DOCTOR OF PHILOSOPHY

The Control of Automatic Imitation - Neural Mechanisms and Individual Differences

Darda, Kohinoor

Award date:
2019

Awarding institution:
Bangor University

DOCTOR OF PHILOSOPHY

The Control of Automatic Imitation - Neural Mechanisms and Individual Differences

Darda, Kohinoor

Award date:
2019

PRIFYSGOL
BANGOR
UNIVERSITY

The Control of Automatic Imitation:

Neural Mechanisms and Individual Differences

Kohinoor Monish Darda

Thesis submitted to the School of Psychology, Bangor University, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Bangor, United Kingdom

July 2019

SoBA Lab
SOCIAL BRAIN IN ACTION

ACKNOWLEDGEMENTS

When I first started my PhD, many people both in and outside of academia told me that it would be challenging and a lot of hard work, and the rewards might not be worth it. I have to say that I have not only found every day of my PhD enjoyable, but also grown, both personally and professionally, irrespective (and perhaps because) of the challenges and obstacles that I have faced. This would not have been possible without the support I have received from everyone around me.

First and foremost, I would like to thank my supervisor Dr. Richard Ramsey who accepted me in his lab in August 2016. Since then, his door has been open for even the smallest of questions, and I am extremely grateful for his guidance every step of the way. I owe him a monumental thank you for encouraging me to do my best, inspiring me with his scientific vision and commitment to rigorous and open science practice, and showing me the good side of academia. I look forward to working and collaborating with him in the future as well! I would also like to thank my second supervisor Prof. Paul Downing and my PhD committee chair Dr. Ken Valyear for their feedback, comments, and guidance throughout my PhD.

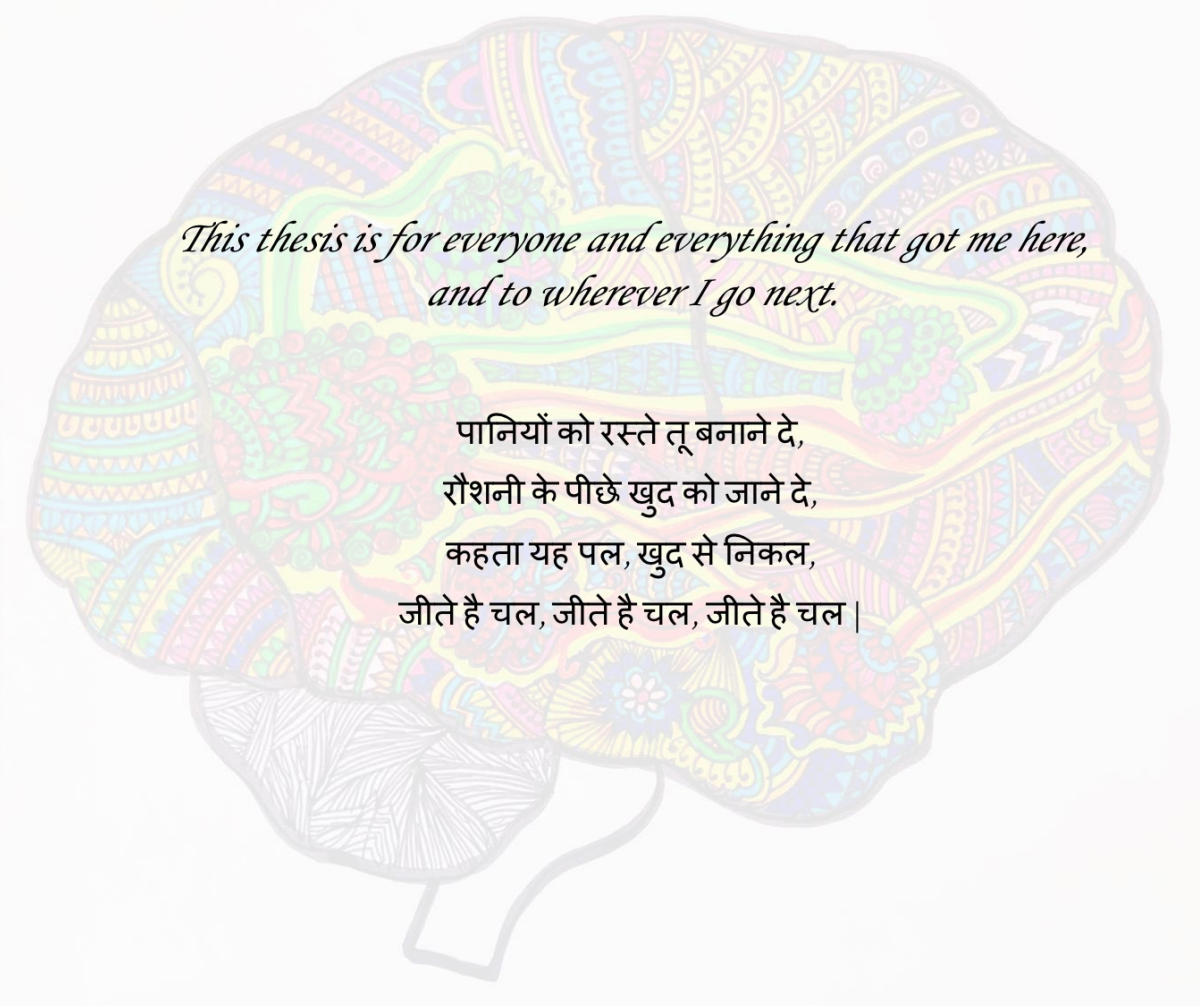
I would also like to express my gratitude to my academic mentor Prof. Emily Cross – she has been an inspiration both personally and professionally, and I am glad to have such a strong example to follow in my academic career. To all members of the Social Brain in Action Laboratory, both past and present – I am grateful for your help and feedback at various stages during my PhD. Specifically, I would like to thank Dr. Emily Butler whose research, data, and support have contributed immensely to my thesis. Of course, a BIG thank you goes to our lab's Social Dognition representative, Bowie, who kept me sane throughout!

I am also extremely grateful to the collaborative environment and exceptional facilities provided by the School of Psychology at Bangor University, in particular the Bangor Imaging Unit, as well as all faculty members and fellow graduate students and friends who provided a stimulating and supportive environment to discuss my research and do my work in.

A monumental thank you also goes to my friends for tolerating my idiosyncrasies and believing in my impossible dreams. Sana, thank you for being my person. I would be lost otherwise. Neerja, वक़्त की कैद में ज़िन्दगी है मगर, चंद घड़ियाँ यही है जो आज़ाद है। Shreya,

you're the calm to our chaos. Zobair, آنچه شما به دنبالش هستید اوبه دنبال شماست, thank you for being the *Rumi* to my *Shams*. Angad – तू सब्र है, मैं सब्र हूँ, तो यह ज़िन्दगी काफी नहीं।

I am forever indebted to my family for all the love and care. A HUGE thank you to my grandparents Sudha and Mangalkumar, my parents Monish and Kalpana, my brothers Mayoor and Noopur, and my four-legged balls of fur Jazz and Cosco. Nothing is possible without your support – you keep me motivated and inspired, and have been a constant source of support throughout my life. I am here only because of you, and I look forward to what the future holds, and all the continued acknowledgements I will be sending your way!



This thesis is for everyone and everything that got me here,
and to wherever I go next.

पानियों को रस्ते तू बनाने दे,
रौशनी के पीछे खुद को जाने दे,
कहता यह पल, खुद से निकल,
जीते है चल, जीते है चल, जीते है चल |

RESEARCH OUTPUT

Research outputs from this thesis:

Darda, K. M., Butler, E. E., & Ramsey, R. (2018). Functional specificity and sex differences in the neural circuits supporting the inhibition of automatic imitation. *Journal of Cognitive Neuroscience,* 30, 914-933.

Darda, K. M., & Ramsey, R. (2019). The inhibition of automatic imitation: a meta-analysis and synthesis of fMRI studies. *NeuroImage,* 170, 320-329.

Darda, K. M., Butler, E. E., & Ramsey, R. (2019, May 23). Individual differences in social and non-social cognitive control. Preprint: https://psyarxiv.com/8b4cz/.

Research outputs related to this thesis:

Ramsey, R., **Darda, K. M.,** & Downing, P. E., (2019). Automatic imitation remains unaffected under cognitive load. *Journal of Experimental Psychology: Human Perception and Performance*, 45(5), 601-615.

Related research outputs in preparation:

Darda, K.M., & Ramsey, R. (in prep). Neural mechanisms of social and non-social control.

SUMMARY

Automatic imitation, or an involuntary tendency to imitate others, is a ubiquitous behaviour that is central to our interactions in the social world. Despite centuries of interest in this phenomenon from philosophers and scientists across different disciplines, many open questions still remain. The current thesis employs approaches from cognitive psychology and social cognitive neuroscience to elucidate the underlying cognitive and neural mechanisms of the control of automatic imitation and how these mechanisms vary as a function of individual differences. The first empirical chapter (Chapter 2) uses functional magnetic resonance imaging (fMRI) across two experiments in order to investigate whether specialised or generalised neural mechanisms underlie the control of automatic imitation. The second empirical chapter (Chapter 3) synthesises and meta-analyses extant neuroimaging literature in order to identify brain regions that are consistently activated across fMRI studies investigating automatic imitation. In Chapter 4, multiple large-sample behavioural approaches are employed to investigate the relationship between individual differences (stable personality traits and biological sex) and social (imitative) control and non-social control. Overall, the results from this thesis unequivocally support the engagement of a domain-general neural network in the control of automatic imitation, and a reduced or altered role for domain-specific processes. More generally, these findings suggest that models of social cognition need to place greater emphasis on the role of domain-general processes, and the interactions between domain-specific and domain-general processes, instead of focusing only on domain-specificity. Further, the control of automatic imitation is largely invariant to stable traits of personality and biological sex. However, the cognitive and neural underpinnings of individual differences in social and non-social control are more complex than what has been previously conceived. In sum, the current findings have important implications for and shed new light on the methodological and theoretical debates surrounding automatic imitation as well as social cognition.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

GENERAL INTRODUCTION

Perhaps one of the greatest questions that humankind has ever contemplated is itself. The prospect of answering this question by trying to understand the complexity of the human brain has been the quest for many philosophers and scientists. It has been claimed that on average, the brain allows an individual to hold up to one million gigabytes of information, process up to 120 bits/second, and interact with 80,000 people in an entire lifetime (Levitin, 2014; Reber, 2010; Vital, 2003). While the empirical evidence behind these claims argues on the exact numbers, the consensus seems to be that humans interact with a large number of people, process an immense amount of information, and seem to do so with relative ease (Adolphs, 1999; McCormick, Salganik, & Zheng, 2010; Mead & Kurzweil, 2006).

In the last few decades, the field of cognitive neuroscience has provided enormous insight into the mechanisms underlying higher level cognitive functions such as attention, memory, problem solving, executive functioning as well as consciousness (Gazzaniga, 2009). These investigations consider the individual as a processor of information and focus on the individual brain in isolation (Singer, 2012). However, human beings exist in an interesting paradox – although we consider ourselves as individual entities, our day-to-day lives are embedded in social interactions with other individuals. Researchers have argued that the ability of humans to form large and complex social groups is what distinguishes them from other animal species, and drives their intellectual and cultural development (Adolphs, 2001; Dunbar, 2009). The nascent field of social cognitive neuroscience allows for the investigation of social processes through the lens of cognitive psychology, social psychology as well as neuroscience (Cacioppo & Berntson, 1992; Ochsner & Lieberman, 2001). Additionally, an important consideration for social cognitive neuroscientists has been to investigate whether social cognition is indeed "special" i.e. whether it relies on special purpose mechanisms, or whether it is just one instance of general-purpose mechanisms such as attention, memory, or executive functioning (Barrett, 2012). Understanding the complex brain and cognitive processes,

whether specialised or generalised, that underlie our ability to navigate in a social world is central to understanding the complexity of the human brain.

In the current thesis, I draw upon approaches from cognitive psychology and social cognitive neuroscience to gain insight into the neural and cognitive mechanisms underlying automatic imitation, including the extent to which such mechanisms exhibit specialised and/or generalised functionality. Imitation is a critical behaviour that has rich social consequences, facilitating our social interactions (Heyes, 2009; 2011). In order to fully comprehend the mechanisms that underlie our capacity to navigate in a complex social world, it is essential to understand non-verbal cues such as automatic copying behaviours that guide social interactions (Hamilton, 2014). Automatic imitation, the involuntary tendency of humans to copy others, has been the focus of much systematic investigation across different domains including developmental psychology, evolutionary biology, cognitive and social psychology, as well as social cognitive neuroscience (Heyes, 2011). However, many open questions still remain – what are the neural mechanisms that underlie our tendency to automatically imitate, are these mechanisms specialised for imitation, and how do these imitative tendencies differ as a function of individual differences? Systematic investigations to answer these questions will help us to better understand and comprehend the nature of automatic imitation, and social interactions more generally.

In the following sections of this chapter, in order to place the focus of my thesis in the context of relevant literature from psychological science, I provide a critical analysis of extant literature on the neural mechanisms of and individual differences in automatic imitation. First, I provide a general background and history of imitation including the different types, definitions, and measurements of automatic imitation (Section 1.1.). I then review behavioural evidence on the factors modulating automatic imitation, and how the tendency to automatically imitate differs between individuals (Section 1.2.). Next, I critically evaluate existing neuroimaging evidence on whether specialised or generalised neural circuits underpin our tendency to automatically imitate (Section 1.3). In the last section of this chapter (Section 1.4), I identify gaps in the literature and methodological issues which my thesis will aim to address.

1.1. Introduction

1.1.1. History and background.

Few areas of current research have provided as much insight into the nature of human social cognition as the fundamental properties and underlying mechanisms of human imitation. Historically, from Plato's and Aristotle's theory of 'mimesis' (the Greek word for imitation) to current investigations of imitative behaviours, there is a whole spectrum of conceptions of imitation. These conceptions range from a capacity to bring about similar behaviour on one end, to the "representation," "transmission," "contagion," "synchrony" or "sharing" of affect states or behavioural patterns between individuals on the other end (for reviews, see Bavelas, Black, Lemery, & Mullett, 1987; Galef, 2013; Zentall, 2006). Indeed, different disciplines with different origins have produced new layers of meaning for the term imitation (Maran, 2017). From a very simple conceptualisation of imitation used in common parlance (e.g. "monkey see, monkey do") that implies copying as a simple motor act with no rich social consequences, it is now generally accepted that imitation is a much more complex and multi-dimensional phenomenon that serves important social functions across multiple species.

While it may seem that imitation is fairly easy to recognise, scientists have debated over the nature of imitation since the 19th century. Imitation has been the focus of much theoretical interest for biologists like Darwin (1871), Wallace (1870), and Romanes (1884) who discussed learning by imitation in animals, and thought of imitation as an evolutionary mechanism that facilitated continuity across generations. The late 19th century marked the emergence of Psychology as a separate scientific discipline, with an increasing interest in imitation behaviours. Many philosophers and scientists in the early history of psychology including pragmatists, experimental psychologists, behavioural scientists, as well as developmental psychologists studied and speculated about the nature of human imitation, arguing over whether imitation was instinctual, reflective, or a combination of both (Berry, 1906; Haggerty, 1909; James, 1961; Kohler, 1925; Morgan, 1900; Piaget, 1952; Thorndike, 1911; Witmer, 1910). These investigations led to divergent definitions and classifications of the term, as well as disagreement over underlying processes, primarily due to the use of differing methodologies (Kymissis & Poulson, 1990).

