priors, then having modular motivation seemed to predict fitness. I confirmed this with a secondary analysis in which any agent with modular motivation was found to outperform any agent with domain-general motivation during these runs with skewed but consistent priors (see appendix 6). Perhaps modular motivation was beneficial as it could have compensated for any cognitive biases that may have emerged in the agents across this run (see section 3.2).

Contrastingly, domain-general agents have the highest fitness when the domains had skewed but inconsistent priors ($p_A = 0.1$, $p_B = 0.9$) when the benefits of cooperation were strong ($b=4$) (fully modular estimate = -1.08, p<0.001; modular motivation estimate = -1.06, p<0.001; modular cognition estimate -0.77, p<0.001). This is perhaps unexpected, as modular agents could adjust their cognition and/or motivation to the inconsistent priors in both domains. In fact, this specialisation can actually explain why modular agents accrued less fitness. On runs where $p_A = 0.1$, $p_B = 0.9$, the modular agents specialised to the demands of both domains and upheld mutual defection. They hunted more when hunting more would be defection and built less when building more would be cooperation.

Domain-general agents could not track the fluctuating priors of multiple domains. The domain-general agents would therefore experience uncertainty (i.e., they were unsure whether hunting or building more would be cooperation, or defection). This uncertainty led to the agents choosing to cooperate as a mistake. Of course, provided that the benefits of cooperation were strong ($b=4$), then mutual cooperation could outperform mutual defection. Cooperation may arise as a 'big mistake' in domain-general agents, but it still pays off.

The fully modular agents and the agents with modular motivation were similar in terms of their behavioural outcomes (see Section 3.1) and phenotype (see Section 3.2). In order to investigate whether there were any meaningful differences in fitness between these two agents, I conducted a series of linear combinations on the predictors of the regression reported in Table 3 (see appendix 7). These linear combinations reveal that fully modular agents were often indistinguishable in terms of fitness from agents with modular motivation only. This suggests that agents may only need modular motivation when making cooperative (or uncooperative) decisions in two distinct domains. These linear combinations also revealed that the agents with modular motivation only always had a higher fitness than agents with modular cognition only. Thus, having modular motivation was consistently more influential in driving agent fitness than having modular cognition (see appendix 6 and 7). When it comes to the decision to cooperate (or defect), what we *want to do* is more important than what we *think we ought to do.*

3.5. Summary of the key results

For ease, I consider the previous results section as highlighting three key findings which shall now be addressed in the discussion:

1. Cooperation could emerge as a 'big mistake'. Specifically, the bar charts in Figure 2 – coupled with the fitness regressions reported in Table 3 – confirm that cooperation emerges as a 'mistake' in domain-general agents, who could not track inconsistent priors across multiple domains ($p_A = 0.1$, $p_B = 0.9$).

2. Motivation could evolve to compensate for cognitive biases. Here, I use the term cognitive bias to refer to runs where the cognitive thresholds adjusted in the opposite direction to what was expected based on the prior probabilities in both domains. For example, if state 1 was likely to be rare ($p_A = 0.1$, $p_B = 0.1$), then agents with *lower* cognitive thresholds would display a cognitive bias to believe in a rare event. Unexpected cognitive biases emerged in a few runs though motivation could evolve to compensate for such biases (as can be seen in the heatmaps in Figure 3).
3. Motivational thresholds may have affected agent fitness equally to, or more than, cognitive thresholds. Fully-modular agents may have been indistinguishable from partly-modular agents with modular motivation only. This suggests that agents do not have to be modular in both psychological components to uphold complex behavioural outcomes.

4. Discussion

This model investigated how (i) allowing cognition and motivation to coevolve and (ii) whether cognition and/or motivation were domain-general (i.e., flexible across many domains) or modular (i.e., specialised to each domain) affected the decision to cooperate when there was uncertainty over which behaviours were cooperative in two distinct domains. This model found that cooperation was only likely to emerge in the domain-general agents, who could not track the inconsistent priors in environmental states across multiple domains. When agents cannot specialise to the conflicting demands of multiple domains, then perhaps cooperation emerges as a ‘mistake’.

This finding has implications for the idea that cooperation results from a mismatch between the modern-day environment and the conditions of our ancestral past

(Burnham & Johnson, 2005; Price, 2008). This field of research– which Chudek et al. (2013) dub the 'mismatch hypothesis'– argues that we only show costly levels of cooperation today due to a mismatched cognitive bias resulting from the tight-knit groups that were common throughout the ancestral past. Some researchers argue that this selection pressure led to a cognitive bias favouring cooperation, which still drives our behaviour today even when cooperation is unlikely to be reciprocated (Delton et al., 2011; Tooby et al., 2006). My findings contradict this assumption. If cooperation is a 'big mistake' (Boyd & Richerson, 2006; Burnham & Johnson, 2005; Chudek et al., 2013), then it arises from a domain-general system.

Domain-general systems may be more likely to lead to mistaken behaviour than modular psychological systems, as domain-general systems cannot be infinitely flexible. There will be some trade-offs when the domain-general system has to balance demands from multiple contrasting domains. This may be similar to the drop in performance that individuals experience when multitasking as they take on more cognitive load (Örün & Akbulut, 2019; Vergauwe et al., 2010). The fact that costly cooperation has been interpreted as an evolved cognitive bias; though in this model, it emerged from a domain-general system, may suggest that the complexity of human cognition is hard to infer from observed cooperative behaviour alone.

The costly levels of one-shot cooperation have been taken by some researchers to evidence a cognitive bias favouring cooperation due to the conditions of our ancestral past (Delton et al., 2011; Krasnow & Delton, 2016). Rather than having a specialised cognitive bias to favour cooperation, my model highlighted that mutual cooperation arose in a domain-general agent who could not adjust to the contrasting demands of multiple conflicting domains ($p_A = 0.1$, $p_B = 0.9$). Mistaken cooperation occurs when domain-general psychology cannot specialise to the distinct demands of multiple domains rather than being due to a specific cognitive bias.

Another issue with the view that cooperation emerges from a cognitive bias is that it may underplay the role of motivation. In my model, motivation consistently predicted agent fitness above the influence of cognition (see Section 3.3 and 3.4). It is worth noting that Evolutionary Psychology is a theory of cognitive psychology. Of course, motivation is also important (Delton et al., 2011; Öhman & Mineka, 2001; Tooby et al., 2006) though cognitive processing and modularity are often discussed more in-depth. The view that cognitive modules drive motivational processing may be broadly consistent with the partly modular agents who had modular cognition but domain-general motivation in my model. However, these agents typically accrued less fitness than agents with modular motivation. In fact, the partly modular agents' modular cognitive thresholds would only adjust in runs with inconsistent prior probabilities of the environmental state in two domains ($p_A = 0.1$, $p_B = 0.9$). Cognitive biases may therefore only reliably emerge when motivation cannot track the demands of multiple, contrasting environments. This shows that, when it comes to cooperation, what we *want to do* is perhaps more important than *what we think we ought to do.*

Motivation could compensate for cognitive biases when the two coevolved. Take runs where state 1 was very unlikely, for example ($p_A = 0.1$, $p_B = 0.1$). It would intuitively make sense for an agent to require a higher degree of evidence to believe in a rare event. Despite this, the agents sometimes developed a smaller cognitive threshold – and thus needed less evidence to believe in a rare event. These agents with cognitive biases to believe in a rare event did not show markedly different behaviour from the unbiased agents, however. Most agents defected on runs with skewed but consistent prior probabilities.

This is because their motivation helped to compensate for cognitive biases. Agents who (erroneously) believed that the state was 1 had a high α motivational threshold and so were motivated to play 1. Although the agent may believe that she was cooperating, playing 1 is in fact more likely to clash with the actual environmental state of 0. For example: I may believe that the game is abundant where I live and so go hunting more, with the intention of sharing some of this meat. However, in reality the game is scarce and so hunting has removed a food source from my neighbours if I do not follow up on sharing with them. To illustrate with an even more terse – though perhaps more illustrative – example, I could hold the biased belief that aliens routinely visit me. If I am never motivated to act upon this strange belief however, then it will have little effect on my behaviour and so may be unproblematic.

If motivation can evolve to compensate for cognition, then a range of psychological architecture may be consistent with any behaviour. This has implications for EMT. For example, consider the sexual overperception bias. According to EMT, men may have evolved a cognitive bias to over-perceive female sexual interest as historically, the cost of slight embarrassment when being rejected was outweighed by the cost of a missed mating opportunity for ancestral men (Haselton & Buss, 2000; Haselton et al., 2015). Picture a man at a bar enthusiastically approaching many women and being rejected multiple times. This man may indeed have a sexual overperception bias as per the predictions of EMT. The man may be biased to approach women as he is overinterpreting the evidence that they are interested in himself. However, this behaviour is also consistent with a number of other psychological phenotypes. Perhaps this man actually has an accurate perception of the women in the bar and knows that they are unlikely to be interested in himself. As he does not want to miss out on a mating opportunity however, then he may be highly motivated to approach these women anyway. Thus, his motivation can compensate for cognitive biases.

Now, let us change the scenario from a bar to a workplace. The man in question does not approach the woman that he works with. Perhaps he is accurately interpreting her lack of interest in himself. Or perhaps he still has the sexual overperception bias; he does believe that the woman he works with is interested. However, the man may be very unmotivated to approach this woman due to the different constraints of this scenario. Of course, this makes inferring support for the sexual overperception bias based on behavioural data alone quite difficult (McKay & Efferson, 2010). Did the man approach the uninterested woman in the bar as he made a cognitive error to overperceive interest, or was he motivated to at least try? The same logic applies the other way. Did the man avoid approaching the woman in the workplace because he accurately assessed that she was not interested, or because he was unmotivated to make a sexual proposition in the public and formal scenario of the workplace? If this model's results will hold in other domains besides cooperation, then it can be difficult to infer the presence of cognitive and/or motivational biases based on agent behaviour alone.

If motivation can evolve to compensate for cognitive biases, then this may help agents to be flexible in their behaviour. This again has implications for EMT. For example, consider the base-rate fallacy in foraging. For prey animals, the cost of wasting some calories responding to a false alarm is greatly outweighed by the costs of failing to run away from a hidden predator. All else being equal, prey animals should have evolved to be hypersensitive to cues of predation and flee when in doubt. However, if the animal was currently starving then they may choose riskier grounds to forage as, in this case, the costs of dying of starvation outweigh the risk of predation (Autzen, 2017; Nettle, 2019). These state-dependent changes could be driven by fluctuating motivational signals from within the animal, as evolved cognitive biases may be harder to update in line with fluctuating situational demands.

Motivation could also correct for any cognitive biases that may favour cooperation. For example, one could hold the stable cognitive belief that cooperation is the correct thing to do. However, fluctuations in motivation may help the individual to compensate for certain situational factors. For example, the individual may be less motivated to cooperate when her resources are low, or when she believes that the other individual is unlikely to cooperate back (Chudek & Henrich, 2011; Delton et al., 2011; Evans & Rand, 2019).

Thus far, my findings have implications for the view that cooperation is a 'big mistake' (any such mistakes favouring cooperation likely arises from a domain-general system as opposed to the modular view emphasised in Evolutionary Psychology) and highlighted that some previous evolutionary theories may have overemphasised the importance of cognitive processing at the expense of motivational systems (Delton et al., 2011). A final issue that my results may address is the 'massive modularity' debate.

Whilst Evolutionary Psychologists see modular cognition as an important prerequisite for human behaviour (Cosmides & Tooby, 1994; Delton et al., 2001), other researchers argue that modularity is constrained to a few peripheral processing demands, such as colour perception (Fodor, 2001). Other researchers argue that the flexibility in human behaviour would be consistent with a fully domain-general psychology (Bolhuis et al., 2011). Of course, the scope of this debate is beyond this current paper. I do however observe that partly modular agents with modular motivation only were indistinguishable in terms of fitness from fully modular agents (see section 3.4). This perhaps offers a middle ground for researchers to consider partly modular agents between the extreme full modularity and full domain-generality viewpoints.

It is interesting to note that modularity was more important in motivation than cognition in my model. This somewhat contradicts the predictions of massive modularity in Evolutionary Psychology. This theory may underemphasise the importance of motivation in preference of cognitive modules (Cosmides & Tooby, 1994; Krasnow & Delton, 2016; Sperber & Mercier, 2018). This model does not seek to invalidate the wealth of evidence on evolved cognitive biases, but instead opens up a future avenue of research to consider the importance of evolved and specific motivational biases, too (see Tooby et al. [2006] for example).

It is interesting that it was the domain-general agents – who cannot track the contrasting pressures over two domains – that cooperate (Figure 2iii). When there is uncertainty over which behaviour is cooperative over multiple domains, then domain-general agents may cooperate by mistake. This may suggest that human psychology is domain-general, as we are cooperative across most societies (Henrich et al., 2001; Henrich & Muthukrishna, 2021). Moreover, this mutual cooperation allowed the domain-general agents to accrue more fitness than modular agent types (see Table 3). Perhaps domain-general agents may be able to out-compete modular agent types, provided that the benefits of mutual cooperation are sufficiently large.

Of course, this claim is purely speculative until this code is adapted so that the four agent types can co-exist (see appendix 3). If this model started with an unbiased distribution split fairly between the four agent types, then it would be interesting to see which agent type – if any – reached evolutionary dominance across multiple simulations. When runs have skewed but consistent prior probabilities of the environmental state ($p_A = 0.1$, $p_B = 0.1$ or $p_A = 0.9$, $p_B = 0.9$), then the four agent types are sufficiently similar. Perhaps drift, or random sampling error during reproduction (Rorabaugh, 2014), would have more of an outcome on which agent type became evolutionary dominant across such runs. When the environment has inconsistent priors however ($p_A = 0.1$, $p_B = 0.9$), then domain-general agents can become evolutionary dominant. This is because domain-general agents cooperate (initially by mistake). If they can then assort into likeminded groups however, then mutual cooperation becomes a much more successful strategy than mutual defection and so can become evolutionary dominant (Gintis, 2000; Ihara, 2011; van Veelen et al., 2012). To investigate this speculative claim, future research must redesign this model to accommodate factors such as assortment and group competition.

Other factors could be added to this baseline model to understand which constraints, if any, lead to mutual cooperation evolving reliably in all four agent types. Factors such as genetic relatedness (Trivers, 1971), repeated interactions (Delton et al., 2011) and reputation (Price, 2008) may underlie the emergence of cooperation but none of these factors were considered in this model. I instead wished to focus on the novel aspect of cooperation in multiple domains, though future work could adjust this code to investigate such parameters.

For example, this code could be adjusted so that all agents play another round of a prisoner's dilemma with a probability, ω. If EMT is supported, then one-shot cooperation would be expected to emerge (Delton et al., 2011; though see Zimmerman & Efferson [2017] for a rebuttal that this depends on making restricted and arbitrary assumptions about the complexity of agent psychology). I also did not allow agents to track what others in the group thought about them and update their strategies accordingly, though reputation may be a factor that was important to the emergence of cooperation during the ancestral past (Bhui et al., 2019; Panchanathan & Boyd, 2004; Price, 2006; Swakman et al., 2016).

To summarise, I investigated how (i) allowing cognition and motivation to coevolve and (ii) whether cognition and/or motivation were domain-general (i.e., flexible across many domains) or modular (i.e., specialised to each domain) affected the emergence of cooperation. If cooperation does arise as a 'mistake', then this only occurs in domain-general agents who must flexibly balance their psychology to the contrasting demands of multiple domains. Care must be taken when inferring modularity from behavioural performance. Moreover, fully modular agents may be difficult to distinguish from partly modular agents with modular motivation only. Finally, motivation drove behaviour and compensated for cognitive biases throughout. Perhaps previous research has focused on the influence of cognitive biases in driving behaviour at the expense of motivation. When it comes to cooperation, what we *want to do* may be more important as what we *think we ought to do.*

5. References

Autzen, B. (2017). Error management, reliability and cognitive evolution. *Biology & philosophy*, *32*(6), 935-950. DOI: https://doi.org/10.1007/s10539-017-9583-1.

Bear, A., Kagan, A., & Rand, D. G. (2017). Co-evolution of cooperation and cognition: the impact of imperfect deliberation and context-sensitive intuition. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1851), 20162326. DOI:https://doi.org/10.1098/rspb.2016.2326.

Bhui, R., Chudek, M., & Henrich, J. (2019). How exploitation launched human cooperation. *Behavioral Ecology and Sociobiology*, *73*(78), 1-14. DOI: https://doi.org/10.1007/s00265-019-2667-y.

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biol*, *9*(7), e1001109. https://doi.org/10.1371/journal.pbio.1001109.

Boyd, R., & Richerson, P. J. (2006). Solving the puzzle of human cooperation. In S.C. Levinson and P. Jaisson (Eds.) *Evolution and culture: A Fyssen foundation system*, pp 105-132. England: MIT Press.

Boyd, R., Richerson, P.J. & Henrich, J. (2011). Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioural Ecology and Sociobiology*, *65*, 431–444. https://doi.org/10.1007/s00265-010-1100-3.

Burke, D. (2014). Why isn't everyone an evolutionary psychologist? *Front. Psychol*. *5*(910), 1-8. doi: 10.3389/fpsyg.2014.00910.

Burnham, T. C., & Johnson, D. D.P. (2005). The biological and evolutionary logic of human cooperation. *Analyse & Kritik*, *27*(1), 113-135. https://doi.org/10.1515/auk-2005-0107.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*(5), 218-226. DOI: https://doi.org/10.1016/j.tics.2011.03.003.

Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation, and the shaping of human social psychology. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 425–457). The MIT Press: Cambrdige, UK.

Colman, A. M. (2006). The puzzle of cooperation. *Nature*, *440*(7085), 744-745. DOI: https://doi.org/10.1038/440744b.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. Hirschfeld and S.Gelman (Eds.) *Mapping the mind: Domain specificity in cognition and culture* (pp. 84-116). New York: Cambridge University Press.

Dawkins, R. (1976). *The selfish gene*. UK: Oxford University Press.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, 108(32), 13335-13340. DOI: https://doi.org/10.1073/pnas.1102131108.

Delton, A. W., Nemirow, J., Robertson, T. E., Cimino, A., & Cosmides, L. (2013). Merely opting out of a public good is moralized: An error management approach to cooperation. *Journal of Personality and Social Psychology, 105*(4), 621–638. https://doi.org/10.1037/a0033495.

Efferson, C., McKay, R., & Fehr, E. (2020). The evolution of distorted beliefs vs. mistaken choices under asymmetric error costs. *Evolutionary Human Sciences, 2,* E27. doi:10.1017/ehs.2020.25.

Evans, A. M., & Rand, D. G. (2019). Cooperation and decision time. *Current opinion in psychology*, *26*, 67-71. https://doi.org/10.1016/j.copsyc.2018.05.007.

Fehr, E. & Fischbacher, U. (2005). Human Altruism – Proximate Patterns and Evolutionary Origins. *Analyse & Kritik, 27*(1), 6-47. https://doi.org/10.1515/auk-2005-0101.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature*, *13*(1), 1-25.

Fodor, J. (2001). Evolution and the human mind: Modularity, language and meta-cognition. *British Journal of Philisophical Sciences, 52*(3), 623-628.

Gintis, H. (2000). Group selection and human prosociality. *Journal of Consciousness studies*, *7*(1-2), 215-219.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and human Behavior, 24*(3), 153-172. https://doi.org/10.1016/S1090-5138(02)00157-5.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78*(1), 81–91. https://doi.org/10.1037/0022-3514.78.1.81.

Haselton, M.G., Nettle, D. & Murray, D.R. (2015). The evolution of cognitive bias. In D.M. Buss (Ed.) *Handbook of Evolutionary Psychology* (pp 968-987). Wiley, United Kingdom. DOI: https://doi.org/10.1002/9781119125563.evpsych241.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review, 91*(2), 73-78. DOI: 10.1257/aer.91.2.73.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogyl, C., Marlowe, F., Tracer, D. & Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767-1770. DOI: 10.1126/science.1127333.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes C. (2016). Blackboxing: social learning strategies and cultural evolution. *Philosophical Transactions of the Royal Society B, 371* 20150369. http://dx.doi.org/10.1098/rstb.2015.0369.

Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature*, *13*(1), 105-128. https://doi.org/10.1007/s12110-002-1016-3.

Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*(12), 3999-4002. https://doi.org/10.1073/pnas.0308738101.

Ihara, Y. (2011). Evolution of culture-dependent discriminate sociality: a gene–culture coevolutionary model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 889-900. https://doi.org/10.1098/rstb.2010.0247.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Krasnow, M. M., & Delton, A. W. (2016). Are humans too generous and too punitive? Using psychological principles to further debates about human social evolution. *Frontiers in psychology*, *7*, 799. https://doi.org/10.3389/fpsyg.2016.00799.

Laland, K. N. (1993). The mathematical modelling of human culture and its implications for psychology and the human sciences. *British Journal of Psychology*, *84*(2), 145-169. DOI: https://doi.org/10.1111/j.2044-8295.1993.tb02471.x.

Lu, F. & Wirth, C. (2011). Conservation Perceptions, Common Property, and Cultural Polarization among the Waorani of Ecuador's Amazon. *Human Organization*, *70*(3), 233-243. DOI: https://doi.org/10.17730/humo.70.3.l3504v21wrgt2607.

McKay, R., & Efferson, C. (2010). The subtleties of error management. *Evolution and human behavior*, *31*(5), 309-319. https://doi.org/10.1016/j.evolhumbehav.2010.04.005.

Molleman, L., Kölle, F., Starmer, C., & Gächter, S. (2019). People prefer coordinated punishment in cooperative interactions. *Nature human behaviour*, *3*(11), 1145-1153. https://doi.org/10.1038/s41562-019-0707-2.

Nettle, D. (2019). State-dependent cognition and its relevance to cultural evolution. *Behavioural processes*, *161*, 101-107. DOI: https://doi.org/10.1016/j.beproc.2018.01.018.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483.

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, *92*, 216-229. https://doi.org/10.1016/j.chb.2018.11.027.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, *432*(7016), 499-502. https://doi.org/10.1038/nature02978.

Pietraszewski, D., & A. E., Wertz (2021). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science* DOI: 10.1177/1745691621997113. [preprint].

Power, M. L., & Schulkin, J. (2013). *The evolution of obesity*. US: John Hopkins University Press.

Price, M. E. (2006). Monitoring, reputation, and 'greenbeard'reciprocity in a Shuar work team. *Journal of Organizational Behavior,*, *27*(2), 201-219. DOI: 10.1002/job.347.

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Social Justice Research*, *21*(2), 228-240. DOI: https://doi.org/10.1007/s11211-008-0064-1.

Rorabaugh, A. N. (2014). Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *Journal of Archaeological Science*, *49*, 255-264. https://doi.org/10.1016/j.jas.2014.05.016.

Safin, V., Arfer, K. B., & Rachlin, H. (2015). Reciprocation and altruism in social cooperation. *Behavioural processes*, *116*, 12-16. https://doi.org/10.1016/j.beproc.2015.04.009.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and society*, *7*(1), 58-92. DOI: https://doi.org/10.1177/1043463195007001004.

Sperber, D., & Mercier, H. (2018). Why a modular approach to reason?. *Mind & Language*, *33*(5), 533-541. DOI: 10.1111/mila.12208.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nature Reviews Neuroscience*, *18*(9), 559-567. DOI: https://doi.org/10.1038/nrn.2017.76.

Swakman, V., Molleman, L., Ule, A., & Egas, M. (2016). Reputation-based cooperation: empirical evidence for behavioral strategies. *Evolution and Human Behavior, 37*(3), 230-235. DOI: https://doi.org/10.1016/j.evolhumbehav.2015.12.001.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for n-person exchange: the evolutionary roots of organizational behavior. *Managerial and Decision Economics*, *27*(2-3), 103-129. https://doi.org/10.1002/mde.1287.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology*, *46*(1), 35-57. https://doi.org/10.1086/406755.

Van den Berg, P., Molleman, L., & Weissing, F. J. (2012). The social costs of punishment. *Behavioral and Brain Sciences, 35*(1), 42-43. DOI:10.1017/S0140525X11001348.

Van Veelen, M., García, J., Rand, D. G., & Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences, 109*(25), 9929-9934. https://doi.org/10.1073/pnas.1206694109.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource?. *Psychological science*, *21*(3), 384-390. https://doi.org/10.1177/0956797610361340.

Vogt, S., Efferson, C., Berger, J., & Fehr, E. (2015). Eye spots do not increase altruism in children. *Evolution and Human Behavior*, *36*(3), 224-231. DOI: https://doi.org/10.1016/j.evolhumbehav.2014.11.007.

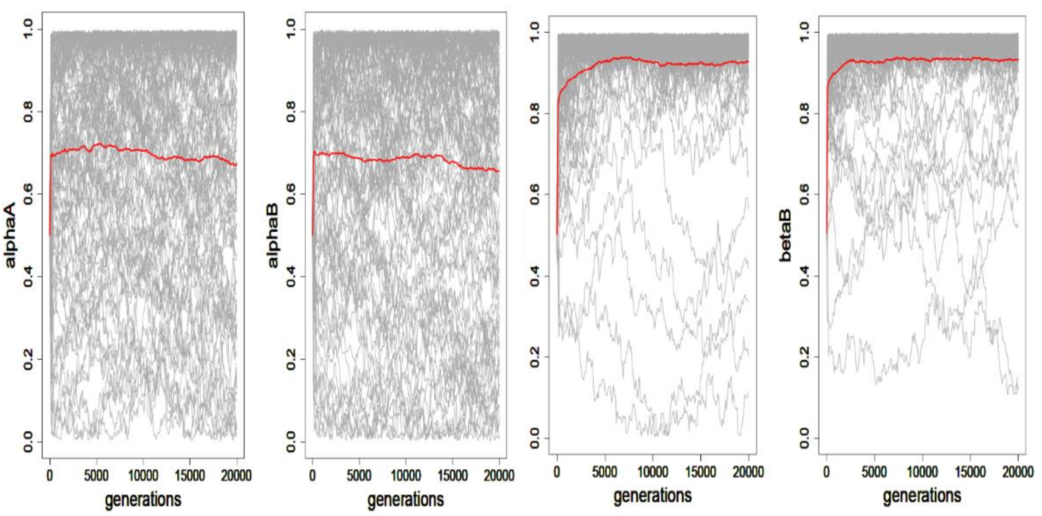
Zimmerman, J. & Efferson, C. (2017). One-shot reciprocity under error management is unbiased and fragile. *Evolution and Human Behaviour*, *38*(1), 39-47. https://doi.org/10.1016/j.evolhumbehav.2016.06.005.

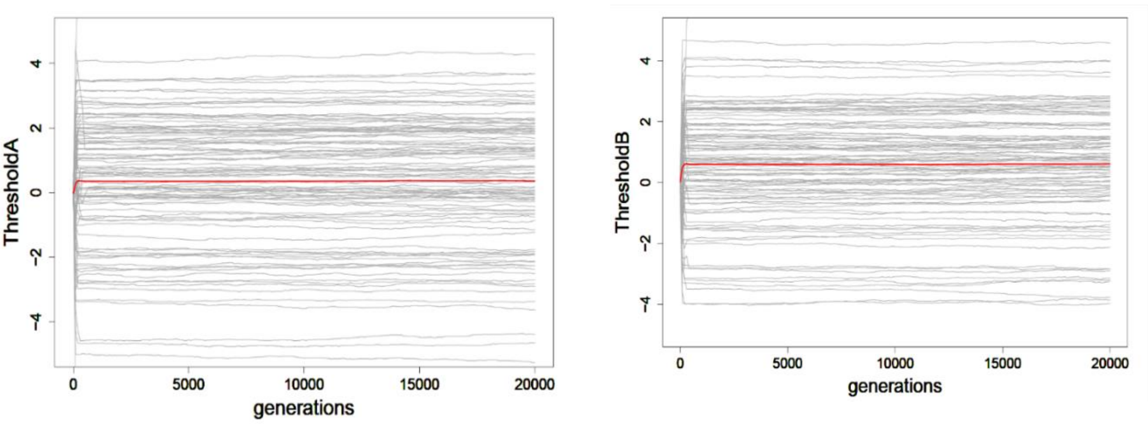
6. Appendices

Appendix 1: The line graphs confirming that these runs converged on a stable psychological architecture.

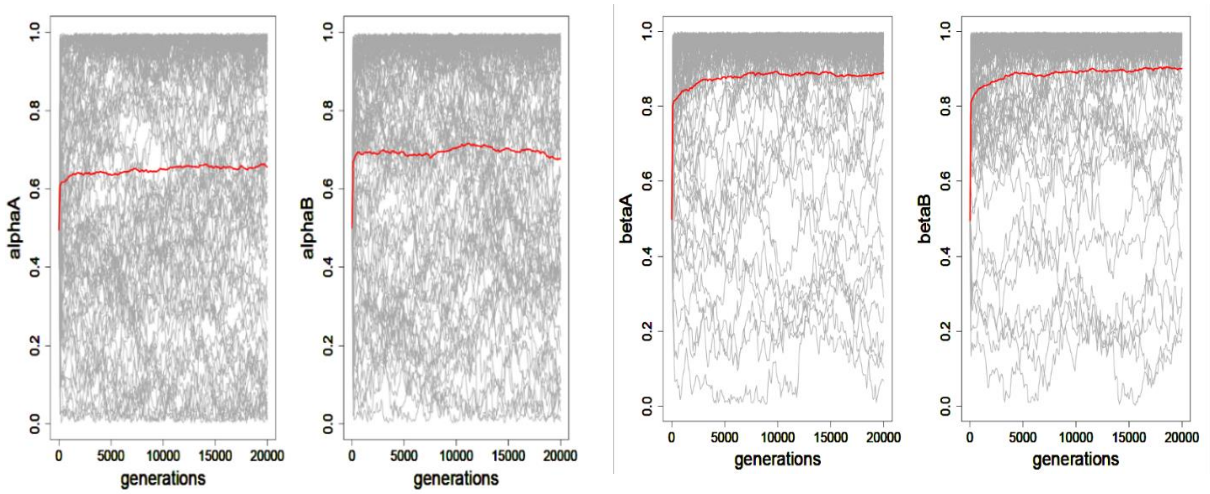
Appendix 1A: The line graphs for runs where p_A =0.1 and p_B =0.1

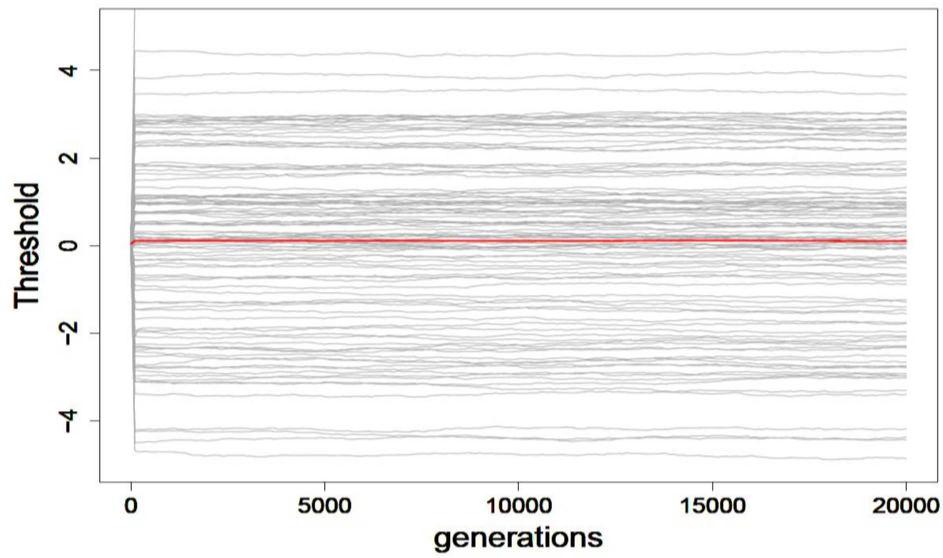
i) Fully modular agents



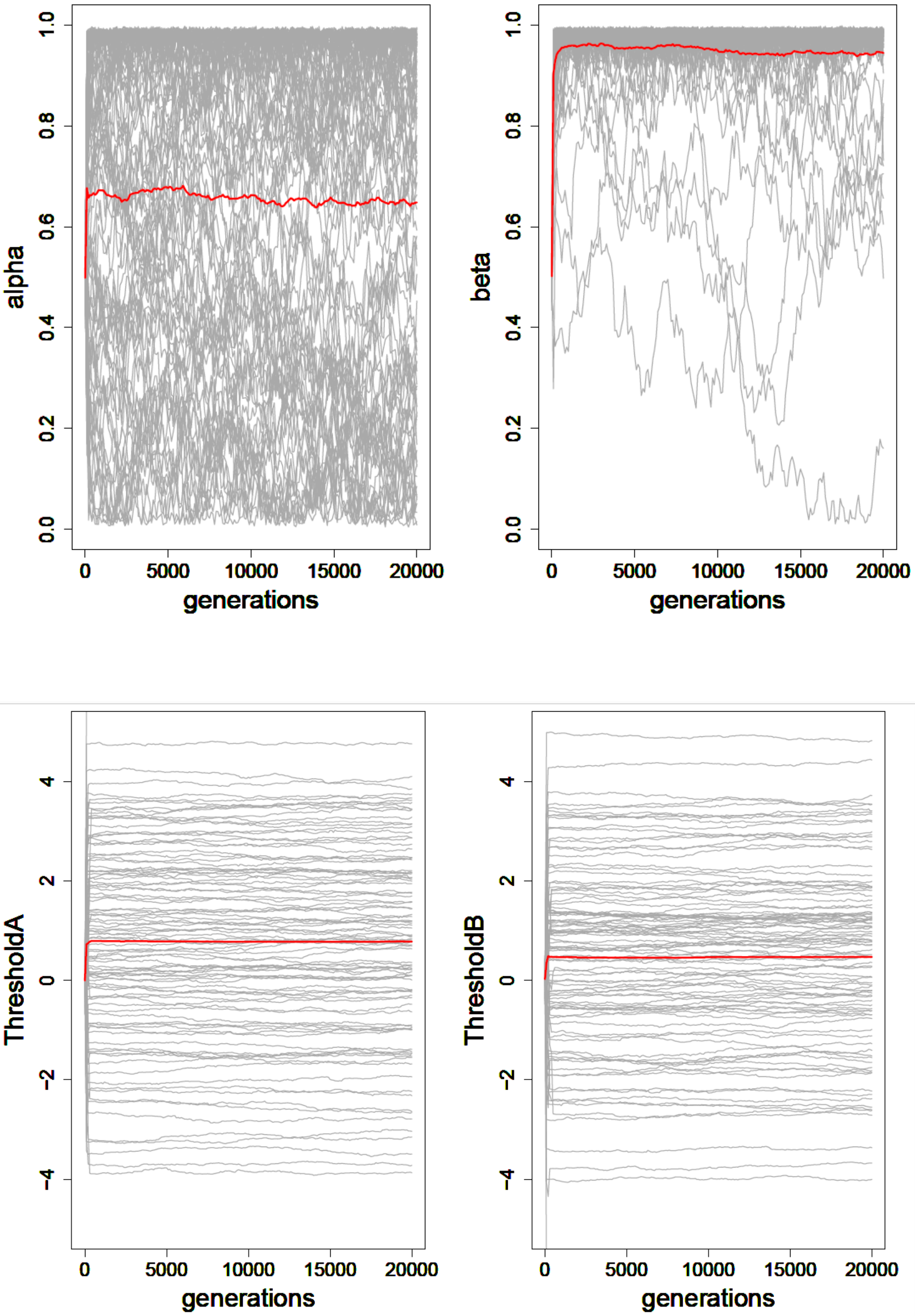


ii) Agents with modular motivation only

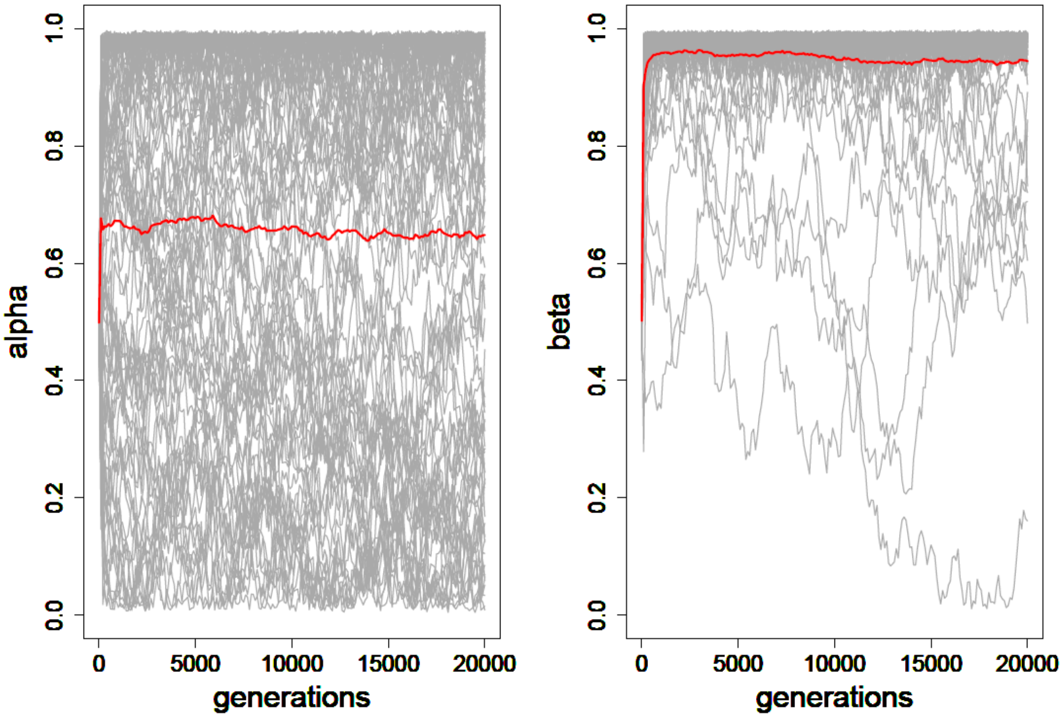


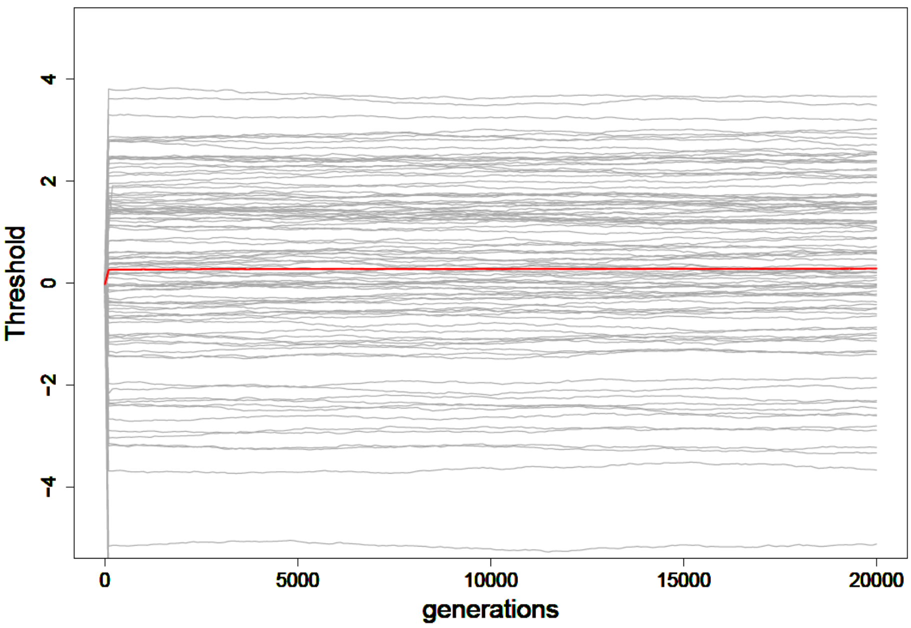


iii) Partly modular agents with modular cognition only



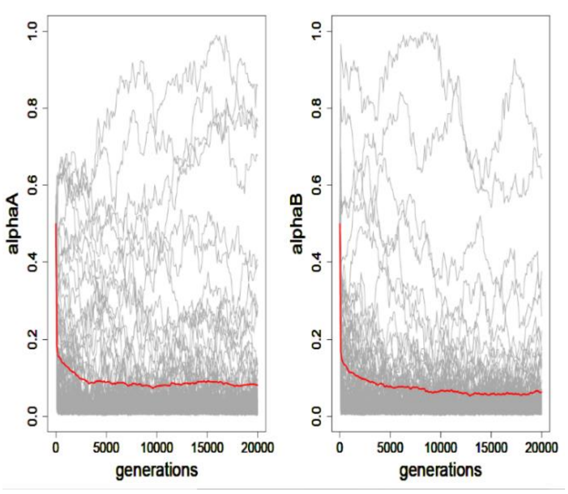
iv) Domain-general agents

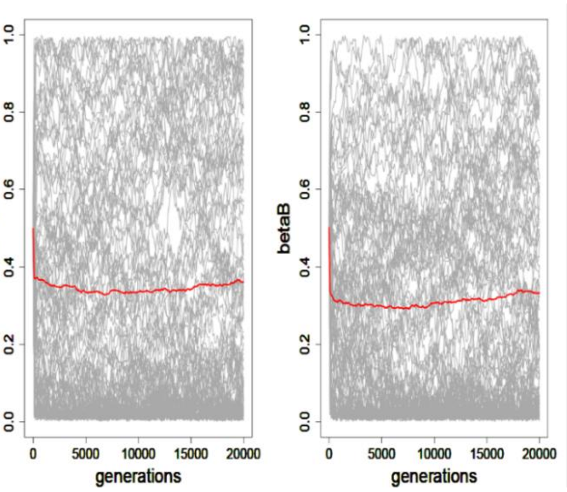


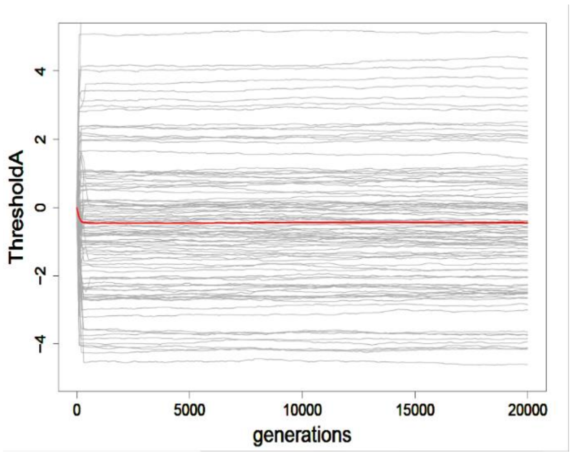


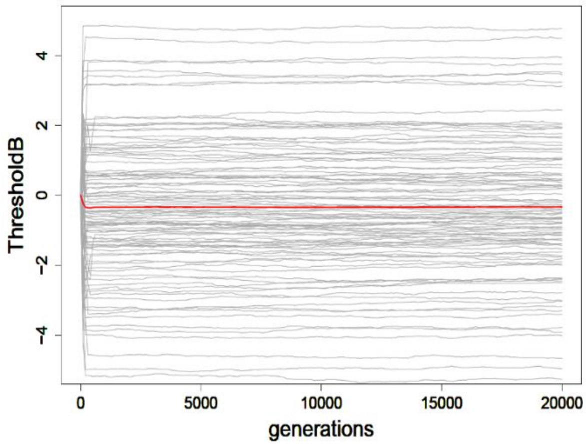
Appendix 1B: The line graphs for runs where p_A =0.9 and p_B =0.9

i) Fully modular agents

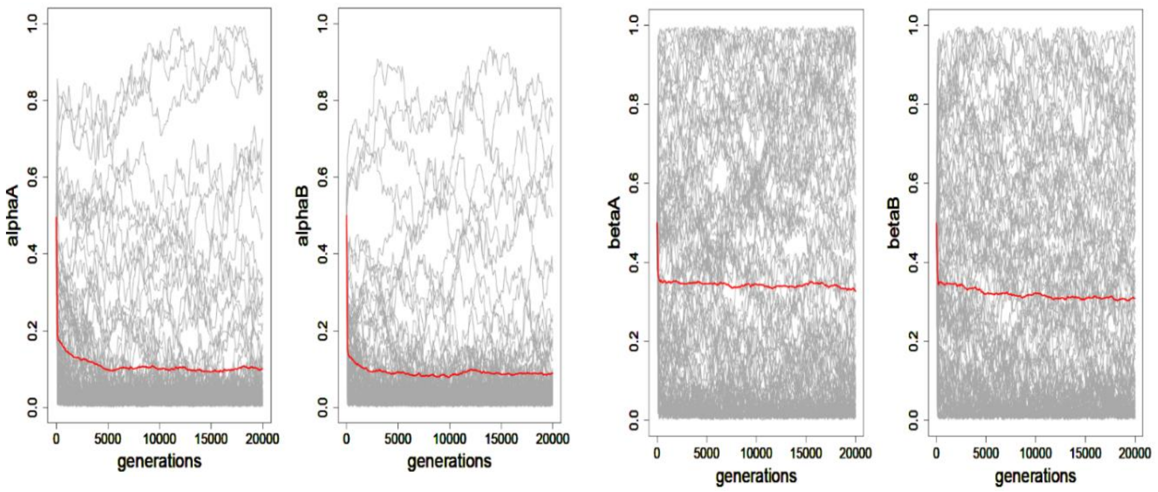


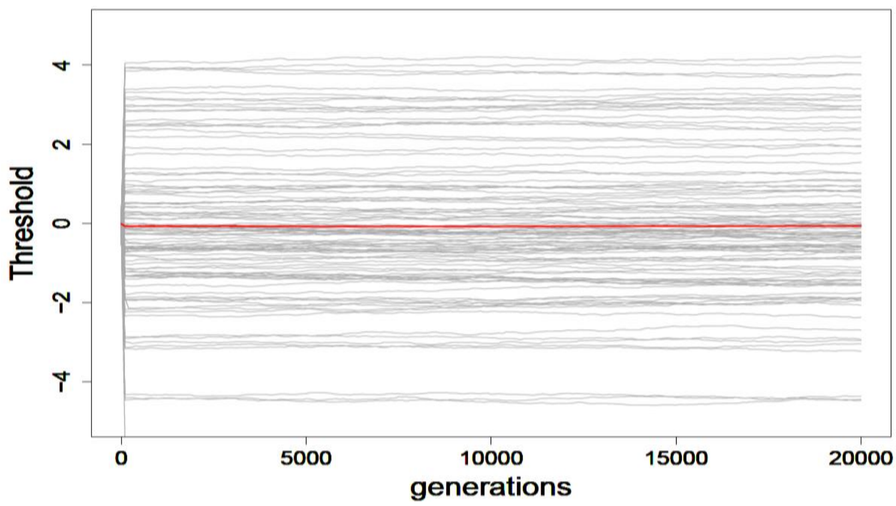




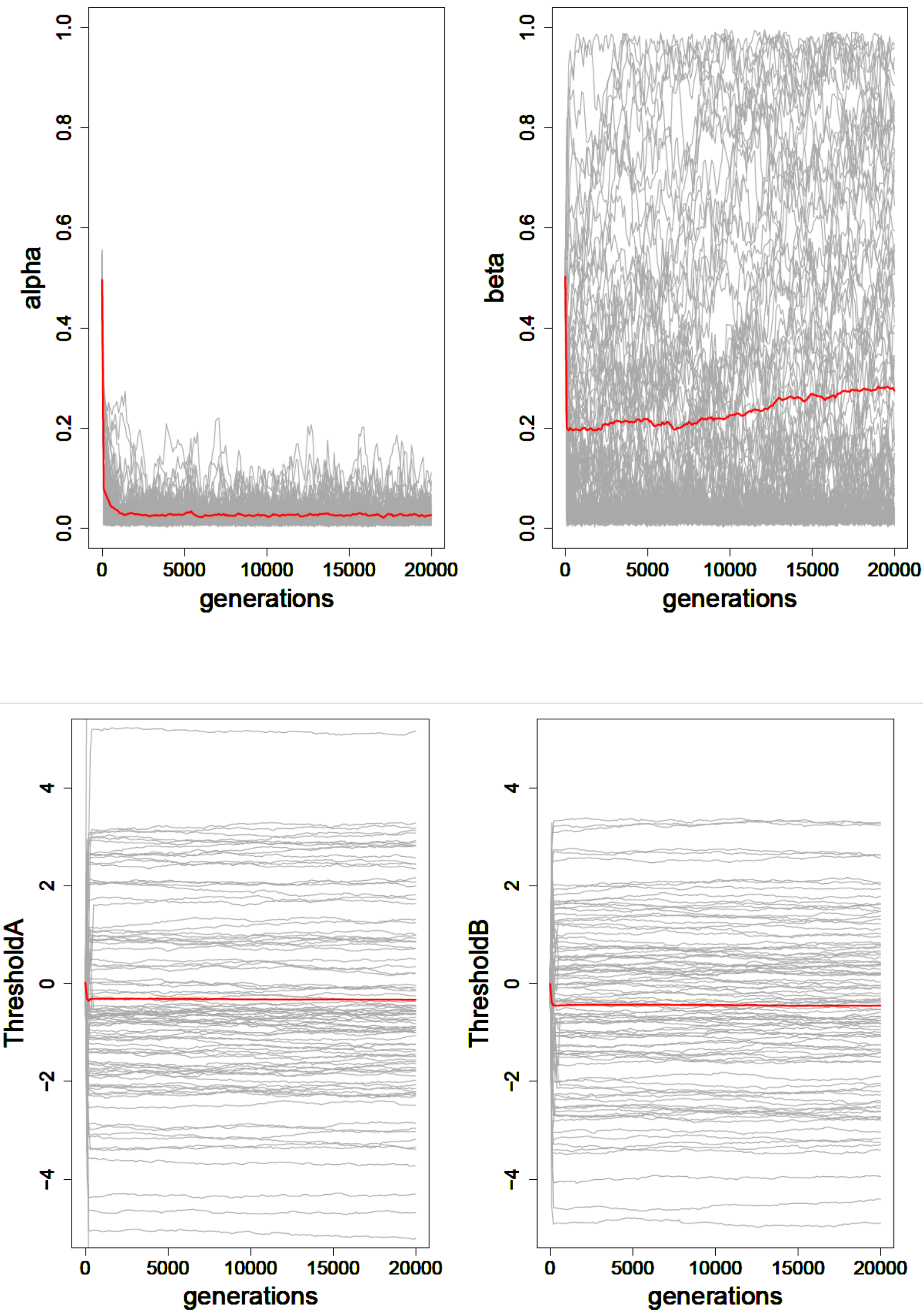


ii) Partly modular agents with modular motivation

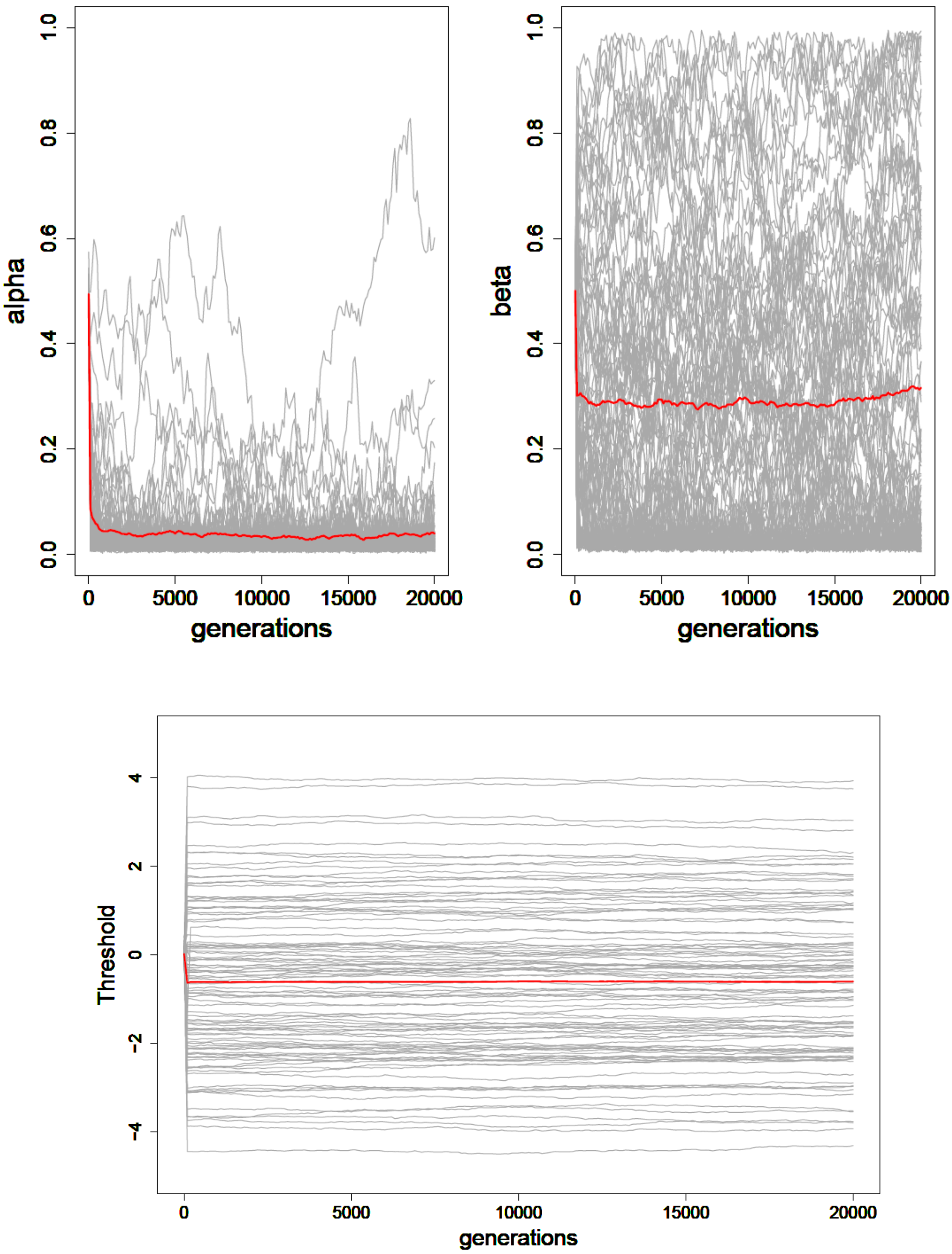




iii) Partly modular agents with modular cognition only

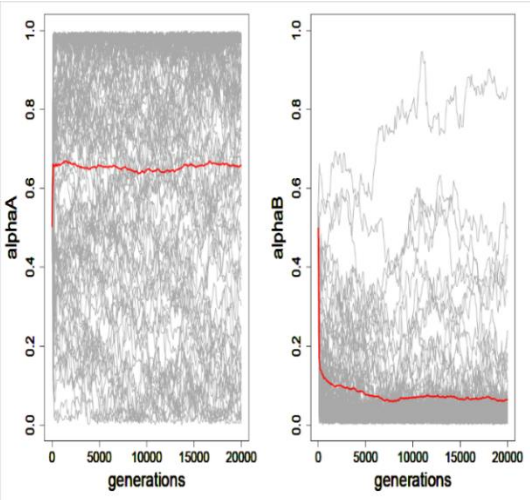


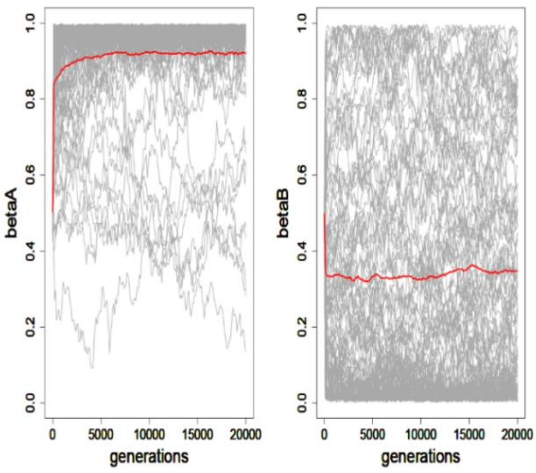
iv) Domain-general agents

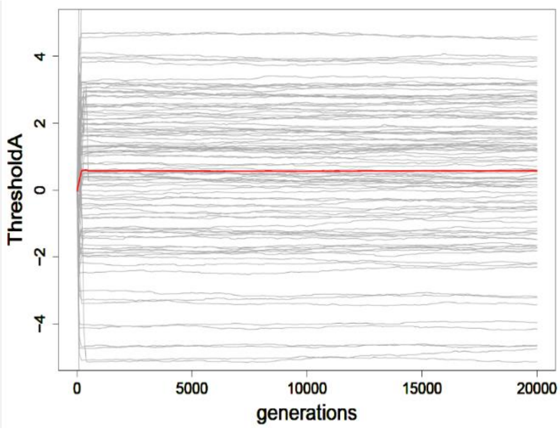


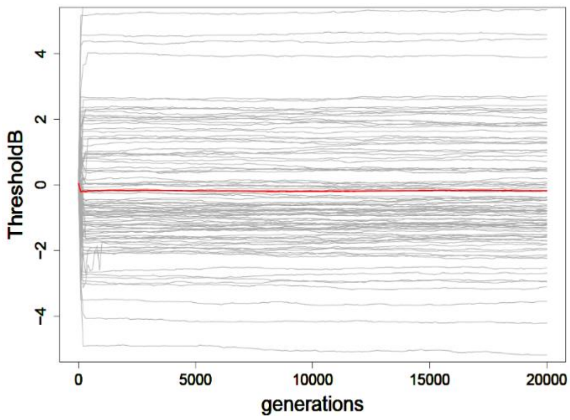
Appendix 1C: The line graphs for runs where p_A =0.1 and p_B =0.9

i) Fully modular agents

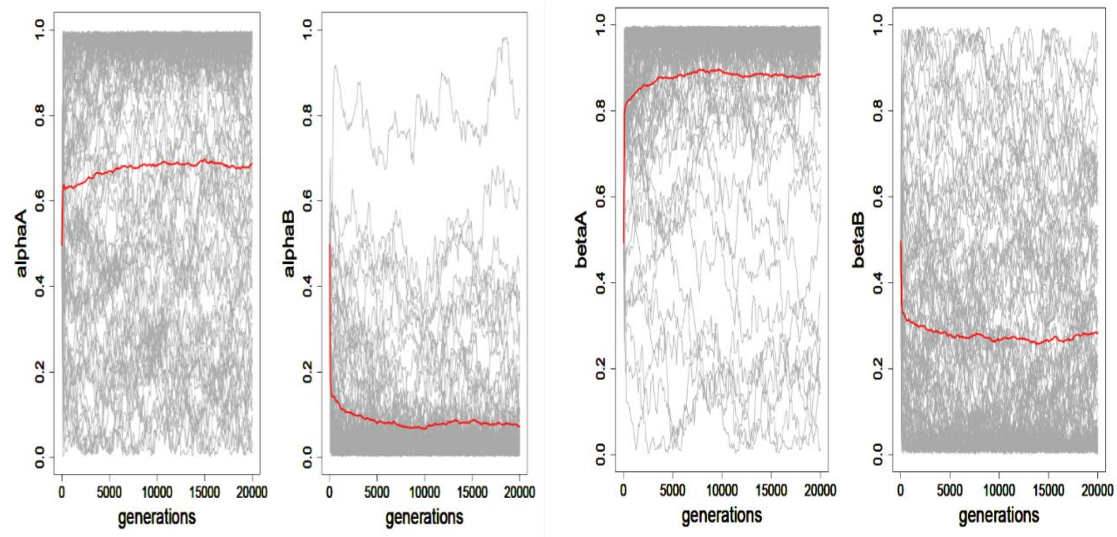


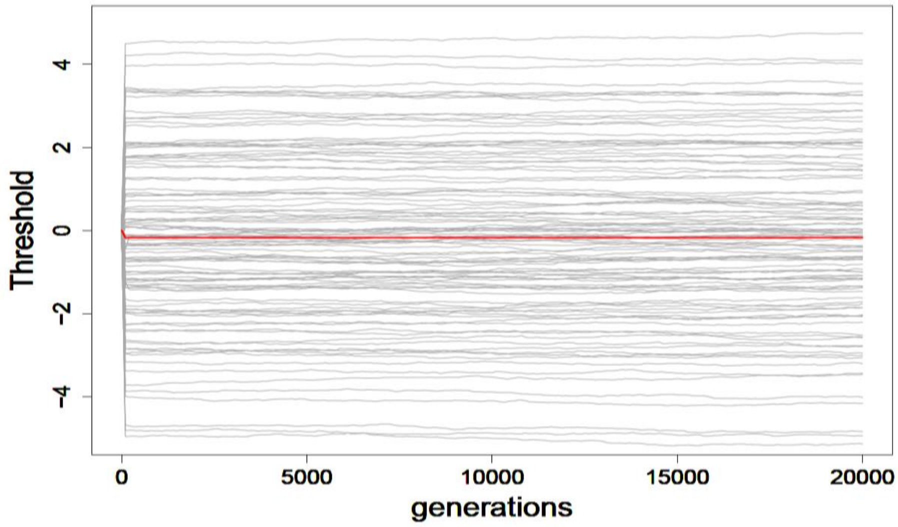




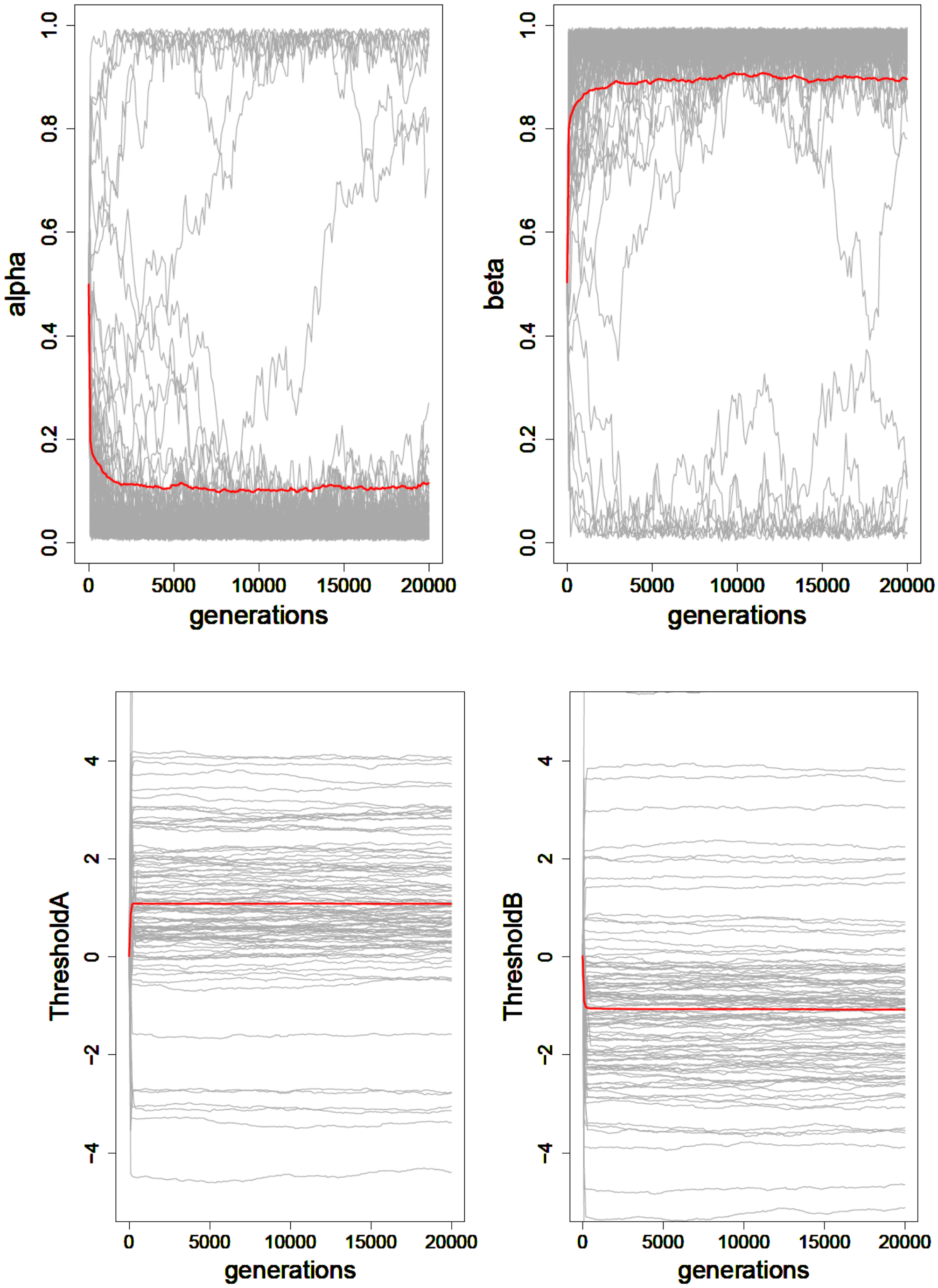


ii) Partly modular agents with modular motivation only

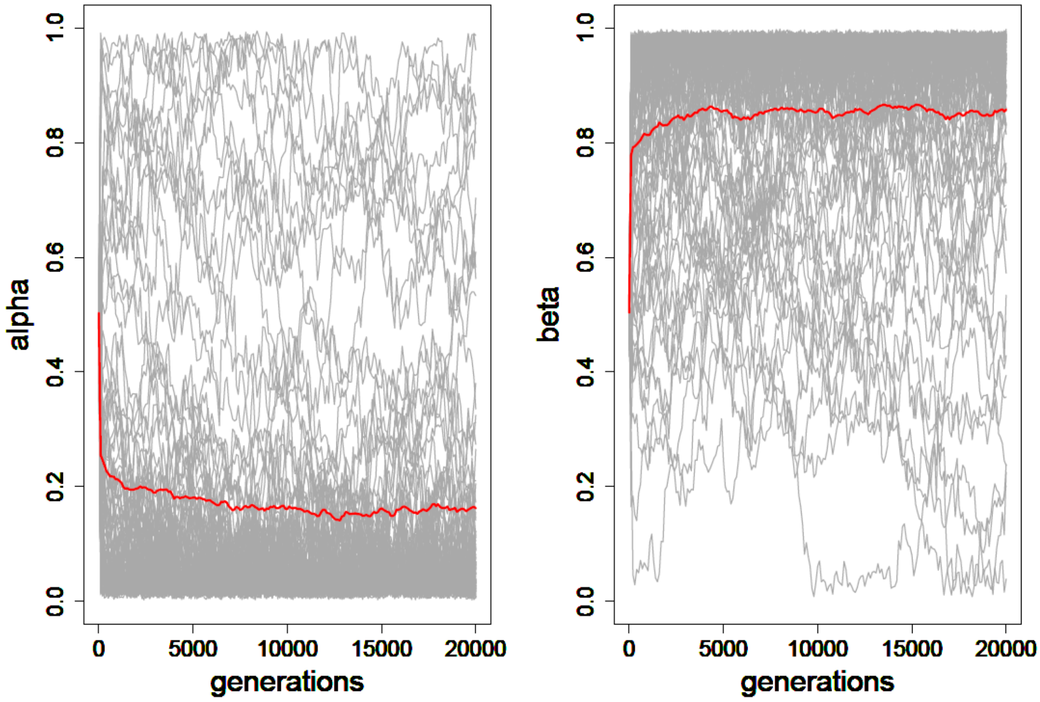


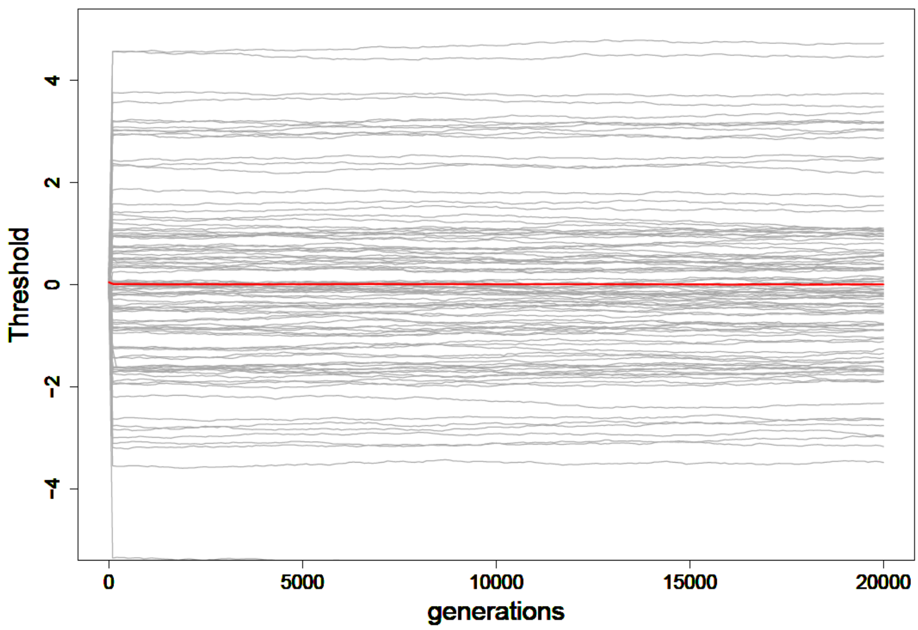


iii) Partly modular agents with modular cognition only



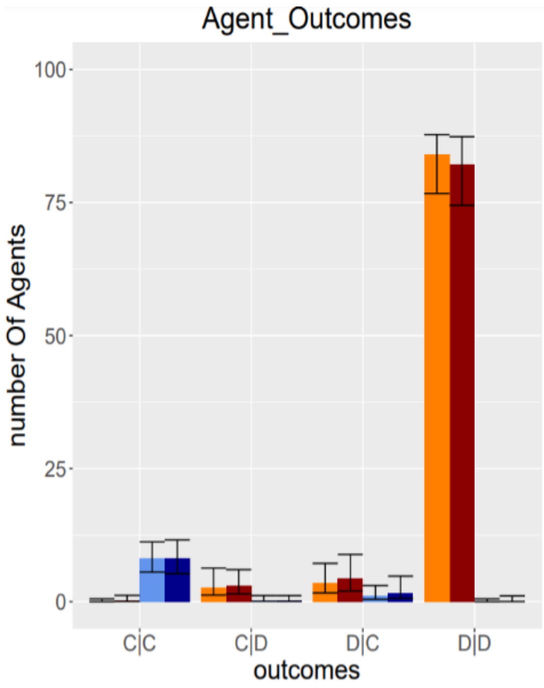
iv) Domain-general agents



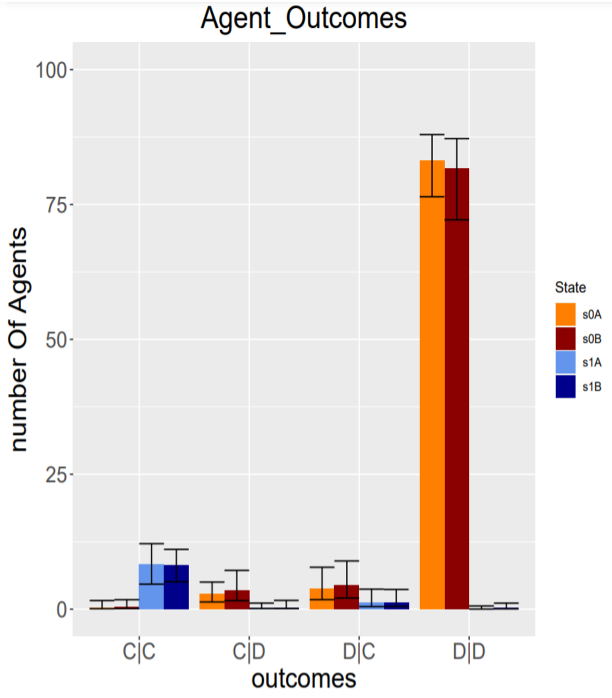


Appendix 2: The bar charts displaying the agent strategies on runs where *b*=2

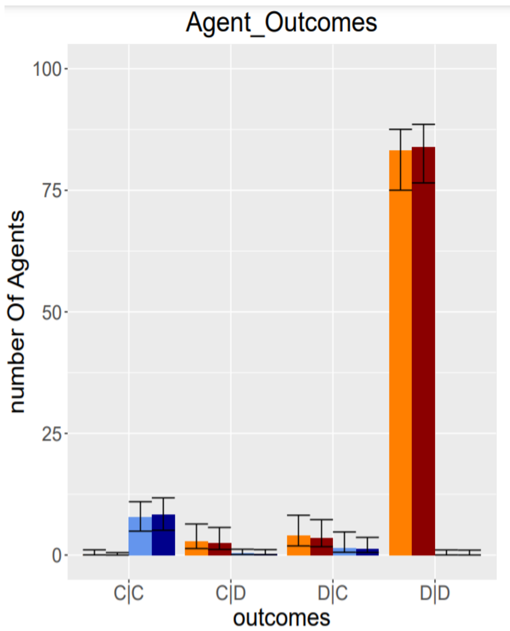
Appendix 2(i): $p_A = 0.1$, $p_B = 0.1$.