With the discovery of the mirror neuron system (MNS) in the late 20th century, imitation became a topic of investigation not only within social, cognitive, and

experimental psychology, but also the more nascent fields of social and cognitive neuroscience. The MNS refers to areas in the brain which contain mirror neurons – neurons that fire during the observation and execution of the same actions (Rizzolatti & Craighero, 2004; Iacoboni, 2009). Although the MNS was initially discovered in macaque monkeys by neurophysiologists using single-unit recordings in the ventral premotor cortex (area F5; de Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), growing evidence suggests that it exists in humans as well (Kilner, Neal, Weiskopf, Friston, & Frith, 2009; Molenberghs, Cunnington, & Mattingley, 2012). Given that the mirror neurons possessed both visual and motor properties, the discovery of the MNS had a revolutionary impact on investigation of perception-action links (Ramachandran, 2000). Researchers purported links between the MNS and imitation, action understanding, empathy, as well as language development (Gallese, 2001; Rizzolatti & Airbib, 1998). Consequently, research on imitation increased, as it had the potential to play an important role in investigating the functions of the MNS.

Even after almost two centuries of research using different approaches, however, scientists have not yet agreed on clear-cut definitions and classifications of different types of imitation (for different terminologies used and types of imitation, see Table 1). Broadly speaking, imitation refers to the copying or reproducing of observed behaviours or actions of another individual (Heyes, 2011). Although there is disagreement about definition, there is wide consensus that imitation is an important and complex phenomenon, and is a fundamental aspect of human behaviour. It has traditionally been suggested that humans imitate to serve two main functions – learning, and communicating mutuality with others (Uzgiris, 1981). Indeed, imitation is argued to be the foundation of skill and language acquisition, and learning novel behaviours (Flynn & Smith, 2012; Tomasello, Savage-Rumbaugh, & Kruger, 1993; Tomasello, 1999). For example, a dancer learns a new dance routine by observing and copying the movements shown by the dance instructor. Not only does imitation function as a mechanism for learning new skills, it is also thought to be the cornerstone of social learning, development, and connectedness throughout the lifespan, as well as the transfer of cultural knowledge (Over & Carpenter, 2013; Uzgiris, 1981; Lakin, Jefferis, Cheng, & Chartrand, 2003; Heyes, 2009; 2011). For example, when we travel to a new place, it is likely that we adjust our behaviour to that of those around us.

1. GENERAL INTRODUCTION

Research investigating imitation behaviours in humans, non-human primates, as well in other animals has provided some grounds for distinguishing between different forms of imitation. One important distinction is based on intentionality – imitation can be intentional/conscious (voluntary) or automatic/non-conscious (involuntary) (Hamilton, 2014; Whiten, Horner, Litchfield, & Marshall-Pescini, 2004). Imitation is generally intentional when learning a new skill, while in the case of social interactions, it is often involuntary, and enhances social connectedness (Chartrand & Bargh, 1999; R. Cook, Bird, Lünser, Huck, & Heyes, 2012; Hamilton, 2014; Heyes, 2011). The focus of the current thesis is on automatic imitation: imitative behaviour that occurs without the conscious awareness of either of the interacting partners.

Table 1

Different Definitions and Types of Imitation.

	Voluntary			**Involuntary**	
Term	**Definition**	**Example**	**Term**	**Definition**	**Example**
(True) Imitation	Copying the means and end of a goal-directed action (Whiten et al., 2004)	A student learning how to play the piano watches their instructor play a sequence of keys and then imitates the exact sequence	**Automatic Imitation* (Cognitive Psychology)**	Unintentional copying of actions or behaviours (Heyes, 2011)	SRC effect in which observing a compatible action facilitates performance and observing an incompatible action interferes with performance
Emulation	Copying the end or outcome by not necessarily the same means (Tomasello, 1990)	A child sees an adult throw a ball with one hand but throws the ball herself with two hands	**Mimicry* (Social Psychology)**	Unintentional copying of actions or behaviours (that usually occurs in social interactions; Chartrand & Bargh, 1999)	When conversing with one another, one friend adopts the body posture of the other without realising it
Over-imitation	Copying "extra" actions even though they are not necessary for the action goal (Hamilton, 2014)	A novice guitarist observes an instructor shake their fingers before playing the guitar and copies the same action			

*Automatic imitation is considered to be the laboratory equivalent for mimicry. It is assumed that these terms mean the same thing, but are named differently in cognitive and social psychology. *N.B.* SRC = Stimulus Response Compatibility.

1.1.2. What is automatic imitation?

Automatic or spontaneous imitation is defined as the involuntary tendency to copy others' actions, speech patterns, facial expressions, as well as other behaviours without the conscious awareness of either of the interacting partners (Chartrand & Bargh, 1999; Heyes, 2011). A friend recently pointed out that I imitate the accent of the person I am talking to – I speak with an English accent when conversing with friends in the United Kingdom, whereas my accent shifts to an Indian accent when conversing with my family. Indeed, automatic imitation is ubiquitous, and a critical behaviour for non-verbal communication (Chartrand & Bargh, 1999). While research on automatic imitation has become a key focus in the last two decades in both social psychology as well as cognitive neuroscience, theoretical interest in this phenomenon dates as far back as the 10th century in the work of Indian philosopher and aesthete, Abhinavagupta.

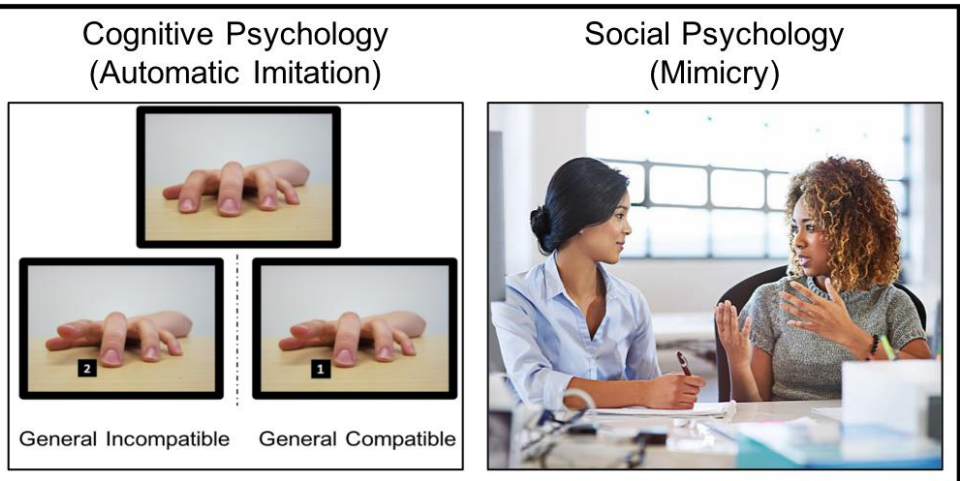
Abhinavagupta was intrigued by the capacity of humans to smile when others smile, clap when others clap, and feel a prick when they see someone else pricking their finger (Abhinavagupta, 1986 (*IPV)* as cited in Chakrabarti & Weber, 2015). Referred to as his complex "mirroring" theory of perception, he suggested that just as a sound made by another person is perceived because perceivers generate an echo within themselves, so does an action by another person generates an "echo" of that action in the observer, thus leading the observer to imitate this action spontaneously. This "mirroring", Abhinavagupta proposed, was the foundation for empathy, interpersonal connectedness, as well as (more transcendentally) the underlying unity of all sentience. As such, it is perhaps not that big an inductive leap to assume that he was unwittingly forecasting the mirror-neuron debates of the 21st century. In a similar vein, in the 18th century, philosopher and thinker Adam Smith (1759) purported a link between the tendency to spontaneously imitate and the capacity of humans to empathise with others.

Despite the long-standing theoretical interest in automatic imitation, scientists in the past focused on the role of imitation in learning, with less focus on the involuntary tendency of human beings to imitate others. More recently, however, there has been a call for investigating automatic imitation as it is considered to be a process central to social cognition, with important social consequences (Brass & Heyes, 2005; Hamilton, 2014; Heyes, 2011). Automatic imitation is thought to facilitate social interactions by increasing liking and affiliation between interacting partners, as well as increasing pro-social behaviour (Chartrand & Bargh, 1999; Chartrand & Lakin, 2013; van Baaren,

Janssen, Chartrand, & Dijksterhuis, 2009). Thus, in order to elucidate the mechanisms that underlie non-verbal communication as well as social interactions more generally, it is critical to study automatic imitation.

1.1.3. How is automatic imitation measured?

In the last two decades, there has been much investigation on automatic imitation in both social psychology as well as cognitive neuroscience. Both disciplines have used different approaches and methodologies (see Figure 1) in order to elucidate the nature of automatic imitation, its causes, consequences, as well as underlying mechanisms (Heyes, 2009).



	Stimulus response compatibility measure	Live social interaction measure
Paradigm	Laboratory-based	Naturalistic
Ecological Validity	Low	High
Reliability	High	Low
Time	Simultaneous inhibition of copying	Delayed copying
Measure	Reaction times or movement kinematics indexing covert copying behaviour	Observed overt copying behaviour

Figure 1. Measures of automatic imitation in cognitive and social psychology.

Social psychology measures. In the field of social psychology, automatic or involuntary imitation behaviour has been termed as "mimicry." Scientists have undertaken systematic investigation into the nature of mimicry – why we mimic, how we mimic, who we mimic, as well as the factors that affect mimicry. Social psychologists

generally use naturalistic paradigms that involve recording overt or visible copying behaviours (Figure 1; Ray & Heyes, 2011). Participants are often asked to engage in a simple task with another person (a confederate). For example, in one of the seminal social psychology studies evidencing mimicry, Chartrand & Bargh (1999, Experiment 1) asked participants to engage in a photo description task with the confederate. The confederate's actions were manipulated by the experimenters – for instance, the confederate either touched their face or shook their foot when the interaction was taking place, and the interactions were recorded with hidden video cameras. These interactions were later observed by coders blind to the experimental conditions, and it was found that participants substantially copied the movements of the confederate. They reported no awareness of these behaviours suggesting that it was an automatic or involuntary behaviour. The authors called this phenomenon as the "chameleon effect" – just as a chameleon changes colours to match the environment, we change our mannerisms to match those of others around us. This behaviour has been further evidenced not just for actions, gestures and body postures, but also facial expressions and speech patterns (Bailey & Henry, 2009; Bock, 1986; Chartrand & Bargh, 1999; Dimberg, 1982; Levelt & Kelter, 1982; Sheflen, 1964; Stel & Vonk, 2010).

Cognitive psychology measures. With the advent of the cognitive revolution in the 1950s, researchers in the domain of social psychology started adopting methodologies from cognitive psychology in order to provide experimental control over phenomena of interest (Lambert & Scherer, 2013). In imitation research, cognitive psychologists developed measures of automatic imitation based on stimulus-response compatibility (SRC) paradigms (e.g. Brass, Bekkering, & Wohlschager, 2000). In contrast to the more naturalistic social psychology paradigms, SRC tasks of automatic imitation measure the covert processes involved in controlling our tendency to automatically imitate using reaction time (RT) measures or kinematics (Heyes, 2011). These tasks are thought to be the laboratory equivalent of mimicry and therefore allow for more experimental control (Figure 1).

In typical stimulus-response compatibility paradigms, participants are asked to respond to stimuli which are either compatible or incompatible to their response. In automatic imitation research, specialised SRC paradigms were initially developed by Brass et al. (2000), Stürmer, Aschersleben, and Prinz (2000), and Kilner, Paulignan, and Blakemore (2003) in order to measure automatic imitation. Since then, different versions

of the initially developed SRC paradigms have been widely used in order to index automatic imitative processes, and usually involve finger movements (e.g. Bertenthal, Longo, & Kosobud, 2006; Brass et al., 2000; Catmur & Heyes, 2011), hand opening/closing movements (Heyes, Bird, Johnson, & Haggard, 2005; Press, Bird, Walsh, & Heyes, 2008) or arm movements in a vertical or horizontal plane (Kilner et al., 2003; Stanley, Gowen, & Miall, 2007).

In a typical SRC task indexing automatic imitation involving finger movements, participants are instructed to lift their index finger when they see a number '1' on the screen and lift their middle finger when they see a number '2' on the screen. Simultaneously, they either view an index finger lift movement or a middle finger lift movement on the screen. In compatible trials, the finger movement they execute is the same as the finger movement they observe (e.g. lifting their index finger when they see the number '1' on the screen and simultaneously observing an index finger lift movement). In incompatible trials, participants execute a finger movement which is different to the observed finger movement (e.g. lifting their index finger when they see the number '1' on the screen and simultaneously observing a middle finger lift movement). Participants are slower to respond to incompatible trials as the observed movement interferes with their movement execution. The difference between reaction times on incompatible and compatible trials is thus said to be a measure of automatic imitation (Heyes, 2011). This compatibility effect has been demonstrated not only with finger lifting/tapping movements, but also with finger abduction (Catmur & Heyes, 2011), hand opening/closing (Heyes et al., 2005), as well as mouth opening/closing movements (Leighton & Heyes, 2010).

Similarly, along with reaction times, this effect has also been demonstrated using movement kinematics. For example, in a study by Kilner et al. (2003), participants were instructed to move their arms in a vertical or horizontal plane while observing another person moving their hand in a vertical or horizontal plane. The movement trajectory of participants contained more variability when they were observing a movement incompatible to their own, compared to when the interacting partner performed a compatible movement (Bouquet, Gaurier, Shipley, Toussaint, & Blandin, 2007; Kilner et al., 2003; Stanley et al., 2007).

Comparison between social and cognitive psychology measures. Social psychology researchers use naturalistic paradigms which are more ecologically valid i.e.

they are directly related to "real-life" behaviours. In these paradigms, copying behaviour typically occurs after a delay of a few seconds (Bailenson & Yee, 2005). These tasks, however, are not amenable to experimental control in a laboratory setting, and can therefore be less reliable (Genschow et al., 2017). In contrast, SRC tasks used in cognitive psychology show higher reliability but low ecological validity. In these tasks, participants have to resolve conflict by inhibiting the representation of the observed action in order to perform the task. Typically, this behaviour occurs simultaneously to the observed action. Thus, whereas social psychology paradigms measure overt copying, SRC paradigms index the covert processes involved in the *inhibition* or *control* of automatic imitation (Figure 1; Heyes, 2011).

1.1.4. Is automatic imitation a specialised form of SRC?

A key issue in the investigation of automatic imitation is whether automatic imitation is a specialised form of stimulus-response compatibility. For instance, researchers have investigated whether and how much spatial processes contribute to the SRC measure of automatic imitation. In the task used by Brass et al. (2000), left hand images were used on the screen while participants responded with their right hand. When participants performed a compatible movement (e.g. lifted their index finger while watching an index finger lift movement on the screen), the observed and executed finger movements were not just of the same finger but also on the same side of space. Thus, the measure of automatic imitation in this task can be a combination of imitative as well as spatial compatibility effects (Brass, Zysset, & von Cramon, 2001). Researchers have controlled for spatial effects in a number of ways in order to measure automatic imitation independent of the spatial compatibility effect (see Figure 2; Aicken, Wilson, Williams, & Mon-Williams, 2007; Brass et al., 2001; Boyer, Longo, & Bertenthal, 2012; Jansson, Wilson, Williams, & Mon-Williams, 2007; Jimenez et al., 2012; Press et al., 2008).