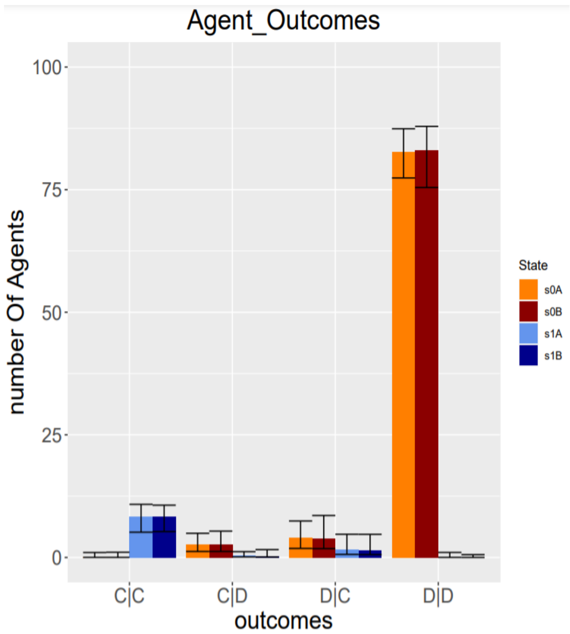
Fully modular



Modular motivation

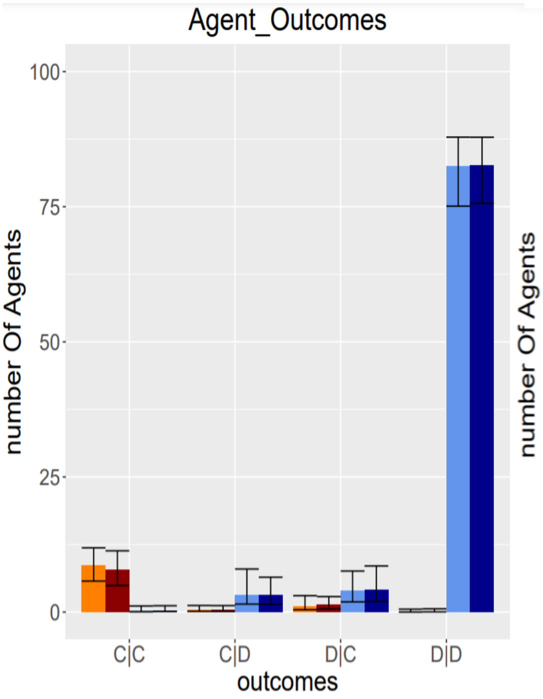


Modular cognition

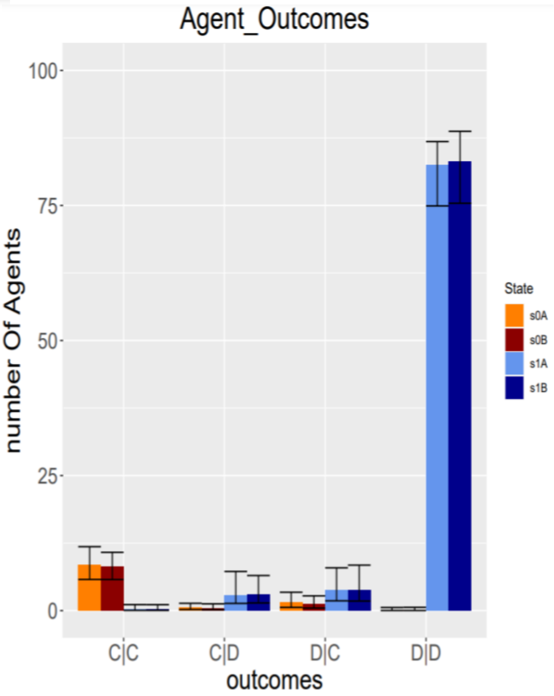


Domain-general

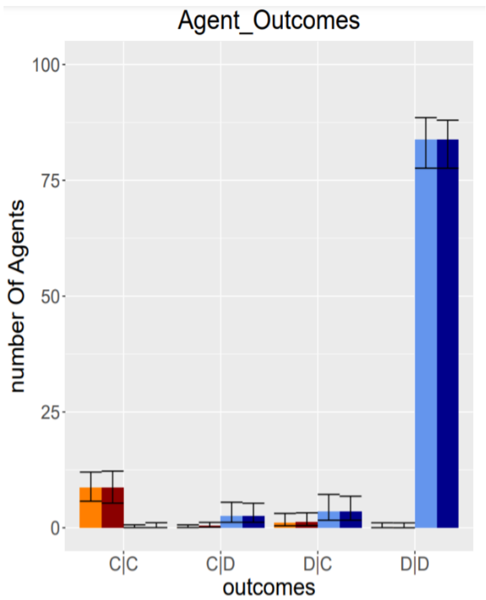
Appendix 2ii) $p_A = 0.9$, $p_B = 0.9$.



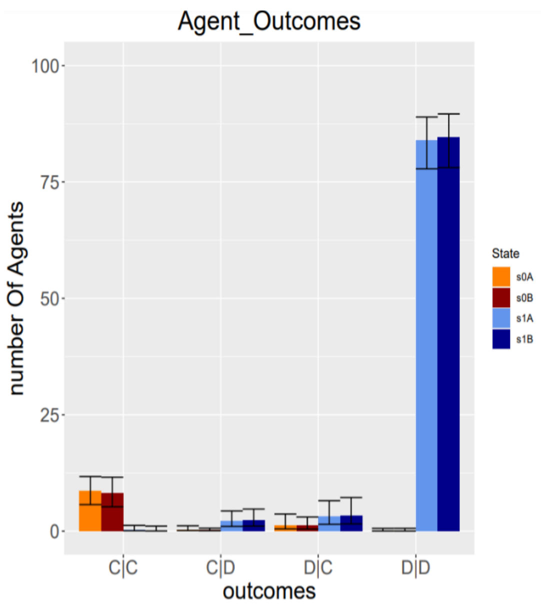
Fully modular



Modular motivation

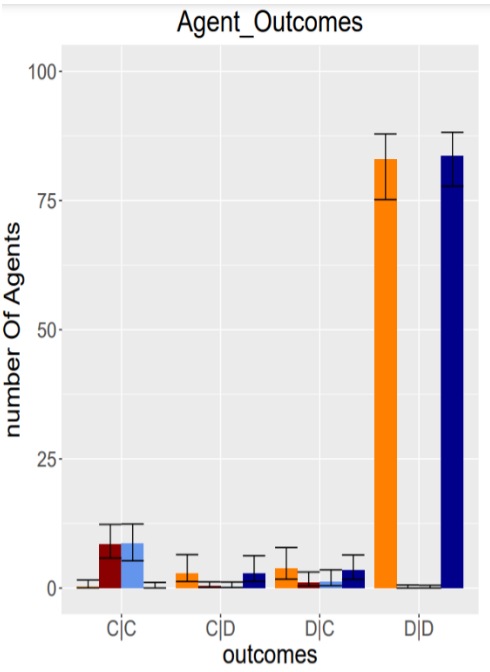


Modular cognition

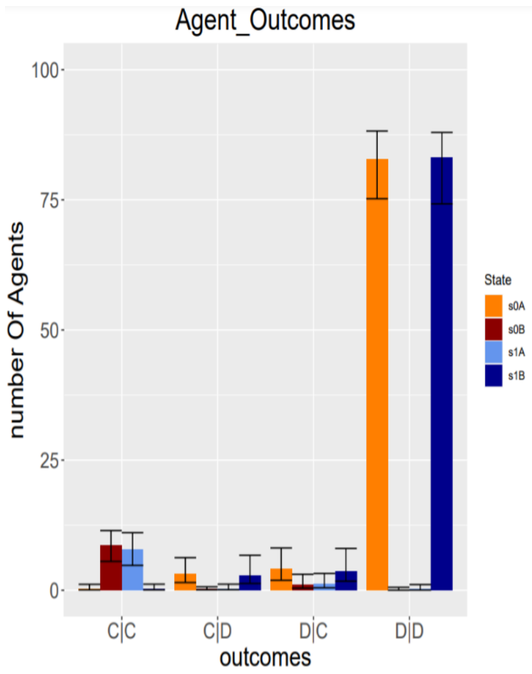


Domain-general

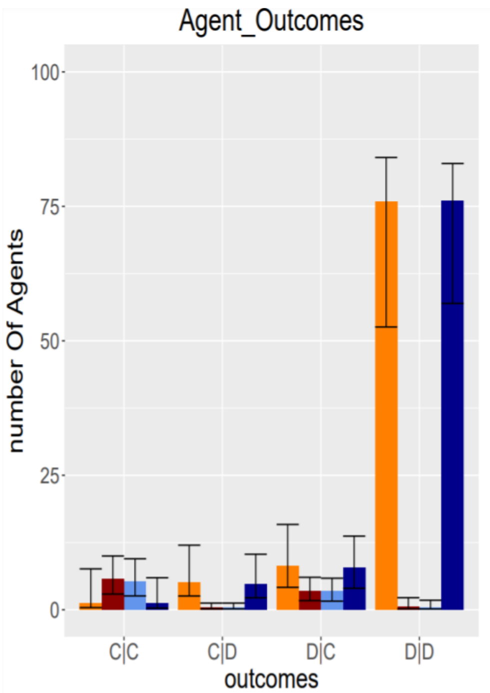
Appendix 2iii) $p_A = 0.1$, $p_B = 0.9$.



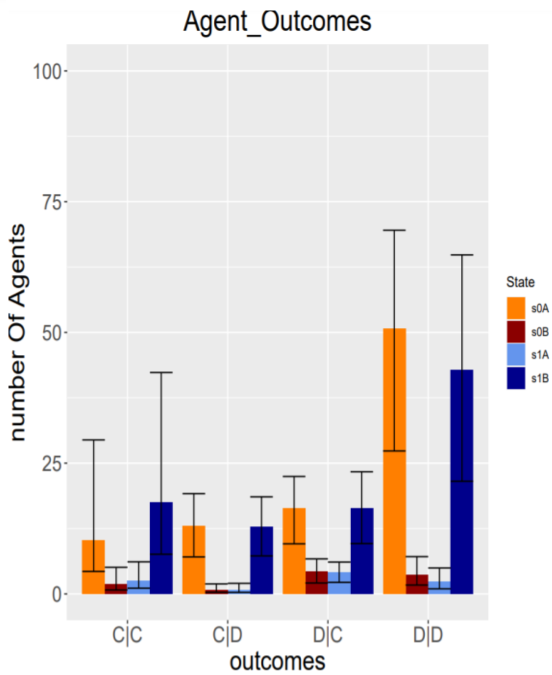
Fully modular



Modular



Modular cognition



Domain-general

Appendix 2. A series of bar charts showing the distribution of cooperation and defection outcomes. This is the outcomes for each of the four agent types on runs with (2i) skewed but consistent priors of state 1, where $p_A = 0.1$, $p_B = 0.1$ and (2ii) when $p_A = 0.9$, $p_B=0.9$ and (2iii) for runs with inconsistent priors of state 1 in both domains ($p_A = 0.1$, $p_B=0.9$). Note the x axis represents the four possible behaviours from the agent when playing a prisoner's dilemma: C|C is cooperation conditional on partner cooperation , C|D is cooperation conditional on partner defection, D|C is defection conditional on partner cooperation and D|D is defection conditional on partner defection. The y axis gives the number of agents who adopt each outcome. These bar charts are for environments with a weak pressure to cooperate: (b=2). Clustered standard error bars represent 95% bootstrapped confidence intervals sampled across the 100 simulations.

Appendix 3: The scripts used to run the models and analysis.

As my agent-based models and analysis scripts were divided into one of the four agent types (fully modular, partly modular agents with modular cognition only, partly modular agents with modular motivation only and domain-general agents), I divide my OSF repositories accordingly. Each repository contains the agent-based model (consisting of a h file, an m file and a main file to compile in Objective C), random number generators for Objective C, the bash file to run the simulations for each combination of the parameters of interest. Plus, each individual repository has some R analysis scripts: one for generating the line graphs in appendix 2, one for generating the clustered bar charts seen in section 3.1 and one for generating the heatmaps and regression by psychological components in sections 3.2 and 3.3. These are:

- Full modular agents: https://osf.io/pj3yd/?view_only=a601dd784d7e4d47ad2cdd0837ba6bae
- Partly modular agents with modular cognition only: https://osf.io/yza8t/?view_only=9e008cba85cb459aa99bfa7168c076ea
- Partly modular agents with modular motivation only: https://osf.io/wcvtz/?view_only=c2eebfdbe1f24d0586b10414f1cf9052
- Domain-general agents: https://osf.io/3yv46/?view_only=bc24ed058eb748f6ab68af16da67cfad
- The raw data: https://www.dropbox.com/sh/5rcv8dzea01rtpg/AABL-bpWdpc849NQzcdYVba6a?dl=0

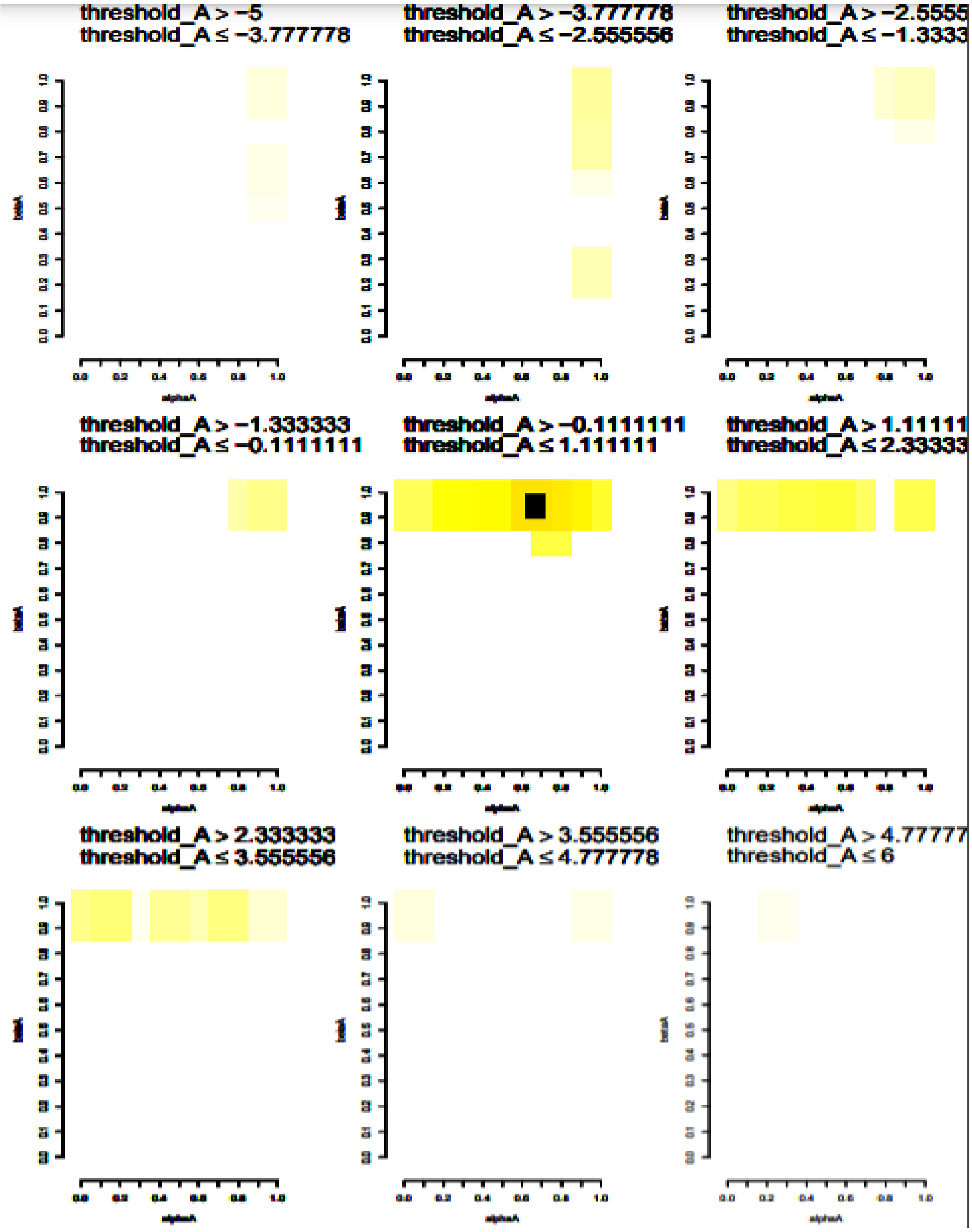
For the raw data, there are multiple files uploaded, one for each of the parameter combinations. The files called *disagg** contain the data for the final generation of agents. These values were used to create the heatmaps in section 3.2, and the regressions in section 3.3. The file called *popDataAgentTypes** represented the cognitive and motivational values averaged over the four agent types of interest over the generations. This was used to make the line graph in appendix 1. The files called *popDataAgentTypesCorrExt** give the agent strategies for the agents at the extreme end of the modularity scale (e.g., fully modular, and domain-general agents). The files called *popDataAgentTypesCorrMod** give the agent strategies for the partly-modular agents (e.g., agents with modular motivation only and agents with modular cognition only). These datasets were used to make the clustered barchart in section 3.1. Finally, the *popDataMot** and *popDataThresh** files give data aggregated over the agent types (for the motivational thresholds and the cognitive thresholds respectively) but were unused in the analysis.

Finally, I also include an overall repository where I attach the script for the regression reported in section 3.4 which compared each agent type, and the supplementary materials showing the full analysis of the full parameter space: https://osf.io/zdyvh/?view_only=f2d9263c1807450da58ef122ebcdf627

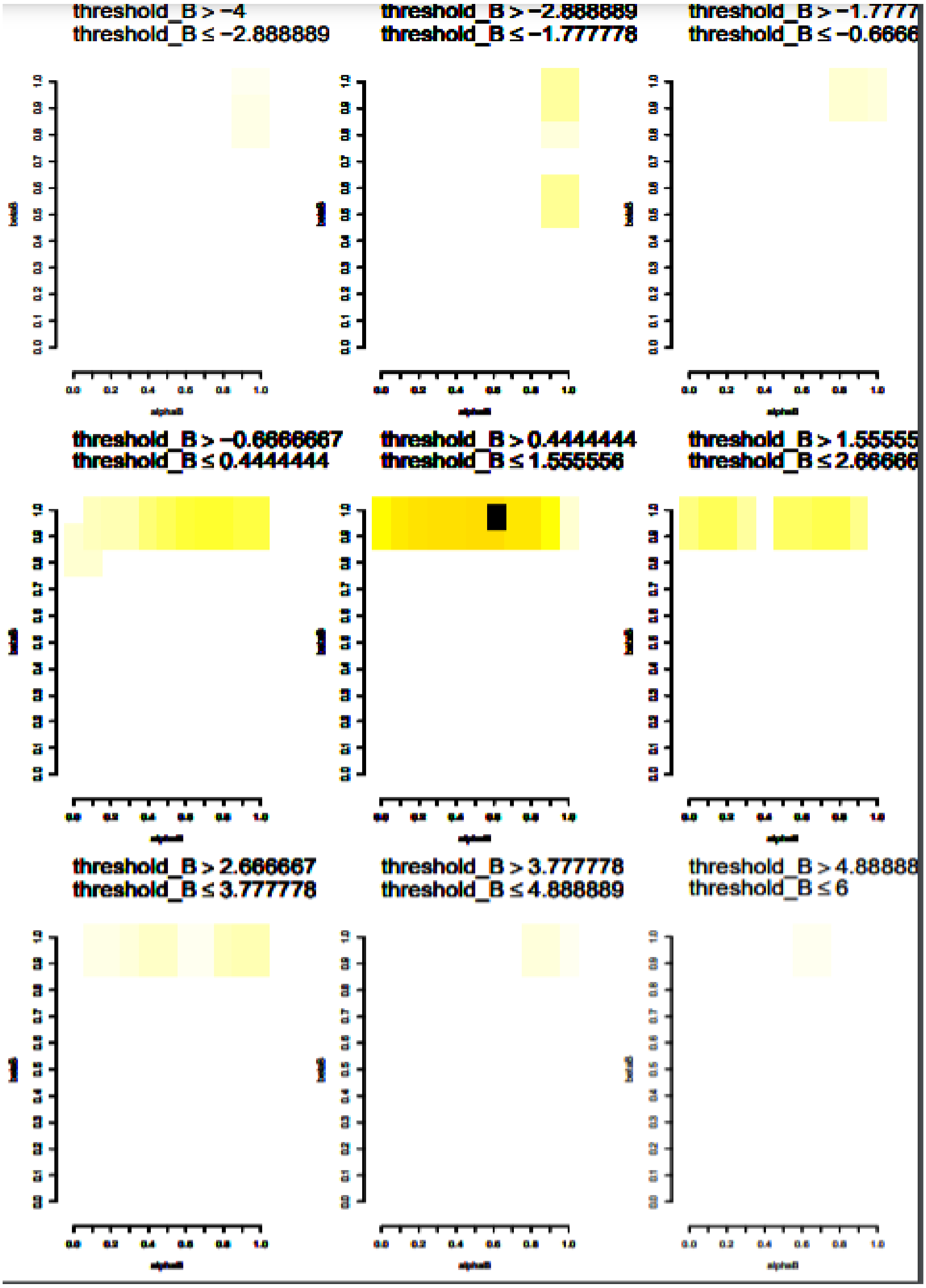
Appendix 4: The heatmaps of psychological architecture for runs where *b*=2.

Appendix 4A: The heatmaps for runs where p_A =0.1 and p_B =0.1 and *b*=2.

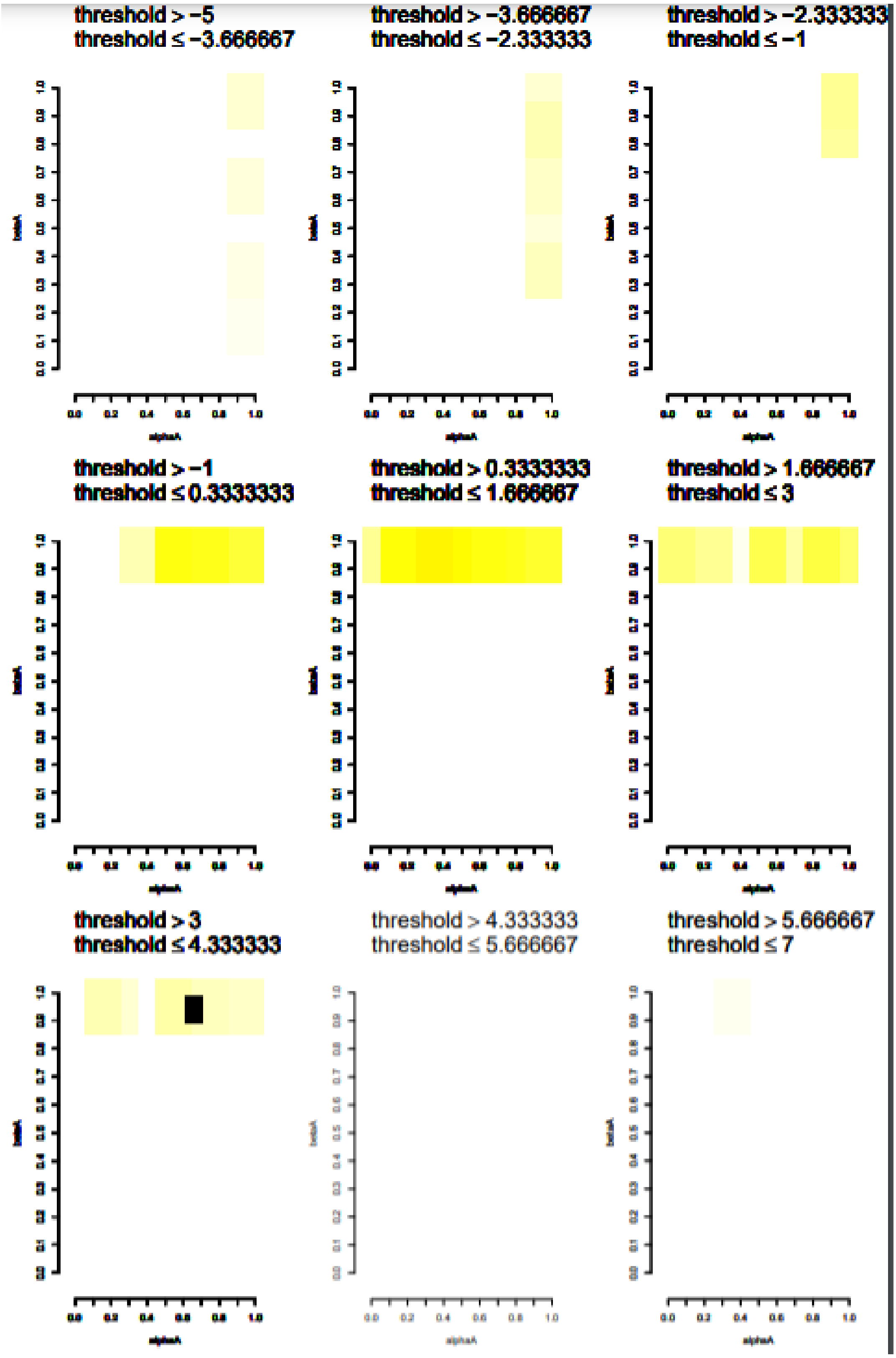
Appendix 4A: p_A = 0.1, p_B = 0.1. Fully modular, Domain A.



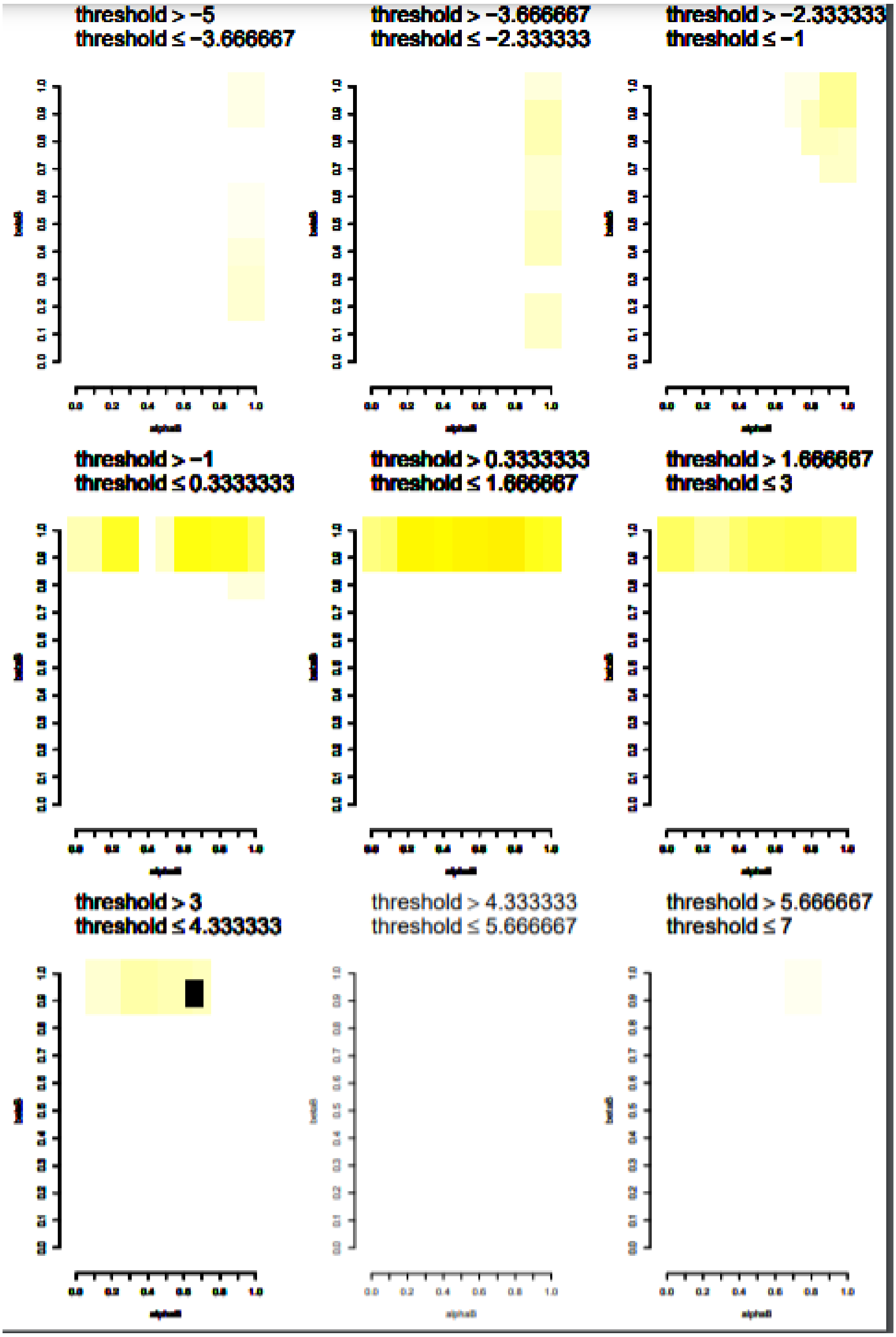
Appendix 4A: p_A = 0.1, p_B = 0.1. Fully modular, Domain B.



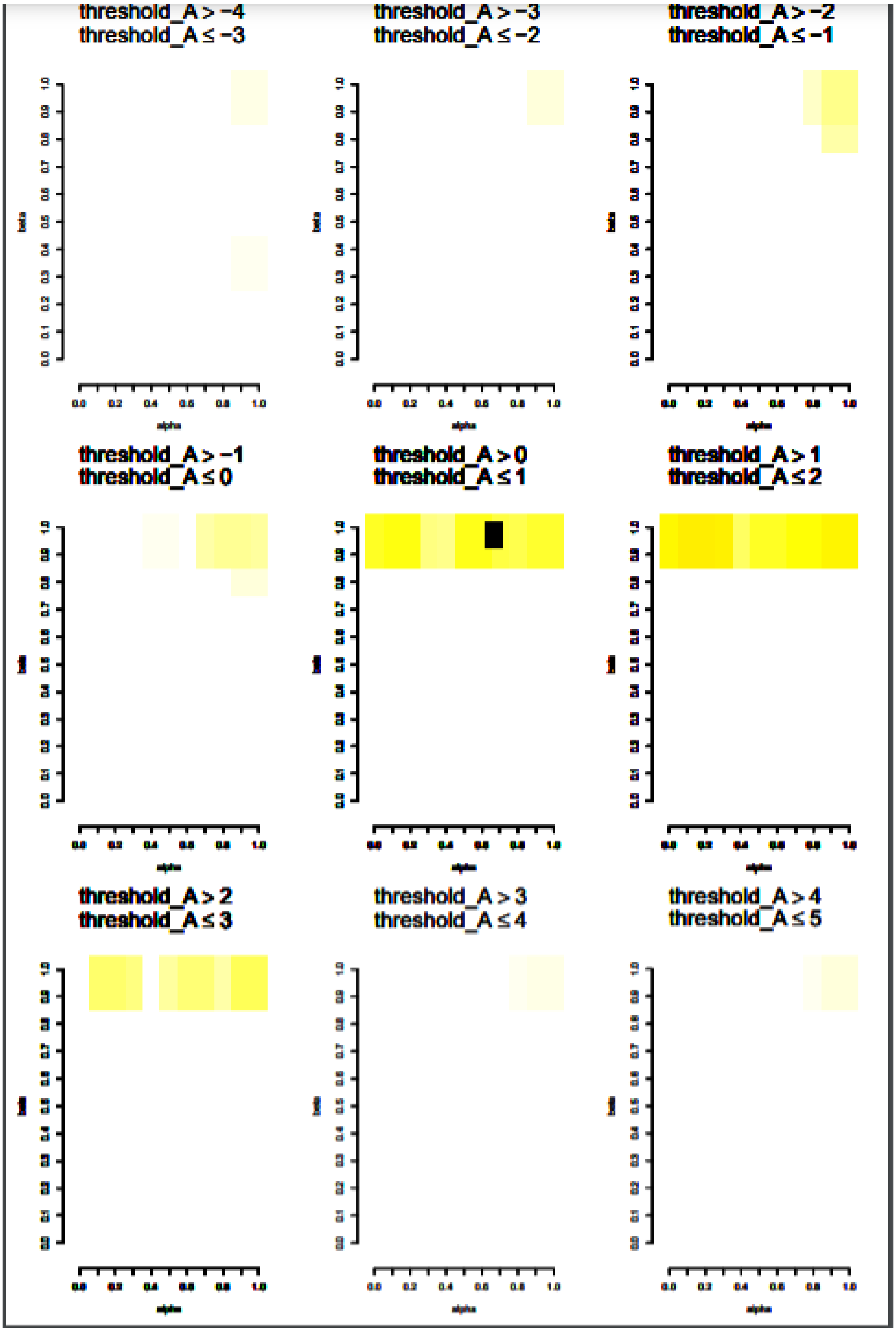
Appendix 4A: p_A = 0.1, p_B = 0.1. Modular motivation, Domain A



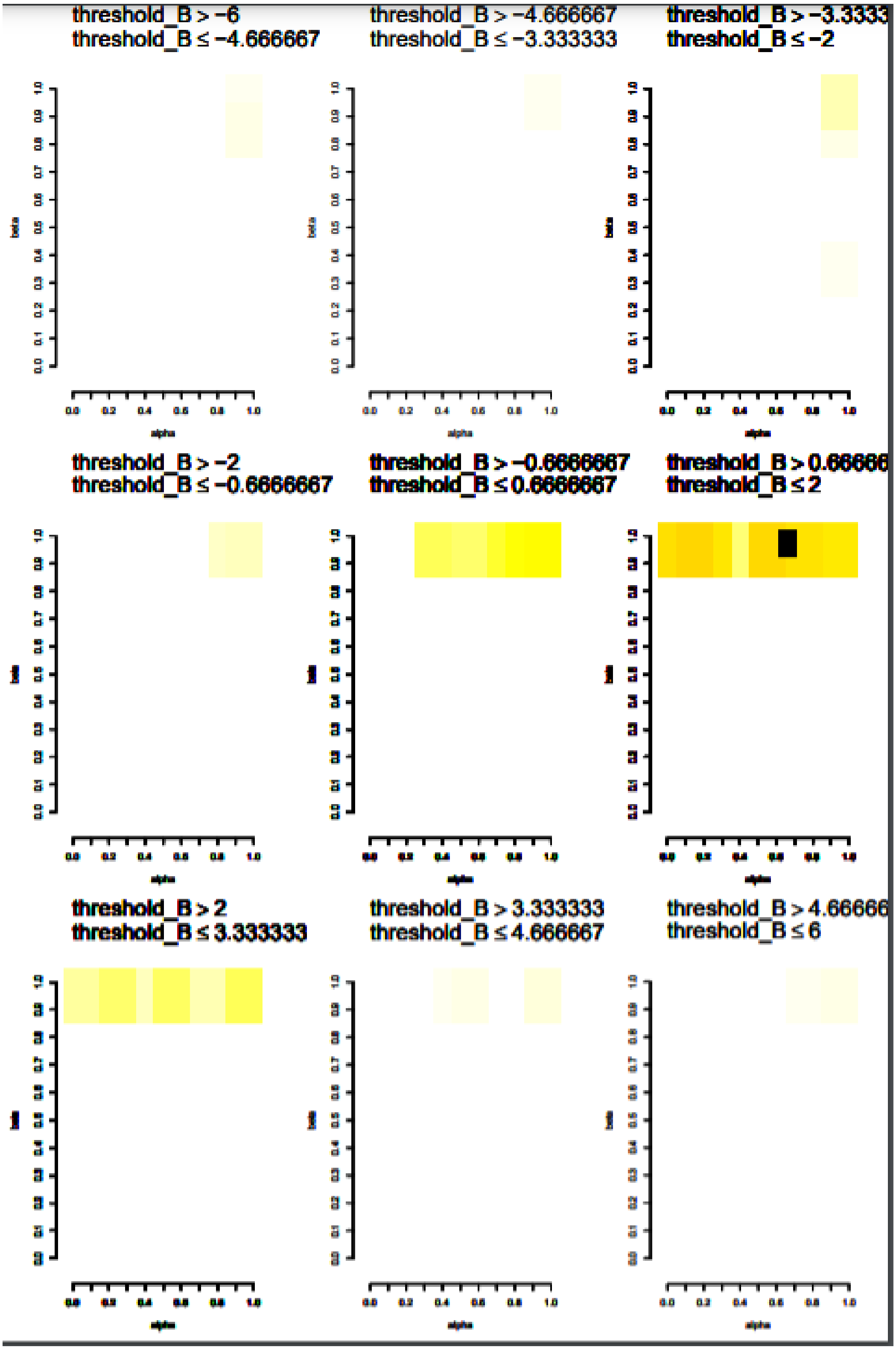
Appendix 4A: p_A = 0.1, p_B = 0.1. Modular motivation, Domain B