Bertenthal and colleagues (2006) displayed the stimuli orthogonal to the participant's hand and also evidenced an automatic imitation effect (Figure 2A). While the presentation of orthogonal stimuli reduces spatial compatibility effects on the left-right axis, they do not rule out the possibility of orthogonal spatial compatibility effects i.e. the propensity of participants to show an advantage for an up-right and down-left pairing (Cho & Proctor, 2003; Weeks, Proctor, & Beyak, 1994). For instance, when stimuli were presented orthogonal to the response hand, the index finger was always below the

middle finger, and the participant's index finger was to the left side of space. A preference for responding to "up" stimuli with a right response and "down" stimuli with a left response may be observed along with imitative effects on the automatic imitation task used by Bertenthal and colleagues (Bertenthal et al., 2006). Thus, the measure of automatic imitation can be a combination of both imitative and orthogonal spatial effects. Therefore, instead of displaying stimuli orthogonal to the response hand, Catmur & Heyes (2011) used both left- and right-hand stimuli – this manipulation allowed for the measurement of the imitative compatibility effect independent of the spatial compatibility effect (Figure 2B). Further, in some studies, symbolic gestures were used which cannot be categorised on a spatial dimension (Figure 2C; Bortoletto, Mattingley, & Cunnington, 2013; Cracco, Genschow, Radkova, & Brass, 2018a; Liepelt, Prinz, & Brass, 2010). Since all of these methods elicited automatic imitation, there is wide consensus that spatial compatibility processes may influence automatic imitation, but automatic imitation cannot be reduced completely to spatial processes (Heyes, 2011).

Instead of spatial compatibility, SRC measures of automatic imitation are thought to be a combination of movement and effector compatibility (Heyes, 2011). Effector compatibility relates to the body part observed and the body part that needs to be moved by the participant (e.g. index or middle finger). Movement compatibility refers to the type of movement observed, and the movement made by the participant (e.g. tapping or lifting). Automatic imitation is therefore thought to be a largely automatic process that is driven by effector and movement compatibility, and is influenced by spatial compatibility (Heyes, 2011; Cracco et al., 2018b).

Even though automatic imitation is not reducible to spatial compatibility, a key question in cognitive psychology is whether the control of automatic imitation is governed by specialised cognitive control mechanisms which are involved only in social/imitative control (Bertenthal & Scheutz, 2012; Brass, Ruby, & Spengler, 2009), or by domain-general cognitive control systems that operate across a whole range of cognitive control tasks including spatial compatibility (Catmur & Heyes, 2011; Cooper, Catmur, & Heyes, 2012).

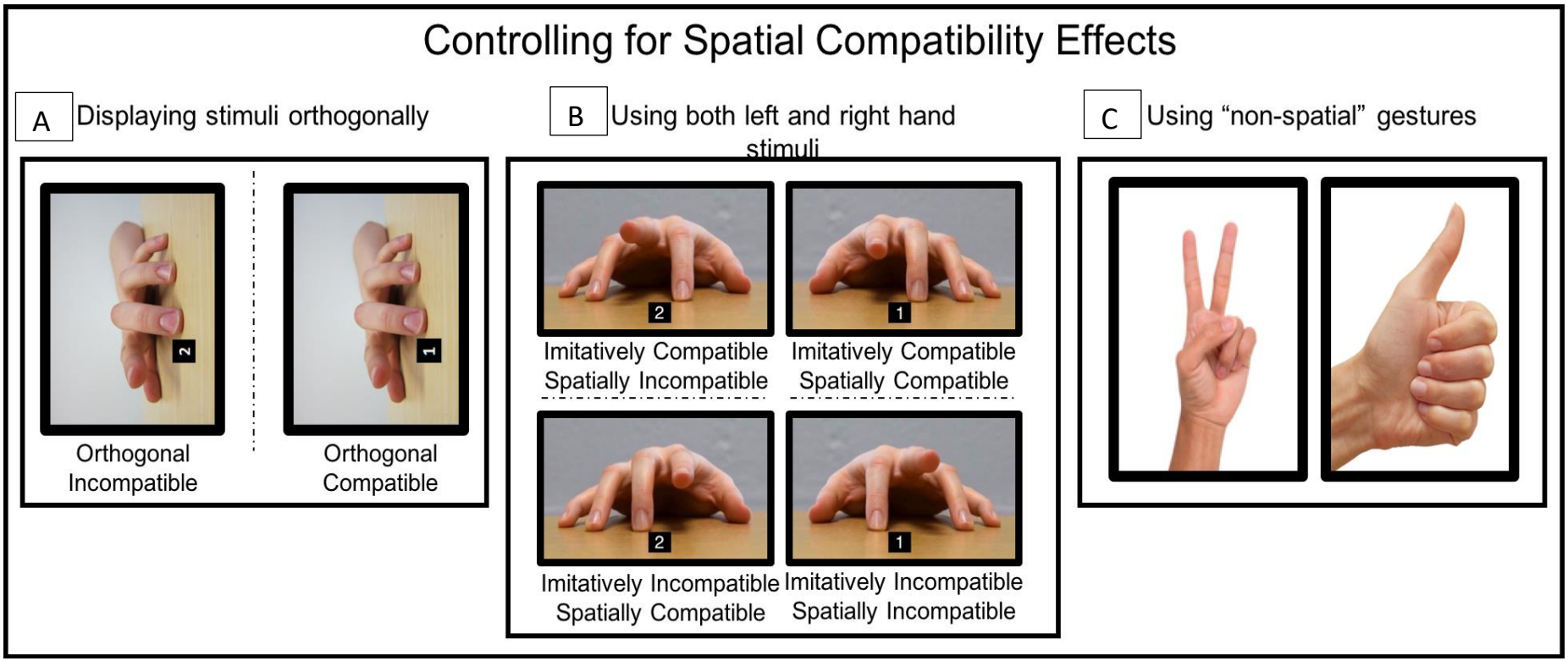


Figure 2. Tasks measuring automatic imitation have controlled for spatial compatibility effects in different ways: A) stimuli are displayed orthogonal to the participant's response, B) both left- and right-hand stimuli are used, allowing for measurement of imitative compatibility independent of spatial effects, C) gestures that cannot be easily classified on a spatial dimension are used. Images in panel (A) and (B) are taken from empirical work (Chapter 4) from the current thesis and are the actual images used in the experiment. Images in panel (C) are based on the description by Cracco et al., 2018a, and Bortoletto et al., 2013 and are not the actual images used in the experiment.

1.1.5. Summary.

Imitation has been the focus of systematic investigation for almost two centuries, with an explosion in interest after the discovery of the mirror neuron system. Across multiple domains including neurophysiology as well as social and cognitive psychology and neuroscience, there is now wide consensus that imitation is a complex and multi-dimensional phenomenon, with rich social consequences. While most investigation has focused on intentional imitation and imitative learning, more recently, there has been a call for research on automatic imitation in order to fully elucidate the mechanisms underlying non-verbal communication and social interactions.

Automatic imitation refers to the involuntary tendency to copy others' actions, speech patterns, gestures, and other behaviours, and has been the focus of investigation in social and cognitive psychology as well as neuroscience. Researchers in social psychology measure overt copying behaviours using naturalistic paradigms that are more ecologically valid i.e. they are directly related to "real-life" behaviours. In contrast, cognitive psychology researchers use SRC paradigms in order to measure the control of automatic imitation. These paradigms have greater reliability and afford better experimental control than naturalistic paradigms.

A key consideration in cognitive psychology is whether the control of automatic imitation as measured by SRC tasks reflects a special type of SR compatibility i.e. whether it relies on specialised cognitive mechanisms or is brought about by domain-general cognitive systems that operate across a range of SRC tasks of cognitive control including spatial compatibility. In Chapters 2, 3 and 4 (which are the empirical chapters of this thesis), we use variations of the SRC finger movement task described in section 1.1.3 in order to measure the control of automatic imitation, investigate its neural mechanisms, as well as individual differences in the tendency to automatically imitate.

In the following sections, I first review behavioural evidence on the factors modulating automatic imitation, and how the tendency to automatically imitate differs between individuals (Section 1.2.). I then review existing neuroimaging evidence on the neural circuitry that underpins the control of automatic imitation (Section 1.3.). Finally, I identify gaps in the literature and methodological issues which my thesis will aim to address (Section 1.4.).

1.2. What factors modulate automatic imitation?

Existing evidence to date from the domains of both social psychology and cognitive psychology suggests that although automatic imitation is automatic and ubiquitous, it is sensitive to many factors including the environment and context in which the interaction occurs, as well as characteristics of both the imitator and the imitatee. Recent research in both these domains has therefore focused on the social and contextual antecedents to automatic imitation in order to gain better insight about this phenomenon.

Evidence from social psychology suggests that automatic imitation or mimicry serves an important function in enhancing social affiliation and interactions. There is increasing evidence that mimicry has important social consequences like increasing positive rapport, affiliation, prosocial orientation, and helping behaviours not only between the interacting dyad (Bailenson & Yee, 2005; Chartrand & Bargh, 1999; Maurer & Tindall, 1983; Stel et al., 2010; Van Baaren, Holland, Steenaert, & van Knippenberg, 2003; Van Baaren, Holland, Kawakami, & van Knippenberg, 2004), but also towards members outside of the dyad (Ashton-James, van Baaren, Chartrand, & Decety, 2007). Further, mimicry is also modulated by characteristics of the imitator as well as the imitatee. For example, a goal to affiliate, higher perspective-taking abilities, as well as a positive mood were all related to higher mimicking behaviours (Chartrand and Bargh, 1999; Lakin & Chartrand, 2003; Likowski et al., 2011; van Baaren, Fockenberg, Holland, Janssen, & van Knippenberg, 2006) whereas mimicry was reduced when individuals held negative stereotypes of dislike about interaction partners, perceived them to be of a lower status compared to themselves, or when interacting partners were members of the out-group based on race, ethnicity, or political membership (Chartrand & van Baaren, 2009; Cheng & Chartrand, 2003; McHugo, Lanzetta, & Bush, 1991; Stel et al., 2010).

In cognitive psychology, after establishing the basic SRC paradigm to measure the control of automatic imitation, subsequent research on automatic imitation has provided insight into the factors that influence automatic imitation (Heyes, 2011; Cracco et al., 2018a). For instance, factors like animacy, sensorimotor experience, pro-sociality, as well as characteristics of both the interacting partners have been found to modulate the tendency to automatically imitate (e.g. Kilner et al., 2003; Catmur, Walsh, & Heyes, 2007; Cook & Bird, 2011, Wang, Newport, & Hamilton, 2011a; Ainley, Brass, & Tsakiris, 2014).

Below, I evaluate these antecedents to automatic imitation and individual differences in more detail.

1.2.1. Animacy.

Automatic imitation is modulated by both bottom-up and top-down cues to animacy. Prior research has found increased automatic imitation for human agents compared to non-human agents (e.g., Bird, Leighton, Press, & Heyes, 2007; Brass et al., 2001; Chaminade, Franklin, Oztop, & Cheng, 2005; Gowen, Bolton, & Poliakoff, 2016; Gowen, Stanley, & Miall, 2008; Kilner et al., 2003; Press, Bird, Flach, & Heyes, 2005; Press, Gillmeister, & Heyes, 2006, 2007). Further, participants' beliefs about animacy also influenced automatic imitation i.e. automatic imitation increased when participants believed the observed stimuli came from a human agent as compared to a non-human agent (Gowen et al., 2016; Klapper, Ramsey, Wigboldus, & Cross, 2014; Longo & Bertenthal, 2009; Liepelt & Brass, 2010).

1.2.2. Sensorimotor Experience.

Several studies using the SRC measure of automatic imitation have shown that sensorimotor experience can influence automatic imitation. For example, when participants are trained to lift an index finger when watching a middle finger lift, and lift the middle when watching an index finger lift, there was a decrease in automatic imitation (Catmur et al., 2007; Heyes et al., 2005). Thus, counter-imitation training or simply even asking participants to respond counter-imitatively led to the elimination and even reversal of the automatic imitation effect (Bardi, Bundt, Notebaert, & Brass, 2015; Catmur et al., 2007; Heyes et al., 2005).

1.2.3. Pro-sociality.

Similar to findings from social psychology, using SRC measures of automatic imitation have also found that priming participants with a pro-social orientation increases automatic imitation (Cook & Bird, 2011, 2012; De Coster et al., 2014; Leighton, Bird, & Heyes, 2010; Wang & Hamilton, 2013). For example, Leighton and colleagues (2010) exposed participants to scrambled sentences that denoted a prosocial attitude (e.g. "let us be together") as well as an anti-social attitude (e.g. "I am now single"). Participants showed increased automatic imitation when they were primed with a pro-social state than with an anti-social state. The administration of oxytocin in order to induce a pro-social state also increased automatic imitation as compared to the administration of a placebo (De Coster et al., 2014). More recent evidence, however, does

not show a link between pro-sociality and automatic imitation and suggests that effect of prosocial priming on automatic imitation may be smaller and less robust than what prior research has suggested (Newey, Koldewyn, & Ramsey, 2019).

1.2.4. Characteristics of the imitatee.

Some prior studies using SRC measures of automatic imitation have investigated how facial signals of the interacting partner like eye-gaze or expression, as well as group membership influences non-emotional imitative actions (Crescentini, Mengotti, Grecucci, & Rumiati, 2011; Grecucci, Koch, & Rumiati, 2011; Grecucci et al., 2013; Rauchbauer, Majdandžić, Hummer, Windischberger, & Lamm, 2015; Wang et al., 2011a; Wang & Hamilton, 2014a). For instance, there is some evidence (although suggestive, from studies with relatively small sample sizes), that both facial imitation and SRC measures of imitation have been found to increase when the interacting partner is an in-group member compared to an out-group member based on race, ethnicity, and arbitrary group assignment (Gleibs, Wilson, Reddy, & Catmur, 2016; Mondillon, Niedenthal, Gil, & Droit-Volet, 2007; Rauchbauer et al., 2015).

In addition, in a study by Wang et al. (2011a), direct eye gaze facilitated imitative responses compared to averted eye gaze. In a study by Butler and colleagues, automatic imitation was greater for smiling faces as compared to angry and neutral facial expressions (Butler, Ward, & Ramsey, 2016). Whereas Rauchbauer and colleagues found a similar increase in automatic imitation for happy faces compared to angry faces, some studies did not show evidence of any link between facial expressions like sad, fearful, and angry, and automatic imitative tendencies compared to a neutral expression (Crescentini et al., 2011; Grecucci et al., 2013; Rauchbauer et al., 2015).

Taken together, these few studies suggest that facial signals that elicit more pro-sociality or affiliation may increase the tendency to automatically imitate. However, sample sizes and effect sizes across these studies have been generally low, and for a robust estimate of the true population effect size, we would require more experiments and meta-analytical approaches.

1.2.5. Characteristics of the imitator (individual differences).

The above studies investigating social and contextual antecedents to automatic imitation have typically used an experimental method, which measures the average influence of a manipulation across a group of participants, rather than a differential approach that measures differences across individuals. As in research practice in

psychology more generally, even imitation research has seen a division between experimental and differential psychology, with a focus on experimental methods (Cronbach, 1957). Although a confluence of both these methods has been suggested for a unification and progress of psychological science as a whole (Cronbach, 1975; Eysenck & Eysenck, 1985; Eysenck, 1997), these two streams of thought have remained largely autonomous (Cronbach, 1957; Cramer, Waldorp, van der Maas, & Borsboom, 2010). To aid cross-pollination between experimental and differential approaches, more recent imitation research has started to take an individual differences approach by investigating how atypical populations, stable personality traits, and biological sex influence automatic imitation.

Atypical populations. Imitative abilities have been known to vary in individuals with autism spectrum disorder (ASD) and schizophrenia (Park, Matthews, & Gibson, 2008; Williams, Whiten, & Singh, 2004; Williams, Whiten, Suddendorf, & Perrett, 2001). Atypical imitative abilities in ASD and schizophrenia have been linked to a dysfunctional mirror-neuron system (MNS), also known as the "broken mirror" theory of autism (Ashton, Paunonen, Helmes, & Jackson, 1998; Oberman & Ramachandran, 2007; Thakkar, Peterman, & Park, 2014; Williams et al., 2001). However, some others have argued against the broken mirror theory and suggest that automatic imitative tendencies are relatively intact in individuals with ASD and schizophrenia (Bird et al., 2007; Cook & Bird, 2012; Hamilton, Brindley, & Frith, 2007; Press, Richardson, & Bird, 2010; Sowden, Koehne, Catmur, & Dziobek, 2015; Spengler, Bird, & Brass, 2010). Thus, theories purporting links between these disorders and atypical automatic imitative tendencies have been recently challenged (Hamilton, 2013; Southgate & Hamilton, 2008).

Stable personality traits. Imitative behaviour has been argued to vary across stable personality traits such as empathy (Chartrand & Bargh, 1999), narcissism (Hogeveen & Obhi, 2013; Obhi, Hogeveen, Giacomin, & Jordan, 2013), interoceptive awareness (Ainley, Brass, & Tsakiris, 2014) as well as alexithymia (Sowden, Brewer, Catmur, & Bird, 2016). Empathy is a personality trait that is characterised by both emotional and cognitive components that include the ability to regulate emotions as well as understand others' emotions and perspectives (Gerdes, Lietz, & Segal, 2011). It has been suggested that automatic imitation is increased in individuals who score higher on empathy compared to those who score lower (Chartrand & Bargh, 1999). Similarly, individuals who have a grandiose sense of self and lack emotional and cognitive empathy

are higher on the narcissism scale and show decreased imitative tendencies (Hogeveen & Obhi, 2013; Obhi et al., 2013). In addition, individuals who experience tactile sensations on their body when observing tactile stimulation on other individuals or objects (a phenomenon known as mirror-touch synaesthesia) also show heightened automatic imitation (Santiesteban, Bird, Tew, Cioffi, & Banissy, 2015). Individuals who are less aware of the internal state of their body i.e. those with low interoceptive awareness, as well as those who score higher on alexithymia (an inability to identify and describe their emotions) show a decrease in automatic imitative tendencies (Ainley et al., 2014; Sowden et al., 2016).

Such claims, however, are limited due to the small number of studies reported to date, relatively small sample sizes, and a lack of powerful replications. Moreover, studies which used considerably larger sample sizes, have not been able to replicate the moderating influence of personality variables on automatic imitation (Butler et al., 2015; Cracco et al., 2018b). For example, in the study by Butler and colleagues (2015), personality variables such as empathy, narcissism, extraversion, agreeableness, as well as autistic-like and schizotypal traits, did not predict automatic imitation as measured on the SRC task. Thus, the evidence for the link between stable personality traits and automatic imitation is mixed and less robust than previously suggested.

Biological sex. Although Butler and colleagues (2015) showed an invariance of automatic imitation to stable personality traits, they demonstrated that females showed a greater compatibility effect compared to males. This sex difference was also replicated by another study using the SRC measure of automatic imitation (Genschow et al., 2017). Therefore, it is possible that biological sex is a factor to consider further when attempting to understand how cognitive mechanisms supporting imitation vary across individuals. However, the extent to which sex differences operate in imitative behaviour has received minimal attention to date. For example, no sex differences have been found on the automatic imitation of actions or gestures, but facial mimicry studies have shown that females automatically imitate more than males (Chartrand & Bargh, 1999; Dimberg, 1990; Larsen, Overbeek, Granic, & Engels, 2010; Sonnby-Borgström, Jönsson, & Svensson, 2008). Given relatively small sample sizes and mixed results, the investigation of sex differences in automatic imitation requires further consideration (Hyde, 2014; Miller & Halpern, 2014). Although there is no clear and consistent empirical evidence, there is theoretical reason to think that sex differences may exist in imitative behaviour.

Indeed, females have been found to be more empathetic than males (Baron-Cohen & Wheelwright, 2004; Christov-Moore et al., 2014). Empathy has been associated with a variety of paradigms investigating imitation (Chartrand & Bargh, 1999; Müller, Van Leeuwen, Van Baaren, Bekkering, & Dijksterhuis, 2013; Sonnyby-Borgstrom, 2002). Females could thus be more pro-social than males in general, which may lead to more imitative tendencies in particular social contexts.

1.2.6. Summary.

A typical assumption in cognitive psychology is that all individuals rely on a common set of cognitive mechanisms, despite obvious individual differences (de Schotten & Shallice, 2017). However, understanding how cognitive mechanisms operate across all individuals, as well as how different individuals vary from these general patterns, is essential for understanding the structure of social cognition (Fischer-Baum, Kook, Lee, Ramos, & Vannucci, 2018). Therefore, in imitation research specifically, it is important to integrate methodological approaches from experimental and differential psychology to investigate the extent to which cognitive systems relating to the control of automatic imitation differ between individuals. While evidence suggests that automatic imitation is modulated by factors individual differences such as personality variables and biological sex, studies have been limited, statistically underpowered, with small sample sizes, and no (if any) direct replications. It is as yet unclear whether and how cognitive mechanisms underlying automatic imitation differ as a function of individual differences, and whether these differences are solely tied to imitative control or operate across different types of cognitive control tasks.

As in cognitive psychology, cognitive neuroscience research has also focused on the control of automatic imitation and whether it relies on domain-general or domain-specific neural mechanisms. In the sections below, I review evidence of investigations into the neural mechanisms of automatic imitation and its control from the domain of cognitive neuroscience.

1.3. Neural mechanisms of automatic imitation.

1.3.1. The role of the mirror neuron system.

In the last two decades, automatic imitation has been widely studied with an attempt to interconnect different disciplines like cognitive science, social psychology, evolutionary biology, as well as social and cognitive neuroscience (Byrne & Russon,

1998; Chartrand & Bargh, 1999; Prinz & Meltzoff, 2002). Prior research in social and cognitive neuroscience suggests that imitation relies on the mirror-neuron system, a neural network engaged in perceiving and performing actions (Iacoboni, 2009; Iacoboni et al., 1999; Rizzolatti & Craighero, 2004).

The proposed link between imitation and the MNS is supported by functional magnetic resonance imaging (fMRI) studies. fMRI relies on the core assumption that cerebral blood flow and neuronal activation are coupled – the engagement of brain regions can be measured by detecting changes in blood flow (the blood oxygen dependent level or BOLD signal). In a study by Iacoboni and colleagues (1999) using fMRI, the authors demonstrated that imitation is modulated by the left inferior frontal gyrus (IFG) and the right inferior parietal lobule (IPL) – the putative MNS in humans, as well as the right superior temporal sulcus (STS) which is thought to be the input to the MNS. In a similar vein, a number of other fMRI as well as transcranial magnetic stimulation (TMS) studies that investigated the neural mechanisms of imitation suggested that an MNS mechanism could explain imitative learning as well as automatic imitation (e.g. Blakemore & Frith, 2005; Buccino et al., 2004; Grezes et al., 2004; Heiser, Iacoboni, Maeda, Marcus, & Mazziotta, 2003; Hickok, 2013; see Figure 3, Model 1).

1.3.2. Neural circuits underlying the control of automatic imitation.

Imitation is unlikely to rely on a single system. To some extent, automatic imitation is a prepotent response i.e. we tend to automatically imitate when we observe an action. However, although automatic imitation may have many benefits for our social interactions, we do not imitate indiscriminately. In many situations, imitation can be maladaptive, and it is essential to circumvent the tendency to automatically imitate (Cross & Iacoboni, 2014; Cross, Torrisi, Losin, & Iacoboni, 2013; van Schie, van Waterschoot, & Bekkering, 2008; Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007). For example, if someone throws a ball at you and you copy their "throwing" action, you would never be able to catch the ball. This need to regulate imitative tendencies indicates the existence of a selection mechanism that prioritises alternative actions and inhibits unwanted ones (Brass et al., 2009). Thus, imitation control can be divided into at least two component processes – action representation and action selection or control (Figure 3). We observe an interaction partner and their actions (representation), and then select the action that needs to be executed (control).

To date, several neuroimaging studies have investigated the neural mechanisms of the control of automatic imitation. In social cognition research more generally, a central question has been whether unique cognitive processes are engaged during social cognition, or whether socio-cognitive processes represent one instance of general-purpose cognition (Blakemore, Winston, & Frith, 2004; Spunt & Adoplhs, 2017). In a similar vein, neural models of automatic imitation have also focused on domain-specific and domain-general neural processes that underpin the control of our tendency to automatically imitate. Domain-specific processes operate on particular types of stimuli or aspects of cognition while domain-general processes operate across a range of stimuli and tasks (Barrett, 2012).

One of the prevailing theories of automatic imitation proposes that imitation control relies on a domain-specific neural circuit related to social cognition (Figure 3, Model 2; Brass et al., 2009). Evidence supporting this theory comes mainly from patient and neuroimaging studies and points to the engagement of two key candidate regions – the anterior medial prefrontal cortex (mPFC) and the right temporoparietal junction (rTPJ) (Brass and Heyes, 2005; Brass et al., 2009).

Some neuroimaging studies investigating the control or inhibition of automatic imitation in humans have found the engagement of mPFC and rTPJ (Brass et al., 2001, 2005; 2009; Spengler et al., 2009; Wang et al., 2011). A dissociation of roles for mPFC and rTPJ has also been proposed - the rTPJ distinguishes between self- and other-generated actions, and the mPFC enforces the self-generated action (Brass et al., 2009). Patient studies have demonstrated that patients with frontal lobe lesions show disrupted imitation inhibition behaviour as well as an increased tendency to automatically imitate even when they are clearly instructed to not do so (Brass et al., 2003; Lhermitte et al., 1986; Spengler et al., 2010). rTPJ involvement is further supported by neurostimulation studies – in studies inhibiting the activity in the rTPJ by transcranial magnetic stimulation (TMS), impaired imitation inhibition was found (Hogeveen et al., 2014; Sowden & Catmur, 2015). Irrespective of the method used, it is worth noting that there have only been a small number of studies implicating mPFC and rTPJ in the control of automatic imitation, and these studies have used relatively small sample sizes (N=10-25) without any direct replications. Therefore, the sum of evidence for mPFC and rTPJ engagement during imitation control seems to be suggestive and not compelling.

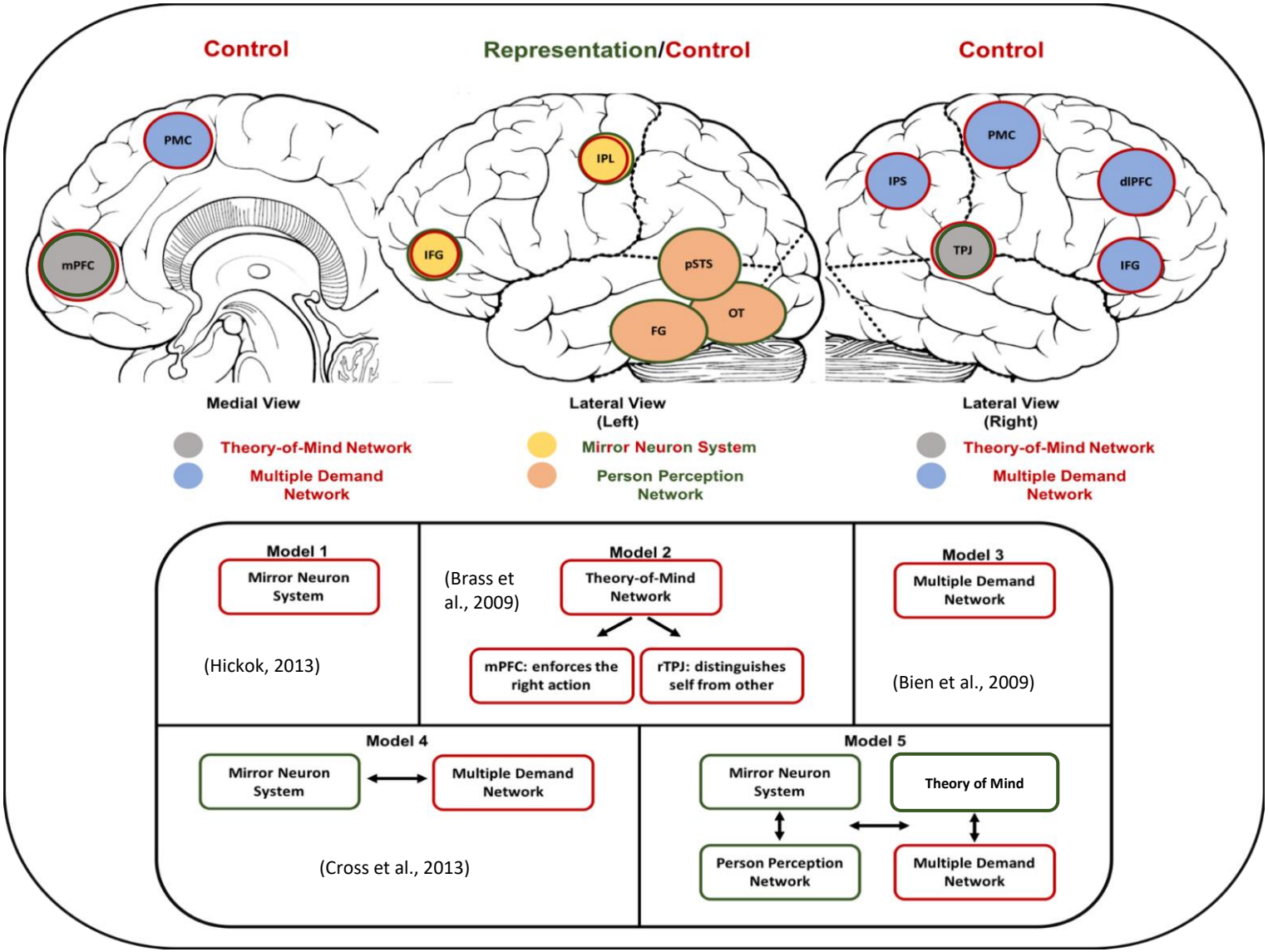


Figure 3. Graphical representation of neural models proposed as underlying mechanisms of automatic imitation. Red outline indicates networks engaged for "control" and green outline indicates networks engaged for "representation" of the action (the mirror neuron system; MNS) and the person (the person perception network; PPN). Different models have been proposed in the literature – imitation control relies on the MNS (Model 1), the theory-of-mind network (ToM; Model 2), the multiple demand network (MD; Model 3), an interaction between MNS and MD networks (Model 4), and interactions between domain-specific socio-perceptual circuits (MNS, PPN), with the ToM modulating the MD network during the control of automatic imitation (Model 5). *N.B.* Abbreviations: IPL = inferior parietal lobule, IFG = inferior frontal gyrus; pSTS = posterior superior temporal sulcus; OT = occipito-temporal cortex; FG = fusiform gyrus, mPFC = medial prefrontal cortex; PMC = primary motor cortex; dlPFC = dorsolateral prefrontal cortex; TPJ = temporo-parietal junction; IPS = intra-parietal sulcus.