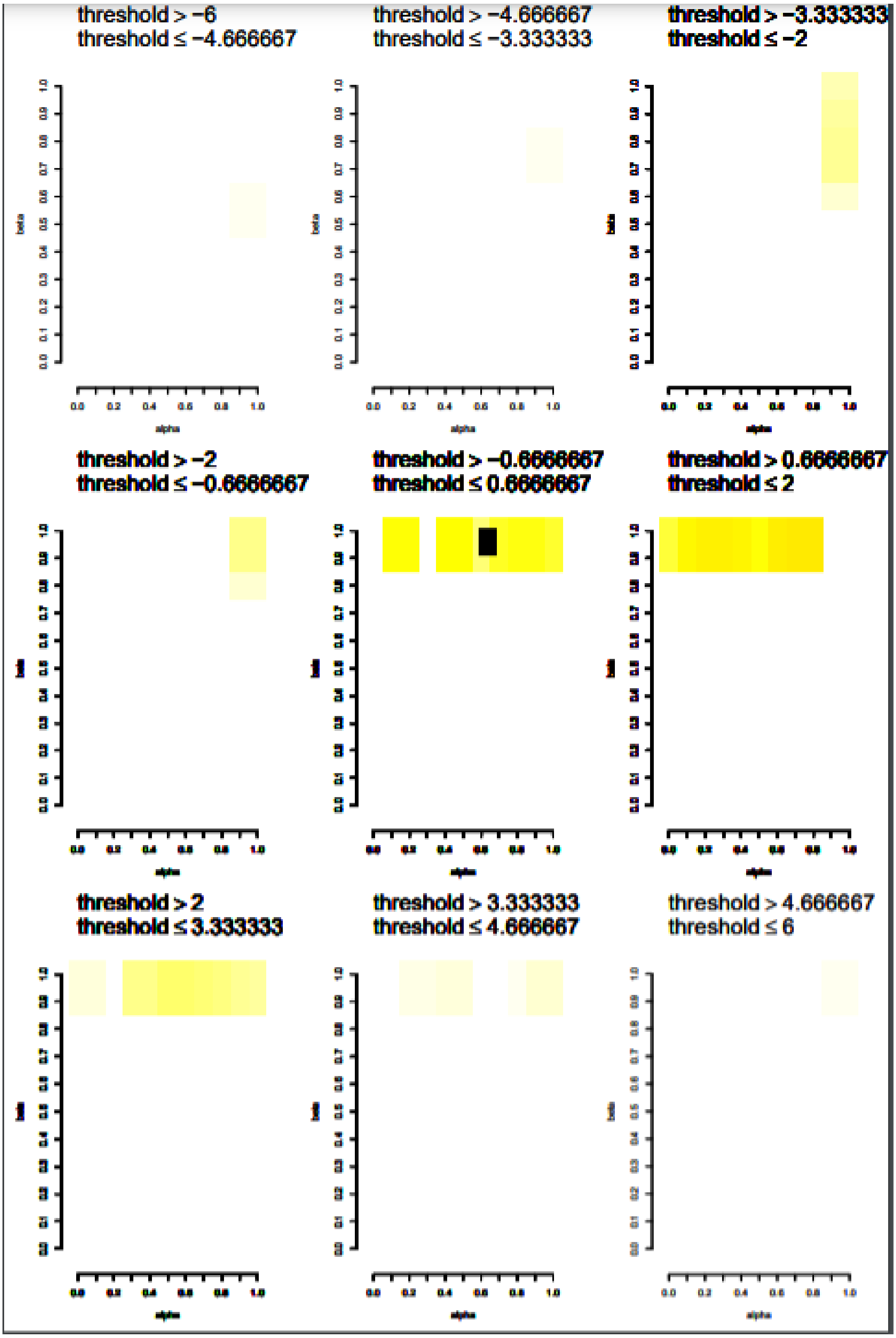
Appendix 4A: $p_A = 0.1$, $p_B = 0.1$. Modular cognition, Domain A



Appendix 4A: p_A = 0.1, p_B = 0.1. Modular cognition, Domain B

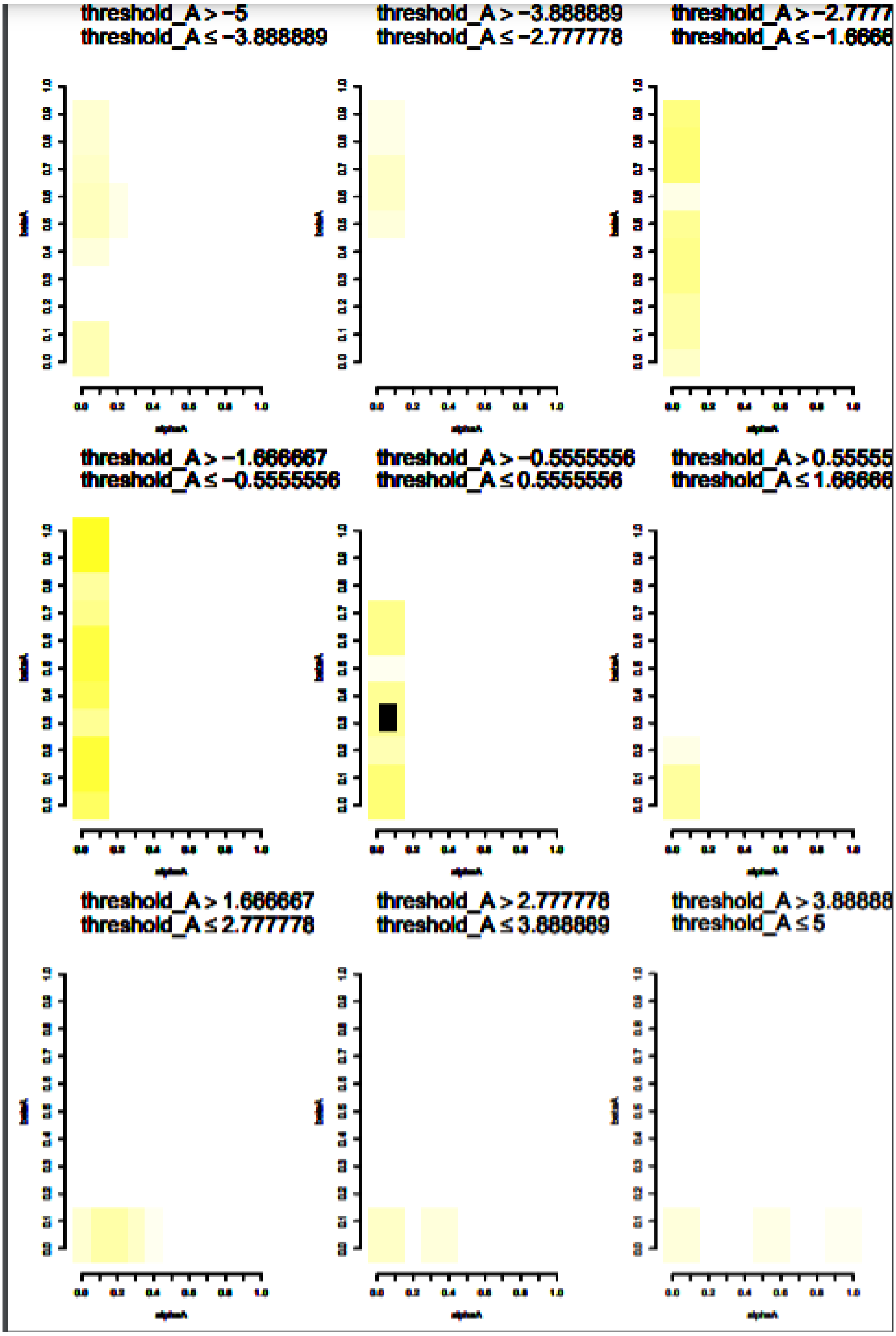


Appendix 4A: $p_A = 0.1$, $p_B = 0.1$. Domain-general

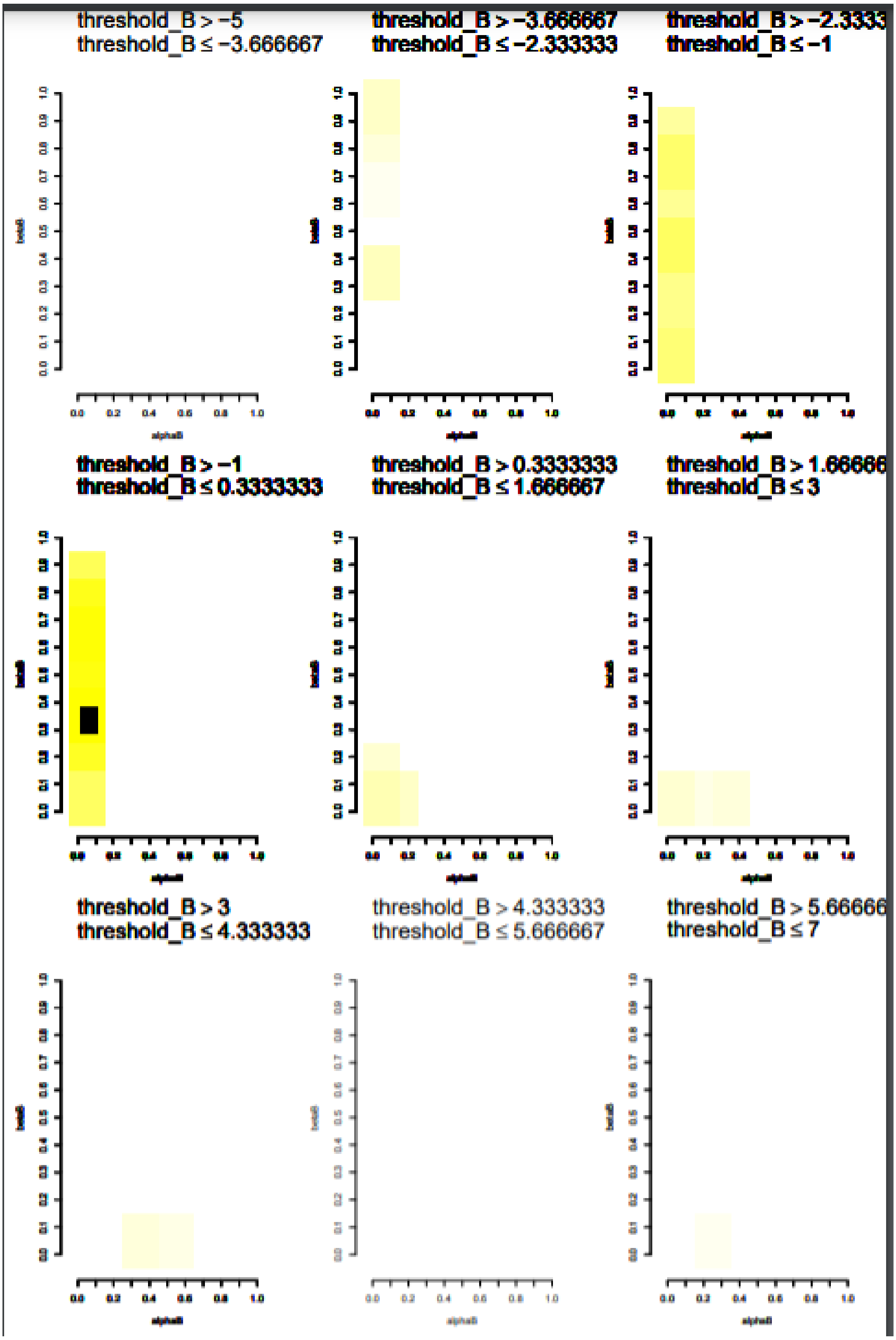


Appendix 4B: The heatmaps for runs where p_A =0.9 and p_B =0.9 and b=2

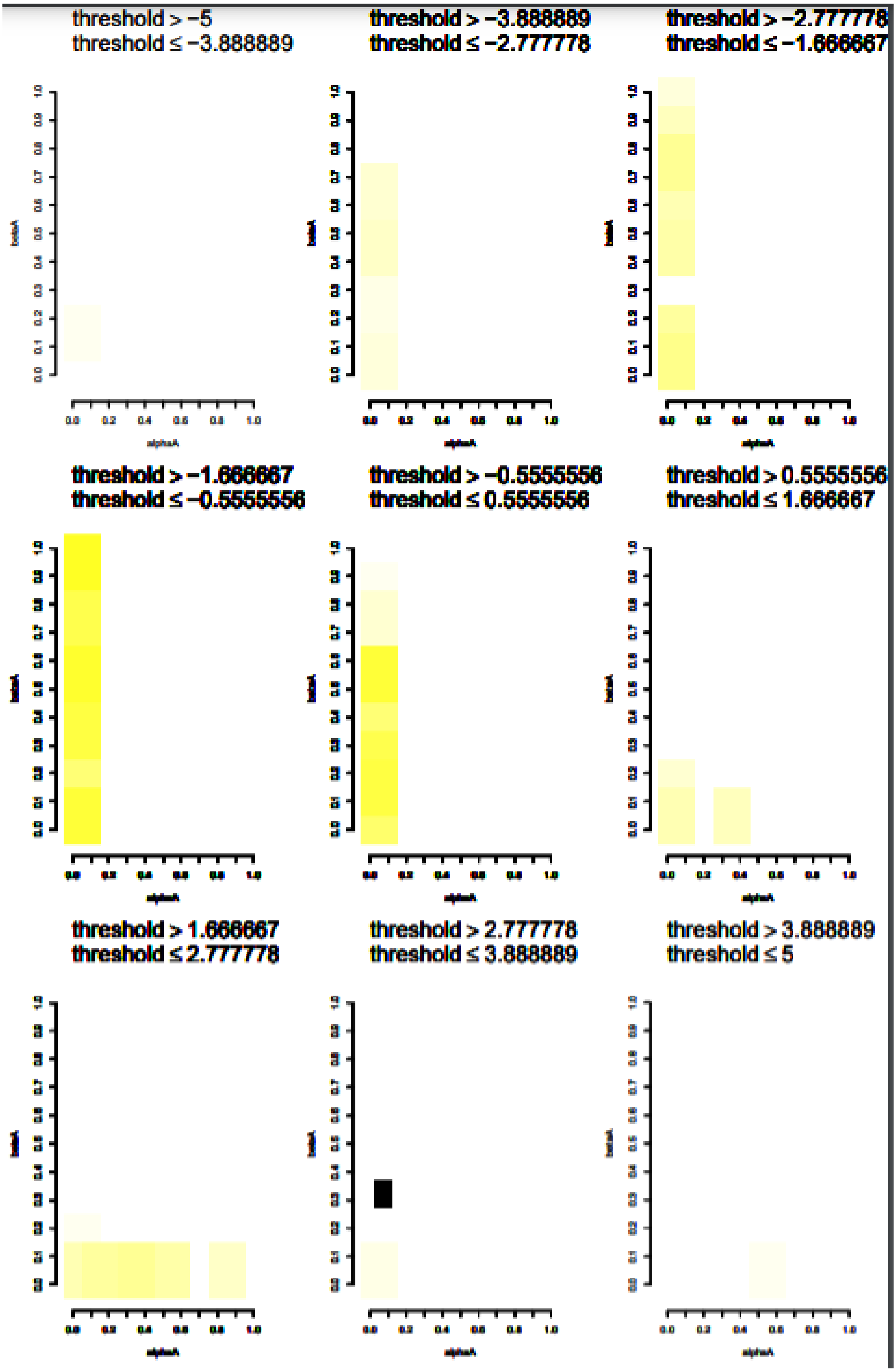
Appendix 4B: p_A = 0.9, p_B = 0.9. Fully modular, Domain A



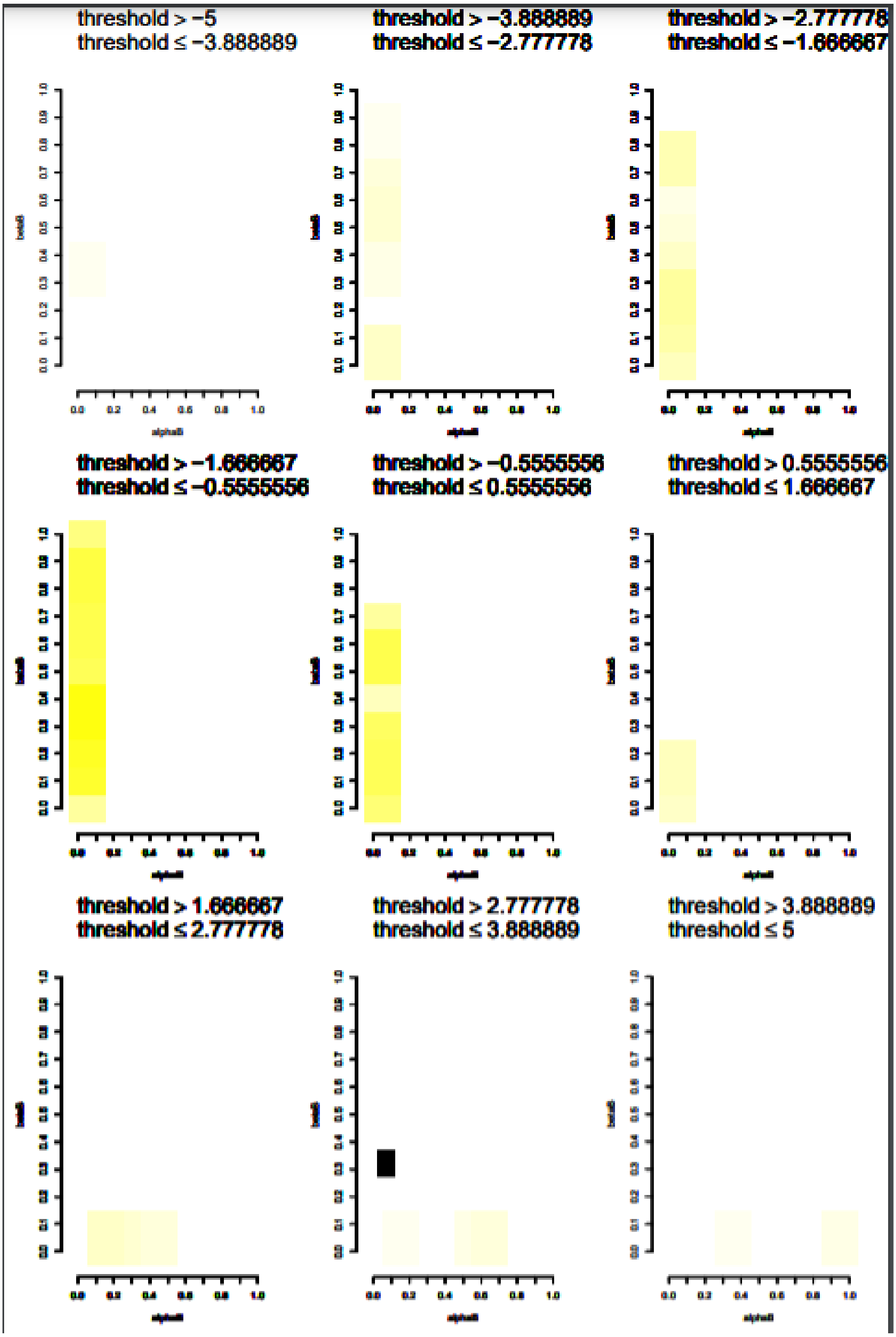
Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Fully modular, Domain B



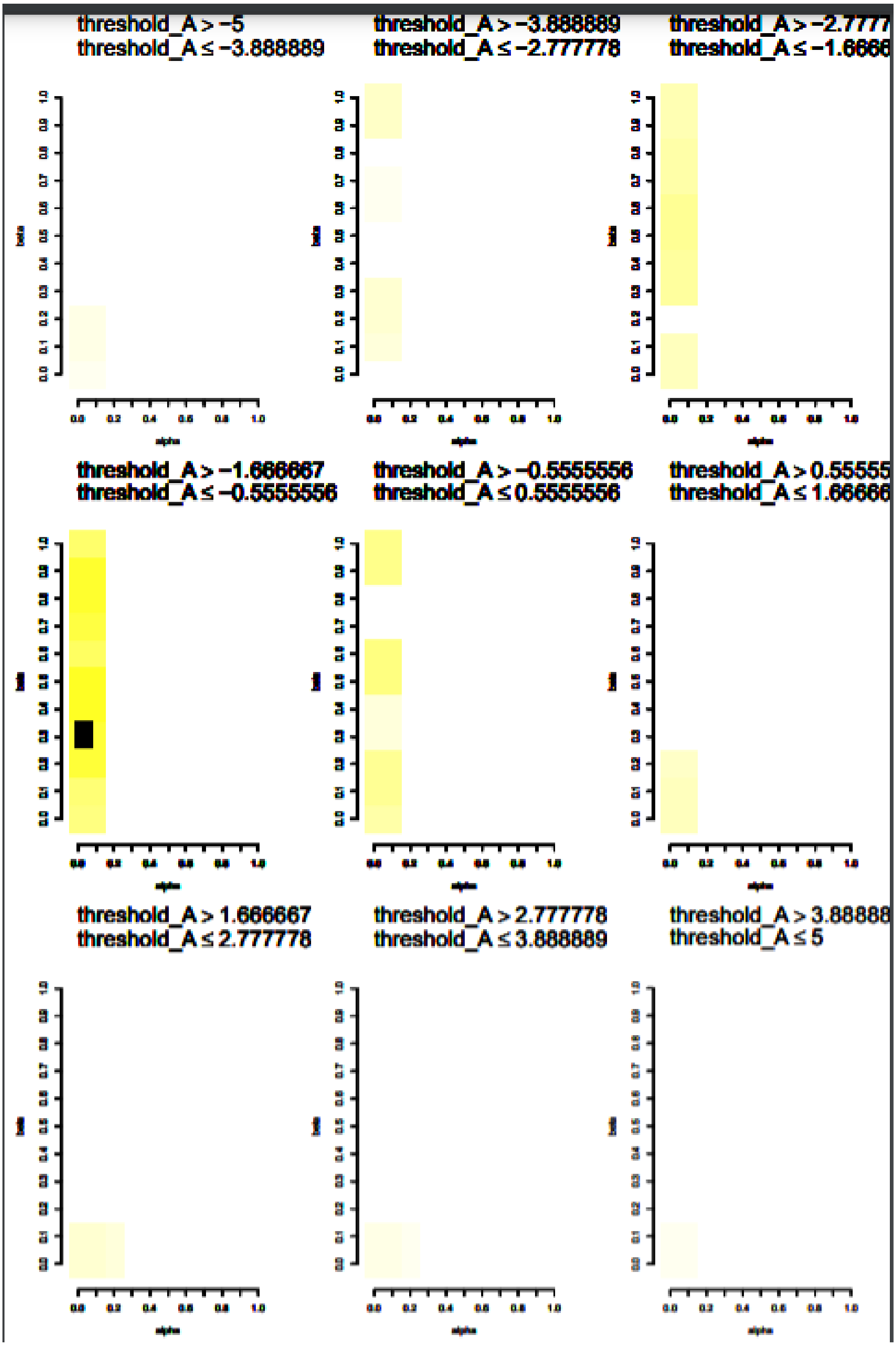
Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Modular motivation, Domain A



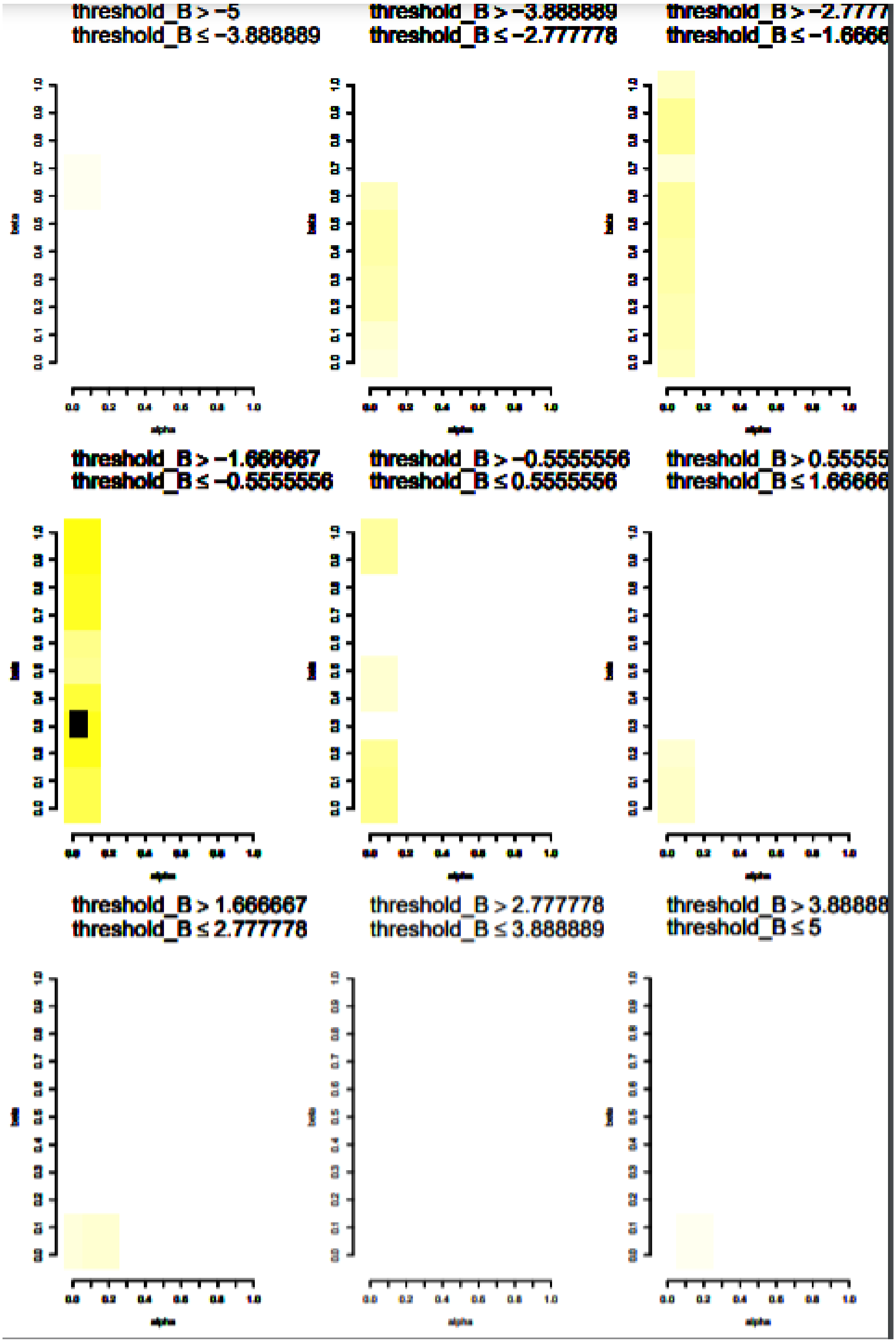
Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Modular motivation, Domain B



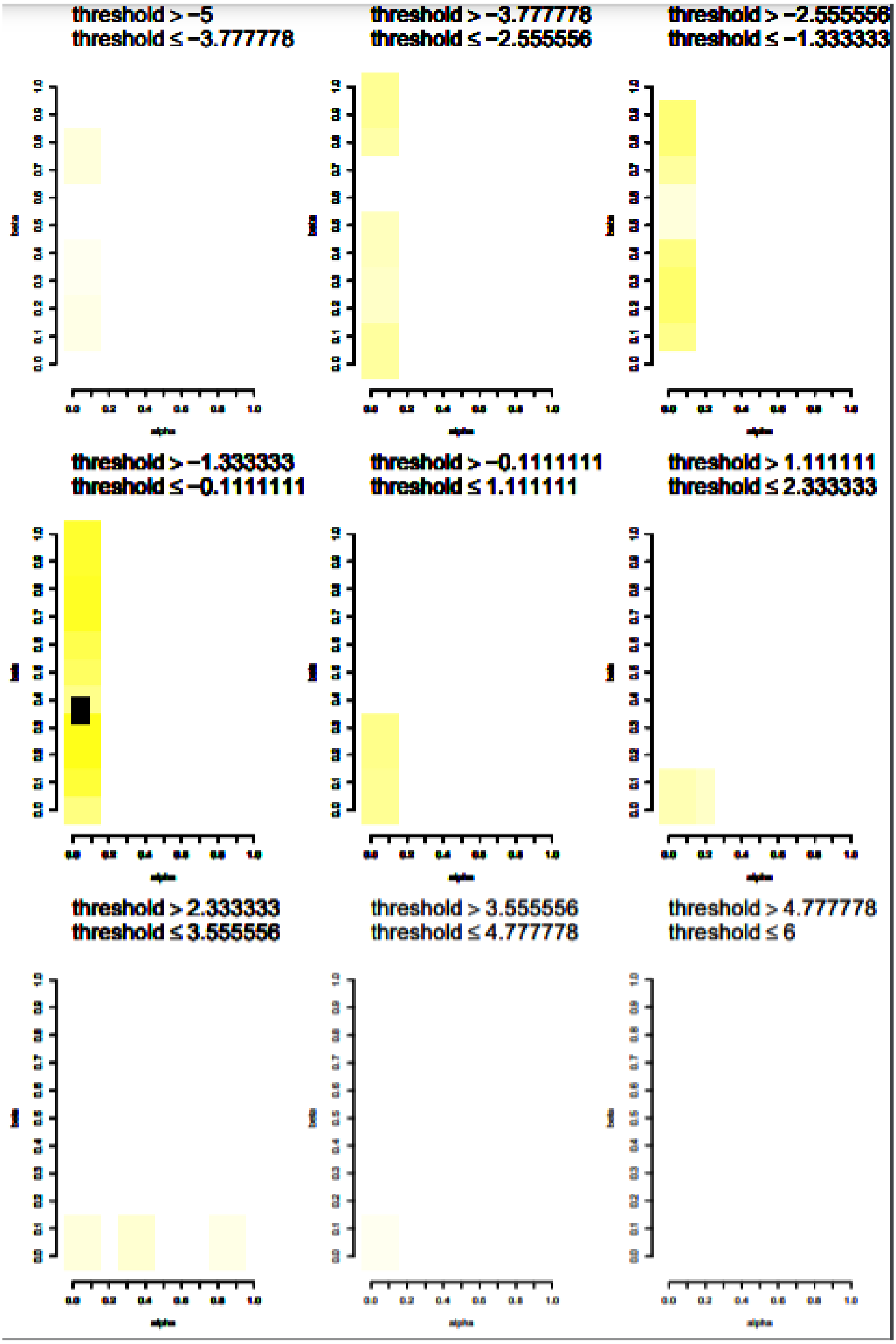
Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Modular cognition, Domain A



Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Modular cognition, Domain B

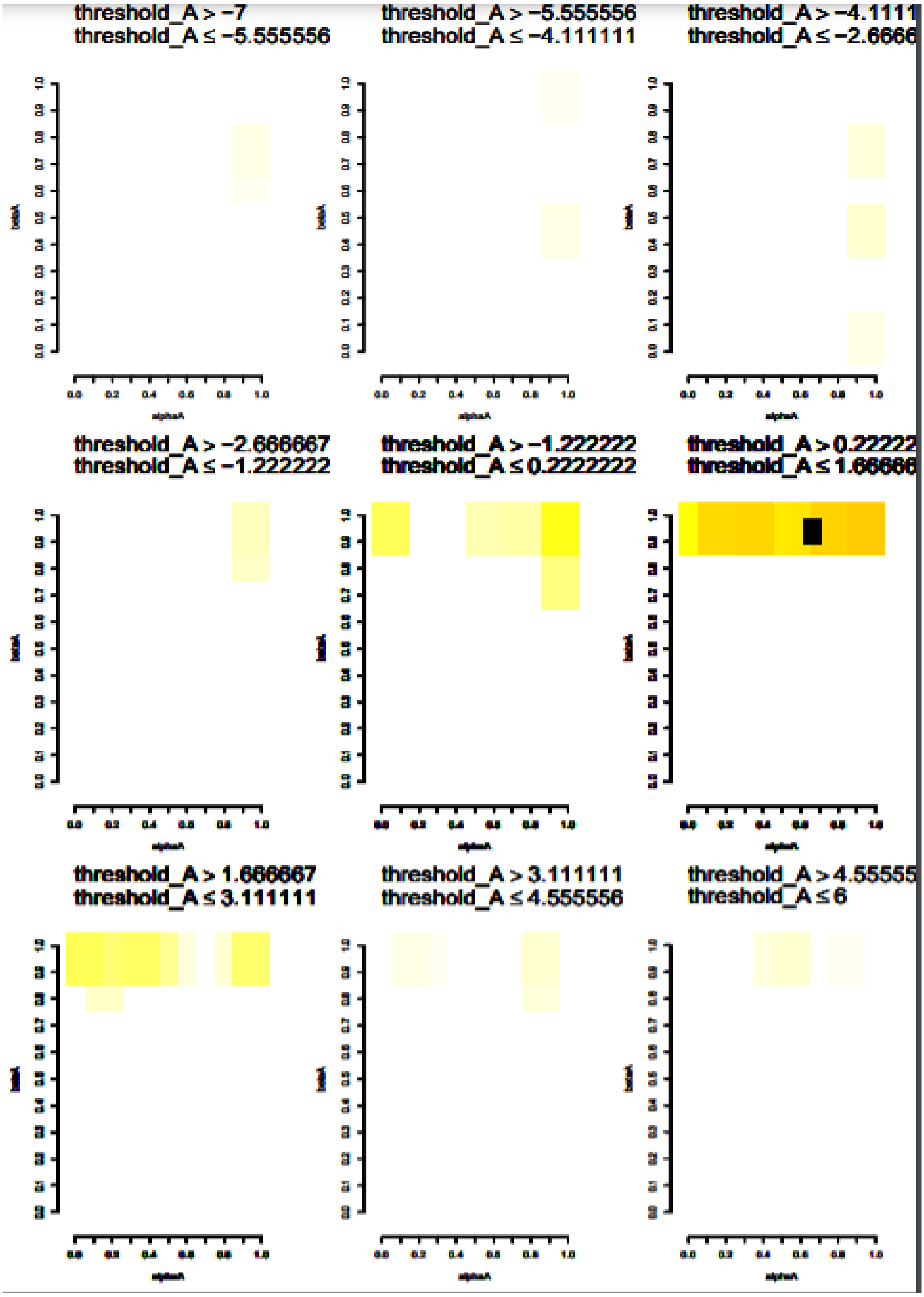


Appendix 4B: $p_A = 0.9$, $p_B = 0.9$. Domain-general

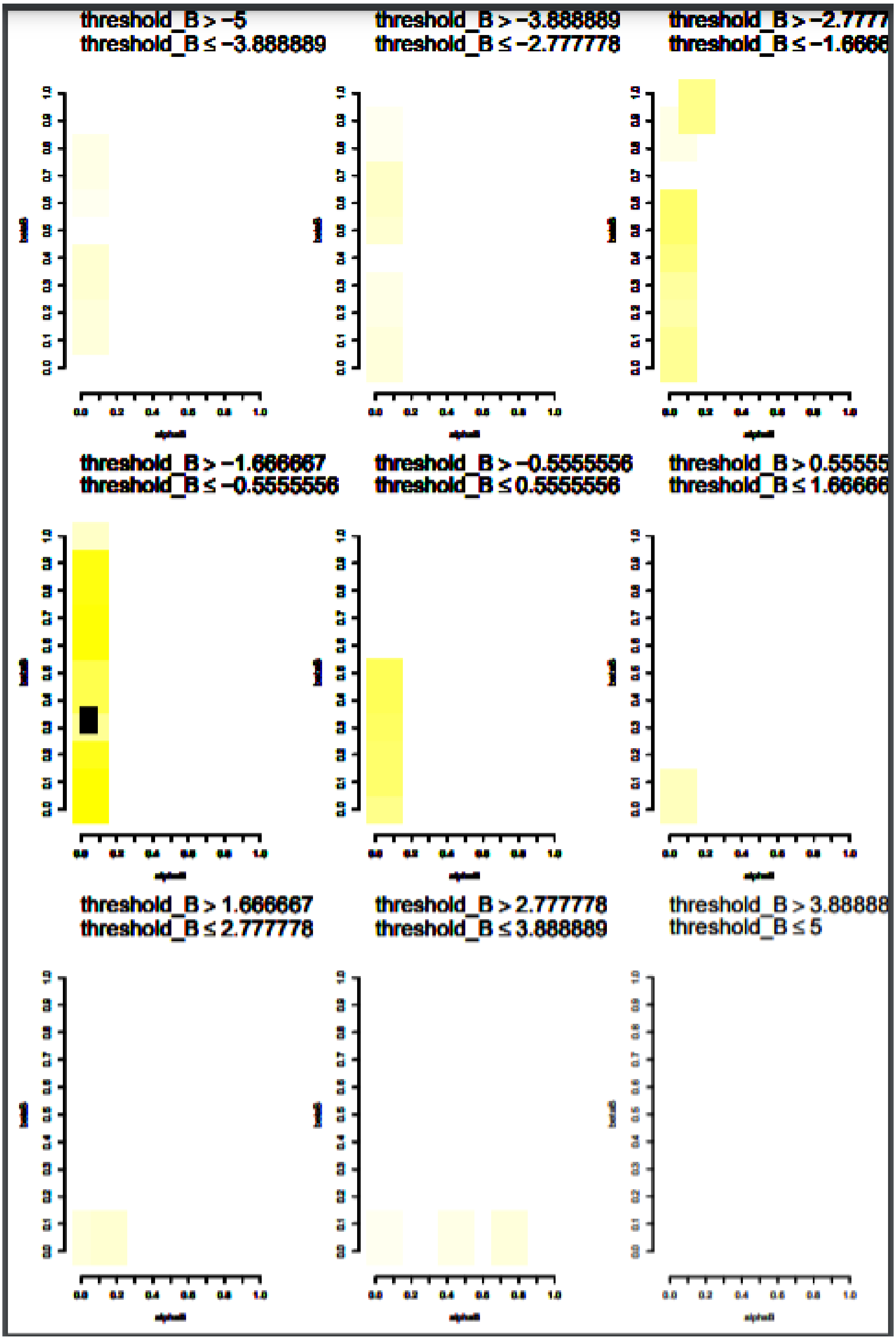


Appendix 5: The heatmaps of psychological architecture for runs where p_A =0.1 and p_B =0.9, and b=2