Along with the control of automatic imitation, mPFC and rTPJ are also engaged in a variety of other socio-cognitive tasks associated with theory of mind (ToM), These include perspective taking, attributing beliefs, desires and attitudes to others, and distinguishing between self and the other (Aichhorn et al., 2006; Amodio and Frith, 2006; Brass et al., 2009; Decety et al., 2002; Gallagher et al., 2000; Ruby and Decety, 2001; Santiesteban et al., 2012; Spengler et al., 2010). Based primarily on this evidence, self-other control processes have been proposed as a candidate mechanism for a range of socio-cognitive functions i.e. it has been proposed that self-other processes are necessary when empathising with others, taking their perspective, or engaging a successful theory-of-mind (de Guzman et al., 2016; Sowden and Shah, 2014). Atypical self-other control has also been linked to disorders like autism and schizophrenia which are characterised by atypical socio-cognitive functioning (Cook and Bird, 2012; Ferri et al., 2012).

In contrast to this view of imitation control, however, other theories of imitation suggest that the inhibition or control of automatic imitation does not differ from any other pre-potent tendencies or general cognitive functions (Figure 3, Model 3; Heyes, 2011; Cooper et al., 2012). Cognitive control tasks such as the Stroop, Simon, and Flanker tasks require the inhibition of automatic overlearned response tendencies as well. These tasks have been found to engage dorsolateral fronto-parietal cortices – regions which are collectively called the multiple demand network (Aron, Robbins, & Poldrack, 2014; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Hazeltine, Poldrack, & Gabrieli, 2007; Nee, Wager, & Jonides, 2007; Wager et al., 2005;). This domain-general multiple demand (MD) network is so called because it is engaged across multiple mental operations (Duncan, 2010). Across studies that investigate imitation inhibition, while some have found engagement of the mPFC and rTPJ (Brass et al., 2001, 2005; 2009; Spengler et al., 2009), others show engagement of the MD network (Bien, Roebroeck, Goebel, & Sack, 2009; Crescentini, Mengotti, Grecucci, & Rumiati, 2011; Cross & Iacoboni, 2013; Marsh, Bird, & Catmur, 2016; Mengotti, Corradi-Dell'Acqua, & Rumiati, 2012). The collective evidence for either of the networks, however, is mixed, and needs further investigation.

The above review of literature suggests two possible neural mechanisms as being crucial to action selection during the control of automatic imitation. On one hand, during imitation control, mPFC and rTPJ are engaged and they work by inhibiting the

representation of the observed person's action and enforcing the self-generated action. On the other hand, the selection mechanism may be guided by a domain-general neural network i.e. the MD network. In both possible mechanisms, the input to the selection system is the same i.e. the observed person and action. This perception may engage domain-specific socio-perceptual neural circuits involved in person and action perception (the person perception network and the MNS; Caspers et al., 2010; Kanwisher, 2010). However, the selection or control mechanism that underlies the inhibition or control of automatic imitative tendencies which finally leads to consequent behaviour may either be domain-general or domain-specific. However, as most previous fMRI studies have been limited by low statistical power and small sample sizes, the evidence demonstrating the extent to which SRC paradigms of the control of automatic imitation engage domain-general or domain-specific neural networks remains unclear.

1.3.3. Modulation of the neural systems underlying the control of automatic imitation.

While it is unclear whether domain-general or domain-specific neural networks underpin the control of automatic imitation, some prior neuroimaging studies suggest that the ToM network (i.e. mPFC and rTPJ) can be involved in the modulation of automatic imitation. For example, Klapper et al. (2014) found a higher response in rTPJ when an interaction partner looked human and was believed to be human compared with when neither of these animacy cues were present. Wang and colleagues (2011) demonstrated that mPFC had a top–down influence on other brain circuits during social modulation of imitation via direct gaze.

These studies suggest that mPFC and/or rTPJ may play a regulatory role, be modulated by social context, and can functionally interact with other regions during the control of automatic imitation. In support of this, regions that do not show direct engagement in a cognitive process of interest have been known to have a regulatory influence on other regions that are directly engaged (Burnett & Blakemore, 2009). In line with this proposal, Cross and colleagues (2013) suggested that imitation control involves top-down regulation between a domain-general cognitive control network, and a domain-specific network relevant for imitation (in this case, the MNS; see Figure 3, Model 4). More generally, research from other domains of social cognition shows growing evidence for higher complexity and functional interplay within and between so-called domain-specific and domain-general networks (Baetens, Ma, Steen, & Van

Overwalle, 2014; Quadflieg et al., 2011; Spunt & Adolphs, 2015; Zaki, Hennigan, Weber, & Ochsner, 2010). Much like social cognition in general, therefore, imitation control may be best explained by interactions between component functional circuits, which themselves need not be domain-specific (see Figure3, Model 5; Spunt & Adolphs, 2017). However, before making any conclusions about the interactions between systems, it is essential to know whether the fundamental system underlying the control of automatic imitation is domain-general or domain-specific (see Figure 4).

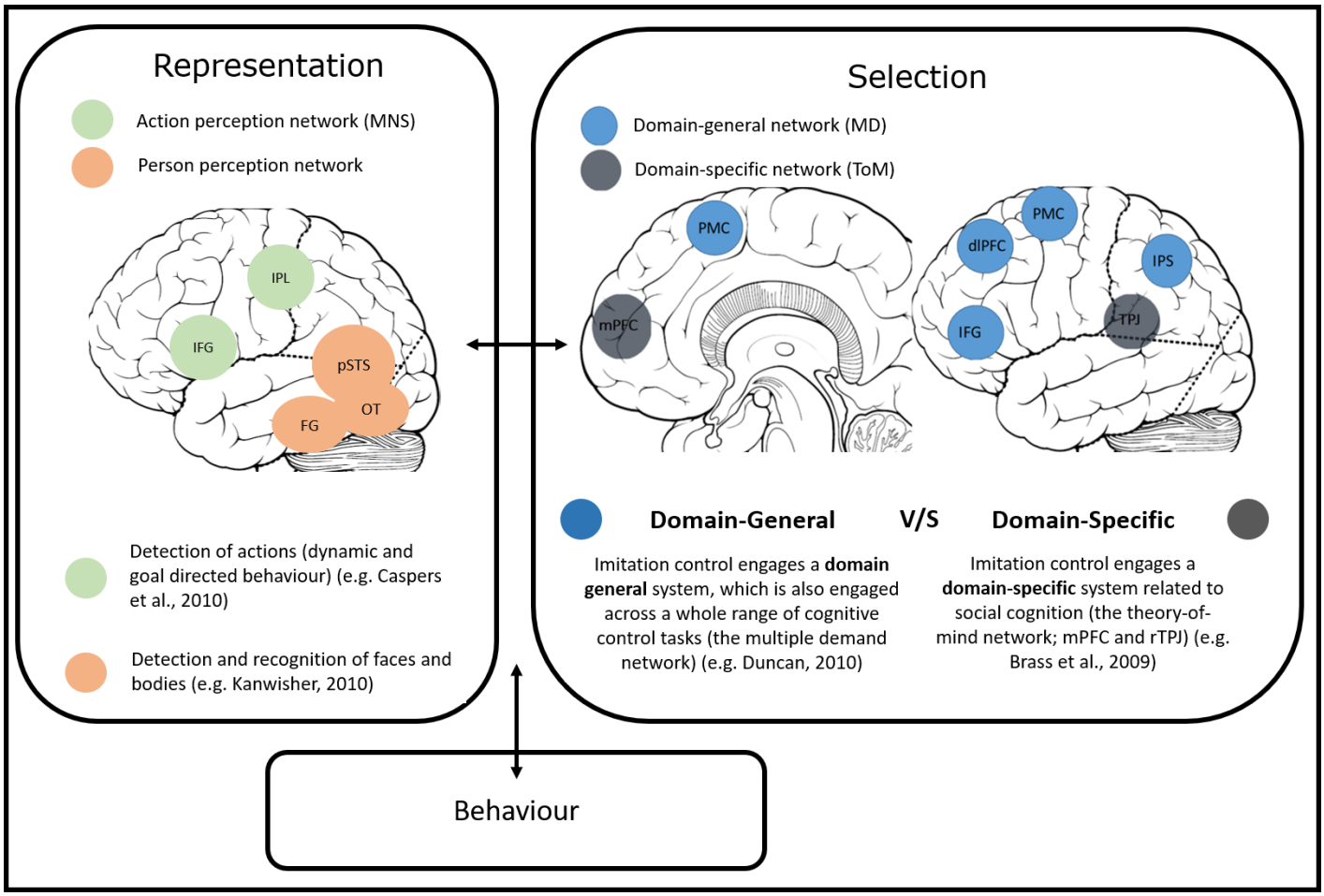


Figure 4. Brain networks associated with the control of automatic imitation. This figure is taken from Darda and Ramsey (2019). This graphical representation divides imitation control into two constituent processes – representation of the person and their action, and the selection (control) of the right action to be executed. In the context of automatic imitation, the representation system consists in face, body, biological motion, and action perception. The neural substrates for person and action perception span the fusiform gyrus, occipitotemporal cortex, and posterior superior temporal sulcus, as well as the mirror neuron system (Kanwisher, 2010; Caspers, et al., 2010). The control or selection system consists in a brain network that is either domain-general (i.e. the multiple demand network) or domain-specific (i.e. the theory-of-mind network). *N.B.* Abbreviations: MNS = mirror neuron system; IPL = inferior parietal lobule, IFG = inferior frontal gyrus; pSTS = posterior superior temporal sulcus; OT = occipito-temporal cortex; FG = fusiform gyrus, MD = multiple demand network; ToM = theory-of-mind network; mPFC = medial prefrontal cortex; PMC = primary motor cortex; dlPFC = dorsolateral prefrontal cortex; TPJ = temporo-parietal junction. The bidirectional arrow indicates links between the different nodes of imitation control.

1.3.4. Summary.

In cognitive neuroscience, research on automatic imitation has seen a divide between two prevailing neural mechanisms – one suggests that the control of automatic imitation relies on a domain-specific neural circuit unique to social cognition i.e. the ToM network. The other suggests that a domain-general neural circuit which operates across multiple cognitive control tasks underpins the control of our tendency to automatically imitate. However, evidence pointing toward one or the other is limited – most prior research has been underpowered, with small sample sizes. Therefore, although research has investigated how neural systems underlying automatic imitation may be modulated by different factors including eye-gaze and animacy, the basic underlying neural mechanisms are not fully understood. It is important to investigate whether and to what extent domain-general and domain-specific neural mechanisms are engaged during the control of automatic imitation. Further, no prior research has investigated individual differences in the neural mechanisms of imitation control. In order to fully understand the mechanisms involved in automatic imitation, and social cognition more generally, it is essential to understand not only how these mechanisms operate across all individuals, but also how different individuals vary from these general patterns.

1.4. Methodological issues and research questions.

In recent years, a key question in science has been the validity and replicability of reported findings (Begley, 2013; Open Science Collaboration, 2015). These concerns have touched almost all domains of science, including psychology and neuroscience (Open Science, 2015; Button et al., 2013; Munafo et al., 2017; Pashler, Coburn, & Harris, 2012; Vazire, 2018) with estimates of replicability ranging between 25 and 75% (Camerer et al., 2018; Marsman et al., 2017; Matzke et al., 2015; Nosek & Lakens, 2014). Prevalence of underpowered studies with low sample sizes as well as questionable research practices like *p-hacking*[1] and multiple analyses seem to be the main contributing factors to low reproducibility (Button et al., 2013; Open Science Collaboration, 2015). It has been suggested that a diverse set of practices including pre-registration, meta-analyses, power analyses, large sample studies, as well as multi-experiment replications

[1] *P-hacking* refers to a type of bias where scientists perform several statistical analyses or data selection procedures, and then selectively report "significant" results (Head et al., 2015).

need to be common practice in psychological research in order to increase the robustness and credibility of findings (Munafo et al., 2017; Nelson et al., 2018).

In the previous sections of my thesis I have provided a critical analysis of extant literature on automatic imitation in the domains of social psychology, cognitive psychology, and cognitive neuroscience. In each summary section, I have highlighted specific methodological issues and gaps in the literature. Below, I outline how my thesis will take several steps to resolve these problems, fill in the gaps in knowledge, and improve methodological approaches in order to elucidate the cognitive and neural mechanisms of automatic imitation, and how automatic imitative tendencies vary as a function of individual differences.

1.4.1. Neuroimaging studies.

In Section 1.3., I reviewed extant evidence on the neural mechanisms of the control of automatic imitation. Prior studies have provided mixed evidence as to whether imitation control relies on domain-specific (ToM) or domain-general (MD) neural networks, and have typically been underpowered with small sample sizes and few (if any) direct replications (e.g. Bien et al., 2009; Brass et al., 2009; Cross & Iacoboni, 2013; n = 10 to 25). Thus, given the mixed findings in prior fMRI studies as well as in psychology and neuroscience more generally, fMRI studies in imitation research need to consider reliability and replicability as key concerns (Button et al., 2013; Open Science Framework, 2015). In addition, there has also been no investigation into individual differences in the neural systems that underpin the control of automatic imitation, although behavioural evidence suggests automatic imitation varies as a function of individual differences.

All above mentioned fMRI studies investigating automatic imitation have either used an anatomical ROI or a whole-brain analysis approach, but no prior study has used a functional region of interest (fROI) approach. Broadly, fMRI analysis methods either focus on the whole-brain or specific regions of interest (ROIs). ROIs can be defined anatomically i.e. based on sulci and gyri or stereotaxic co-ordinates, or functionally i.e. based on consistent patterns of functional response profiles (Saxe et al., 2006). Whereas whole-brain approaches seek to identify overlap across participants, a functional region-of-interest (fROI) approach allows for investigation of the response profile of functionally defined regions of interest in individual participants (Fedorenko et al., 2013; Kanwisher, 2010). In contrast to whole-brain analyses, fROI analyses also typically have higher

statistical power due to the constrained search volume in the brain (Fedorenko et al., 2010; Saxe et al., 2006). Further, whole-brain and anatomical ROI based approaches are susceptible to reverse inference i.e. the inference from activation to mental functions (Poldrack, 2006). This kind of inference does not take into account how selectively a particular region is activated by the mental process under investigation. The fROI approach allows for directly testing hypotheses regarding the role of functionally defined ROIs and therefore minimises the reliance on reverse inference to infer cognitive function based on anatomical localisation (Poldrack, 2006).

Further, there has been no existing quantitative synthesis of fMRI literature on the control of automatic imitation. Single studies are likely to be underpowered leading to missed or spurious results (Button et al., 2013) and it is difficult to generalize across differing experimental paradigms and design choices (Carp, 2012). It is, therefore, unclear whether consistent patterns of activation exist across existing fMRI studies, and whether these patterns support the engagement of a domain-specific or domain-general network in imitation control.