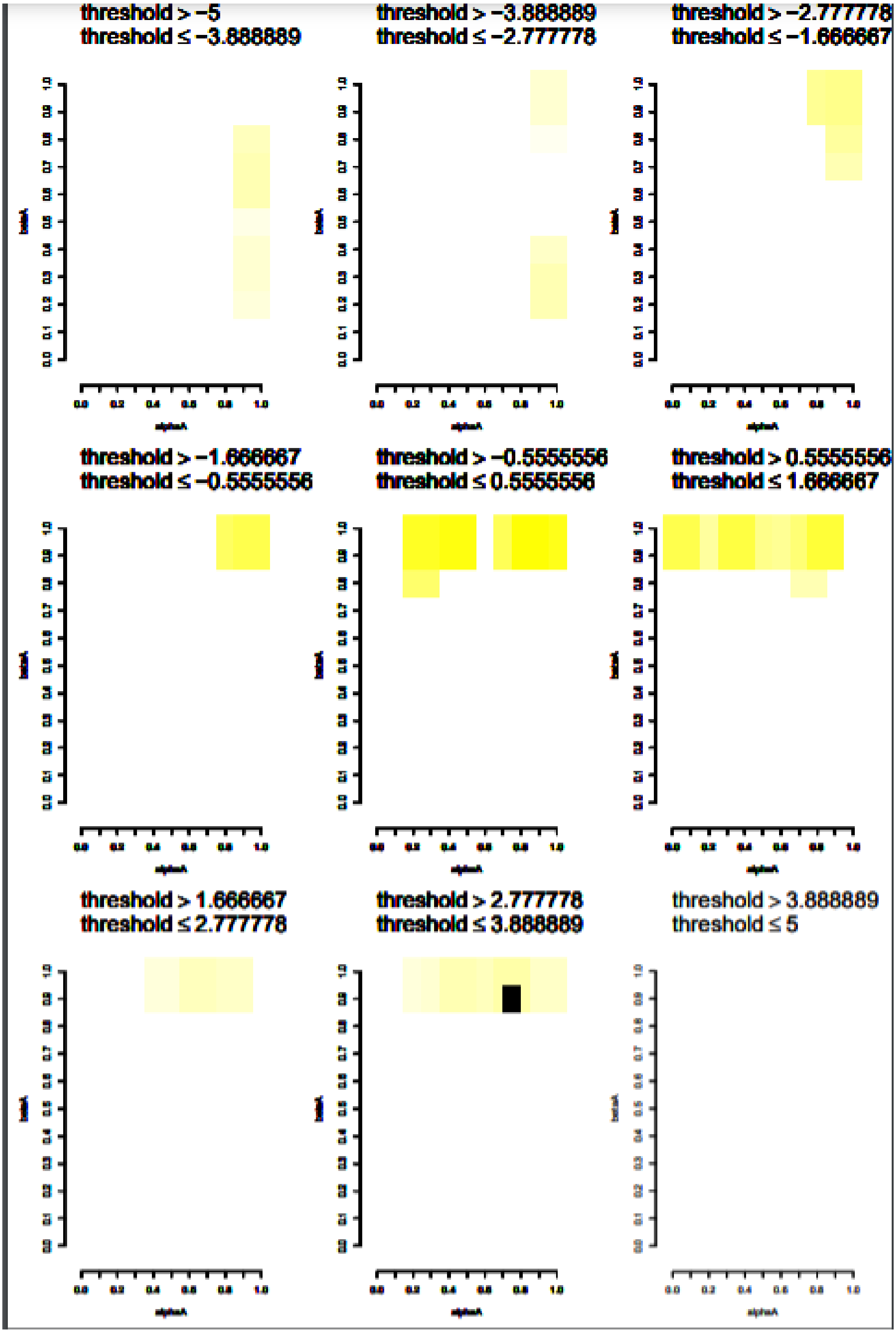
Appendix 5: p_A = 0.1, p_B = 0.9. Fully modular, domain A



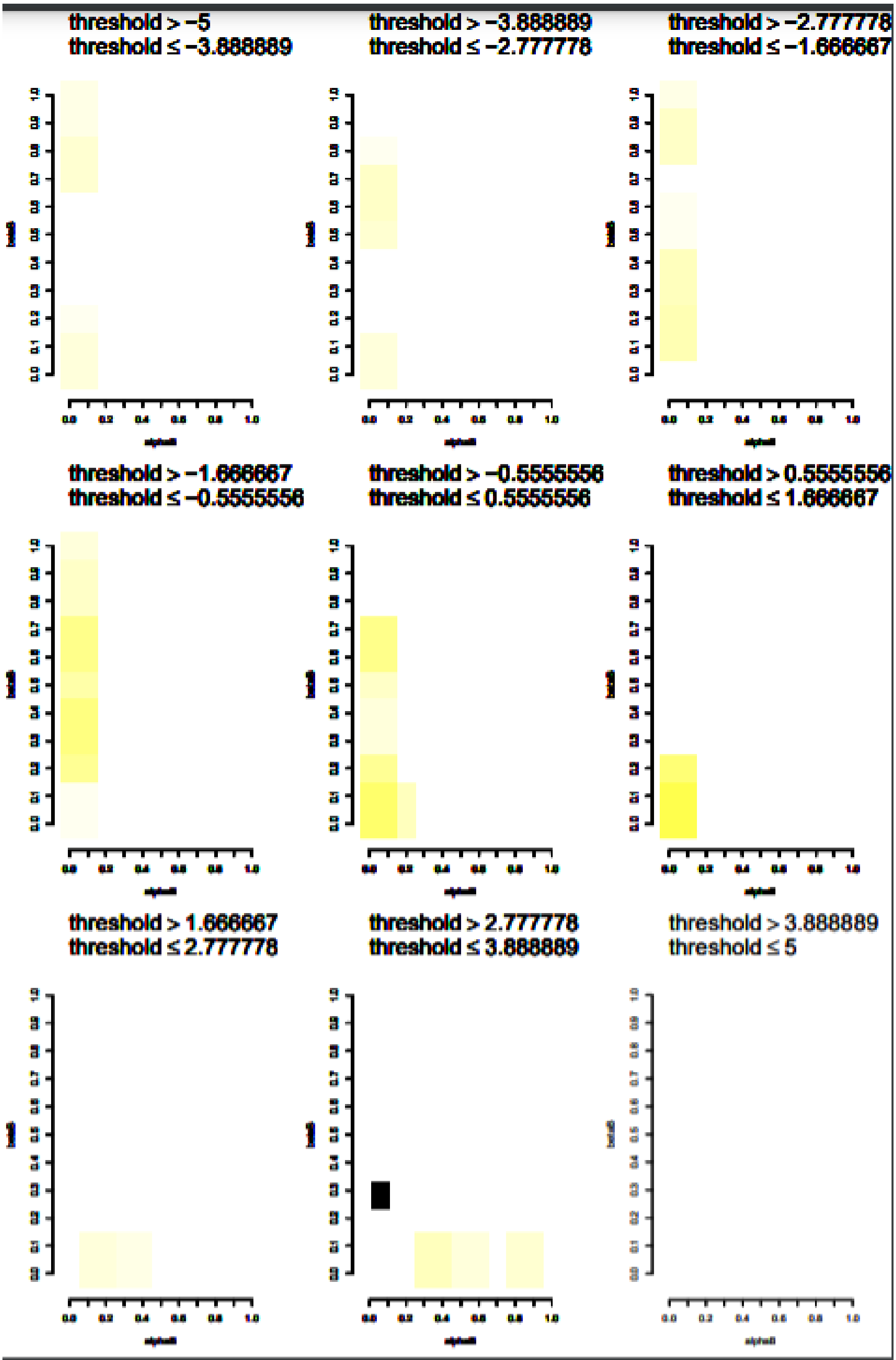
Appendix 5: $p_A = 0.1$, $p_B = 0.9$. Fully modular, domain B



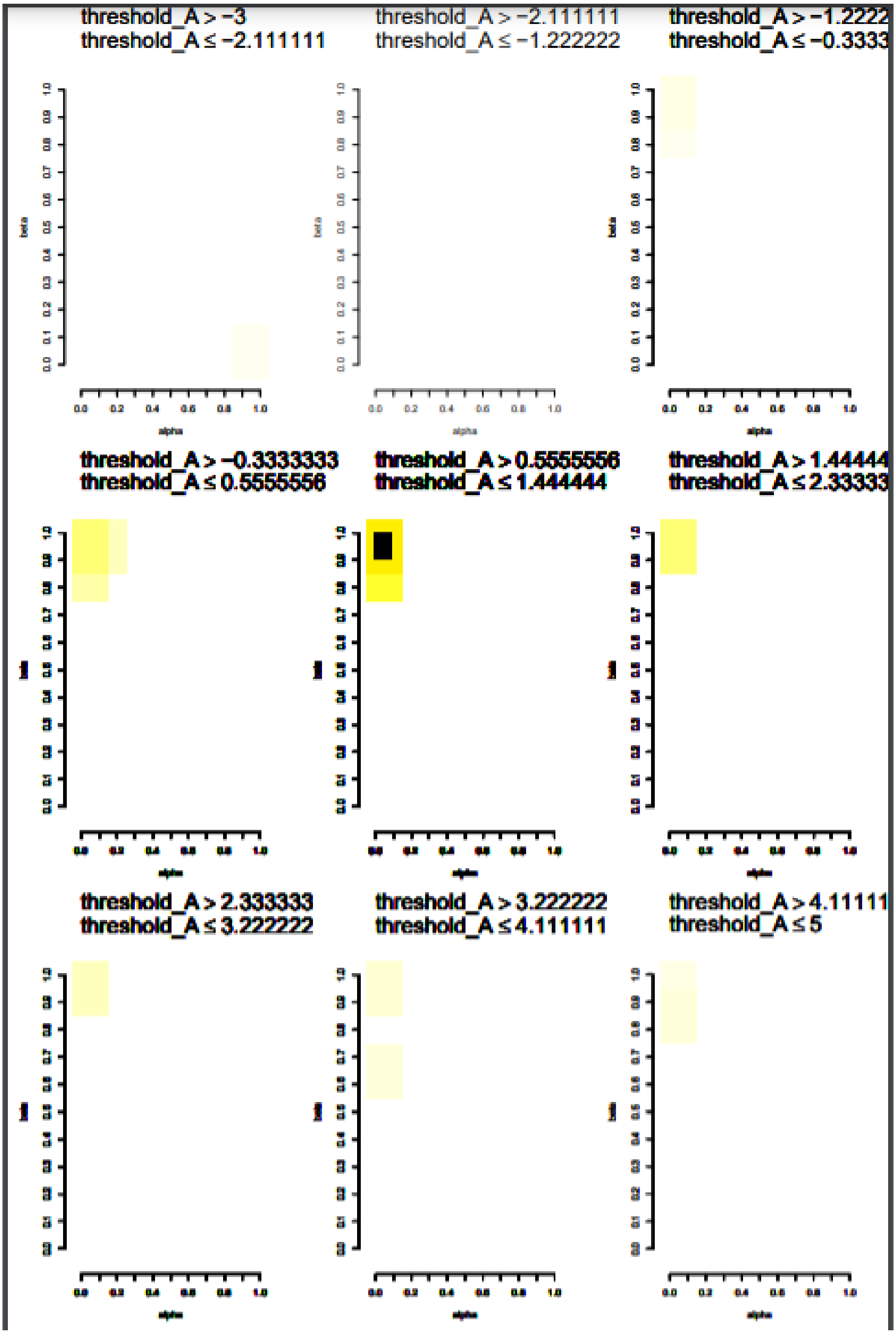
Appendix 5: $p_A = 0.1$, $p_B = 0.9$. Modular motivation, domain A



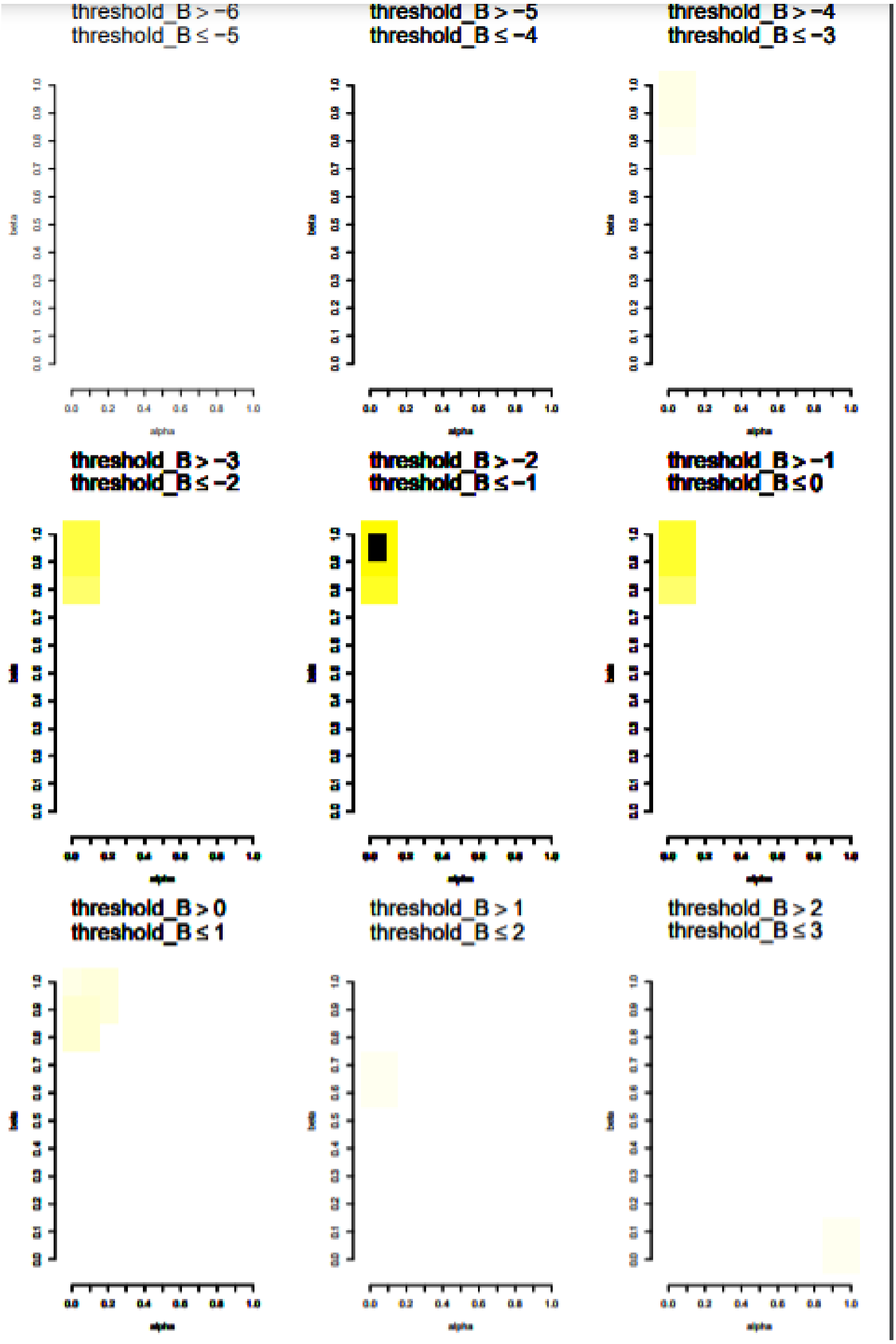
Appendix 5: $p_A = 0.1$, $p_B = 0.9$. Modular motivation, domain B



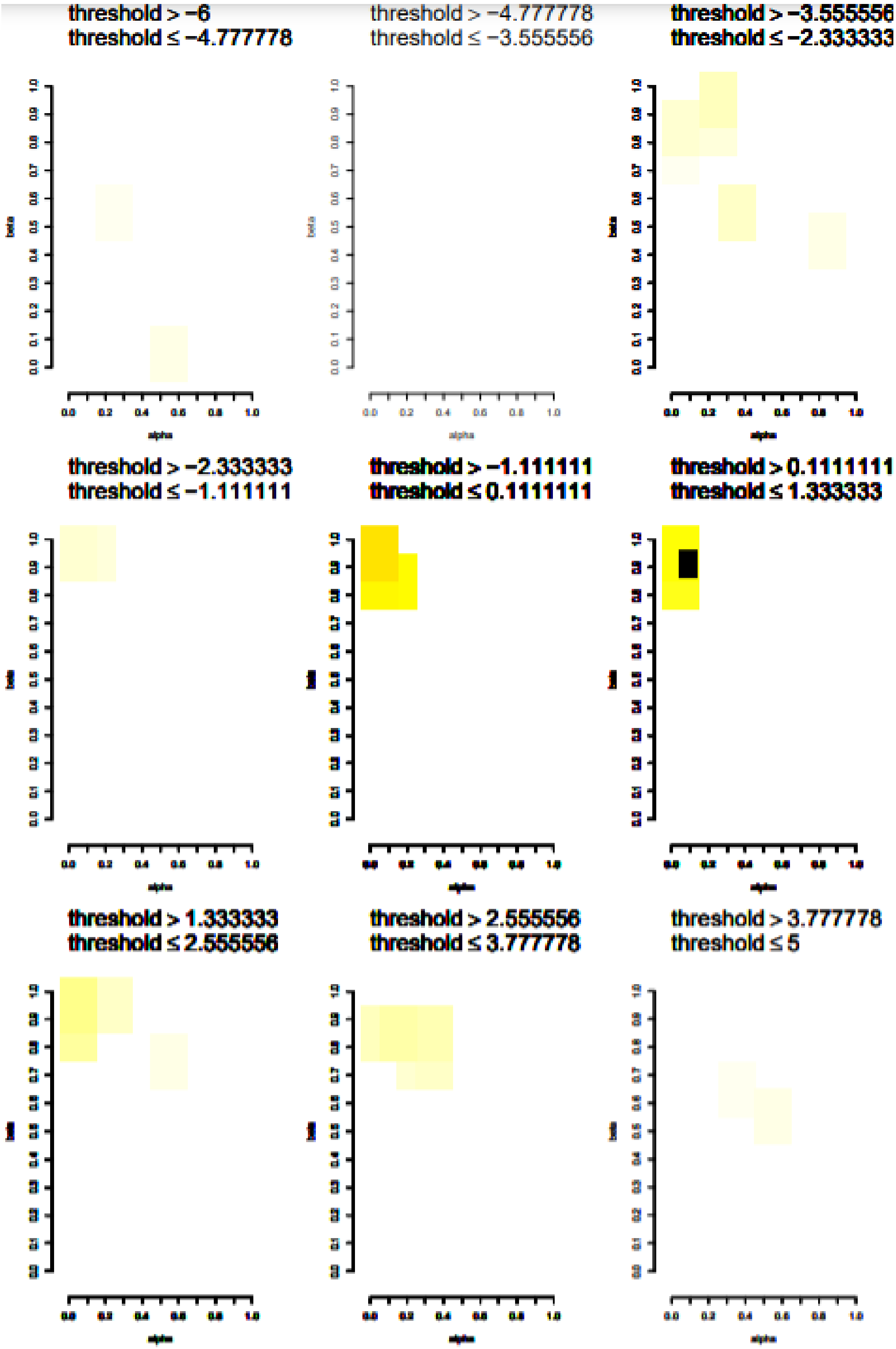
Appendix 5: p_A = 0.1, p_B = 0.9. Modular cognition, domain A



Appendix 5: $p_A = 0.1$, $p_B = 0.9$. Modular cognition, domain B



Appendix 5: $p_A = 0.1$, $p_B = 0.9$. Domain-general



Appendix 6: A regression predicting agent fitness based on whether they had modular or domain-general cognition, and modular versus domain-general motivation

This regression backs up the points argued in section 3.4. When the priors were skewed but consistent ($p_A = 0.1$, $p_B = 0.1$) and ($p_A = 0.9$, $p_B = 0.9$), there was no difference in fitness between agents with modular or domain-general cognition, though modular motivation was more fit than domain-general motivation. Finally, for the runs with skewed but inconsistent priors ($p_A = 0.1$, $p_B = 0.9$), we see that both agents with modular motivation and agents with modular cognition accrue *less* fitness than their domain-general counterparts, which supports my view that domain-general psychology will be selected for as it allows for mutual cooperation to occur by mistake (which results in a higher overall fitness for both cooperating agents).

Appendix 6. The regression results displaying any differences between domain-general and modular cognition, versus domain-general and modular motivation, in environments with skewed but consistent prior probabilities such as ($p_A = 0.1$, $p_B = 0.1$) and ($p_A = 0.9$, $p_B = 0.9$) and environments with inconsistent prior probabilities of state 1 in the two domains ($p_A = 0.1$, $p_B = 0.9$). The results in italic text depict runs where the benefits of cooperation were weak (b=2), whilst the results in non-italic text depict runs where the benefits of cooperation were strong (b=4). The domain-general agents were the omitted category of this regression as they were dummy coded as 0.

Prior Probabilities	**Estimates for regression predicting fitness**		
	Intercept	**Cognitive threshold dummy (0 = domain-general, 1 = modular)**	**Motivation threshold dummy (0 = domain-general, 1 = modular)**
$p_A = 0.1$, $p_B = 0.1$	*3.047 **** *(0.004)* 3.58 *** (0.011)	*-0.0007* *(0.004)* -0.016 (0.013)	*-0.006* *(0.004)* 0.0547 *** (0.013)
$p_A = 0.9$, $p_B = 0.9$	*3.037 **** *(0.0035)* 3.589 *** (0.01)	*<0.001* *(0.004)* -0.012 (0.013)	*-0.002* *(0.004)* 0.059 *** (0.013)
$p_A = 0.1$, $p_B = 0.9$	*3.133 **** *(0.005)* 4.48 *** (0.013)	*-0.007* *(0.006)* -0.391 *** (0.016)	*-0.096 **** *(0.006)* -0.687 *** (0.016)

The asterisks denote the significance of our p values with the following key:
** = 0.05*
*** = 0.01*
**** <0.001*
. = trend

Appendix 7: The linear combinations performed between all the modular agent types.

These back up the points argued in section 3.4, showing that both fully modular agents and partly modular agents with modular motivation only consistently had a higher fitness than partly modular agents with modular cognition only. Moreover, partly modular agents with modular motivation only were often indistinguishable in terms of fitness to fully modular agents. This suggests that if modularity does evolve, it may only be needed in the motivational system when it comes to cooperation.

Appendix 7. The linear combinations compare the fitness of all the modular agent types, for the regression reported in Table 3, section 3.4 of the main text. The environments have skewed but consistent prior probabilities such as (p_A = 0.1, p_B = 0.1) and (p_A = 0.9, p_B = 0.9) or inconsistent prior probabilities of state 1 in the two domains (p_A = 0.1, p_B = 0.9). The results in italic text depict runs where the benefits of cooperation were weak (b=2), whilst the results in non-italic text depict runs where the benefits of cooperation were strong (b=4).

Prior Probabilities	**Linear Combination**		
	Fully modular - modular motivation	**Fully modular - modular cognition**	**Modular motivation - modular cognition**
p_A = 0.1, p_B = 0.1	*F(1,39996)=0.98, p=0.32* F(1,39996)=0.56, p=0.45	*F(1,39996)=0.02, p=0.89* F(1,39996)=9.36, p=0.002	*F(1,39996)=0.76, p=0.38* F(1,39996)=15.85, p<0.001
p_A = 0.9, p_B = 0.9	*F(1,39996)=0.07, p=0.79* F(1,39996)=4.08, p=0.04	*F(1,39996)=0.004, p=0.95* F(1,39996)=2.35, p=0.13	*F(1,39996)=0.127, p=0.72* F(1,39996)=15.19, p<0.001
p_A = 0.1, p_B = 0.9	*F(1,39996)=0.05, p=0.82* F(1,39996)=0.51, p=0.48	*F(1,39996)= 252.15, p<0.001* F(1,39996)=204.46, p<0.001	*F(1,39996)=167.47, p<0.001* F(1,39996)=208.64, p<0.001

Chapter 8:
Conclusion

8.1: Thesis overview

Throughout this thesis, I aimed to investigate both the flexibility of human social learning behaviour and the complexity of human psychological processing from a gene-culture coevolutionary perspective as per the predicitions of the dual-inheritence framework. Specifically, I had two main research aims: (i) to investigate how flexible an individual's decision to conform is in response to social information about the group from whom she learns and (ii) to investigate whether complex decision-making is likely to be underlaid by a domain-general, or modular, processing system. These two research aims were explored via empirical studies and agent-based models respectively. These aims were linked by investigating three areas of decision-making: asocial skills, social norms, and cooperation. These were investigated using a game against nature, a coordination game, and a Prisoner's Dilemma game respectively. In this overview, I summarise the main findings of each chapter with reference to these two overarching research aims.

To begin addressing the first research aim, **Chapter 3** investigated whether the participants would adjust their frequency-dependent social learning strategies to social information about the group from whom they learned when mastering both an asocial skill (game against nature) or a social norm (coordination game). This social information included: (i) the frequency of choices amongst a group; (ii) whether these group members were identified as deciding in a similar or different decision-making environment to the social learner and (iii) the reliability of these similarity signals (reliably incorrect, uninformative, and reliably correct). This study was the first to my knowledge to find that individuals can adjust their social learning strategies to a third-order complexity. Most participants clearly tried to adjust to the reliability signals, though few could do so optimally. There may be a trade-off in the depth of social information that the social learners could process. They found it easier to master optimal skillsets and coordinate on optimal social norms when learning from groups of reliably similar others. This addresses my first research aim by showing flexibility on the part of the social learners, though there is likely to be an upper limit to the amount of social information that the learners can process.

Having established that individuals flexibly chose frequency-dependent social learning strategies when mastering both asocial skills and social norms, **Chapter 4** applied a similar methodological design to cooperation, by utilising a Prisoner's Dilemma**.** Again, the individuals had to base their decision to conform, to follow the majority in a linear fashion or to follow the minority on social information about the group from whom they could learn. This information included: (i) the frequency of choices (i.e., cooperation or defection) amongst the group; (ii) whether this group learned to cooperate in a similar or different decision-making environment to the participants and (iii) the reliability of these similarity signals. The social learners could adjust their social learning strategies up to a third-order complexity. They followed the majority of reliably similar others and followed the minority of those identified as similar with a reliably incorrect signal. Beyond this, the social learners did not process the reliability information from those identified as different. Again, there was a trade-off in how flexible a social learner was when choosing her strategies.

The second research aim addressed the flexibility of human decision-making processes. **Chapter 5** investigated how domain-general, partly modular, and fully modular agents came to master asocial skills in a game against nature. I found that domain-general psychology was sufficient when mastering two similar skillsets, though modular psychology was needed to master disparate skillsets over distinct domains. I also found that modular cognition experienced a greater evolutionary shift over the model's runs than motivation did.

I then used a coordination game to investigate how domain-general, partly modular, and fully modular agents came to master social norms. **Chapter 6** revealed that domain-general psychology was sufficient when the agents mastered two similar social norms, but modular psychology was necessary to master two widely different social norms over two distinct domains. Both cognition and motivation were influential when mastering a social norm. Modular cognition could track the different probabilities of social coordination occurring over two distinct domains. Modular motivation could respond to disparate fitness pressures in terms of which behaviour was the most fit to coordinate on across two distinct domains. The models revealed that the agents' cognition and motivation coevolved over the model's runtime to facilitate coordination, though whether they coordinated on the optimal or suboptimal social norms as a strategy was largely decided by drift.

Finally, **Chapter 7** investigated how domain-general, partly modular, and fully modular agents learned to cooperate (or defect) over two domains by using the Prisoner's Dilemma game. Interestingly, domain-general agents outperformed modular agents in this chapter. As domain-general agents struggled to track the probabilities and fitness pressures over multiple distinct domains, then they started to cooperate by accident. Mutual cooperation could arise as a 'big mistake' but only in domain-general agents. Motivation was also more important than cognition in influencing agent behaviour. Even if modular agents developed a cognitive bias to favour cooperation, their motivation could evolve in the opposite direction to compensate. Conversely, if the domain-general agents developed a cognitive bias to favour defection, then their motivation evolved to compensate for these cognitive dispositions. In sum, **Chapter 7** suggests that a domain-general psychology underlies the unique scale and complexity of human cooperation (Henrich & Muthukrishna, 2021). Motivation was more influential than cognition when driving behaviour. When it comes to cooperation, what we want to do may be more important than what we think we ought to do.

The following sections of my conclusion aim to break down and explore these findings further. Section 8.2 synthesises my social learning studies and relates these to the field of gene-culture coevolution. Section 8.3 synthesises my agent-based models and relates these to the fields of gene-culture coevolution and Evolutionary Psychology (EP). Section 8.4 then synthesises both the studies and the model findings together, and places these in a broader evolutionary context. Section 8.5 outlines some key strengths and implications of my work, and section 8.6 outlines some key limitations and directions for future research. Finally, section 8.7 concludes my thesis.

8.2. Research aim 1: Placing the studies on the flexibility of social learning into a gene-culture coevolutionary context

This thesis aimed to investigate the flexibility with which people chose to follow the majority– or minority–as a function of certain social information about the group from whom they learned. Across all three key areas of decision-making, the participants in **Chapters 3** and **4** showed a flexible use of frequency-dependent social learning strategies. All participants adjusted to (i) frequency information; (ii) a signal informing the participants whether they learned in a similar or different environment to the group and (iii) the reliability of this similarity signal.

In the introduction of my thesis (**Chapter 1,** section 1.2**)**, I discussed a view in the gene-culture coevolutionary field that social learning strategies, such as conformity, operate like rules-of-thumb (Gintis, 2003; Henrich, 2004; Henrich & Boyd, 2001;

Molleman et al., 2014). Applying these rules-of-thumb indiscriminately may lead to participants upholding incorrect behaviour (Asch, 1955) and to spillover effects (Brand et al., 2020; Marks et al., 2019). My empirical work in **Chapters 3** and **4** contradicted this viewpoint. Most participants tried to adjust to both the similarity and reliability information, rather than blindly copying any majority. This has implications for how we understand the function of conformity (Kendal et al., 2018). Rather than functioning like a simple rule, the decision to conform is responsive to social information about the group from whom we learn. Perhaps frequency-dependent social learning strategies function to help us to respond to the range of groups that we are likely to encounter throughout our everyday lives and that our forebears may have encountered throughout the ancestral past. These groups likely require flexible and complex social learning strategies to navigate (Boyd & Richerson, 1985 [chapter 7]; Deffner et al., 2020; Efferson et al., 2008b).

Despite this impressive flexibility, my participants showed an asymmetric adjustment across all three decision-making games, though to differing extents. All participants could adopt the same option as the group from whom they learned when they saw a reliably correct signal indicating that they learned from similar others (see Figure 4). This finding may suggest an in-group preference (Efferson et al., 2016). This contradicts a viewpoint that I discussed in the introduction of this thesis (**Chapter 1,** section 1.2); that individuals are completely flexible in their choice of social learning strategies (Evans et al., 2018; Kendal et al., 2018; Rendell et al., 2011). Across all three economic games– and regardless of whether they were British or Indian– the participants were more likely to copy the same behaviour as the majority of the group provided that this group were signalled to be reliably similar. With the exception of a few participants in each study who reported that they did not use social information, most participants showed this asymmetric adjustment when they tried to adjust to all three orders of social information. This contrasts with research suggesting that some participants always conform, and others choose to be flexible (Efferson et al., 2008a; McElreath et al., 2008). Rather than having a preference that was fixed within each individual, my analysis revealed that most participants did genuinely adjust to the three orders of social information (see appendix 1). I now compare the social learning strategies utilised across all three game types (see figure 4).

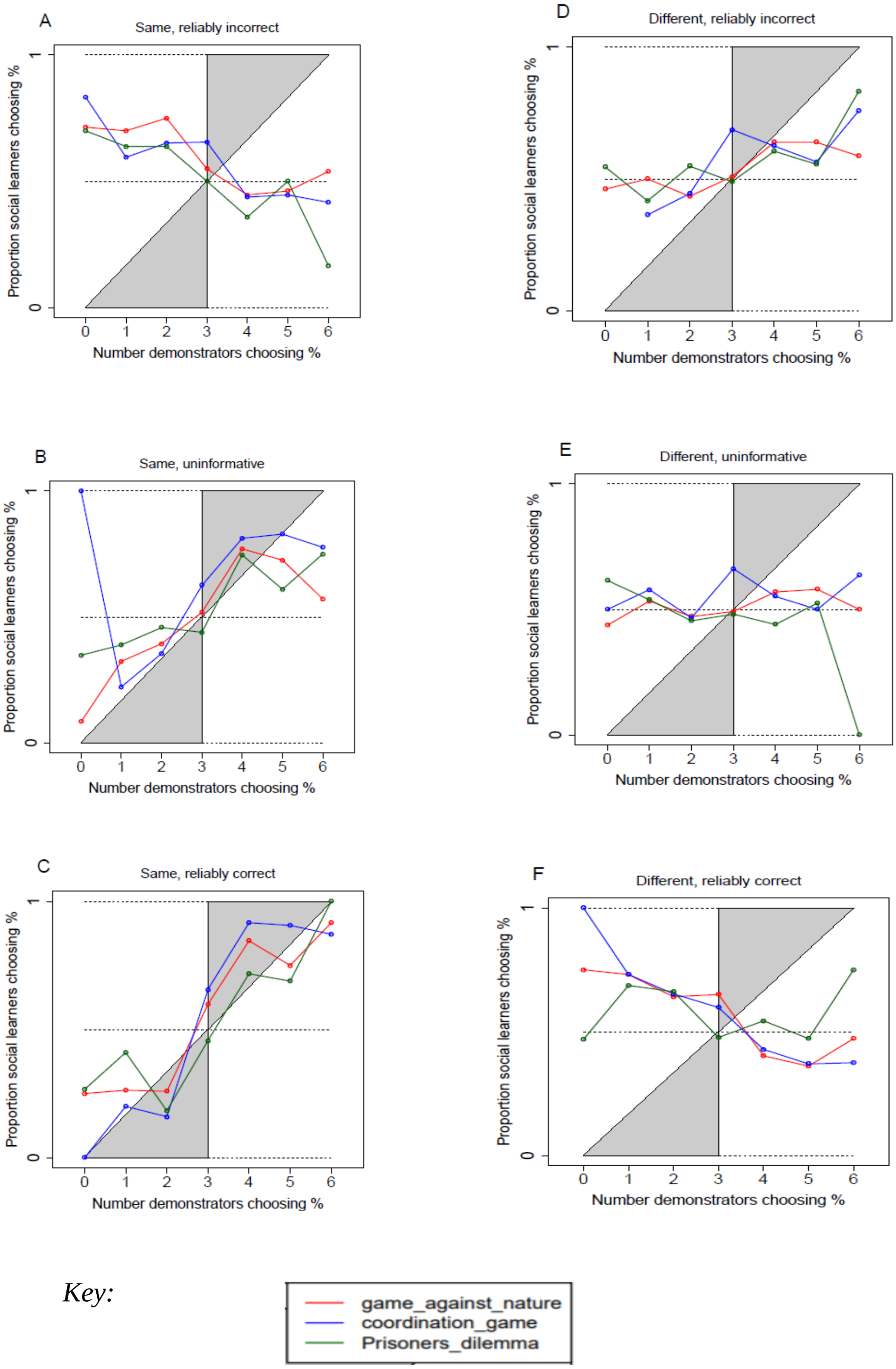


Figure 4. The proportion of social learners who chose % based on the number of demonstrators who chose %. The different panels show the social learners' responses to frequency-dependent social information by each level of the second- and third-order

social information. The regions shaded in grey depict where the social learners' data would fall if they used a conformist strategy, while the dashed lines give points of reference for proportions of learners choosing % at 0, 0.5, and 1. The colour denotes the games being played, with game against nature (asocial skills) in red, coordination game (social norms) in blue and Prisoner's Dilemma (cooperation) in green.

Starting with reliably correct signals of similarity, the social learners typically followed the majority on these blocks. As their preference to follow the majority around reliably similar others was the most pronounced social learning strategy, then the social learners were more likely to master the optimal norms and skillsets– and to cooperate– when learning from groups of reliably similar others than any other group. This asymmetric adjustment may suggest that the function of conformity– while flexible– is still tailored to learning scenarios that one would have been likely to encounter throughout one's everyday life, or throughout the ancestral past. If it is common to learn from reliably similar others, who were likely to have been our in-group members (Efferson et al., 2016; McElreath et al., 2003); then perhaps our social learning abilities reflect this commonality.

This bias to learn from reliably similar others may be genetic and/or cultural in nature. If there has been a high level of assortment into like-minded groups throughout the ancestral past, then there would have been higher payoffs for following the skillsets and norms of these like-minded others. Perhaps the preference to follow the majority of reliably similar others is therefore a genetically-evolved in-group preference (Henrich & Boyd, 1998). Alternatively, social learning rules may themselves be socially learned (Kendal et al., 2018; Heyes, 2016; Mesoudi et al., 2016). Perhaps the social learners found it easier to respond to reliably similar others in the current studies as they were used to copying reliably similar in-group members in their everyday lives.

Note that this thesis cannot inform whether this bias is genetic or cultural in nature, though this would be an interesting avenue for further work to explore.

When responding to reliably similar others (Figure 4C), there was a particularly pronounced conformity effect on the coordination game (social norms). As reliably similar others are likely to be in-group members, then this finding corroborates previous gene-culture coevolutionary work which suggests that the social norms of one's own group must be preserved with a high fidelity (Legare, 2019; Henrich & Muthukrishna, 2021; Molleman et al., 2013a). Indeed, these findings corroborate previous work which suggests that conformity should be more pronounced when learning a social norm than an asocial skill (Clegg & Legare, 2016; Legare, 2017; Legare & Nielsen, 2015; Wen et al., 2019).

Across all three economic games, the social learners were more likely to choose the same option as that which the minority of demonstrators chose, in response to similar others with reliably incorrect signals (see Figure 4A). The participants understood that those who were unlikely to be playing the same game as themselves were in fact likely to be playing a different game. In this case, the participants should choose the opposite option to the majority of this group. There is uncertainty as to whether this was a true preference to follow a minority per se, or merely a caveat of the task being binary. With only two options to choose from, the social learners may have just chosen the opposite option to that which the majority of demonstrators had chosen under both reliably different and unreliably similar signals (see section 8.6.2). Regardless, the fact that social learners could respond to these signals may suggest that signals of group membership have not always been reliable (Smaldino et al., 2018) and perhaps could have been faked (Sosis et al., 2007).

Moreover, the preference to follow the minority of similar others– or to choose the opposite option to the majority of similar others– with a reliably incorrect signal was the most pronounced during the Prisoner's Dilemma which measured cooperation. This suggests that the ability to process third-order reliability information may be key in one's response to free-rider infiltration. To return to an example that I have used throughout this thesis, picture an extremely cooperative group who share their wealth equally amongst their group members. Now imagine that these group members are identifiable as they all wear red. It would be relatively easy for an outsider to acquire red clothing and infiltrate this group. As this outsider is likely to be a free-rider who only intends to take resources from the group and give nothing in return, then it is prudent that such individuals are identified and can be ousted from the group. If, for example, the exact shade of red that the group wears is hard to replicate then any free-riders infiltrating the group may end up wearing clothes that are an off-shade of red. In this case, the off-shade of red becomes a reliable signal of who is faking group membership (and is therefore the equivalent of a reliably incorrect signal of similarity in my study).

Precisely because cooperative groups are at risk of free-rider infiltration, it may make sense that those participants who learned cooperative norms were ultra-aware of who could be faking a similar group membership. This supports previous work which suggests that the reliability of group signals can be questionable, especially due to free-rider infiltration of cooperative groups (Smaldino, 2019; Smaldino et al., 2018; Sosis et al., 2007). Taken together, our social learning strategies have a flexible function in response to groups that should be copied (e.g., reliably similar others) and to follow the minority behaviour in response to groups who should not be copied (e.g., reliably incorrect signals indicating the participants learned from similar others). In the latter case, the social learning strategies are a failsafe to protect against learning from those infiltrators with faked signals. In fact, these results may even be indicative of opposing those with faked signals. To return to an example that I have used throughout this thesis, consider the cooperative group who signal their group membership by wearing red clothing. If the agent is aware that the person with the off-shade of red clothing is a free-rider, then they may do the opposite of them on principle. For example, the focal agent will cooperate (with others) whenever the infiltrator defects (with others). This easier on tasks with two options to choose from, such as 'cooperate' or 'defect'.

The social learners followed the majority of different others with reliably incorrect signals (Figure 4D). The tendency to follow the majority around different others with reliably incorrect signals did not match the strong preference to follow the majority around reliably similar others, despite the fact that these conditions were equivalent. While it may be common to encounter reliably similar others– and it may unfortunately be common to encounter outsiders who fake a signal of similar group membership to oneself– it is perhaps rarer to encounter those that would pretend to be different. After all, failing to coordinate on a social norm would entail significant costs (Chudek & Henrich, 2011; Molleman et al., 2019b). Likewise, a failure to cooperate when others do would attract negative attention (Molleman et al., 2019b; Price & Johnson, 2011; van den Berg et al., 2012). Cooperating when everyone else defects would result in a substantial loss of resources (Capraro, 2013). Thus, the social learners did not respond to a reliably incorrect signal of difference.

This asymmetric adjustment may be indicative of social learning. If the participants had not experienced people who pretend to be different to themselves enough times, then they are unlikely to have formed a social learning rule regarding how to react around unreliably different others (Heyes, 2018). Alternatively, this asymmetric adjustment may suggest that social learners could only process signals that were likely to have an ancestral analogue. If it was rare for others to pretend to be different to each other throughout the ancestral past, then there would not be a strong selection pressure to have shaped social learner cognition to be able to respond to reliably incorrect signals of difference in the current study. If free-riders often faked group membership to infiltrate cooperative groups however, then social learner cognition would be shaped towards processing reliably incorrect signals of similarity only (Smaldino et al., 2018).

The coordination game in blue shows the strongest response to the reliability signal given alongside different others. Individuals may employ the opposite social norms to different groups on principle (Chudek & Henrich, 2011). This pressure is not necessarily equivalent for other decision-making tasks. For example, the social learners did not respond as strongly to the reliability signal given alongside different others when learning asocial skills, in red. Perhaps we do not need to do the opposite of a group of different others, as the optima of different others are often irrelevant to oneself and do not necessarily influence what one should do when acquiring one's own skills (Boyd & Richerson, 2005).

The social learners' response to uninformative signals is interesting (see Figure 4B and 4E). The social learners who saw an uninformative signal were maximally unsure as to whether they were playing the same or different game to the groups from whom they learned. Uninformative signals thus rendered all social learning strategies equivalent in terms of expected payoffs (Efferson et al., 2016). This left only preferences over the payoff-irrelevant characteristics of the games to have an effect. The social learners may just have just chosen @ or % randomly, and so their response to frequency information may have cancelled to chance at the aggregate level.

Alternatively, the social learners may have chosen to follow the majority– or the minority– in response to these uninformative signals simply because they had a preference for this social learning strategy.

Interestingly, there was no response to frequency information when the participants learned from different others with an uninformative signal (see Figure 4E). This suggests that the participants just chose randomly, and this cancelled out to no response at the aggregate level. However, there was a preferred response to uninformative signals of similarity (Figure 4B). Across all three economic games, the social learners followed the majority of similar others with uninformative signals. This confirms a preference to copy similar in-group members. This strategy may make sense as uninformative signals are not necessarily faked signals. For example, some groups signal their loyalties subtly, as they do not want to attract negative attention from outgroup members (Smaldino, 2019; Smaldino et al., 2018).

When responding to groups of reliably different others, the participants followed the minority when learning both an asocial skill (game against nature; red) and a social norm (coordination game; blue). Social norms of different others are typically opposed so that we can signal a different ideal to an outgroup (Chudek & Henrich, 2011). Likewise, some asocial skills will be opposed when they come to be associated with an outgroup. For example, Smaldino and colleagues suggest that some technologies are slower to spread within a population when the technology gets associated with one specific subgroup as others wish to be seen as different to this subgroup (Smaldino et al., 2017; Smaldino & Jones, 2021). However, this effect was not as pronounced when learning asocial skills as it was for social norms.

Thus far, there has been a remarkable similarity between how the predominantly Indian participants in **Chapter 3** and the predominantly British participants in **Chapter**

4 used social information when playing these games. This is interesting, as a large amount of gene-culture coevolutionary work suggests that there would be differences in the social learning strategies preferred by individualistic and collectivist participants (Bond & Smith, 1996; Molleman & Gächter, 2018). One reason my thesis found something different may be that the specific social information shown, and the economic games used, were more influential in driving frequency-dependent social learning strategies than the participants' cultural orientation. Alternatively, Indian and British participants may have more similar social learning preferences than those from highly-collectivist countries, such as China, who have been more often studied in past research (Bond & Smith, 1996; Molleman & Gächter, 2018). This is particularly likely, as **Chapter 3** recruited from the PUNE lab in FLAME University and **Chapter 4** recruited in the EconLab at RHUL. Both of these universities may have students who are more educated, rich, and democratic than the general population of India and Britain as a whole (Henrich, 2020). Thus, their social learning preferences may be quite similar due to both sets of participants being undergraduate students.

While the participants in **Chapter 3** (who learned asocial skills and social norms) clearly followed the minority around reliably different others, the social learners' response to reliably different others fluctuated around chance when they learned to play the Prisoner's Dilemma from **Chapter 4** (see figure 4F). Of course, this could have been due to differing preferences between my British sample who played the Prisoner's Dilemma game, and my Indian sample who played the other two games. Individualistic participants are less likely to use social information than collectivist participants (Mesoudi et al., 2015; Molleman & Gächter, 2018), which may explain why the individualistic British participants did not use the social information from reliably different others in **Chapter 4** of my thesis**.**

Alternatively, this differing response to reliably different others may be due to the structure of the Prisoner's Dilemma. The Prisoner's Dilemma is fundamentally different to the game against nature and the coordination game. The game against nature and the coordination game each have an objectively correct answer, in the sense that one option is expected to give the highest payoff if everyone chooses it. Of course, this blanket statement does not cover individual preferences for coordination where some individuals feel intrinsically rewarded for not coordinating on a popular social norm (Efferson et al., 2020a). The payoff matrix of the Prisoner's Dilemma does not give a clear optimum. Defection is said to be the Nash equilibrium, but mutual cooperation out-pays mutual defection (Holt & Roth, 2004). Thus, the Prisoner's Dilemma is likely to be influenced by social preferences that are heterogeneous across participants.

Moreover, it is perhaps clear that one should uphold different skillsets and social norms to those who make decisions in a different decision-making environment to oneself. However, the response to reliably different others may be less clear cut when learning to cooperate. Sometimes, participants do the opposite to a group of different others merely on principle (Chudek & Henrich, 2011; Efferson et al., 2016; Smaldino et al., 2017; Smaldino & Jones, 2021). Other times, the participants may use social information about how well a different cooperative group are doing in order to kickstart cooperation in their own group (a group replacement strategy that is referred to as assimilation; Bell et al., [2009]; Burton-Chellew & West, [2012]). In the latter case, then different others should be copied. Note that the participants merely knew the *expected* points from a payoff matrix. I did not allow the social learners access to information about the *actual* points made by the demonstrators. I did this as I wished to isolate frequency-dependent social learning strategies and investigate these alone. However, the participants may have needed information about the payoffs of cooperation to commit to a social learning strategy when it came to learning cooperative behaviours from groups of reliably different others. If a group of different others cooperate, and their groupwide payoffs exceed the group-wide payoffs of our own group, which is defective; then the social learners may attempt to assimilate cooperation into their own interactions (Molleman et al., 2013b).

To recap, these studies showed that on the whole, the three decision-making tasks of interest were learned similarly. Across all three games, the participants made an asymmetric adjustment as they were more likely to uphold the same behaviours as a group of reliably similar others. Tying back to research aim 1, the social learners' choice of frequency-dependent social learning strategies was flexible but there was an upper limit to this flexibility. Most social learners could process three orders of social information, but few could respond to the reliability signals completely symmetrically.

Perhaps these findings are best explained by Heyes' (2016) dual-system approach to social learner cognition. As discussed in the introduction of my thesis (**Chapter 1,** section 1.2), individuals employ a System 1 or a System 2 thinking process. System 1 is fast and effortless and may entail following a rule-of-thumb, such as 'always copy the majority'. System 2 processing is slow and effortful and may allow the individual to adjust their social learning behaviour to a wide variety of situations. Perhaps the asymmetric adjustment in my **Chapters 3** and **4** occurred as the participants could use a System 1 rule to copy the reliably similar others but had to engage in more effortful System 2 processing to respond to the other signals. As social learner cognition is unlikely to be fully flexible in the amount of information that one can process (Krafft et al., 2021), then the social learners may have struggled to completely engage System 2 thinking which may explain the asymmetric adjustment found in this study.

A failure to update to System 2 processing may explain how conformity upholds subpar behaviour. Consider an environmental shift which means that the majority of the group uphold a subpar behaviour (Deffner et al., 2020). Pure-conformist learners will only ever uphold this subpar behaviour. To return to an example from my introduction **(Chapter 1,** section 1.2**)**, imagine a group that often dance as a group ritual. A sudden shift in the environment might result in a loss of food sources. In this case, dancing would be costly as it wastes calories. However, if all of the individuals within this group are conformist learners, then the costly dance ritual will be maintained.

Chapters 3 and **4** found that a small subset of participants always chose the opposite social learning strategy to what was expected, and some reported that they ignored social information completely. Such individuals should be studied more intensively in the future. Perhaps these individuals who did not use frequency information– or used it in a suboptimal fashion– may help to prevent scenarios like the above where a costly behaviour would be maintained if the group consisted of all conformists. Perhaps these individuals are selected for as they can help 'break' cycles of suboptimal behaviour being upheld at the group level after a spatial or temporal shift in the environment (Deffner et al., 2020). This does not imply that these individuals avoid conformity in an attempt to be cooperative. These individuals are unlikely to understand that their preference to avoid conformity will affect group-level dynamics (Krafft et al., 2021). However, this thesis did not study the evolutionary trajectory of these different social learning strategies in a population and so any comments on the evolutionary mechanisms behind the social learning preferences observed in **Chapters 3** and **4** are purely speculative.

To summarise, the social learners in this thesis were impressively flexible in their choice of conformity and other social learning strategies. They based their strategies on (i) frequency information; (ii) a signal informing them whether the group from whom they learned had made decisions in a similar or different environment to themselves and (iii) the reliability of this similarity signal. The participants displayed an asymmetric adjustment across all three games, as all participants were more likely to learn an asocial skill, a social norm, and cooperative behaviour from a group of reliably similar others. These studies show that social learning strategies are flexible to a higher order of social information than previous studies have investigated, though there is likely to be an upper limit to one's flexibility when processing social information.

8.3. Research aim 2: Placing the models of the flexibility of agent decision-making into a gene-culture coevolutionary and Evolutionary Psychology context

The second research aim of my thesis was to investigate the flexibility of the psychological systems that underlie our ability to master asocial skills, social norms, and cooperative behaviours. To this end, **Chapters 5-7** compared agents with domain-general and modular psychology on their ability to master asocial skills (**Chapter 5**), social norms **(Chapter 6)** and cooperative behaviours respectively **(Chapter 7).**

While these models had quite different findings, it is worth noting that one overriding similarity was that human cognition– be it domain-general or modular– was not completely flexible. In **Chapter 5,** a skill with a lower payoff out of two may be adopted. In **Chapter 6,** a surprising number of runs resulted in groups which coordinated on a suboptimal social norm. In **Chapter 7,** some agents cooperated purely by mistake. There is likely to be an upper limit to the flexibility with which individuals can process the contrasting demands of different decision-making environments (Krafft et al., 2021). Interestingly, this upper limit arose without a fitness cost imposed on increasingly complex cognition. This suggests that another process, such as drift, can explain these trade-offs in suboptimal decision-making.

The main aim of these modelling chapters was to compare the learning of the three key areas of decision-making investigated throughout this thesis in domain-general, partly modular, and fully modular agents. Fully modular agents had cognitive and motivational processes that could specialise in each domain; partly modular agents had one psychological component (cognition or motivation) that could specialise in each domain, while the other component was left domain-general. Domain-general agents had psychological components which had to process the information over two domains simultaneously.

For asocial skills in **Chapter 5,** I found that domain-general psychology was sufficient in cases where the two asocial skills being learned were similar. Modular psychology was instead necessary when mastering two different skillsets, with different fitness payoffs and different probabilities of the skills being needed. **Chapter 6** presents a similar finding for social norms. When agents coordinated over two similar domains, then domain-general psychology was sufficient. When agents coordinated over two distinct domains (as in the probability of coordination, and the fitness pressures tied to coordination, differed widely over the two domains), then modular psychology was needed. **Chapter 7** provides contrasting results. It was the domain-general agents who cooperated by mistake on runs where the payoffs to mutual cooperation were sufficiently high. This allowed the domain-general agents to accrue more fitness.

The findings of **Chapter 7** are clearly disparate to the findings of **Chapters 5** and **6.** Human societies are extremely cooperative (Chudek et al., 2013; Gintis et al., 2008; Henrich et al., 2001) and differ in the domains of cooperation (Chudek & Henrich, 2011; Henrich & Muthukrishna, 2021). Domain-general cognitive systems, which flexibly processes the cultural rules regarding cooperative behaviours, may explain this work (Bolhuis et al., 2011; Mesoudi, 2008a). Indeed, the findings of **Chapter 7** corroborate this research. It seems that mutual cooperation does rise from a domain-general system, though this interestingly arose as a 'big mistake' in my model. Domain-general agents could not track the contrasting fitness pressures over both domains to distinguish between cooperation and defection when uncertain as to which behaviour was cooperative across two domains.

As well as investigating domain-general versus modular psychology, another strength of my thesis is that I investigated both cognitive and motivational processing and so I could explore the relative impact of both systems on the behavioural outcomes of my agents. In **Chapter 5,** modular cognition was clearly more influential in driving skill-learning than motivation. This was seen by the relatively large shifts in modular cognitive processing in response to widely different domains, and by the fact that partly modular agents with modular cognition only continuously gained higher fitness than partly modular agents with modular motivation only.

Cognition and motivation were equally influential in driving agents' social coordination behaviour in **Chapter 6**. This has implications for Evolutionary Psychology as this paradigm mostly focuses on explanations of cognitive processing. For example, social norms were argued by Sperber and colleagues to be the result of modules that are organised in the brain with hierarchical complexity (Sperber, 1994; Sperber & Hirschfeld, 1999; 2004). The increased role of motivation that I found in **Chapter 6** would suggest that, even if modular cognition is important, it is only part of the story.

In **Chapter 7,** motivation was more influential in driving the agents' cooperative behaviours than their cognition was. This can be best seen in the heatmaps (see figure 3, section 3.2 of **Chapter 7**). Whenever the agent experienced a shift in their cognitive thresholds which may have led them to believe that cooperation was common among their group (regardless of how accurate this belief actually was), then their motivation evolved to compensate for this cognitive bias. To walk through an example, consider modular agents who typically defected. Imagine they have a cognitive bias to believe that hunting more is cooperative, perhaps as the meat can be shared with one's group. However, this modular agent actually lived in an environment where game was scarce and so over-hunting would be defection. The agent is extremely motivated to hunt more. Despite the fact that the agent *believes* that hunting more is cooperative in this case, she is in fact defecting when one considers the actual state of the environment, where animal prey may have been overexploited and so one must hunt in sustainable ways (Safin et al., 2015).

The results of **Chapter 7** may contradict the Evolutionary Psychology explanation of cooperation. First, cooperation is assumed in Evolutionary Psychology to arise from a cognitive bias in agents as, during the Pleistocene environment more than10,000 years ago; our ancestors would have lived in small, tight-knit groups where cooperation would be encouraged and could be personally tracked (Curtin et al., 2020; Price, 2008; Price & Johnson, 2011). Although cooperation would make sense for our ancestors, it would be costly today as societies are larger and more anonymous (Price, 2008). My findings in **Chapter 7** cast doubt on this as, if cooperation is a 'big mistake', then it can only come from domain-general psychology and not modular cognition, as Evolutionary Psychologists typically predict.

Second, another explanation for costly cooperation in the evolutionary literature is the Error Management Theory (EMT). As discussed in-depth in **Chapters 1** and **7,** EMT is the idea that one should lean on the side of caution whenever one is uncertain about the likely payoffs of one's behaviour (Haselton et al., 2015). Specifically, whenever we meet new people, there is uncertainty over whether we will meet this individual again. Throughout the ancestral past, meeting somebody once suggested that you would meet them again (Krasnow et al., 2013). In this case, it is costly to falsely assume that one will not meet this person again, as then the agent may defect which could anger this new individual and scupper the chances for a long-lasting, mutually beneficial relationship. This is more costly than falsely assuming you will meet another person again and giving a one-off donation of resources that is never reciprocated (called one-shot cooperation). In this case, one-shot cooperation may arise as a cognitive bias on the part of the agents as it is usually the least costly error when one is uncertain about the longevity of any future interactions (Delton et al., 2011; Delton et al., 2013; Krasnow & Delton, 2016). The results of **Chapter 7** contrast EMT, as motivation was always more important than cognition in driving the agent's cooperative behaviour.

The findings of my models thus have implications for Evolutionary Psychology which, as a theory of cognitive processing (Tooby et al., 2006), may have focused on cognition at the expense of motivation in previous work (Delton et al., 2011). Future research is needed to clarify why the distinction between cognitive and motivational processes is important. As asocial skills could be mastered privately, then perhaps cognitive processing is sufficient to master these. As cooperation does not have a true optimum, and as one has to accept a personal cost in order to provide a benefit to another; then perhaps one's desire to help becomes increasingly more important than trying to 'rationally' assess the situation.

An important aspect of this thesis was that I did not commit to a modular or domain-general viewpoint but instead investigated how both types of agents came to make decisions, and investigated a partly-modular agent, too. If modular cognition is important, as Evolutionary Psychology predicts, then my models suggest that it was likely to only be important when mastering asocial skills (see **Chapter 5)**. This may be why the most consistent evidence for modularity comes from individual processes, such as colour perception (Schalk et al., 2017), and face detection (Rosenthal et al., 2017). Even cheater-detection, a social task, is usually measured via an individual's ability to solve logical 'if-then' rules. Therefore, cognition– modular or otherwise– will be more influential in driving participant performance on this logic task (Cosmides et al., 2010; van Lier et al., 2013). Domain-general psychology, particularly motivation, was instead important when deciding to cooperate (see **Chapter 7).**

Perhaps domain-general psychology was also important when mastering asocial skills and coordinating on social norms, but the design of my models did not capture this caveat. If the agents were allowed periods of trial-and-error or social learning, then these generic learning processes may have become influential in skill and norm learning (Bolhuis et al. 2011; Heyes, 2016; Mesoudi, 2008a). It is worth noting that domain-general processing was sufficient when mastering both asocial skills and social norms, provided that the domains were similar in terms of fitness demands and probabilities of occurring.

Modular psychology was only necessary for skill and norm learning over two distinct domains. By modelling the probabilities of certain environments independently across the two domains, then these models may have increased the likelihood of finding modular agents. In reality, certain domains are likely to become linked. For example, those that cooperate to hunt are likely to cooperate in other domains as well (Hill, 2002). Groups who are extremely cooperative when hunting tend to play economic games cooperatively (Henrich et al., 2001; 2006). These cooperative domains themselves may become linked to social norms regarding morality (Hill & Gurven, 2004) and rituals (Henrich & Muthukirshna, 2021; Wen et al., 2020). Perhaps if I allowed the environments in certain domains to become probabilistically linked to reflect this caveat, then I may have found a role for domain-general processing when mastering asocial skills and social norms as well as cooperation.

In summary, human decision-making was relatively complex. Both cognition and motivation coevolved to influence human decision-making, suggesting that previous work which focuses on the two processes in isolation may be limited in scope (Delton et al., 2011). The models gave conflicting findings as to whether psychology is likely to be domain-general or modular, and as to whether cognition or motivation were more influential in driving agent behaviour. Perhaps the exact nature of human decision-making is influenced more by the behaviour being learned. Even so, a consistent finding across my models was that human decision-making was flexible but not completely so.

8.3.1: The importance of drift

Finally, I note another interesting finding in my agent-based models. There was a large role of drift seen in **Chapters 5** and **6.** It is pivotal to learn a skill, but whether we adopted an optimal or a suboptimal asocial skill seems dependent on drift. In the context of **Chapter 5,** this was given by the example of a run where the most commonly needed skill was not the one with the highest payoff. Behaviour (0 or 1) that matches the environment was always selected for in this model. But the agent strategy underlying this behaviour may have been to lean towards acquiring the common skill with a lower fitness payoff; or towards a skill which is rarely needed but, when the opportunity does present itself, will lead to a higher fitness payoff.

To illustrate with an example in the hunting domain, it would be useful to learn how to build traps. These traps can only catch small animals however and so provide a low calorific resource. Even though it is rarer to encounter a large animal, it may pay off to master long-range weapons for this task. In such an environment, it seemed relatively down to drift (or random fluctuations via the course of evolution; Rorabaugh [2014]) whether the final generations were more favourable of the commonly-needed skill with lower payoffs (e.g., trap-building) or the rarer skill with higher payoffs (e.g., mastering the bow-and-arrow for larger prey hunted over a longer distance). It is interesting to note that both modular and domain-general agents experienced drift in their strategies.

In **Chapter 6,** all generations came to coordinate. Coordination was the behaviour selected for, but the strategy underlying coordination varied between runs. Some populations coordinated on the optimal social norm, and some coordinated on the suboptimal social norm. The exact strategy for social coordination may have been largely driven by drift.

The role of drift may be due to the imbalance in evolutionary trajectories that favour small but non-trivial error costs (Efferson et al., 2020b). That is, the difference between no payoff (as would occur if one did not master a skill or could not coordinate on any social norm) and a small payoff (as would occur if one mastered a suboptimal skill or coordinated on a suboptimal social norm) is perhaps more influential than the difference in payoffs between suboptimal behaviour and optimal behaviour. This corroborates previous research suggesting a large role of drift in some cultural trends, like baby names and pottery design (Bentley et al., 2004; Billiard & Alvergne, 2018).

Interestingly, there was less of a role of drift when it came to cooperation. The typical behaviour that was selected for in **Chapter 7** was defection, and so mutual defection was the most common outcome. This was especially true among modular agents. The domain-general agents, who could not balance their cognitive and motivational thresholds over the contrasting demands of multiple domains, had more noise in their psychological architecture. This noise may have led to more varied behaviour on the part of the agent, and this may have given the opportunity for drift to act on the agents' strategies. Eventually, this drift may have led to a shift away from unconditional defection towards unconditional cooperation in the domain-general agents' strategies.

If drift is influential in the behaviour that we come to coordinate on, then perhaps an interesting extension to this model would be to consider individualism versus collectivism orientation (Hofstede, 1980). A wealth of research has considered the difference in social learning strategies between individuals of the two different orientations (Bond & Smith, 1996; Mesoudi et al., 2015; Molleman & Gächter, 2018), but perhaps one's individualism-collectivism orientation also influences how one would respond to drift in social norms (Muthukrishna & Schaller, 2020). Cultural groups with tighter social networks may experience a more pronounced effect of cultural drift (Jung et al., 2021).

Perhaps cultural orientation mixed with drift could account for how different social groups may come to converge on different maladaptive social norms (Boyd & Richerson, 2007). Groups may converge on norms such as female genital cutting (Efferson et al., 2020a), foot-binding (Gavrilets, 2020) and suicide trends (Mesoudi, 2009). Behaviours which are typically associated with a cost to the individual may still be upheld as a social norm in cases where this behaviour is common, as then it becomes more individually-beneficial to coordinate on the same behaviour as the rest of the social group (Efferson et al., 2020a). Perhaps these groups uphold maladaptive social norms as some random noise influenced their evolutionary trajectory towards a strategy of coordinating on a suboptimal social norm.

Now that I have summarised the findings of both my empirical studies (section 8.2) and my agent-based models (section 8.3) in relation to the gene-culture coevolutionary framework, section 8.4 will now tie these two research aims together and place my thesis in a broader evolutionary framework.

8.4. Synthesis of social learning studies and agent-based model findings in the field of gene-culture coevolution

The following section will now tie the main findings of my social learning studies and my agent-based models together. I start by comparing the three economic games to highlight if there are any differences in decision-making and/or social learning when mastering asocial skills, social norms, and cooperative behaviours (section 8.4.1). I then discuss why both the social learning studies and the agent-based models imply an upper limit to the flexibility of social learning behaviour and the psychology of decision-making (section 8.4.2). Finally, I discuss the implications of my findings in supporting a norm psychology effect (section 8.4.3).

Section 8.4.1: Comparing the three economic games

Overall, there was a remarkable similarity between how asocial skills, social norms and cooperative behaviours were acquired. Both the social learning studies and the agent-based models suggested that individuals were flexible when adapting these behaviours, but there was an upper limit to this flexibility.

One unique finding was the suboptimal coordination observed in the coordination game. In **Chapter 3,** the social learners would conform to an uninformative signal of similarity. This enabled them to coordinate with their partners, though they were more likely to coordinate on the suboptimal option. Likewise, the agents in **Chapter 6** evolved to coordinate. Though it was mostly down to drift whether the agents tried to coordinate on the optimal or suboptimal social norm as a strategy. One explanation for this may be due to a cost asymmetry implied in the structure of coordination games specifically. Although coordinating on the socially suboptimal norm gives lower payoff to coordinating on the socially optimal norm, it is still much preferable to miscoordination (Kets et al., 2021). Indeed, Efferson et al. (2020b) argue that cost asymmetries do not need to be large to influence evolutionary outcomes. They merely have to be non-trivial. The difference between no payoff (or even negative payoffs) of miscoordination during a coordination game, and the small payoff of suboptimal coordination is non-trivial and thus may be more meaningful than even a large difference in payoffs between suboptimal coordination and optimal coordination.

The behaviour which was the most unique throughout my thesis was cooperation (as measured with the Prisoner's Dilemma). In **Chapter 4,** the participants who learned to cooperate did not respond to reliably different others, though the participants who learned asocial skills and social norms in **Chapter 3** would follow the minority of reliably different others. In section 8.2, I hypothesised that this different finding may be due to a unique role of different others when learning to cooperate. Sometimes we oppose the cooperative ideals of different others on principle (Wen et al., 2020), and other times we may have to assimilate cooperation from different others when we realise that this strategy pays off (Bell et al., 2009; Burton-Chellew & West, 2012; Molleman et al., 2013b).

Likewise, **Chapter 7** was the only model in which domain-general agents were selected for, and the only chapter where motivation was more important than cognition in driving behaviour. When it came to cooperation, what we wanted to do was more important than what we thought we ought to do. If domain-general motivation was important in explaining cooperation in **Chapter 7,** then this may suggest a role of flexible cultural processes in shaping our cooperative preferences (Kroneisen & Bell, 2021). When synthesising these findings, this may explain why cooperative behaviour was socially learned differently to both social norms and asocial skills in the empirical studies. Perhaps there is a unique role of culturally-learned rules, and motivation to cooperate, that must be considered when learning this behaviour.

It is particularly interesting that the Prisoner's Dilemma produced different social learning behaviour to the coordination game, and that the psychology underlying decision-making in this model was different to the coordination game. A growing body of literature suggests that cooperation is learned like a social norm (Dawes et al., 2007; Henrich et al., 2001; Henrich & Muthukrishna, 2021). My findings instead suggest that there were subtle differences between how a social norm and cooperative behaviours are upheld. Perhaps this is due to the structure of the games. In the coordination game, there is an objectively 'correct' answer to choose as one behaviour will have a higher payoff than the other when the players coordinate (Kets et al., 2021). For the Prisoner's Dilemma, defection gives more at the individual level and so is considered the Nash equilibrium (Holt & Roth, 2004). However, as the payoffs to mutual cooperation outweigh the payoffs to mutual defection (Holt & Roth, 2004), then the game can produce a lot of variation in the strategies that different agents prefer (Zimmerman &

Efferson, 2017). That is, the choices made during a Prisoner's Dilemma may be largely driven by personal preferences. This translated to different social learning strategies, and a different role of motivation, in both the studies and models respectively.

8.4.2: Social learning and psychological processing are flexible, but they are not equally flexible to all decision-making scenarios

In the empirical studies, the social learners tried to adjust to all three levels of social information, but they made an asymmetric adjustment as they found it easier to learn the three key decision-making games from groups of reliably similar others (**Chapters 3** and **4)**. These findings may be echoed by my agent-based models. Although the agents were complex in the behaviours that they upheld, there were trade-offs in their ability to master each task. Occasionally, the agents would adapt a sub-par skill (**Chapter 5).** Occasionally, the agents would coordinate on a suboptimal social norm (**Chapter 6).** Occasionally, the agents would cooperate by mistake (**Chapter 7).**

Taken together, the results of these agent-based models and my empirical studies may support the notion that cultural evolution can be 'blind' (Boyd & Richerson, 2007; Mesoudi, 2008b). That is, cultural evolution may sometimes uphold suboptimal behaviour in a similar way to how genetic evolution can sometimes lead to poorly-designed systems as evolution must co-opt the structures that already exist. For example, the human eye appears to be designed backwards (Dawkins, 1996). In a similar fashion, cultural evolutionary processes may stabilise social information that is actually harmful in cases where the majority of the group have already come to coordinate on this behaviour via drift.

Indeed, the current thesis suggests that suboptimal behaviour may be the result of a psychological system that is not fully flexible to processing contrasting environmental demands over multiple domains, or social learning strategies that are not fully flexible to a range of social information about the groups from whom we can learn. Of course, the influence of both cognitive and social learning trade-offs is likely to be key in explaining the suboptimal behaviours that are upheld at a group level. For example, groups uphold costly behaviour such as Female Genital Cutting (FGC; Efferson et al., 2020a); witchcraft beliefs (Tanaka et al., 2009) and mob behaviour (Raafat et al., 2009). Section 8.4.3 now turns to how these findings may be explained by a norm psychology effect.

8.4.3: Norm psychology effects in this thesis

When linking the findings of my empirical studies to the agent-based models, I may highlight a norm psychology effect. To begin, **Chapter 4** merely told the participants the frequency of @ and % choices amongst the group from whom they learned. They did not have information regarding whether the group upheld cooperation or defection. This was done to reflect realistic uncertainty as to which behaviour is considered cooperative (Chudek & Henrich, 2011). Interestingly, the results revealed that the participants typically cooperated whenever the majority of the group had and that they typically defected whenever the majority of the group had. This implies that it was the *act of trying to coordinate* with a partner or group that was more influential in driving social learning than the specific behaviour being learned. This corroborates previous work which suggests that children can even learn to punish cooperative behaviours, as they merely focus on the behaviour which is being punished, rather than crediting the cooperative intentions underlying this behaviour (Abbink et al., 2017; Bhui et al., 2019; Salali et al., 2015).

Further reinforcing this need for coordination, consider the response to uninformative signals when playing the coordination game in **Chapter 3.** The participants followed the majority of similar others with uninformative signals during a coordination game. This strategy clearly helped the social learners to coordinate, though they were more likely to coordinate on the suboptimal behaviour. This is supported by the agent-based model in **Chapter 6,** as the agent evolved to coordinate. However, whether the agents adapted a strategy of coordinating on the optimal option, or the suboptimal option, was largely down to drift (Rorabaugh, 2014). Social coordination is important. On some runs, the agents came to coordinate on the behaviour with the highest payoff. On other runs, the agents came to coordinate on the behaviour which was common, regardless of the payoffs. These latter runs may end in populations that coordinate on a suboptimal social norm simply via drift.

These findings together may support a norm psychology effect (Gintis, 2003; 2004; Henrich & Muthukrishna, 2021). Perhaps we merely have an evolved preference to coordinate, but the exact behaviour that we coordinate on may be shaped by social institutions. For example, a police force introduces punishment for those that fail to stick to the most fundamental social norms (Chudek & Henrich, 2011), while religion may increase social coordination by offering supernatural rewards for those that conform to certain ideals, and supernatural punishments for those that do not (Gray & Watts, 2017; Lenfesty & Morgan, 2019).

These results have important implications in how we understand the prevalence of maladaptive social norms. For example, consider Female Genital Cutting (FGC; Efferson et al., 2015; 2020a; Novak, 2020). This social norm has received a lot of research in policy change, not only because of the harmful nature of the tradition but because some societies have not reached a pure equilibrium (Efferson et al., 2015;

Novak, 2020). That is to say, even amongst ‘pro-cutting’ countries, there are very few societies where all of the girls are cut. This suggests that the populations of these countries currently exist in a two-strategy equilibrium, where some girls are cut, and some girls are not cut.