Taken together, these findings open up an important consideration and critical research question pertaining to the neural correlates of the control of automatic imitation – to what extent do neural mechanisms of automatic imitation consistently rely on domain-general and/or domain-specific neural networks, and how do these mechanisms differ as a function of individual differences?

1.4.2. Behavioural studies.

In Section 1.2., I reviewed extant evidence on individual differences in automatic imitation. Prior studies that have suggested that automatic imitation varies as a function of personality traits have typically used small sample sizes and are underpowered (Ainley et al., 2014; Chartrand & Bargh, 1999; Hogeveen & Obhi, 2013; Obhi et al., 2014; Santiesteban et al., 2015; n=18 to 50). In contrast, studies with a larger sample size have not been able to find evidence for such a link (Butler et al., 2015; Cracco et al., 2018b). Thus, before making any firm conclusions, it is essential that studies purporting different patterns of relationship between automatic imitation and individual differences in general across the population (Chartrand & Bargh, 1999; Hogeveen & Obhi, 2013; Obhi et al., 2013) need to be replicated in order to enable a cumulative science to develop (Munafo et al., 2017; Zwaan et al., 2018). With regards to sex differences, only two studies to date have investigated and found a sex difference on the SRC measure of automatic

imitation (Butler et al., 2015; Genschow et al., 2017). It is also unclear whether these studies demonstrate a difference as a function of biological sex solely in imitation control as the SRC task was a combined measure of spatial and imitative effects.

Taken together, these findings suggest that that little is known about individual differences in automatic imitation and only a limited number of studies have used large sample sizes, run multi-experiment replications, or performed quantitative meta-analyses. In order to fully understand the mechanisms of social (and non-social) cognition, it is essential that we not only investigate how these faculties work across individuals, but also how different individuals vary compared to average patterns. Therefore, an important consideration and aim is to investigate whether individual differences in automatic imitation exist, and whether these differences are specialised (i.e. solely tied to imitation control) or are generalised across different types of cognitive control.

Therefore, in my thesis, using large-sample, multi-experiment behavioural and fMRI experiments, I aim to investigate to what extent specialised or generalised cognitive and neural systems underlie the control of automatic imitation, and how these systems differ as a function of individual differences.

1.5. Overview of the thesis.

In the current Chapter (Chapter 1), I provided a critical analysis of existing behavioural and neuroimaging literature on the cognitive and neural mechanisms of automatic imitation, and how these vary between individuals. In the following chapters, I empirically test and investigate pertinent questions regarding the control of automatic imitation using fMRI (Chapter 2), meta-analytical (Chapter 3), and behavioural (Chapter 4) approaches. I then discuss the findings from the empirical chapters and point to broader implications of this work (Chapter 5).

Chapter 2 investigates functional specificity and sex differences in the neural circuits underlying the control of automatic imitation using fMRI. This is the first study to date to use a functional ROI approach and explore sex differences in automatic imitation. Using functional localisers and an analysis pipeline that bolsters sensitivity, this study (N=50) investigates the response profile of domain-specific (ToM) and domain-general (MD)

neural networks in the inhibition of automatic imitation, and how these profiles differ between males and females.

Chapter 3 investigates the consistency of the neural networks that underlie the control of automatic imitation using a multi-level kernel density analysis (MKDA). This is the first study to date that quantitatively meta-analyses existing neuroimaging literature investigating neural mechanisms of automatic imitation. We used MKDA in order to provide a combined quantitative estimate of many individual studies, and avoid problems associated with interpreting individual studies.

Chapter 4 investigates individual differences in automatic imitation using multiple large-sample behavioural studies (N>∼600 across 3 experiments). This study integrates experimental and differential psychology approaches in order to investigate the extent to which cognitive systems related to social (imitative) and non-social control differ between individuals, and whether these differences are domain-general or solely tied to imitative control.

Chapter 5 summarises the findings from the empirical chapters (Chapters 2, 3, and 4). I discuss the broader implications of the work included in this thesis, highlight how the current work addresses previous methodological issues and gaps in our knowledge, and point to new and exciting directions for future research.

CHAPTER 2

Functional Specificity and Sex Differences in the Neural Circuits Supporting the Inhibition of Automatic Imitation

This chapter investigates functional specificity and sex differences in the neural circuits underlying the control of automatic imitation using functional magnetic resonance imaging (fMRI). This is the first study to date to use a functional region-of-interest (fROI) approach and explore sex differences in automatic imitation. Using functional localisers and an analysis pipeline that bolsters sensitivity, this study (N=50) investigates the response profile of domain-specific (Theory-of-Mind) and domain-general (Multiple Demand) neural networks in the inhibition of automatic imitation, and how these profiles differ between males and females.

This chapter is published as it is from:
Darda, K. M., Butler, E. E., & Ramsey, R. (2018). Functional specificity and sex differences in the neural circuits supporting the inhibition of automatic imitation. *Journal of Cognitive Neuroscience*, 30, 914-933.

N.B. Emily Butler designed and collected data for Experiment 1 of this paper.

Acknowledgements
The authors thank Andrew Fischer for providing Level 2 cover during scanning for Experiment 2.

Abstract

Humans show an involuntary tendency to copy other people's actions. Although automatic imitation builds rapport and affiliation between individuals, we do not copy actions indiscriminately. Instead, copying behaviours are guided by a selection mechanism, which inhibits some actions and prioritizes others. To date, the neural underpinnings of the inhibition of automatic imitation and differences between the sexes in imitation control are not well understood. Previous studies involved small sample sizes and low statistical power, which produced mixed findings regarding the involvement of domain-general and domain-specific neural architectures. Here, we used data from Experiment 1 (N = 28) to perform a power analysis to determine the sample size required for Experiment 2 (N = 50; 80% power). Using independent functional localisers and an analysis pipeline that bolsters sensitivity, during imitation control we show clear engagement of the multiple-demand network (domain-general), but no sensitivity in the theory-of-mind network (domain-specific). Weaker effects were observed with regard to sex differences, suggesting that there are more similarities than differences between the sexes in terms of the neural systems engaged during imitation control. In summary, neurocognitive models of imitation require revision to reflect that the inhibition of imitation relies to a greater extent on a domain-general selection system rather than a domain-specific system that supports social cognition.

Introduction

Human social interactions are guided by nonverbal cues, such as copying behaviours. In the last two decades, much research has investigated the involuntary tendency to copy other's actions—a phenomenon known as automatic imitation (Heyes, 2011). Automatic imitation is thought to be beneficial in social situations because it develops affiliative attitudes, better cooperation, and feelings of closeness between interacting partners (Chartrand & Lakin, 2013). Prior neuroscience research has shown that imitation is supported by the mirror neuron system, a neural network engaged in perceiving and performing actions (Iacoboni, 2009; Rizzolatti & Craighero, 2004; Iacoboni et al., 1999). Imitation, however, is unlikely to rely on a single cognitive or brain system (Southgate & Hamilton, 2008). For example, in many circumstances, imitation is maladaptive and requires inhibition (Cross & Iacoboni, 2014; Cross, Torrisi, Losin, & Iacoboni, 2013; van Schie, van Waterschoot, & Bekkering, 2008; Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007). In such situations, a selection mechanism is required to suppress the tendency to imitate and prioritize alternative actions (Brass, Ruby, & Spengler, 2009). To date, studies investigating the neural mechanisms of imitation control have been limited by small sample sizes and low statistical power, which has produced mixed findings (Table 1). Furthermore, no neuroscience research has investigated how individual differences such as sex modulate imitation control, even though behavioural research has shown that imitative tendencies vary as a function of sex (Butler, Ward, & Ramsey, 2015; Sonnby-Borgström, Jönsson, & Svensson, 2008; Dimberg & Lundquist, 1990). Across two fMRI experiments, which had higher statistical power and functional sensitivity than prior studies, we investigated the extent to which imitation inhibition relies on a domain-specific or domain-general neural network, which varies its response as a function of sex.

Much like cognitive science in general (Kanwisher, 2010; Hirschfeld & Gelman, 1994), inhibitory control research has focused on a neat division between domain-general and domain-specific mental operations. Domain-general inhibitory systems, which operate across multiple tasks, have been identified in dorsal frontoparietal cortices (Aron, Robbins, & Poldrack, 2014; Hazeltine, Poldrack, & Gabrieli, 2007; Nee, Wager, & Jonides, 2007; Wager et al., 2005; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002). This brain circuit has been labelled the multiple demand (MD) network because of its engagement in a diversity of mental operations (Duncan, 2010). By contrast, evidence from fMRI, neurostimulation, and

neuropsychological patient studies has suggested that a domain-specific circuit in an anterior portion of medial prefrontal cortex (mPFC) and right temporo-parietal junction (rTPJ) operates during the inhibition of imitation (Bardi, Gheza, & Brass, 2017; Santiesteban, Banissy, Catmur, & Bird, 2012, 2015; Sowden & Catmur, 2015; Hogeveen et al., 2014; Klapper, Ramsey, Wigboldus, & Cross, 2014; Wang, Ramsey, & Hamilton, 2011; Spengler, von Cramon, & Brass, 2009, 2010; Brass, Derrfuss, Matthes-von Cramon, & von Cramon, 2003; Brass, Zysset, & von Cramon, 2001). Beyond the control of imitation, mPFC and rTPJ have been consistently implicated in a variety of social cognition functions, which require distinguishing between self and other, as well as reasoning about other people's mental states (theory of mind [ToM]; Van Overwalle, 2009; Amodio & Frith, 2006; Saxe & Kanwisher, 2003; Frith & Frith, 1999). These results led to theorizing that a key neural circuit for social cognition also regulates imitative tendencies (Brass et al., 2009).

Although theories of imitation control have been developed that are based on functioning of the ToM network, evidence from fMRI studies that used an RT measure of imitation inhibition have not provided consistent support for the involvement of a domain-specific neural network (Table 1). The RT measure of imitation involves making finger movements while simultaneously watching compatible or incompatible finger movements (Brass, Bekkering, Wohlschlager, & Prinz, 2000; Stürmer et al., 2000). The difference between RTs in these two conditions (i.e., the general compatibility effect) has been argued to index imitative control, as greater cognitive resources are required to inhibit movements that are incompatible to one's own responses (Heyes, 2011; Brass & Heyes, 2005). Approximately half of the fMRI studies using this paradigm failed to find engagement of rTPJ and anterior mPFC. In addition, a number of studies showed engagement of regions associated with the MD network, including dorsal frontoparietal cortex, supplementary motor area (SMA) and anterior insula (Marsh, Bird, & Catmur, 2016; Cross & Iacoboni, 2013; Mengotti, Corradi-Dell'Acqua, & Rumiati, 2012; Crescentini, Mengotti, Grecucci, & Rumiati, 2011; Bien, Roebroeck, Goebel, & Sack, 2009).

Table 1.
fMRI Studies Investigating Imitation Control Using Modified Versions of the Imitation Inhibition Task.

	Sample (Male:Female)	***Dissociation of Imitative and Spatial Processes***	***Analysis***		***Brain Networks***		
					ToM		***MD***
			ROI	***Whole brain***	***mPFC***	***rTPJ***	
Brass et al., 2001	10 (4:6)			✓	✓		✓
Brass et al., 2005	20 (8:12)			✓	✓	✓	✓
Brass et al., 2009	20[a]		✓		✓	✓	
Spengler et al., 2009	18 (9:9)		✓	✓	✓	✓	✓
Bien et al., 2009	15 (5:10)	✓		✓			✓
Crescentini et al., 2011	19 (9:10)	✓		✓			✓
Cross & Iacoboni, 2013	24 (12:12)	✓		✓			✓
Mengotti et al., 2012	22 (10:12)	✓		✓			✓
Cross et al., 2013	25 (5:15)	✓		✓	✓		✓
Klapper et al., 2014	19 (2:17)		✓		✓[b]		
Marsh et al., 2016	24 (7:17)	✓		✓			✓
Wang et al., 2011	20 (5:15)			✓	✓	✓	✓

Evidence that the engagement of mPFC and rTPJ is inconsistent across fMRI studies that investigated imitation control using modified versions of the imitation inhibition task. For all studies, engagement of mPFC or rTPJ is reported only for contrasts that test for inhibiting the urge to automatically imitate. Engagement of the MD network is reported only for whole-brain analyses. Except for Wang et al. (2011), which used hand movements, all other tasks used modified versions of the imitation inhibition tasks involving finger movements (Brass et al., 2000). For a more detailed version of this table, see Supplementary Table S6.
[a]Number of male and female individuals not mentioned.
[b]mPFC showed engagement only at $p < .005$, uncorrected.

Moreover, the most common measure of imitation interference is confounded by spatial compatibility or the tendency to respond faster to a stimulus when it is on the same side of space as the response (e.g., Simon, 1969). To measure imitation interference independent of spatial compatibility effects, spatial and imitative processes need to be dissociated (Gowen, Bolton, & Poliakoff, 2016; Marsh et al., 2016; Boyer, Longo, & Bertenthal, 2012; Cooper, Catmur, & Heyes, 2012; Catmur & Heyes, 2011; Wiggett, Hudson, Tipper, & Downing, 2011; Bertenthal, Longo, & Kosobud, 2006). Therefore, the extent to which imitation inhibition relies on domain-specific and domain-general architectures remains unclear. Indeed, no research to date has dissociated spatial from imitative processes and used a functional ROI (fROI) approach (Fedorenko, Duncan, & Kanwisher, 2013; Kanwisher, 2010). Using a fROI approach enables investigation of how functionally defined brain circuits, such as the MD and ToM networks, operate during the control of imitation.

A further area of imitation research that has received little attention is the extent to which imitative control varies across individuals, especially between the sexes. It has been argued that imitation is modulated by stable individual differences, such as empathy (Chartrand & Lakin, 2013) and sex (Butler et al., 2015; Sonnby-Borgström et al., 2008). Although it has been suggested that women excel across a range of social processes compared to men (Baron-Cohen, 2002), only a limited number of studies have investigated sex differences in social cognition, and the results are often mixed, do not replicate, or are specific to very select contexts or samples (Hyde, 2014; Miller & Halpern, 2014). Furthermore, studies of sex differences in social cognition have mainly focused on emotional expression perception and mental state reasoning with little emphasis placed on imitation (Krach et al., 2009; Russell, Tchanturia, Rahman, & Schmidt, 2007; Rahman, Wilson, & Abrahams, 2004; Campbell et al., 2002; Thayer & Johnsen, 2000).

A recent study that used an RT measure of imitation inhibition (Brass et al., 2000) showed that female individuals showed a greater level of interference than male individuals (Butler et al., 2015). It is possible that this sex difference in imitation control may be mediated by empathy—female individuals have been shown to be more empathetic compared with male individuals (Christov-Moore et al., 2014; Baron-Cohen & Wheelwright, 2004). However, even though empathy has been associated with different types of imitation paradigms (Müller, Leeuwen, Baaren, Bekkering, &

Dijksterhuis, 2013; Sonnby-Borgström, Jönsson, & Svensson, 2003; Sonnby-Borgström, 2002; Chartrand & Bargh, 1999), the evidence to date suggests that there is no link between imitation, as measured by RTs, and empathy (Genschow et al., 2017; Butler et al., 2015). In addition, in the study by Butler and colleagues (2015), it is unclear whether sex modulates the tendency to automatically imitate or the tendency to automatically respond in the same spatial location to the observed action. The former indicates a sex difference that is specifically tied to imitation control, whereas the latter might indicate a sex difference in processes associated with resolving spatial conflict. More recent work also showed a greater interference effect for female individuals compared with male individuals (Genschow et al., 2017), as well as greater error rates for predominantly female samples than male samples (Cracco et al., 2018). The imitation task used by Genschow and colleagues (2017) was controlled for left–right spatial compatibility by presenting the stimulus hand orthogonal to the response. Even though this shows that the sex difference remains when spatial compatibility is reduced, it does not rule out the possibility of orthogonal spatial compatibility (Weeks & Proctor, 1990). More generally, sex differences have been found on a wide range of inhibitory control tasks, including flanker, gaze cueing, arrow cueing, oddball, and Simon tasks, wherein female individuals have been shown to require more cognitive resources than male individuals to inhibit automatic response tendencies (Figure 1; Stoet, 2010, 2017; Clayson, Clawson, & Larson, 2011; Rubia, Hyde, Halari, Giampietro, & Smith, 2010; Bayliss, di Pellegrino, & Tipper, 2005). It is possible, therefore, that a domain-general system may underpin the sex differences observed across these tasks, including during imitation control, but no research to date has directly investigated this proposal.

Across two fMRI experiments, the current study investigated functional specificity and sex differences in imitation control. Several aspects of the experimental design provide grounds to extend current understanding in meaningful and concrete ways. First, this is the first study to use independent functional localisers to identify MD and ToM networks in single subjects and directly test the involvement of these networks in imitation control. By doing so, we can directly test hypotheses regarding the role of functionally defined neural circuits (i.e., MD and ToM networks) and therefore minimize the reliance on reverse inference to infer cognitive function based on anatomical localization (Poldrack, 2006). Second, we used data from Experiment 1 to perform a power analysis to determine the sample size required to achieve a desired level of power

in Experiment 2. Given the inconsistent findings in prior studies, which had relatively small sample sizes, this multi-experiment approach made sure that our key experiment had over 80% power to detect expected effect sizes. Third, to avoid spatial compatibility confounds, in Experiment 2, we used a modified version of the imitation inhibition paradigm that allowed for an independent measure of spatial and imitative compatibility (Catmur & Heyes, 2011). If the inhibition of automatic imitation relies on a domain-specific neural architecture that is associated with social cognition, as proposed by Brass and colleagues (2009), mPFC and rTPJ would be engaged in imitative control. In contrast, engagement of the MD network would suggest that domain-general processes sub-serve imitation control. Furthermore, the sex difference found previously (Butler et al., 2015) may be supported by differences in ToM or MD networks.

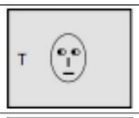
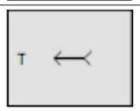
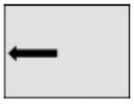
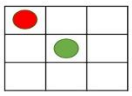
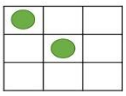
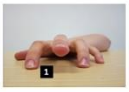
	Task	Sample (Male: Female)	Trials per condition	Sex Difference – Interference	Task Requirements	Conditions	
						Compatible	Incompatible
Bayliss et al. (2005)	Gaze-cueing	80 (40:40)	144	Female > Male	Respond to letter cue		
Bayliss et al. (2005)	Arrow-cueing	40 (20:20)	144	Female > Male	Respond to letter cue		
Rubia et al. (2010)	Oddball	63 (38:25)	160 congruent, 24 oddball	Female > Male	Respond to arrow direction		
Stoet (2010)	Flanker	80 (40:40)	120	Female > Male	Respond to central circle		
Clayson et al. (2011)	Flanker	114 (60:54)	450	Female > Male	Respond to central arrow		
Butler et al. (2015)	Automatic Imitation	230 (97:133)	~30	Female > Male	Respond to number cue		
Stoet (2017)	Simon task	418 (236:182)	51	Female > Male	Respond to arrow direction		

Figure 1. Sex differences in inhibitory control tasks. Female individuals experience greater interference than male individuals in multiple inhibitory control tasks. Images are produced based on figures and description in each experiment apart from Butler et al. (2015), which are the actual images used. Also, in Rubia et al. (2010), the sex difference showed increased interference by the oddball trials rather than the incongruent trials, and this is what is represented by the images. Finally, in Butler et al. (2015), participants completed 60 trials that were 30 ± 2 trials per compatible and incompatible condition.

Methods

Overview of the experimental approach.

Experiment 1 used a group-level whole-brain analysis, which provided the basis for power analyses that set up Experiment 2 as the critical experiment with high statistical power (80%). In Experiment 2, to increase sensitivity and functional resolution, we used independent localisers to identify key functional circuits (i.e., MD and ToM networks), and analyses were performed in single subjects to precisely quantify the consistency of network engagement across individuals (Nieto-Castañón & Fedorenko, 2012; Kanwisher, 2010). Group-level analyses require responses across individuals to overlap in individual voxels. In contrast, the fROI approach allows identification of corresponding functional regions without the requirement of exact voxel overlap across individuals. Therefore, the same voxels need not be active across individuals, as long as voxels within a functionally defined ROI are consistently active across individuals. Consequently, group-level analyses may underestimate functional specificity, whereas fROI analyses can show increased sensitivity (Nieto-Castañón & Fedorenko, 2012). In addition, because of a constrained search volume, fROI analyses typically have higher statistical power than whole-brain analyses (Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010; Saxe, Brett, & Kanwisher, 2006).

Experiment 1

Participants.

Twenty-eight participants (M_{age} = 23.96, SD_{age} = 5.52; 14 women) participated for monetary compensation of £15. Participants gave informed consent in line with the guidelines set by the Research Ethics and Governance Committee of the School of Psychology at Bangor University, were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological damage.

Design and procedure.

All participants performed the imitation task inside the scanner. The participants also did four additional tasks in the same scanning session as part of another experiment. The scanning session started with the imitation task, followed by a run of a face perception task, a flanker task (Eriksen & Eriksen, 1974), another run of the face perception task, a dynamic face localiser (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011), and a ToM localiser (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011).

The order of the tasks was counterbalanced across participants such that, of the 28 participants, 14 participants did the imitation task first, and 14 participants did the flanker task first, with the order of the other tasks remaining the same.

The imitation inhibition task.

The imitation task was based on a stimulus–response compatibility paradigm developed by Brass et al. (2000) consisting of observation and execution of finger-lifting movements during fMRI scanning. Before the task, participants were instructed to hold down the "blue" and "yellow" buttons on the response box with their index and middle fingers of the right hand, respectively. A number cue (either "1" or "2") was presented to participants, and they were asked to lift their index finger on presentation of the number "1" and the middle finger for the number "2." Simultaneously, they also viewed an image of an index or middle finger lift of a left hand viewed from the third-person perspective, such that the fingers extended toward the participants. Thus, there were four trial types in an event-related design that led to two conditions—participants performing the same (congruent) or different (incongruent) finger movement to the observed hand image.

Each trial started with a fixation cross (500 msec) followed by a neutral hand (for a random ISI of 500, 700, or 1000 msec) and a hand image with an index/middle finger lift, which stayed onscreen for 2000 msec, irrespective of when the participant made the response. Sequencing the hand images in such a way led to the appearance of apparent motion of the finger. After 2000 msec, the next trial started immediately with a fixation cross (500 msec). To separately model the influence of individual events in an event-related design, the four trial types were pseudorandomized, such that each trial type was preceded by each other trial type and by itself an equal number of times (Wager & Nichols, 2003; Josephs & Henson, 1999). There were 17 trials in each block. The first trial was used to set up the randomization sequence but excluded from the analysis as it was not preceded by any other trial. The remaining 16 trials within a block were analysed and consisted of eight trials per condition. Each run consisted of five blocks separated by a 3-, 4-, or 5-sec fixation cross. All participants completed one run of the imitation task. Thus, there were 80 trials of interest (40 congruent and 40 incongruent).

Behavioural data analysis.

RT on the imitation inhibition task was measured as the time from number cue onset to when participants made a response. To ensure participants were engaging correctly with the task, participants who had less than 80% accuracy were removed. In

addition, RTs more than 3 *SD*s away from the mean were excluded from the analyses. Furthermore, trials on which participants made an "error" were excluded from the analyses. Errors included an incorrect response, no response, a response after 2000 msec, and pressing an invalid key. The general compatibility effect was calculated as the RT difference between incompatible and compatible trials. A one-sample *t* test was performed to verify the presence of a general compatibility effect. A one-tailed independent sample *t* test was performed to determine if the compatibility effect was greater for female than male individuals. Mean differences, 95% confidence intervals (CI), and Cohen's *d* (Cohen, 1992) are reported for all effects of interest. For the one-sample *t* test, Cohen's d_z was calculated as mean difference divided by the standard deviation of the sample (Lakens, 2013). The 95% CI is reported for the lower bound for a one-tailed *t*test. For the independent samples *t* test, Cohen's *d* was calculated as mean difference between the two groups divided by the pooled standard deviation (Cohen, 1992).

fMRI data analysis.

Data acquisition. Participants were placed supine in a 3-T Philips MRI scanner using a SENSE 32-channel phased array coil. They were requested to avoid head motion during the scanning session and were presented stimuli on a computer screen placed behind the scanner made visible by a mirror attached to the head coil. Responses on the task were recorded with the help of a button box that recorded RTs. Thirty-five axial slices were acquired in an ascending order using a T2*-weighted EPI sequence. The reference slice for slice time correction was the slice acquired in the middle of the sequence (Slice 17). Parameters are as follows: voxel size = 3 × 3 × 4 mm, repetition time = 2000 msec, echo time = 30 msec, flip angle = 90°, slice thickness = 4 mm, slice gap = 0.8 mm, field of view = 230 × 230 × 167 mm^3. One hundred seventy-four volumes were collected for the imitation task.

Four dummy scans collected at the beginning of each run of the task were not included in any analyses. A high-resolution T1-weighted anatomical image was also collected with the following parameters: repetition time = 12 msec, echo time = 3.5 msec, flip angle = 8°, number of axial slices = 170, voxel size = 1 mm^3, field of view = 250 × 250 × 170 mm^3.

Data preprocessing and general linear model. Functional images were pre-processed in SPM-8. Data were realigned, un-warped, and corrected for slice timing. Data

were normalized to the Montreal Neurological Institute (MNI) template with a resolution of 3 mm^3, and images were spatially smoothed (8 mm).

For the imitation task, a design matrix was fit for each participant with three regressors: one each for the correct trials of the two conditions, and one for the "new" trials (i.e., the first trial of each block). The new trials were not used in any further analyses. Stimulus onsets were time-locked to the presentation of the number cue with a duration of 0 sec and convolved with the standard hemodynamic response function.

Whole-brain analyses. Contrast images (incompatible > compatible) were calculated at the single-subject level for the imitation inhibition task to identify regions of the brain showing a compatibility effect. Group-level contrast images were created from these single-subject contrast images to identify regions that were consistently engaged for the compatibility effect across the sample using one-sample *t* tests. To identify a neural signature of the sex difference in imitation inhibition, a Sex × Compatibility ANOVA was computed (female [incompatible > compatible] > male [incompatible > compatible]) as female individuals have been shown to have a higher compatibility effect than male individuals in the imitation task (Butler et al., 2015). For all analyses, contrast images were taken to the group level and thresholded using a voxel-level threshold of $p < .001$ and a voxel extent of 10 voxels. Correction for multiple comparisons was performed at the cluster level (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994), with clusters that survive correction for multiple corrections using a family-wise error correction ($p < .05$; shown in bold font in Table 2A and B; see Results). This restricts the likelihood of false positives (Eklund, Nichols, & Knutsson, 2016). Clusters of activity were identified with the SPM Anatomy toolbox (Eickhoff et al., 2005).

Experiment 2

Participants.

Fifty-five participants (M_{age} = 22.04, SD_{age} = 3.70; 27 women) were recruited from the Bangor community and were either reimbursed with £15 or three course credits for their participation. Informed consent was obtained in line with the guidelines set by the Research Ethics and Governance Committee of the School of Psychology at Bangor University. All participants were right-handed, did not have dyslexia or dyspraxia, were not on any medication, did not report neurological damage, and had normal or corrected-

to-normal vision. The sample size was determined by a power analysis based on Experiment 1 data (see Results).

Design and procedure.

Each participant performed three tasks inside the scanner—the automatic imitation task, a ToM network localiser task, and an MD network localiser task. The order of the tasks was as follows: two runs of the MD network localiser task were interspersed between three runs of the imitation task to offset boredom. This was followed by two runs of the ToM network localiser task. The ToM task was always presented at the end to reduce the likelihood that belief reasoning during the ToM task would influence performance in the imitation task. The order was the same for all participants. Participants also completed a 50-item International Personality Item Pool questionnaire (Donnellan, Oswald, Baird, & Lucas, 2006; Goldberg, 1992; unrelated to the current study) and a stimulus rating form where they were asked to rate the hand stimulus from the imitation task as either male, female, or neutral. The entire session lasted approximately 1.5 hr, with 60 min inside the scanner. All stimulus presentation was coded in MATLAB 2015b and presented with PsychToolBox 3.0.6.

The imitation inhibition task.

The automatic imitation task was similar to the one used in Experiment 1, but with two changes. First, we used a different hand stimulus, which was rated as sex-neutral by observers. The sex of the hand was an important consideration to minimize the possibility of an own-sex bias while exploring sex differences in imitation inhibition. As such, we conducted pilot work that asked observers to evaluate a range of hand stimuli in terms of masculinity and femininity, and we selected the most sex-neutral stimulus (see Supplementary Information, Development of Stimuli). We only used one hand stimulus to simplify the design space. Although using one sex-neutral hand stimulus provided greater experimental control, it may have harmed our ability to study or elicit sex differences. Future work could probe this further by varying the sex of the stimulus and/or by using more sex-typical stimuli.

The second change that we made was to calculate an imitative compatibility effect independent of spatial compatibility (Catmur & Heyes, 2011). To do so, participants viewed an image of an index or middle finger lift of either a right or left hand but always responded with their right hand. Using right- and left-hand images produced eight trial types and four main conditions of interest (see Figure 2A). For example, when cued to lift

their index finger while observing a left-hand index finger lift, the observed movement is both imitatively compatible (same finger), as well as spatially compatible (same side of space to the executed movement). In contrast, when observing a right-hand index finger lift, the participant's response is imitatively compatible (same finger), but it is not on the same side of space (they are spatially incompatible). Thus, participants performed the same (imitatively compatible) or different (imitative incompatible) finger movement on the same (spatially compatible) or different (spatially incompatible) side of space to the observed finger movement, giving rise to the following four conditions:

1. Imitatively and spatially compatible
2. Imitatively and spatially incompatible
3. Imitatively compatible and spatially incompatible
4. Imitatively incompatible and spatially compatible

Sequencing information and pseudo-randomization was the same as Experiment 1. There were 65 trials in each block. The first trial was used to set up the randomization sequence but excluded from the analysis as it was not preceded by any other trial. The remaining 64 trials were analysed, consisting of 16 trials per condition. Each run consisted of two blocks separated by a 3-sec fixation cross. All participants completed three runs of the imitation task. In total, there were 384 trials of interest, 96 per condition. Experiment 2, therefore, had more than twice the number of trials per condition than Experiment 1.