This is very important, as agent-based models like the one presented in **Chapter 6** of my thesis, suggest that any drift– or random noise– that is introduced at this point could have a unifying effect on the strategies selected within a population. If left to chance, then this drift may unfortunately result in societies that come to coordinate on a pure-equilibria strategy of cutting girls. However, nudges and other policy interventions that are introduced in countries existing at this two-strategy split can be effective in tipping the scales towards a pure-equilibrium strategy of not cutting. Researchers like Efferson et al (2015; 2020a) and Novak (2020) have done promising studies into the effects of behavioural nudging and other policy change that can help to reduce pro-cutting attitudes. The findings of **Chapter 6** in the current thesis suggest that this line of research is a valid way to reduce the chances of a group coming to coordinate on a maladaptive social norm. As drift was seemingly influential in the behaviours that the agents upheld throughout my models, then this suggests that work– like behavioural nudging– can harness this drift to drive behaviour towards more desirable end goals.

Taken together, this thesis has found that social learning was flexible, but not fully flexible, when processing three orders of social information (as I was chiefly interested in investigating for my first research aim); and that the psychology underlying complex decision-making was flexible, but not fully flexible when making decisions over multiple domains with many conflicting environmental inputs (as I was chiefly interested in investigating for my second research aim). I have also highlighted the large role that drift plays in influencing the strategies that one uses to acquire skills, or to coordinate on social norms. Taken together, this thesis may support the notion that cultural evolution is blind (Mesoudi, 2008b). My thesis suggests that costly group behaviour can be upheld via a mixture of social learning trade-offs and trade-offs in decision-making structures.

8.5. Strengths and implications of work

I now turn to the strengths and implications of my thesis, by first addressing the strengths of my study chapters (8.5.1) and my modelling chapters (8.5.2) individually, before tying these together in section 8.5.3.

8.5.1. Strengths of the social learning studies

Previous research tends to focus on just one game at a time (e.g., asocial skills; Mesoudi et al., 2015; Miu & Morgan, 2020; Reader, 2003). The only previous work to my knowledge to compare the three key decision-making tasks investigated throughout my thesis was Molleman and Gächter (2018), though they focused on whether the participants would conform, versus follow the highest earner (payoff-based learning), versus using their own trial-and-error. Instead, my empirical work focused on frequency information only which allowed me to isolate whether the individual's choice of frequency-dependent social learning strategies specifically remain flexible to the different social information about the group from whom the participants could learn. This was important for shedding light on the likely function of conformity, and other frequency-dependent social learning strategies (Kendal et al., 2018).

Another strength of my studies is that I move away from testing only WEIRD participants (Henrich, 2020; Henrich et al., 2010). Indeed, I recruit **Chapter 3** in Pune, India. I chose to recruit in India to move away from this bias to recruit only Western undergraduates. By only focusing on British or American undergraduates, most previous research tends to recruit highly-individualistic participants. Some recent research has recruited samples beyond these highly-individualistic participants, by recruiting highly-collectivist participants instead, such as Chinese or Japanese students (Mesoudi et al., 2015; Molleman & Gächter, 2018; Muthukrishna et al., 2020). The issue with this is that it only reflects social learning at the extreme ends of Hofstede's (1980) individualism-collectivism scale. Countries in the middle of this orientation, such as India, may be overlooked. Whilst individualistic participants have been shown to use trial-and-error, and collectivist participants have been shown to use social information (Bond & Smith, 1996; Mesoudi et al., 2015); we cannot necessarily assume that participants from the countries in the middle of Hofstede's (1980) individualism-collectivism scale will fall exactly in the middle of these preferences for social information use. In the current thesis, Indian participants typically learned very similarly to the British participants. Of course, this may be due to the fact that both participant samples were recruited at universities. Future studies should expand beyond student samples, as students are typically more educated, democratic, and rich than the rest of their country (Ekuni et al., 2020; Henrich, 2020).

8.5.2: Strengths of the agent-based models

I found a large component of drift influencing the strategies that my agents upheld throughout all agent-based modelling chapters. The importance of drift and other stochastic (i.e., random) effects have typically been overlooked in traditional gene-culture coevolutionary models (Billiard & Alvergne, 2018), and so my findings open an avenue for future research to explore the importance of this process in more detail.

As well as highlighting an important role of drift, my agent-based models were the first to my knowledge to investigate the coevolution of both cognition and motivation on an equal footing, and to compare these processes for domain-general, partly modular, or fully modular agents. Previous work tends to commit to a domain-general or modular perspective based on purely theoretical grounds (Burke, 2014; Bolhuis et al., 2011; Fodor, 2001; Pietraszewski & Wertz, 2021; Stephen, 2014; Stokes & Bergeron, 2015).

By investigating both processes, the findings of my model may be unique in suggesting that modular cognition may be important for underlying skillsets only, while domain-general psychology (specifically, motivation) is likely to uphold the costly levels of cooperation upheld across human societies (Henrich & Muthukrishna, 2021; Kroneisen & Bell, 2021). These models may shed light on which behaviours are most likely to be upheld by cultural evolutionary mechanisms (cooperation) and which behaviours may be partly explained by modular cognitive biases (asocial skills). Modular biases may be genetically evolved (Cosmides & Tooby, 1994b) or may emerge over the course of our development (Müller, 2007; Reader, 2006). Future work should seek to corroborate whether modular cognition is important for mastering skillsets and, if so, when and how these modules are likely to emerge.

8.5.3. Overall strengths and implications

An overarching strength of this thesis is that the results of the empirical studies in **Chapters 3** and **4,** and the results of the agent-based models in **Chapters 5-7,** are comparable as I used the same type of economic games throughout (game against nature, coordination game and Prisoner's Dilemma). Using these games gave a certain amount of structure to social interactions that are typically messy to disentangle which allows the researcher to draw sounder conclusions (Haselhuhn & Mellers, 2005; Pisor et al., 2020; Thielmann et al., 2021).

By utilising both empirical studies and agent-based models, this work has impact in suggesting that any arbitrary or maladaptive behaviour that persists via cultural evolutionary processes may be influenced by both social learning trade-offs, and trade-offs in the psychological components underlying decision-making themselves. Clearly, both processes must be investigated together in order to fully understand the range of skillsets and behaviours upheld over a diverse array of societies (Henrich & Muthukrishna, 2021). This is particularly true for social norms, as both social learning biases (**Chapter 3)** and drift **(Chapter 6)** may increase the chances of coordinating on suboptimal social norms. This finding has implications in supporting the use of behavioural 'nudging' policies and other interventions to help reduce negative social behaviour (Berger, 2021, Efferson et al., 2020a).

8.6: Limitations and directions for future research

I now turn to the limitations of my thesis, by first addressing the limitations of my study chapters (8.6.1) and my modelling chapters (8.6.2) individually, before tying these together in section 8.6.3.

8.6.1. Limitations of social learning studies

While the inclusion of Indian participants is on the whole a strength of my thesis, one limitation is that the participants who played the game against nature and the coordination game in **Chapter 3** were from the FLAME University in Pune, India while the participants who were recruited to play the Prisoner's Dilemma in **Chapter 4** were from Royal Holloway University (RHUL) in the UK. This unfortunately means that I cannot be sure whether the Prisoner's Dilemma was learned differently to the other two game types due to the unique nature of this game's structure (Holt & Roth, 2004), or due to a genuine difference in social learning preferences between the samples. After all, individualistic British participants tend to use social information less than collectivist participants (Mesoudi et al., 2015; Molleman & Gächter, 2018). Perhaps the lack of response to reliably different others on the part of the British participants playing the Prisoner's Dilemma was not due to the nature of this game per se, but rather due to an individualistic preference to use social information less than the Indian participants did.

To address this, future work should replicate the study in **chapters 3** and **4** but with a full cultural comparison (i.e., so that British participants also play the game against nature and the coordination game, while Indian participants also play the Prisoner's Dilemma). This could clearly establish any differences in the chosen social learning strategies between highly-individualistic participants and participants who are from a country in the middle of Hofstede's (1980) individualism-collectivism scale.

This work focused in-depth on the flexibility of frequency-dependent social learning strategies, in a way which previous literature has not. However, it is worth noting that trial-and-error is important in skill learning (Mesoudi, 2008a; 2011b; Miu & Morgan, 2020) and that payoff-based learning– or the lack thereof– is important when learning cooperation (Burton-Chellew et al., 2015; Molleman et al., 2013a). Future extensions of this empirical work may wish to apply the three orders of social information to the study of payoff-based and prestige-based social learning strategies, too.

Another limitation of my studies may be in the conceptualisation of the similarity and difference signals. I directly told the participants whether they were playing the same or a different game to the group from whom they learned. When previous work discussed a bias to copy 'similar' others, this usually referred to similarity on an observable trait. For example, demonstrators were typically similar in age or gender to the individual being studied (House et al., 2013; Salali et al., 2015; Shutts et al., 2010). Previous research thus tended to focus on those that *look* similar to the social learner. Based on literature regarding ethnic markers, I took those that look similar to the social learner as a short-hand to represent those that made decisions in the same environment as the social learner (Deffner et al., 2020; Efferson et al., 2008b; McElreath et al., 2003; Richerson et al., 2016). This justified the use of the signal in the current study, which directly conveyed information about shared decision-making environments to the participants. As this was the first study to test frequency-dependent social learning strategies to a third-order complexity, it was important to keep the similarity signals as straightforward as possible, to confirm that the effect exists before exploring this in more depth.

Of course, there still may be a bias to copy only those that look similar to oneself (Smaldino et al., 2018). Indeed, any biases to learn from those that *look* similar to ourselves– and to do the opposite to those that *look* different– would be interesting for future research to investigate as these processes are likely to be influenced by ethnocentrism (Hales & Edmonds, 2019) and in-group preferences versus out-group prejudices (Efferson et al., 2008b; Konrad & Morath, 2012). One way to test this conceptualisation in a minimal study design would be to use an arbitrary symbol to investigate ethnic markers. Efferson et al. (2008b) have already confirmed a preference to play a coordination game with those who have the same on-screen avatar as oneself, provided that these arbitrary avatars can become linked to the game being played in non-trivial ways. Future work should extend this idea, to test the social learning of asocial skills and cooperative behaviours, too. The avatar should only become reliably linked to the person's in-game preferences on some rounds. On other rounds, they should remain uninformative or become a reliably incorrect signal of the game type being played. This could allow us to understand how prejudices against ethnic markers emerge differently in cases where the markers could be easily faked and thus could become unreliable.

8.6.2. Limitations of agent-based models

The ultimate aim of my models was to investigate the evolutionary trajectory of decision-making in the three key areas across domain-general, partly modular, and fully modular agents. If one wished to make more definitive statements about whether human decision-making is likely to be modular or domain-general, then the code for the models can be updated to start with an equal split between all four agent types (fully modular, partly modular with modular cognition, partly modular with modular motivation and fully domain-general). These four agent types can then enter a state of competition, to see which agent would likely become evolutionary dominant across the runs.

When modelling competition, the structure of the populations within the model would have a huge impact on which agents are likely to become evolutionary dominant. This is especially true for the Prisoner's Dilemma modelled in **Chapter 7.** If the populations are left unstructured, then the domain-general agents who cooperate by mistake will be vulnerable to the free-riding modular agents. In a group-structured population with competition between groups, then the mistaken cooperation by domain-general agents would be mutually beneficial and thus any groups of domain-general agents would be likely to outcompete the free-riding groups of modular agents. Some previous models have considered neighbourhood structures where agents can assort themselves into a preferred group (Junikka et al., 2017), and future extensions to these models may wish to adapt this idea.

The modularity or domain-generality of the agent psychology may also be influenced by certain physiological costs which were not considered in my models. Throughout our ancestral past, as cognitive processes became more specialised there was likely to be a cost attached to this process (van Schaik et al., 2012). Massive modularity would require more neural tissue to ensure there is space for each module and an increased supply of oxygenated blood (Krohs, 2009). As modules may come with a physiological cost, then future models should recreate this where some cost is opposed on to a degree of modularity.

Once modular systems have emerged however, then they may be less costly to maintain than domain-general ones (Ellefsen et al., 2015). This is because modularity can 'compartmentalise' our decision-making whereas domain-general systems must build new associative links constantly both in order to learn new skills and to update older ones (Ellefsen et al., 2015). Future extensions to this model may wish to consider the relative costs of evolving modularity versus the costs of maintaining domain-generality.

8.6.3. Overall limitations and direction for future research

Research aim one of my thesis focused on the flexibility of social learning behaviours, which I addressed in **Chapters 3** and **4.** Research aim two focused on the flexibility of agent psychology that upholds decision-making, which I investigated in **Chapters 5-7.** Future work should investigate the flexibility of our decision-making processes and social learning behaviours together. The most obvious extension to this thesis would have been to add a period of social learning to my models. This would be important in clarifying my suggestions as to why social learning was found to be flexible but with asymmetric adjustments. Indeed, models are a great way to check that the logic of the assumptions made by the researchers holds up in a theoretically-evolving population (Muthukrishna & Henrich, 2019). Moreover, adding social learning to the agent-based models may also influence the type of psychological processing that underlies decision-making. The current findings suggest that modular cognition is important when mastering asocial skills (**Chapter 5),** and that fully modular psychology is important to master social norms across disparate domains (**Chapter 6).** If social learning was modelled, there may instead be a role for domain-general process in these behaviours (Mesoudi, 2008a).

I only considered economic games with two options to choose from (@ or % in **Chapters 3** and **4,** and 0 or 1 in **Chapters 5-7).** Future studies and models should investigate tasks with three or more behavioural options. These tasks are likely to be much more complex. To illustrate why, consider an example of a multi-choice social learning problem. Imagine that I am trying to decide what to order at a new restaurant that my friends take me to. Of my seven friends, three order fish, two order steak, one orders pasta and one orders a stir-fry. If I was a conformist learner, should I get the fish as well? Technically, this would not be conformity as the true majority avoids ordering fish (4/7 of my friends order something else). This becomes even more difficult if I wished to follow the minority choice. Would the minority decision in this case be to order the pasta, or the stir-fry? Both only have one vote each, and so there are two minorities. Or, if I'm truly committed to being the minority, perhaps the best course of action would be to order something different entirely: but then which menu item should

I get? This task is already sufficiently complex to visualise from an empirical perspective and ordering off a menu is a simple task.

To illustrate an even more complex example, think of the arrowhead design task employed by Mesoudi and colleagues (Mesoudi et al., 2015). This task had multiple equilibria which could change in subtle and unexpected ways as the environment shifted over the course of the study, and even included irrelevant options (changing the arrow's colour did not influence success). Moreover, the arrowhead design task is clearly more important than the previous example of ordering food in a restaurant. If the arrowhead is designed sub-optimally, I am relatively unlikely to make a kill when hunting. This is not only problematic for me but may have huge social consequences if I am hunting to provide for my family or tribe.

It may be hard to pinpoint any exact frequency-dependent social learning strategies in tasks with multiple equilibria, and perhaps the role of trial-and-error (Mesoudi, 2008a; 2011a) or payoff-based and success-based social learning (McElreth et al., 2008; Mesoudi, 2011a) becomes more important. Of course, it was important that my study investigated the flexibility of both frequency-dependent social learning and agent psychology in a simple two-choice task, as these investigations are novel and so should be tested in the most straightforward case possible. I merely note that tasks with more than two options require much more complex and flexible processing to master, and this would be very interesting for future research to explore.

Another limitation of the economic games that I employ throughout both my empirical studies and modelling chapters is that these games model the behaviour of two people. While this is the typical structure of a coordination game (Bernard et al., 2020; Wilson & Rhodes, 1997) and a Prisoner's Dilemma (Capraro, 2013), this takes a limited approach. After all, the majority of an entire social group must coordinate on certain social norms, and cooperative behaviours. Future replications may wish to use *n*-person coordination games and Prisoner's Dilemmas to address this caveat (Takezawa & Price, 2010; Tooby et al., 2006).

Finally, there are many other economic games– particularly those based on cooperation– which I could not include here for brevity. Future work could attempt to replicate the social learning studies and/or agent-based models when considering an ultimatum game (Henrich et al., 2001; 2006), or a Public Goods Game (PGG; Baum et al., 2012; Price, 2006; Wang et al., 2020), to name just two examples.

8.7. Final conclusions

This thesis aimed to investigate both the flexibility of social learning strategies, and the flexibility of the psychological processing underlying human decision-making, when making decisions in three key areas: asocial skills, social norms, and cooperation. In more detail, I found that the participants in the social learning studies adjusted to: (i) frequency information; (ii) similarity information and (iii) reliability signals, though they all made asymmetric adjustments as they found it easier to learn the same behaviours as groups of reliably similar others. In my agent-based models, modular cognition may have been influential for learning asocial skills only. The role of motivation was more important for social norms and particularly cooperation, with domain-general agents being the most likely to uphold the costly levels of cooperation that are observed across human societies. These findings suggest that modularity in Evolutionary Psychology may be relegated to asocial skill acquisition only, though domain-general processes which remain flexible to the input of cultural rules are necessary to explain the unique, costly levels of cooperation that are seen across human societies. This supports a gene-culture coevolutionary perspective on cooperation, which suggests that this behaviour can be best understood by a dual-inheritence framework.

Taken together, these results imply that individuals are flexible in both their social learning strategies and psychological processing, but there is an upper limit to the amount of information that one can process when making decisions. This has implications for the notion that cultural evolution is blind. Perhaps maladaptive behaviour proliferates due to trade-offs in the individual's social learning behaviours, or decision-making structures; or both. This thesis helps to shed some light on how and why maladaptive behaviours may be upheld by a group; via a mixture of social learning trade-offs and trade-offs in psychological processing underlying decision-making which in turn may be affected by drift. These processes should be studied by future research in order to understand how we can reduce the prevalence of harmful behaviours that may be transmitted by cultural evolution.

References

Abbink, K., Gangadharan, L., Handfield, T., & Thrasher, J. (2017). Peer punishment promotes enforcement of bad social norms. *Nature communications*, *8*(609), 1-8. https://doi.org/10.1038/s41467-017-00731-0.

Acerbi, A., Mesoudi, A., & Smolla, M. (2020). Individual-based models of cultural evolution: A step-by-step guide using R. OSF Preprints. doi:10.31219/osf.io/32v6a.

Acerbi, A., & Tehrani, J. J. (2018). Did Einstein Really Say that? Testing Content Versus Context in the Cultural Selection of Quotations, *Journal of Cognition and Culture*, *18*(3-4), 293-311. doi: https://doi.org/10.1163/15685373-12340032.

Amodio, P., Boeckle, M., Schnell, A. K., Ostojíc, L., Fiorito, G., & Clayton, N. S. (2019). Grow smart and die young: why did cephalopods evolve intelligence?. *Trends in ecology & evolution*, *34*(1), 45-56. https://doi.org/10.1016/j.tree.2018.10.010.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98-108. https://doi.org/10.1016/j.labeco.2014.05.008.

Aoki, K., Wakano, J. Y., & Lehmann, L. (2012). Evolutionarily stable learning schedules and cumulative culture in discrete generation models. *Theoretical Population Biology*, *81*(4), 300-309. https://doi.org/10.1016/j.tpb.2012.01.006.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 31-35.

Avarguès-Weber, A., Lachlan, R., & Chittka, L. (2018). Bumblebee social learning can lead to suboptimal foraging choices. *Animal behaviour*, *135*, 209-214. https://doi.org/10.1016/j.anbehav.2017.11.022.

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, *24*(1), 3-25.https://doi.org/10.1177/002200278002400101.

Barrett, H.C. (2015). Culture. In H.C. Barett (Ed.) *The shape of thought: How mental adaptations evolve* (pp. 224-243). New York; Oxford University Press

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review, 113*(3), 628–647. https://doi.org/10.1037/0033-295X.113.3.628.

Bartlett, F. C. (1920). Some experiments on the reproduction of folk-stories. *Folklore*, *31*(1), 30-47. https://doi.org/10.1080/0015587X.1920.9719123.

Baum, W. M., Paciotti, B., Richerson, P., Lubell, M., & McElreath, R. (2012). Cooperation due to cultural norms, not individual reputation. *Behavioural processes*, *91*(1), 90-93. https://doi.org/10.1016/j.beproc.2012.06.001.

Bell, A. V., Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences, 106*(42), 17671-17674. https://doi.org/10.1073/pnas.0903232106.

Beja-Pereira, A., Luikart, G., England, P. R., Bradley, D. G., Jann, O. C., Bertorelle, G., Chamberlain, A.T., Nunes, T.P., Metodiev, S., Ferrand, N. & Erhardt, G. (2003). Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nature genetics*, *35*(4), 311-313. doi:10.1038/ng1263.

Bennett, K. (2018). Environment of evolutionary adaptedness (EEA). *Encyclopedia of personality and individual differences*, *1*(1627), 1-3. DOI**:** https://doi.org/10.1007/978-3-319-28099-8_1627-1.

Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1547), 1443-1450. https://doi.org/10.1098/rspb.2004.2746.

Berger, J. (2021). Social Tipping Interventions Can Promote the Diffusion or Decay of Sustainable Consumption Norms in the Field. Evidence from a Quasi-Experimental Intervention Study. *Sustainability*, *13*(6), 3529. https://doi.org/10.3390/su13063529.

Bernard, A., Bredeche, N., & André, J. B. (2020). Indirect genetic effects allow escape from the inefficient equilibrium in a coordination game. *Evolution letters*, *4*(3), 257-265. https://doi.org/10.1002/evl3.155.

Bhui, R., Chudek, M., & Henrich, J. (2019). How exploitation launched human cooperation. *Behavioral Ecology and Sociobiology*, *73*(78), 1-14. https://doi.org/10.1007/s00265-019-2667-y.

Billiard, S., & Alvergne, A. (2018). Stochasticity in cultural evolution: a revolution yet to happen. *History and philosophy of the life sciences*, *40*(9), 9. https://doi.org/10.1007/s40656-017-0173-y.

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biology*, *9*(7), e1001109. https://doi.org/10.1371/journal.pbio.1001109.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119*(1), 111–137. https://doi.org/10.1037/0033-2909.119.1.111

Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical population biology*, *65*(1), 17-28. https://doi.org/10.1016/j.tpb.2003.07.001.

Boyd, B. (2009). Evolution, Cognition and Fiction. In B. Boyd's *On the origin of stories* (pp.129-209). Massachusetts: Harvard University Press.

Boyd, R., & Richerson, P. J. (1985). Frequency-dependent Bias and the Evolution of Cooperation. In R. Boyd & P.J. Richerson (Eds.) *Culture and the Evolutionary Process*, (p 204-241). Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford University Press: Oxford, UK.

Boyd, R., & Richerson, P. J. (2006). Solving the puzzle of human cooperation. In S.C. Levinson and P. Jaisson (Eds.) *Evolution and culture: A Fyssen foundation system*, pp 105-132. England: MIT Press.

Boyd, R., & Richerson, P. J. (2007). Cultural adaptation and maladaptation: Of Kayaks and Commissars. In S.W. Gangestad & J.A. Simpson (Eds.) *The evolution of the mind: Fundamental questions and controversies* (p327-332). New York: Guilford Press.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10918-10925. https://doi.org/10.1073/pnas.1100290108.

Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, *149*(4), 746–756. https://doi.org/10.1037/xge0000673.

Brand, C. O., Heap, S., Morgan, T. J. H., & Mesoudi, A. (2020). The emergence and adaptive use of prestige in an online social learning task. *Scientific reports, 10(*12905), 1-11. https://doi.org/10.1038/s41598-020-68982-4.

Burdett, E. R.R., Dean, L. G., & Ronfard, S. (2018). A diverse and flexible teaching toolkit facilitates the human capacity for cumulative culture. *Review of Philosophy and Psychology*, *9*(4), 807-818. https://doi.org/10.1007/s13164-017-0345-4.

Burke, D. (2014). Why isn't everyone an evolutionary psychologist? *Front. Psychol*. *5*(910), 1-8. doi: 10.3389/fpsyg.2014.00910.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences, 113*(5), 1291-1296. https://doi.org/10.1073/pnas.1509740113.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1853), 20170067. https://doi.org/10.1098/rspb.2017.0067.

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences, 282*(1801), 20142678. https://doi.org/10.1098/rspb.2014.2678.

Burton-Chellew, M. N., & West, S. A. (2012). Pseudocompetition among groups increases human cooperation in a public-goods game. *Animal Behaviour, 84*(4), 947-952. https://doi.org/10.1016/j.anbehav.2012.07.019.

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological inquiry*, *6*(1), 1-30. https://doi.org/10.1207/s15327965pli0601_1.

Capraro, V. (2013). A model of human cooperation in social dilemmas. *PloS one*, *8*(8), e72427. https://doi.org/10.1371/journal.pone.0072427.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*(5), 218-226. https://doi.org/10.1016/j.tics.2011.03.003.

Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation, and the shaping of human social psychology. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 425–457). US: The MIT Press.

Claidière, N., Amedon, G. K. K., André, J. B., Kirby, S., Smith, K., Sperber, D., & Fagot, J. (2018). Convergent transformation and selection in cultural evolution. *Evolution and Human Behavior*, 39(2), 191-202. https://doi.org/10.1016/j.evolhumbehav.2017.12.007.

Clegg, J. M., & Legare, C. H. (2016). A cross-cultural comparison of children's imitative flexibility. *Developmental Psychology*, *52*(9), 1435–1444. https://doi.org/10.1037/dev0000131.

Coletto, M., Aiello, L. M., Lucchese, C., & Silvestri, F. (2017). Adult content consumption in online social networks. *Social Network Analysis and Mining*, *7*(28), 1-21. https://doi.org/10.1007/s13278-017-0449-y.

Colman, A. M. (2006). The puzzle of cooperation. *Nature*, *440*(7085), 744-745. DOI: https://doi.org/10.1038/440744b.

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M. G., Perilloux, C., & Buss, D. M. (2010). Evolutionary psychology: Controversies, questions, prospects, and limitations. *American Psychologist*, *65*(2), 110–126. https://doi.org/10.1037/a0018413.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences, 107*(Supplement 2), 9007-9014. https://doi.org/10.1073/pnas.0914623107.

Cosmides, L. & Tooby, J. (1994a). Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. *Cognition, 50*(1-3), 41-77. https://doi.org/10.1016/0010-0277(94)90020-5.

Cosmides, L., & Tooby, J. (1994b). Origins of domain specificity: The evolution of functional organization. In L. Hirschfeld and S.Gelman (Eds.) *Mapping the mind: Domain specificity in cognition and culture* (pp. 84-116). New York: Cambridge University Press.

Cosmides, L., & Tooby, J. (2007). *Evolutionary psychology: A primer.* Retrieved January 9th 2022 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.615.8167&rep=rep1&type=pdf

Curtin, C. M., Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M.T., Fitzpatrick, S., Gurven, M., Kanovsky, M., Kushnick, G., Laurence, S., Pisor, A., Scelza, B., Stich, S., von Rueden, C. & Henrich, J. (2020). Kinship intensity and the use of mental states in moral judgment across societies. *Evolution and Human Behavior*, *41*(5), 415-429. https://doi.org/10.1016/j.evolhumbehav.2020.07.002.

Davis, D.D. & Holt, C.A. (1993). Some procedural and design considerations. In D.D. Davis & C.A. Holt's (Eds.), *Experimental economics* (pp. 20-33). US: Princeton University Press.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature, 446*(7137), 794-796. https://doi.org/10.1038/nature05651.

Dawkins, R. (1996). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company: London, UK.

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734.

Delton, A. W., Cosmides, L., Guemo, M., Robertson, T. E., & Tooby, J. (2012). The psychosemantics of free riding: Dissecting the architecture of a moral concept. *Journal of Personality and Social Psychology, 102*(6), 1252–1270. https://doi.org/10.1037/a0027026.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences, 108*(32), 13335-13340. DOI: https://doi.org/10.1073/pnas.1102131108.

Delton, A. W., Nemirow, J., Robertson, T. E., Cimino, A., & Cosmides, L. (2013). Merely opting out of a public good is moralized: An error management approach to cooperation. *Journal of Personality and Social Psychology, 105*(4), 621–638. https://doi.org/10.1037/a0033495.

Dhum, M., Herwig, U., Opialla, S., Siegrist, M., & Brühl, A. B. (2017). Evolutionary and modern image content differentially influence the processing of emotional pictures. *Frontiers in human neuroscience, 11*, 415. https://doi.org/10.3389/fnhum.2017.00415.

Downes, S. M. (2015). Evolutionary psychology, adaptation and design. In T. Heams, Hunemann, P., Lecointre, G. & Silberstein, M. (Eds). *Handbook of evolutionary thinking in the sciences* (pp. 659-673). Springer, Dordrecht. DOI: https://doi.org/10.1007/978-94-017-9014-7_31.

Efferson, C., Lalive, R., Cacault, M. P., & Kistler, D. (2016). The evolution of facultative conformity based on similarity. *PloS one*, *11*(12), e0168551. https://doi.org/10.1371/journal.pone.0168551.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008a). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, *29*(1), 56-64. https://doi.org/10.1016/j.evolhumbehav.2007.08.003.

Efferson, C., Lalive, R., & Fehr, E. (2008b). The coevolution of cultural groups and ingroup favoritism. *Science*, *321*(5897), 1844-1849. DOI: 10.1126/science.1155805.

Efferson, C., Vogt, S., Elhadi, A., Ahmed, H. E. F., & Fehr, E. (2015). Female genital cutting is not a social coordination norm. *Science*, *349*(6255), 1446-1447. DOI: 10.1126/science.aaa7978.

Efferson, C., Vogt, S., & Fehr, E. (2020a). The promise and the peril of using social influence to reverse harmful traditions. *Nature human behaviour*, *4*(1), 55-68. DOIhttps://doi.org/10.1038/s41562-019-0768-2.

Efferson, C., McKay, R., & Fehr, E. (2020b). The evolution of distorted beliefs vs. mistaken choices under asymmetric error costs. *Evolutionary Human Sciences*, *2*, E27. doi:10.1017/ehs.2020.25.

Ekuni, R., de Souza, B. M. N., Agarwal, P. K., & Pompeia, S. (2020). A conceptual replication of survey research on study strategies in a diverse, non-WEIRD student population. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000191.

Ellefsen, K. O., Mouret, J. B., & Clune, J. (2015). Neural modularity helps organisms evolve to learn new skills without forgetting old skills. *PLoS computational biology*, *11*(4), e1004128. https://doi.org/10.1371/journal.pcbi.1004128.

Eriksson, K., & Coultas, J. C. (2014). Corpses, Maggots, Poodles and Rats: Emotional Selection Operating in Three Phases of Cultural Transmission of Urban Legends. *Journal of Cognition and Culture*, 14(1–2), 1–26. https://doi.org/10.1163/15685373-12342107.

Ertac, S., & Kotan, E. (2020). z-Tree in VLab: A Method for Running Online Economic Experiments. *Available at SSRN 3756019*. http://dx.doi.org/10.2139/ssrn.3756019.

Evans, C. L., Laland, K. N., Carpenter, M., & Kendal, R. L. (2018). Selective copying of the majority suggests children are broadly "optimal-" rather than "over-" imitators. *Developmental science*, *21*(5), e12637. https://doi.org/10.1111/desc.12637.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241. https://doi.org/10.1177/1745691612460685.

Fagiolo, G., Moneta, A., & Windrum, P. (2007). A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics*, *30*(3), 195-226. DOI 10.1007/s10614-007-9104-4.

Feldman, M. W., Aoki, K., & Kumm, J. (1996). Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropological Science, 104*(3), 209-231. https://doi.org/10.1537/ase.104.209.

Feldman, M. W., & Laland, K. N. (1996). Gene-culture coevolutionary theory. *Trends in ecology & evolution*, *11*(11), 453-457. https://doi.org/10.1016/0169-5347(96)10052-5.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature, 13*(1), 1-25.

Fehr, E. & Fischbacher, U. (2005). Human Altruism – Proximate Patterns and Evolutionary Origins. *Analyse & Kritik, 27*(1), 6-47. https://doi.org/10.1515/auk-2005-0101.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*, 171–178. https://doi.org/10.1007/s10683-006-9159-4.

Fodor, J. A. (1983). *The modularity of mind*: *An essay on faculty psychology*. England, MIT press.

Foramitti, J. (2021). AgentPy: A package for agent-based modeling in Python. *Journal of Open Source Software*, *6*(62), 3065. DOI: 10.21105/joss.03065.

Frankenhuis, W. E., & Ploeger, A. (2007). Evolutionary psychology versus Fodor: Arguments for and against the massive modularity hypothesis. *Philosophical Psychology*, *20*(6), 687-710. https://doi.org/10.1080/09515080701665904.

Gangestad, S. W., Haselton, M. G., & Buss, D. M. (2006). Evolutionary foundations of cultural variation: Evoked culture and mate preferences. *Psychological Inquiry*, *17*(2), 75-95. DOI: 10.1207/s15327965pli1702_1.

Gavrilets, S. (2020). The dynamics of injunctive social norms. *Evolutionary Human Sciences*, *2*, 60-80. doi:10.1017/ehs.2020.58.

Gelfand, M. J., Caluori, N., Jackson, J. C., & Taylor, M. K. (2020). The cultural evolutionary trade-off of ritualistic synchrony. *Philosophical Transactions of the Royal Society B*, *375*(1805), 20190432. https://doi.org/10.1098/rstb.2019.0432.

Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D.M. & Thomas, M. G. (2011). Evolution of lactase persistence: an example of human niche construction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 863-877. doi:10.1098/rstb.2010.0268.

Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of theoretical biology*, *220*(4), 407-418. https://doi.org/10.1006/jtbi.2003.3104.

Gintis, H. (2004). The genetic side of gene-culture coevolution: internalization of norms and prosocial emotions. *Journal of economic behavior & organization*, *53*(1), 57-67.

Gintis, H. (2011). Gene–culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 878-888. https://doi.org/10.1098/rstb.2010.0310.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*(2), 241-253. https://doi.org/10.1007/s11211-008-0067-y.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of theoretical biology*, *213*(1), 103-119. https://doi.org/10.1006/jtbi.2001.2406.

Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual sequences. *Cognitive psychology*, *65*(2), 284-320. https://doi.org/10.1016/j.cogpsych.2012.04.002.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the royal society of London. Series B. Biological Sciences*, *205*(1161), 581-598. https://doi.org/10.1098/rspb.1979.0086.

Gray, R. D., & Watts, J. (2017). Cultural macroevolution matters. *Proceedings of the National Academy of Sciences*, *114*(30), 7846-7852. https://doi.org/10.1073/pnas.1620746114.

Hale, N. E. (2017). Are we Genetically Adapted to the Neolithic Diet?. *Journal of Evolution and Health: A joint publication of the Ancestral Health Society and the Society for Evolutionary Medicine and Health*, *2*(3), 1-10. DOI: 10.15310/2334-3591.1068.

Hales, D., & Edmonds, B. (2019). Intragenerational cultural evolution and ethnocentrism. *Journal of Conflict Resolution*, *63*(5), 1283-1309. DOI: 10.1177/0022002718780481

Hancock, A. M., Witonsky, D. B., Ehler, E., Alkorta-Aranburu, G., Beall, C., Gebremedhin, A., Sukernik, R., Utterman, G., Pritchard, J., Coop, G. & Di Rienzo, A. (2010). Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of*

Sciences, *107*(Supplement 2), 8924-8930.
https://doi.org/10.1073/pnas.0914625107.

Hansen, R. G., & Samuelson, W. F. (1988). Evolution in economic games. *Journal of Economic Behavior & Organization*, *10*(3), 315-338. https://doi.org/10.1016/0167-2681(88)90053-4.

Hao, J., Sun, J., Chen, G., Wang, Z., Yu, C., & Ming, Z. (2017). Efficient and robust emergence of norms through heuristic collective learning. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, *12*(4), 1-20. https://doi.org/10.1145/3127498.

Harris, C. R. (2004). The evolution of jealousy: Did men and women, facing different selective pressures, evolve different" brands" of jealousy? Recent evidence suggests not. *American scientist*, *92*(1), 62-71.

Haselhuhn, M. P., & Mellers, B. A. (2005). Emotions and cooperation in economic games. *Cognitive brain research*, *23*(1), 24-33. https://doi.org/10.1016/j.cogbrainres.2005.01.005.

Haselton, M.G., Nettle, D. & Murray, D.R. (2015). The evolution of cognitive bias. In D.M. Buss (Ed.) *The handbook of Evolutionary Psychology* (pp 968-987). Wiley, United Kingdom. DOI: https://doi.org/10.1002/9781119125563.evpsych241.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behaviour and Organisation*, *53*(1), 3–35. doi:10.1016/S0167-2681(03)00094-5.

Henrich, J. (2015). *The secret of our success*: *How culture is driving human evolution, domesticating our species and making us smarter*. USA: Princeton University Press.

Henrich, J. (2020). *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous*. New York: Farrar, Straus and Giroux.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and human behavior*, *19*(4), 215-241. https://doi.org/10.1016/S1090-5138(98)00018-X.

Henrich, J, & Boyd R. (2001). Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology, 208*(1), 79– 89. https://doi.org/10.1006/jtbi.2000.2202.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review*, *91*(2), 73-78. DOI: 10.1257/aer.91.2.73.

Henrich, J., Chudek, M., & Boyd, R. (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1683), 20150013. https://doi.org/10.1098/rstb.2015.0013.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and human behavior*, *22*(3), 165-196. https://doi.org/10.1016/S1090-5138(00)00071-4.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29. https://doi.org/10.1038/466029a.

Henrich, J., & McElreath, R. (2007). Dual inheritance theory: the evolution of human cultural capacities and cultural evolution. In L. Barrett and R.Dunbar (Eds.) *Oxford handbook of evolutionary psychology*, 555-570. New York: Oxford University Press. DOI: 10.1093/oxfordhb/9780198568308.013.0038.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. & Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767-1770. DOI: 10.1126/science.1127333.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106.

Heyes C. (2016). Blackboxing: social learning strategies and cultural evolution. *Philosophical Transaction Royal Society. B, 371*(1693), 20150369. http://dx.doi.org/10.1098/rstb.2015.0369.

Heyes C. (2018). Enquire within: cultural evolution and cognitive science. *Philosophical Transaction Royal Society. B 373*: 20170051. http://dx.doi.org/10.1098/rstb.2017.0051.

Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature*, *13*(1), 105-128. https://doi.org/10.1007/s12110-002-1016-3.

Hill, K., & Gurven, M. (2004). Economic experiments to examine fairness and cooperation among the Ache Indians of Paraguay. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr & H. Gintis (Eds.) *Foundations in Human Sociality:*

Economic experiments and ethnographic evidence from fifteen small-scale societies, pp 382-412. Oxford: University Press.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values.* Beverly Hills, CA: Sage.

Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*(12), 3999-4002. https://doi.org/10.1073/pnas.0308738101.

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B.S., McElreath, R. & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, *110*(36), 14586-14591. https://doi.org/10.1073/pnas.1221217110.

Hughes, S. S. (1998). Getting to the point: evolutionary change in prehistoric weaponry. *Journal of Archaeological Method and Theory*, *5*(4), 345-408. https://doi.org/10.1007/BF02428421.

Ihara, Y. (2011). Evolution of culture-dependent discriminate sociality: a gene–culture coevolutionary model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 889-900. https://doi.org/10.1098/rstb.2010.0247.

Itan, Y., Powell, A., Beaumont, M. A., Burger, J., & Thomas, M. G. (2009). The origins of lactase persistence in Europe. *PLoS computational biology*, *5*(8), e1000491. https://doi.org/10.1371/journal.pcbi.1000491.

Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(20), 1-12. https://doi.org/10.1057/s41599-019-0228-7.

Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on twitter. In *Proceedings of the*

7th workshop on social network mining and analysis, 8, (pp. 1-9). https://doi.org/10.1145/2501025.2501027.

Jung, J., Bramson, A., Crano, W. D., Page, S. E., & Miller, J. H. (2021). Cultural drift, indirect minority influence, network structure, and their impacts on cultural change and diversity. *American Psychologist*, *76*(6), 1039–1053. https://doi.org/10.1037/amp0000844.

Junikka, J., Molleman, L., van den Berg, P., Weissing, F. J., & Puurtinen, M. (2017). Assortment, but not knowledge of assortment, affects cooperation and individual success in human groups. *PloS one, 12*(10), e0185859. https://doi.org/10.1371/journal.pone.0185859.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637-651. https://doi.org/10.1016/j.cognition.2013.09.001.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in cognitive sciences*, *22*(7), 651-665. https://doi.org/10.1016/j.tics.2018.04.003.

Ketelaar, T., & Ellis, B. J. (2000). Are evolutionary explanations unfalsifiable? Evolutionary psychology and the Lakatosian philosophy of science. *Psychological Inquiry*, *11*(1), 1-21. DOI: 10.1207/S15327965PLI1101_01.

Kets, W., Kager, W., & Sandroni, A. (2021). The value of a coordination game. *Available at SSRN:* http://dx.doi.org/10.2139/ssrn.3796188.

Konrad, K. A., & Morath, F. (2012). Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology*, *306*, 61-67. https://doi.org/10.1016/j.jtbi.2012.04.013.

Krafft, P. M., Shmueli, E., Griffiths, T. L., Tenenbaum, J. B., & Pentland, A. (2021). Bayesian collective learning emerges from heuristic social learning. *Cognition*, *212*, 104469. https://doi.org/10.1016/j.cognition.2020.104469.

Krasnow, M. M., & Delton, A. W. (2016). Are humans too generous and too punitive? Using psychological principles to further debates about human social evolution. *Frontiers in psychology*, *7*, 799. https://doi.org/10.3389/fpsyg.2016.00799.

Krasnow, M. M., Delton, A. W., Tooby, J., & Cosmides, L. (2013). Meeting now suggests we will meet again: Implications for debates on the evolution of cooperation. *Scientific reports*, *3*(1747), 1-8. https://doi.org/10.1038/srep01747.

Krasnow, M. M., Howard, R. M., & Eisenbruch, A. B. (2020). The importance of being honest? Evidence that deception may not pollute social science subject pools after all. *Behavior research methods*, *52*(3), 1175-1188. https://doi.org/10.3758/s13428-019-01309-y.

Krohs, U. (2009). The Costs of Modularity. In U. Krohs & P. Kroes (Eds.), *Functions in biology and artificial worlds: Comparative philosophical perspectives* (pp 259-276). Cambridge, MA:MIT press.

Kroneisen, M., & Bell, R. (2021). Memory as a cognitive requirement for reciprocal cooperation. *Current opinion in psychology*, *43*, 271-277. https://doi.org/10.1016/j.copsyc.2021.08.008.

Lachlan, R.F., Ratmann, O. & Nowicki, S. (2018). Cultural conformity generates extremely stable traditions in bird song. *Nature Communications*, 9(2417). https://doi.org/10.1038/s41467-018-04728-1.

Laland, K. N. (1993). The mathematical modelling of human culture and its implications for psychology and the human sciences. *British Journal of Psychology*, *84*(2), 145-169. https://doi.org/10.1111/j.2044-8295.1993.tb02471.x.

Laland, K. N. & Brown, G. R. (2011). *Sense and nonsense: Evolutionary perspectives on human behaviour*. England: Oxford University Press.

Laland, K. N., Kendal, J. R., & Brown, G. R. (2007). The niche construction perspective: Implications for evolution and human behaviour. *Journal of Evolutionary Psychology JEP*, *5*(1-4), 51-66. DOI: 10.1556/JEP.2007.1003.

Laland, K. N., Odling-Smee, J., Feldman, M. W., & Kendal, J. (2009). Conceptual barriers to progress within evolutionary biology. *Foundations of science*, *14*(3), 195-216. DOI 10.1007/s10699-008-9153-8.

Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: bringing genetics and the human sciences together. *Nature Reviews Genetics*, *11*(2), 137-148. https://doi.org/10.1038/nrg2734.

Lamba, S. (2014). Social learning in cooperative dilemmas. *Proceedings of the Royal Society B: Biological Sciences, 281*(1787), 20140417. DOI: https://doi.org/10.1098/rspb.2014.0417.

Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W., & Müller, K. R. (2019). Unmasking Clever Hans predictors and assessing what machines really learn. *Nature communications*, *10*(1096), 1-8. https://doi.org/10.1038/s41467-019-08987-4

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences*, *114*(30), 7877-7883. https://doi.org/10.1073/pnas.1620743114.

Legare, C. H. (2019). The development of cumulative cultural learning. *Annual Review of Developmental Psychology*, *1*, 119-147. https://doi.org/10.1146/annurev-devpsych-121318-084848.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in cognitive sciences*, *19*(11), 688-699. https://doi.org/10.1016/j.tics.2015.08.005.

Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*, *142*, 351-361. https://doi.org/10.1016/j.cognition.2015.05.020.

Lenfesty, H.L., & Morgan, T. J.H. (2019). By reverence, not fear: prestige, religion, and autonomic regulation in the evolution of cooperation. *Frontiers in Psychology*, *10*, 2750. https://doi.org/10.3389/fpsyg.2019.02750.

Li, X., Molleman, L., & van Dolder, D. (2021). Do descriptive social norms drive peer punishment? Conditional punishment strategies and their impact on cooperation. *Evolution and Human Behavior, 42*(5), 469-479. DOI: https://doi.org/10.1016/j.evolhumbehav.2021.04.002.

Lotem, A., Halpern, J. Y., Edelman, S., & Kolodny, O. (2017). The evolution of cognitive mechanisms in response to cultural innovations. *Proceedings of the National Academy of Sciences*, *114*(30), 7915-7922. https://doi.org/10.1073/pnas.1620742114.

Mackie, M. E., Surovell, T. A., & O'Brien, M. (2015). Identifying stone alignments created by adults and children: A case study from a Dukha Reindeer herder summer camp, Khövsgöl Aimag, Mongolia. *Sibirica*, *14*(2), 29-44. DOI: https://doi.org/10.3167/sib.2015.140202.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives*, *14*(3), 115-136. DOI: 10.1257/jep.14.3.115.

Markov, A. V., & Markov, M. A. (2020). Runaway brain-culture coevolution as a reason for larger brains: Exploring the "cultural drive" hypothesis by computer modeling. *Ecology and Evolution*, *10*(12), 6059-6077. https://doi.org/10.1002/ece3.6350.

Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharot, T. (2019). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, *188*, 74-84. https://doi.org/10.1016/j.cognition.2018.10.003.

Mason, K., Qadrdan, M., & Jenkins, N. (2021). Investing in generation and storage capacity in a liberalised electricity market: An agent based approach. *Applied Energy*, *294*, 116905. https://doi.org/10.1016/j.apenergy.2021.116905.

McElreath, R. (2004). Social learning and the maintenance of cultural variation: An evolutionary model and data from East Africa. *American Anthropologist*, *106*(2), 308-321. https://doi.org/10.1525/aa.2004.106.2.308.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3515-3528. https://doi.org/10.1098/rstb.2008.0131.

McElreath, M. B., Boesch, C., Kühl, H., & McElreath, R. (2018). Complex dynamics from simple cognition: The primary ratchet effect in animal culture. *Evolutionary Behavioral Sciences, 12*(3), 191–202. https://doi.org/10.1037/ebs0000117.

McElreath, R., Boyd, R., & Richerson, P. (2003). Shared norms and the evolution of ethnic markers. *Current anthropology*, *44*(1), 122-130. https://doi.org/10.1086/345689.

McElreath, R., & Strimling, P. (2008). When natural selection favors imitation of parents. *Current Anthropology*, *49*(2), 307-316. https://doi.org/10.1086/524364.

McGuigan, N., Burdett, E., Burgess, V., Dean, L., Lucas, A., Vale, G., & Whiten, A. (2017). Innovation and social transmission in experimental micro-societies: exploring the scope of cumulative culture in young children. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1735), 20160425. https://doi.org/10.1098/rstb.2016.0425.

McKay, R., & Efferson, C. (2010). The subtleties of error management. *Evolution and human behavior*, *31*(5), 309-319. https://doi.org/10.1016/j.evolhumbehav.2010.04.005.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A.A., Hopkins, E.J., Howard, R.M., Hartshone, J.K., Jennings, M.V., Simson, J., Bainbridge, C.M., Pinker, S., O'Donnell, T.J., Krasnow, M.M. & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468). doi: 10.1126/science.aax0868.

Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and function in human song. *Current Biology*, *28*(3), 356-368. https://doi.org/10.1016/j.cub.2017.12.042.

Mesoudi, A. (2008a). An experimental simulation of the “copy-successful-individuals” cultural learning strategy: Adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*, *29*(5), 350-363. https://doi.org/10.1016/j.evolhumbehav.2008.04.005.

Mesoudi, A. (2008b). Foresight in cultural evolution. *Biology & Philosophy*, *23*(2), 243-255. DOI 10.1007/s10539-007-9097-3.

Mesoudi, A. (2009). The cultural dynamics of copycat suicide. *PLoS One*, *4*(9), e7252. https://doi.org/10.1371/journal.pone.0007252.

Mesoudi, A. (2011a). Variable cultural acquisition costs constrain cumulative cultural evolution. *PloS one*, *6*(3), e18239. https://doi.org/10.1371/journal.pone.0018239

Mesoudi, A. (2011b). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334-342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001.

Mesoudi, A. (2016a). Cultural evolution: Integrating psychology, evolution and culture. *Current Opinion in Psychology*, *7*, 17-22. https://doi.org/10.1016/j.copsyc.2015.07.001.

Mesoudi, A. (2016b). Cultural evolution: a review of theory, findings and controversies. *Evolutionary biology*, *43*(4), 481-497. https://doi.org/10.1007/s11692-015-9320-0.

Mesoudi, A. (2017). Pursuing Darwin’s curious parallel: Prospects for a science of cultural evolution. *Proceedings of the National Academy of Sciences*, *114*(30), 7853-7860. https://doi.org/10.1073/pnas.1620741114.

Mesoudi, A. (2018). Migration, acculturation, and the maintenance of between-group cultural variation. *PloS one*, *13*(10), e0205573. https://doi.org/10.1371/journal.pone.0205573.

Mesoudi, A., Chang, L., Dall, S. R.X., & Thornton, A. (2016). The evolution of individual and cultural variation in social learning. *Trends in ecology & evolution*, *31*(3), 215-225. https://doi.org/10.1016/j.tree.2015.12.012.

Mesoudi, A., Cheng, L., Murray, K., & Lu, H. J. (2015). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20142209. https://doi.org/10.1098/rspb.2014.2209.

Mesoudi, A., & Lycett, S. J. (2009). Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior*, *30*(1), 41-48. https://doi.org/10.1016/j.evolhumbehav.2008.07.005.

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution?. *Proceedings of the Royal Society B*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712.

Mesoudi, A., Whiten, A., & Dunbar, R. (2006a). A bias for social information in human cultural transmission. *British journal of psychology*, *97*(3), 405-423. https://doi.org/10.1348/000712605X85871.

Mesoudi, A., Whiten, A., & Laland, K. (2006b). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, *29*(4), 329-347. doi:10.1017/S0140525X06009083

Mesoudi, A., Whiten, A., & Laland, K. N. (2004). Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the perspective of The Origin of

Species. *Evolution*, *58*(1), 1-11. https://doi.org/10.1111/j.0014-3820.2004.tb01568.x.

Mikhalevich, I., Powell, R., & Logan, C. (2017). Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition. *Interface Focus 7*, 20160121. http://dx.doi.org/10.1098/rsfs.2016.0121.

Miu, E., Gulley, N., Laland, K. N., & Rendell, L. (2020). Flexible learning, rather than inveterate innovation or copying, drives cumulative knowledge gain. *Science Advances*, *6*(23), eaaz0286. DOI: 10.1126/sciadv.aaz0286.

Miu, E., & Morgan, T. J.H. (2020). Cultural adaptation is maximised when intelligent individuals rarely think for themselves. *Evolutionary Human Sciences*, *2*, e43. doi:10.1017/ehs.2020.42.

Molleman, L., & Gächter, S. (2018). Societal background influences social learning in cooperative decision making. *Evolution and Human Behavior, 39*(5), 547-555. https://doi.org/10.1016/j.evolhumbehav.2018.05.007.

Molleman L, Kanngiesser P, van den Bos W (2019a) Social information use in adolescents: The impact of adults, peers and household composition. *PLoS ONE 14*(11): e0225498. https://doi.org/10.1371/journal.pone.0225498

Molleman, L., Kölle, F., Starmer, C., & Gächter, S. (2019b). People prefer coordinated punishment in cooperative interactions. *Nature human behaviour, 3*(11), 1145-1153. https://doi.org/10.1038/s41562-019-0707-2.

Molleman L, Pen I, & Weissing FJ (2013a) Effects of Conformism on the Cultural Evolution of Social Behaviour. *PLoS ONE 8*(7): e68153. https://doi.org/10.1371/journal.pone.0068153.

Molleman, L., Quiñones, A. E., & Weissing, F. J. (2013b). Cultural evolution of cooperation: the interplay between forms of social learning and group selection. *Evolution and Human Behavior*, *34*(5), 342-349. https://doi.org/10.1016/j.evolhumbehav.2013.06.001.

Molleman, L., Van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications, 5*(3570), 1-9. https://doi.org/10.1038/ncomms4570.

Morgan, T. J.H., Acerbi, A., & Van Leeuwen, E. J.C. (2019). Copy-the-majority of instances or individuals? Two approaches to the majority and their consequences for conformist decision-making. *PloS one*, *14*(1), e0210748. https://doi.org/10.1371/journal.pone.0210748.

Morgan, T. J. H., & Laland, K. N. (2012). The biological bases of conformity. *Frontiers in neuroscience*, *6*, 87. DOI: 10.3389/fnins.2012.00087.

Müller, G. B. (2007). Evo–devo: extending the evolutionary synthesis. *Nature reviews genetics*, *8*(12), 943-949. https://doi.org/10.1038/nrg2219.

Mullon, C., Wakano, J. Y., & Ohtsuki, H. (2021). Coevolutionary dynamics of genetic traits and their long-term extended effects under non-random interactions. *Journal of Theoretical Biology*, *525*, 110750. doi: https://doi.org/10.1016/j.jtbi.2021.110750.

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological science*, *31*(6), 678-701. https://doi.org/10.1177/0956797620916782.

Muthukrishna, M., Doebeli, M., Chudek, M., & Henrich, J. (2018). The Cultural Brain Hypothesis: How culture drives brain expansion, sociality, and life history. *PLoS computational biology*, *14*(11), e1006504. https://doi.org/10.1371/journal.pcbi.1006504.

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, *3*, 221-229. https://doi.org/10.1038/s41562-018-0522-1.

Muthukrishna, M., Morgan, T. J.H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior, 37*(1), 10-20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004.

Muthukrishna, M., & Schaller, M. (2020). Are collectivistic cultures more prone to rapid transformation? Computational models of cross-cultural differences, social network structure, dynamic social influence, and cultural change. *Personality and Social Psychology Review*, *24*(2), 103-120. https://doi.org/10.1177/1088868319855783.

Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and human behavior*, *26*(1), 88-105. https://doi.org/10.1016/j.evolhumbehav.2004.08.002.

Nettle, D. (2009). Beyond nature versus culture: cultural variation as an evolved characteristic. *Journal of the Royal Anthropological Institute*, *15*(2), 223-240. https://doi.org/10.1111/j.1467-9655.2009.01561.x.

Novak, L. (2020). Persistent norms and tipping points: The case of female genital cutting. *Journal of Economic Behavior & Organization*, *177*, 433-474. https://doi.org/10.1016/j.jebo.2020.05.016.

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, *92*, 216-229. https://doi.org/10.1016/j.chb.2018.11.027.

Pearl, J. (2015). Causes of effects and effects of causes. *Sociological Methods & Research*, *44*(1), 149-164. DOI: 10.1177/0049124114562614

Pfungst, O. (1911). Clever Hans:(the horse of Mr. Von Osten): a contribution to experimental animal and human psychology. *Journal of Philosophical Psychological Scientific Method*, *8,* 663-666.

Pietraszewski, D., & A. E., Wertz (2021). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science* DOI: 10.1177/1745691621997113. [preprint].

Pisor, A. C., Gervais, M. M., Purzycki, B. G., & Ross, C. T. (2020). Preferences and constraints: the value of economic games for studying human behaviour. *Royal Society open science*, *7*(6), 192090. https://doi.org/10.1098/rsos.192090.

Power, M. L., & Schulkin, J. (2013). *The evolution of obesity*. US: John Hopkins University Press.

Price, M. E. (2006). Judgments about cooperators and freeriders on a Shuar work team: An evolutionary psychological perspective. *Organizational Behavior and Human Decision Processes, 101*(1), 20-35. https://doi.org/10.1016/j.obhdp.2006.06.001.

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Social Justice Research*, *21*(2), 228-240. https://doi.org/10.1007/s11211-008-0064-1.

Price M.E., Johnson D.D.P. (2011) The Adaptationist Theory of Cooperation in Groups: Evolutionary Predictions for Organizational Cooperation. In: Saad G. (eds) *Evolutionary Psychology in the Business Sciences*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-92784-6_5.

Price, M. E., & Van Vugt, M. (2014). The evolution of leader–follower reciprocity: the theory of service-for-prestige. *Frontiers in human neuroscience, 8,* 363. https://doi.org/10.3389/fnhum.2014.00363.

Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in cognitive sciences*, *13*(10), 420-428. https://doi.org/10.1016/j.tics.2009.08.002.

Reader, S.M. (2003). Innovation and social learning: individual variation and brain evolution. *Animal biology*, *53*(2), 147-158. DOI: https://doi.org/10.1163/157075603769700340.

Reader, S.M. (2006). Evo-devo, modularity, and evolvability: Insights for cultural evolution. *Behavioral and Brain Sciences, 29*(4), 361-362. doi:10.1017/S0140525X06399086.

Rendell, L., Fogarty, L., Hoppitt, W. J.E., Morgan, T. J.H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in cognitive sciences*, *15*(2), 68-76. https://doi.org/10.1016/j.tics.2010.12.002.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E.K., Naar, N., Newson, L., Ross, C., Smaldino, P.E., Waring, T.M. & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*, e30. doi:10.1017/S0140525X1400106X.

Richerson, P. J., & Boyd, R. (1978). A dual inheritance model of the human evolutionary process I: Basic postulates and a simple model. *Journal of Social and Biological Structures*, *1*(2), 127-154. https://doi.org/10.1016/S0140-1750(78)80002-5.

Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. US: University of Chicago press.

Richerson, P. J., & Boyd, R. (2020). The human life history is adapted to exploit the adaptive advantages of culture. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190498. https://doi.org/10.1098/rstb.2019.0498.

Richerson, P. J., Boyd, R., & Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences*, *107*(Supplement 2), 8985-8992. https://doi.org/10.1073/pnas.0914631107.

Richerson, P., & Henrich, J. (2012). Tribal Social Instincts and the Cultural Evolution of Institutions to Solve Collective Action Problems. *Cliodynamics*, 3(1), 38-80. http://dx.doi.org/10.21237/C7clio3112453.

Riebel, K., Lachlan, R. F., & Slater, P. J.B. (2015). Learning and cultural transmission in chaffinch song. *Advances in the Study of Behavior*, *47*, 181-227. https://doi.org/10.1016/bs.asb.2015.01.001.

Romero-Mujalli, D., Cappelletto, J., Herrera, E. A., & Tárano, Z. (2017). The effect of social learning in a small population facing environmental change: an agent-based simulation. *Journal of Ethology*, *35*(1), 61-73. DOI 10.1007/s10164-016-0490-8.

Rorabaugh, A. N. (2014). Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *Journal of Archaeological Science*, *49*, 255-264. https://doi.org/10.1016/j.jas.2014.05.016.

Rosenthal, G., Sporns, O., & Avidan, G. (2017). Stimulus dependent dynamic reorganization of the human face processing network. *Cerebral Cortex*, *27*(10), 4823-4834. https://doi.org/10.1093/cercor/bhw279.

Rutz, C., Hunt, G. R., & St Clair, J. J.H. (2018). Corvid technologies: how do new caledonian crows get their tool designs?. *Current Biology*, *28*(18), R1109-R1111. https://doi.org/10.1016/j.cub.2018.08.031.

Safin, V., Arfer, K. B., & Rachlin, H. (2015). Reciprocation and altruism in social cooperation. *Behavioural processes*, *116*, 12-16. https://doi.org/10.1016/j.beproc.2015.04.009.

Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior*, *36*(2), 86-94. https://doi.org/10.1016/j.evolhumbehav.2014.09.004.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and society*, *7*(1), 58-92. DOI: https://doi.org/10.1177/1043463195007001004.

Schalk, G., Kapeller, C., Guger, C., Ogawa, H., Hiroshima, S., Lafer-Sousa, R., Saygin, Z.M., Kamada, K. & Kanwisher, N. (2017). Facephenes and rainbows: Causal evidence for functional and anatomical specificity of face and color processing in the human brain. *Proceedings of the National Academy of Sciences*, *114*(46), 12285-12290. https://doi.org/10.1073/pnas.1713447114.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental science*, *13*(4), 599-610. DOI: 10.1111/j.1467-7687.2009.00913.x.

Smaldino, P. E. (2019). Social identity and cooperation in cultural evolution. *Behavioural processes*, *161*, 108-116. https://doi.org/10.1016/j.beproc.2017.11.015.

Smaldino, P. E., Flamson, T. J., & McElreath, R. (2018). The evolution of covert signaling. *Scientific reports*, *8*(4905), 1-10. DOI: https://doi.org/10.1038/s41598-018-22926-1.

Smaldino, P. E., Janssen, M. A., Hillis, V., & Bednar, J. (2017). Adoption as a social marker: Innovation diffusion with outgroup aversion. *The Journal of Mathematical Sociology*, *41*(1), 26-45. DOI: 10.1080/0022250X.2016.1250083

Smaldino, P., & Jones, J.H. (2021). Coupled dynamics of behaviour and disease contagion among antagonistic groups. *Evolutionary Human Sciences, 3*, E28. doi:10.1017/ehs.2021.22.

Smith, M. L. (2014). The archaeology of urban landscapes. *Annual Review of Anthropology*, *43*, 307-323. https://doi.org/10.1146/annurev-anthro-102313-025839.

Smith, H. F., & Wright, W. (2018). Vestigial organ. In J. Vonk and T.K. Shackelford (Eds.) *Encyclopedia of Animal Cognition and Behavior*. London: Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_406-1.

Sosis, R., Kress, H. C., & Boster, J. S. (2007). Scars for war: Evaluating alternative signaling explanations for cross-cultural variance in ritual costs. *Evolution and Human Behavior*, *28*(4), 234-247. https://doi.org/10.1016/j.evolhumbehav.2007.02.007.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L.A. Hirschfeld & S.A. Gelman (Eds.) *Mapping the mind: Domain Specificity in Cognition and Culture* (pp. 39-87). Australia: Cambridge University Press.

Sperber, D. (2001). In defense of massive modularity. In E. Dupoux (Ed.) *Language, brain and cognitive development: Essays in honor of Jacques Mehler*, (pp. 47-58). England: MIT Press.

Sperber, D., & Girotto, V. (2002). Use or misuse of the selection task? Rejoinder to Fiddick, Cosmides, and Tooby. *Cognition*, *85*(3), 277-290. https://doi.org/10.1016/S0010-0277(02)00125-7.

Sperber, D., & Hirschfeld, L.R. (1999). Culture, cognition, and evolution. In R. Wilson & F.Keil (Eds.), *MIT encyclopedia of the cognitive sciences*, (pp. 111-132). Cambridge, MA: MIT Press.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in cognitive sciences*, *8*(1), 40-46. https://doi.org/10.1016/j.tics.2003.11.002.

Sperber, D., & Mercier, H. (2018). Why a modular approach to reason?. *Mind & Language*, *33*(5), 533-541. DOI: 10.1111/mila.12208.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nature Reviews Neuroscience*, *18*(9), 559-567. DOI: https://doi.org/10.1038/nrn.2017.76.

Stephen, I. D. (2014). Putting the theory before the data: is “massive modularity” a necessary foundation of evolutionary psychology?. *Frontiers in psychology*, *5*, 1158. https://doi.org/10.3389/fpsyg.2014.01158.

Stieglitz, S. & Dang-Xuan, L. (2013) Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior, *Journal of Management Information Systems*, *29*(4), 217-248, DOI: 10.2753/MIS0742-1222290408.

Stokes, D., & Bergeron, V. (2015). Modular architectures and informational encapsulation: A dilemma. *European Journal for Philosophy of Science*, *5*(3), 315-338. https://doi.org/10.1007/s13194-015-0107-z.

Stout, D., & Hecht, E. E. (2017). Evolutionary neuroscience of cumulative culture. *Proceedings of the National Academy of Sciences*, *114*(30), 7861-7868. https://doi.org/10.1073/pnas.1620738114.

Stubbersfield, J. M., Tehrani, J. J., & Flynn, E. G. (2017). Chicken Tumours and a Fishy Revenge: Evidence for Emotional Content Bias in the Cumulative Recall of Urban Legends, *Journal of Cognition and Culture*, *17*(1-2), 12-26. doi: https://doi.org/10.1163/15685373-12342189.

Suchow, J. W., Bourgin, D. D., & Griffiths, T. L. (2017). Evolution in mind: Evolutionary dynamics, cognitive processes, and bayesian inference. *Trends in cognitive sciences*, *21*(7), 522-530. https://doi.org/10.1016/j.tics.2017.04.005.

Szolnoki, A., & Perc, M. (2015). Conformity enhances network reciprocity in evolutionary social dilemmas. *Journal of The Royal Society Interface*, *12*(103), 20141299. https://doi.org/10.1098/rsif.2014.1299.

Takezawa, M., & Price, M. E. (2010). Revisiting "The Evolution of Reciprocity in Sizable Groups": Continuous reciprocity in the repeated n-person prisoner's dilemma. *Journal of theoretical biology*, *264*(2), 188-196. https://doi.org/10.1016/j.jtbi.2010.01.028.

Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, *8*, 37-43. https://doi.org/10.1016/j.copsyc.2015.09.003.

Tanaka, M. M., Kendal, J. R., & Laland, K. N. (2009). From traditional medicine to witchcraft: why medical treatments are not always efficacious. *PLoS One*, *4*(4), e5192. https://doi.org/10.1371/journal.pone.0005192.

Tarantola, A. (2006). Popper, Bayes and the inverse problem. *Nature physics*, *2*(8), 492-494. https://doi.org/10.1038/nphys375.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2405-2415. https://doi.org/10.1098/rstb.2009.0052.

Thielmann, I., Böhm, R., Ott, M., & Hilbig, B. E. (2021). Economic games: An introduction and guide for research. *Collabra: Psychology*, *7*(1), 19004. https://doi.org/10.1525/collabra.19004.

Thomasson, H. R., Edenberg, H. J., Crabb, D. W., Mai, X. L., Jerome, R. E., Li, T. K., Wang, S.P., Lin, Y.T., Lu, R.B. & Yin, S. J. (1991). Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *American journal of human genetics*, *48*(4), 677-681.

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortenson, H.M., Hirbo, J.B., Osman, M., Ibrahim, M., Omar, S.A., Lema, G., Nyambo, T.B., Ghori, J., Bumpstead, S., Pritchard, J.K., Wray, G.A. & Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature genetics, 39*(1), 31-40. DOIhttps://doi.org/10.1038/ng1946.

Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2014). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences, 281*(1776), 20132864. https://doi.org/10.1098/rspb.2013.2864.

Tomasello, M., & Gonzalez-Cabrera, I. (2017). The role of ontogeny in the evolution of human cooperation. *Human Nature, 28*(3), 274-288. DOIhttps://doi.org/10.1007/s12110-017-9291-1.

Tomka, S. (2013). The Adoption of the Bow and Arrow: A Model Based on Experimental Performance Characteristics. *American Antiquity, 78*(3), 553-569. doi:10.7183/0002-7316.78.3.553.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and sociobiology*, *11*(4-5), 375-424. https://doi.org/10.1016/0162-3095(90)90017-Z.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J.H. Barkow, L. Cosmides and J.Tooby's (Eds.) *The adapted mind: Evolutionary psychology and the generation of culture*, (p19-136). New York: Oxford University Press.

Tooby, J., & Cosmides, L. (2005). Conceptual Foundations of Evolutionary Psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). John Wiley & Sons, Inc.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for n-person exchange: the evolutionary roots of organizational behavior. *Managerial and Decision Economics*, *27*(2-3), 103-129. https://doi.org/10.1002/mde.1287.

Uchiyama, R., Spicer, R., & Muthukrishna, M. (2021). Cultural Evolution of Genetic Heritability. *Behavioral and Brain Sciences*, 1-147. doi:10.1017/S0140525X21000893.

van den Berg, P., Molleman, L., & Weissing, F. J. (2012). The social costs of punishment. *Behavioral and Brain Sciences*, *35*(1), 42-43. DOI:10.1017/S0140525X11001348.

van den Bos, R., & de Ridder, D. (2006). Evolved to satisfy our immediate needs: Self-control and the rewarding properties of food. *Appetite*, *47*(1), 24-29. https://doi.org/10.1016/j.appet.2006.02.008.

Van Lier, J., Revlin, R., & De Neys, W. (2013). Detecting cheaters without thinking: Testing the automaticity of the cheater detection module. *PloS one*, *8*(1), e53827. https://doi.org/10.1371/journal.pone.0053827.

Van Schaik, C. P., Isler, K., & Burkart, J. M. (2012). Explaining brain size variation: from social to cultural brain. *Trends in cognitive sciences*, *16*(5), 277-284. https://doi.org/10.1016/j.tics.2012.04.004.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource?. *Psychological science*, *21*(3), 384-390. https://doi.org/10.1177/0956797610361340.

Vergauwe, E., von Bastian, C. C., Kostova, R., & Morey, C. C. (2021). Storage and processing in working memory: A single, domain-general resource explains multi-tasking. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000895. [preprint].

Vogt, S., Efferson, C., Berger, J., & Fehr, E. (2015). Eye spots do not increase altruism in children. *Evolution and Human Behavior*, *36*(3), 224-231. DOI: https://doi.org/10.1016/j.evolhumbehav.2014.11.007.

Wakano, J. Y., & Aoki, K. (2006). A mixed strategy model for the emergence and intensification of social learning in a periodically changing natural environment. *Theoretical population biology*, *70*(4), 486-497. https://doi.org/10.1016/j.tpb.2006.04.003.

Wang, X., Duh, M., & Perc, M. (2020). Pool expulsion and cooperation in the spatial public goods game. *Physics Letters A*, *384*(18), 126391. https://doi.org/10.1016/j.physleta.2020.126391.

Wen, N. J., Clegg, J. M., & Legare, C. H. (2019). Smart conformists: Children and adolescents associate conformity with intelligence across cultures. *Child development*, *90*(3), 746-758. DOI: 10.1111/cdev.12935.

Wen, N. J., Willard, A. K., Caughy, M., & Legare, C. H. (2020). Watch me, watch you: ritual participation increases in-group displays and out-group monitoring in children. *Philosophical Transactions of the Royal Society B, 375*(1805), 20190437. https://doi.org/10.1098/rstb.2019.0437.

Whitehead, H., Laland, K. N., Rendell, L., Thorogood, R., & Whiten, A. (2019). The reach of gene–culture coevolution in animals. *Nature Communications*, *10*(2405), 1-10. https://doi.org/10.1038/s41467-019-10293-y.

Whiten, A. (2019). Conformity and over-imitation: An integrative review of variant forms of hyper-reliance on social learning. *Advances in the Study of Behavior*, *51*, 31-75. https://doi.org/10.1016/bs.asb.2018.12.003.

Wilson, B.J. (2016). The meaning of deceive in experimental economic science. In G.F. DeMartino & D.N. McClosekey (Eds.), *The Oxford handbook of professional economics ethics* (pp.317-329). US: Oxford University Press.

Wilson, R.K. & Rhodes, C.M. (1997). Leadership and Credibility in N-Person Coordination Games. *Journal of Conflict Resolution, 41*(6), 767-791. DOI: https://doi.org/10.1177/0022002797041006003.

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, *33*(4), 341-356. https://doi.org/10.1016/j.dr.2013.08.002.

Wrangham, R. (2009). *Catching fire: How cooking made us human.* UK: Profile Books.

Zimmerman, J. & Efferson, C. (2017). One-shot reciprocity under error management is unbiased and fragile. *Evolution and Human Behaviour*, *38*(1), 39-47. https://doi.org/10.1016/j.evolhumbehav.2016.06.005.

Appendices

Appendix 1: See this link (https://osf.io/u49nv/?view_only=01afea04127d41e187eab2755bd99103) for the R script used to create Figure 4 in Chapter 8 (the coloured line graph showing the social learning strategies across the three economic games of interest). The script is called *plotBoot_masterGraph_allPlot_LineGraph.R.*

Appendix 2: The same OSF link above also contains the supplementary materials for Chapter 5 and Chapter 6. As the model of asocial skills and social norms was similar in nature, then the full model specifications are given in the same text file. Note that the cooperation in Chapter 7 takes a different structure, and so the supplementary materials for that chapter are contained within that appendix. The supplementary material for Chapters 5 and 6 is attached in the link above, and is called *M1_M2_SMjoint.doc*

Appendix 3: The same OSF link above also contains a glossary of key words used throughout my thesis for ease of reference. See the file titled: *Glossary of key words used in the thesis.doc.*