Localiser tasks.

The MD network localiser. To identify regions of the MD network, a verbal working memory (WM) task was used (Fedorenko, Behr, & Kanswisher, 2011). Participants were asked to remember the sequence in which either four (easy condition) or eight (hard condition) digit sequences were presented on screen (see Figure 2B). After each trial, participants had to choose between two sequences presented numerically, one of which matched the sequence in which the digits were presented as words. Feedback was provided as to whether they answered correctly or incorrectly. The hard > easy contrast has been found to robustly activate regions of the MD network (Fedorenko et al., 2011, 2013). Each run consisted of 10 experimental blocks (each 34 sec long) and 6 fixation blocks (each 16 sec long). The total run lasted for 436 sec. Each participant completed two runs of the WM task.

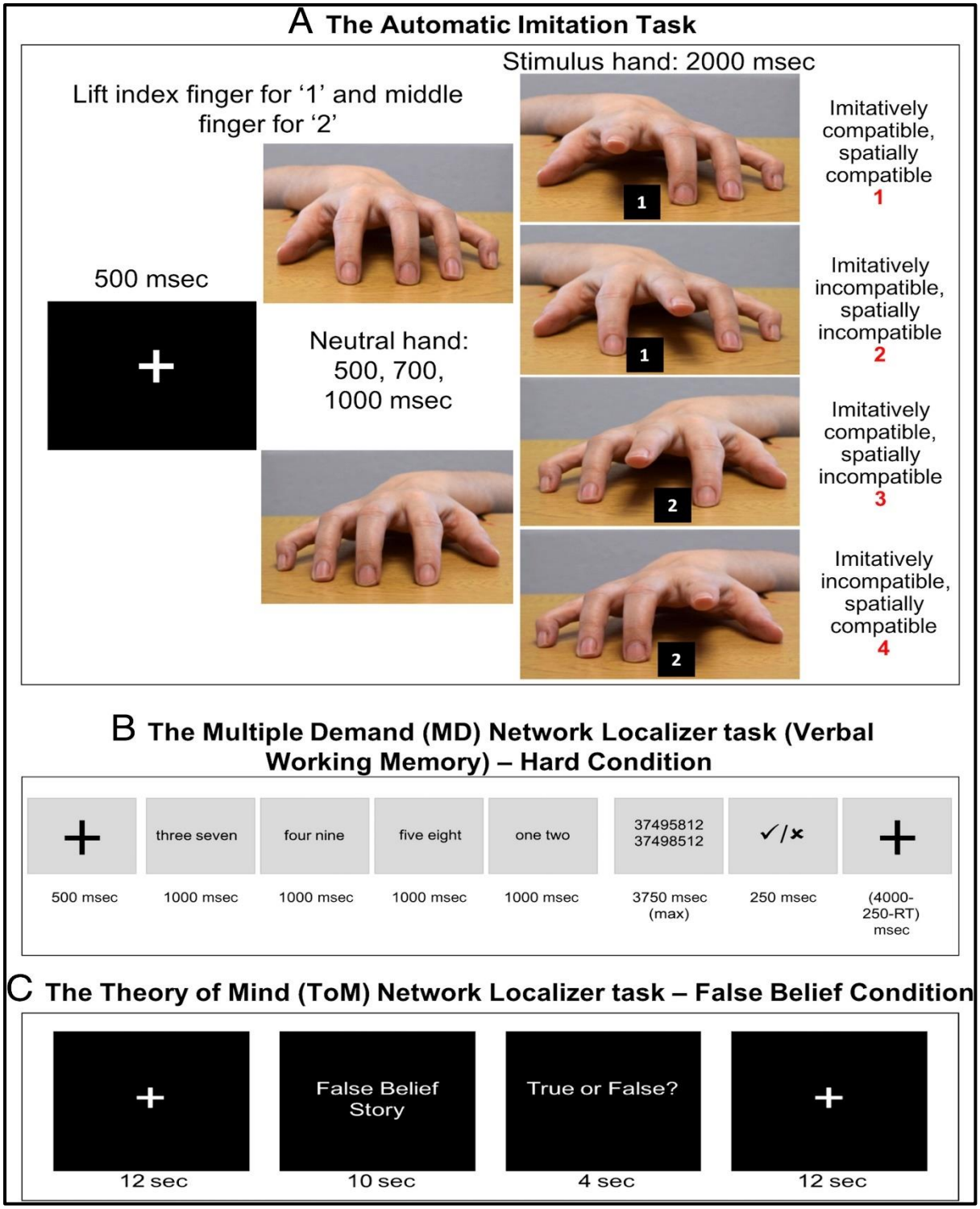


Figure 2. Stimuli for the imitation inhibition and functional localiser tasks. Stimuli and trial design for the imitation inhibition task (A), the Multiple Demand (MD) network localiser task (B), and the Theory-of-Mind network localiser task (C). For the automatic imitation task, spatial compatibility was calculated as spatial incompatible (2 + 3) minus (1 + 4), and imitative compatibility was calculated as imitatively incompatible (2 + 4) minus imitative compatible (1 + 3) trials.

The ToM localiser. To localise brain regions involved in mental state reasoning, we used a paradigm developed by Dodell-Feder and colleagues (2011; saxelab.mit.edu/superloc.php). This localiser task (see Figure 2C) includes 20 stories, each describing a false representation. Ten stories included out-of-date beliefs (the false belief condition), and the other 10 included out-of-date physical representations (photographs/maps; the false photograph condition). The false belief > false photograph contrast has been shown in prior work to robustly activate regions involved in mentalizing (Dufour et al., 2013). All trials consisted of a story (10 sec), followed by a true or false question (4 sec). Each story was separated by a 12-sec rest period. The order of the stories and conditions was the same for all participants. Each participant completed two runs of this task, with five trials per condition presented in each run.

Behavioural data analysis.

RT and accuracy were recorded in the same way as Experiment 1. Compatibility effects were calculated as follows: spatial compatibility = spatially incongruent trials − spatially congruent trials; imitative compatibility = imitatively incongruent trials − imitatively congruent trials. Behavioural data were analysed in the same fashion as Experiment 1, only separately for imitative and spatial compatibility effects. The main aim of the experiment was to test for the presence of imitative and spatial compatibility effects, as well as for differences between the sexes (female > male). Hence, we used a one-sample *t* test to verify the presence of spatial and imitative compatibility effects and a one-tailed independent samples *t* test to test whether female individuals showed a higher spatial/imitative compatibility effect than male individuals.

fMRI data analysis.

Data acquisition. Data acquisition procedures were the same as Experiment 1. There were 249 volumes collected for the imitation task, 219 for the MD network localiser, and 136 for the ToM localiser for each run.

Data pre-processing and general linear model. All MRI data were pre-processed in SPM-8. Data were realigned, unwarped, and corrected for slice timing. Data were normalized to the MNI template with a resolution of 3 mm^3. Normalizing to a common space instead of the individual's native anatomical space allows for comparisons with previous studies (relying on the common space) and is preferred when definition of fROIs

is based on group-constrained functional data (Nieto-Castañón & Fedorenko, 2012). Images were spatially smoothed (8 mm).

For the imitation task, a design matrix was fit for each participant with five regressors: one each for the correct trials of the four conditions and one for "new" trials (i.e., the first trial of each block). Stimulus onsets were time-locked to the presentation of the number cue with a duration of 0 sec and convolved with the standard hemodynamic response function. Contrast images were calculated for each individual participant to identify regions of the brain showing a spatial (spatially incompatible > spatially compatible) or imitative (imitatively incompatible > imitatively compatible) compatibility effect.

For the localiser tasks, the design matrix consisted of regressors for each experimental condition ("Belief" and "Photo" for the ToM localiser and "Hard" and "Easy" for the MD localiser). The onset and duration of each condition was specified and convolved with the standard hemodynamic response function. Contrast images were then calculated for each individual subject to identify regions that responded to cognitive demand (hard > easy) and mentalizing (belief > photo).

Definition of group-constrained subject-specific analyses. For the group-constrained subject-specific (GSS) analyses, the spm_ss toolbox was used, which runs in SPM using MATLAB (web.mit.edu/evelina9/www/funcloc.html). The GSS approach developed by Fedorenko et al. (2010) and Julian, Fedorenko, Webster, and Kanwisher (2012) was used to define fROIs for each participant. These fROIs were defined using (1) each individual's activation map for the localiser tasks and (2) group-constraints or masks. These masks refer to a set of "parcels," which demarcate areas in the brain where prior work has been shown to exhibit activity for the localiser contrasts.

Two sets of fROIs were defined (Figure 3): MD network fROIs that have been known to exhibit activity for a variety of cognitive control tasks (Fedorenko et al., 2013; Duncan, 2010) and ToM network fROIs that support mentalizing and have been specifically implicated for imitation inhibition (Brass et al., 2009; Saxe & Kanwisher, 2003). For the ToM network, four parcels were derived from a group-level map from 462 participants for the false belief > false photograph contrast (Dufour et al., 2013). These regions included the dorsal, middle, and ventral mPFC (DMPFC, MMPFC, and VMPFC, respectively) and the rTPJ. For the MD network, we used 16 parcels derived from a set of functional parcels created by Idan Blank based on a probabilistic overlap

map from 197 participants (available at https://evlab.mit.edu/funcloc/download-parcels). These included areas in bilateral superior and inferior parietal lobules (SPL and IPL, respectively), intra-parietal sulcus (IPS), inferior and middle frontal gyri (IFG, MFG), precentral gyrus (PrecG), insula, and the SMA. These areas were chosen for two reasons: (1) they were part of the MD network (Fedorenko et al., 2013) and (2) they have been shown to respond in prior work to the specific type of interference control of relevance to the current study (Marsh et al., 2016; see also Experiment 1).

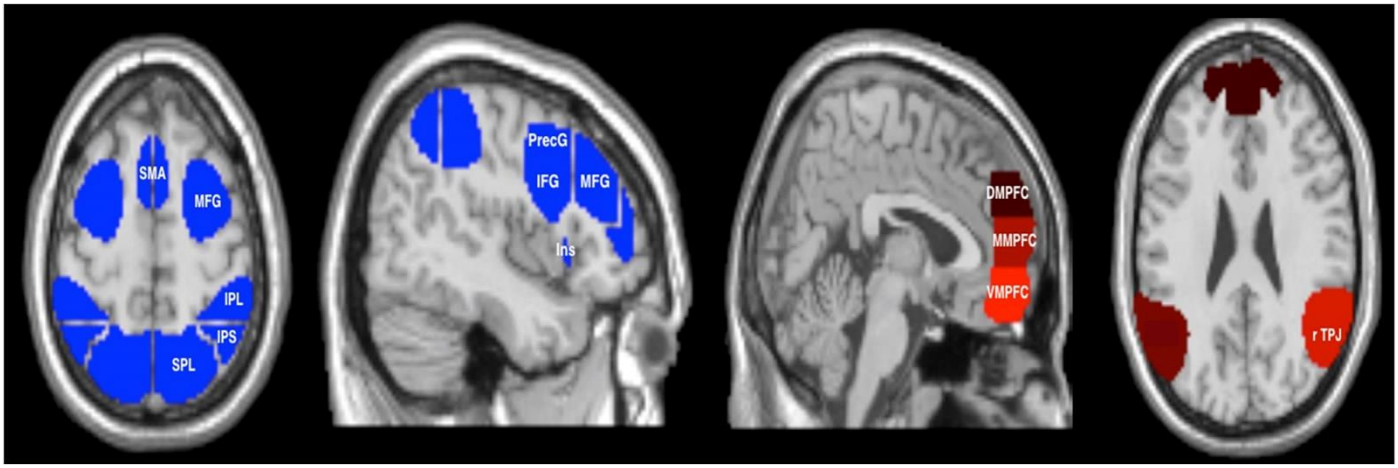


Figure 3. Graphical representation of the parcels used to define the MD and ToM network fROIs. The MD network consisted in 16 parcels, and the ToM network included 4 parcels.

For each individual, these masks were used to constrain the selection of subject-specific fROIs. For each individual, for the ToM network mask, the belief > photo contrast was used, and the top 10% of voxels (based on *t* values) within each parcel were defined as that individual's fROI. Similarly, for the MD network mask, each individual's top 10% of voxels (based on *t* values) in the hard > easy contrast were defined as that individual's fROI. Using the top 10% of voxels, rather than a fixed threshold (e.g., all voxels with $p < .001$), ensures a constant size of each fROI across individuals (Blank, Kanwisher, & Fedorenko, 2014). We also ran the analyses using a fixed threshold ($p < .001$, uncorrected) and found the same pattern of results (see Supplementary Tables S1.1 and S1.2). All analyses reported below are based on the top 10% of voxels that were generated from the localiser data fROIs. Percent signal change values were extracted from all fROIs. For the main analysis, all runs of the localiser tasks were used to define

fROIs in each individual. Responses in these fROIs were estimated for spatial and imitative compatibility effects.

In a supplementary analysis, responses to the localiser contrasts were also estimated to ensure that all the fROIs showed the expected response with respect to the localiser contrasts. This supplementary analysis ensured that the ToM network showed a robust belief > photo and the MD network showed a robust hard > easy effect. For these localiser analyses, an across-run cross-validation approach was used (Nieto-Castañón & Fedorenko, 2012) to ensure that data used for defining fROIs were independent of data used for estimating response (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009).

As implemented in GSS, statistical tests were performed on the percent signal change values using standard Student's *t* tests. One-sample *t* tests were performed to investigate the response of the MD and ToM network fROIs to spatial and imitative compatibility effects. Based on prior behavioural findings, which showed greater RT interference for female than male individuals during imitation inhibition (Butler et al., 2015), we expected to observe sex differences in those regions that also show simple compatibility effects. That is, we expected brain regions that were generally involved in spatial and/or imitative control to show sex differences. As such, we only investigated sex difference in those fROIs that showed spatial or imitative compatibility effects. To do so, one-tailed independent samples *t* tests were performed that tested for greater engagement for female than male individuals. False discovery rate (FDR) multiple-comparison correction ($p < .05$) was used to correct for the number of fROIs in each functional network.

Results

Experiment 1

Behavioural results.

A one-sample *t* test confirmed a general compatibility effect (mean = 80.02, *SE* = 8.19), $t(27) = 9.77$, $p \leq .001$, 95% CI (63.22, 96.82), Cohen's $d_z = 1.85$. A one-tailed independent samples *t*test showed no differences between male individuals (mean = 70.94, *SE* = 13.30) and female individuals (mean = 89.10, *SE* = 9.43), $t(26) = 1.114$, $p = .138$, 95% mean difference = 18.16, 95% CI (−9.64), Cohen's $d = 0.42$. All participants had >80% accuracy; hence, all were included in the analysis. Trials on which participants made an incorrect response (0.95%) did not make a response or responded after 2000

msec (0.52%) or pressed an invalid key or responded too fast (0.09%) were excluded from the analyses.

fMRI results.

In a whole-brain analysis, compatibility effects (general incompatible > general compatible) were observed in dorsomedial frontal cortex and bilaterally in dorsolateral frontal and parietal cortices (Figure 4A; Table 2A). A small volume correction (SVC) using MD and ToM network parcels was performed to restrict the search area to ToM and MD networks. Using the MD network SVC, results showed widespread activation of frontal and parietal regions, which survived correction for multiple comparisons (Figure 4A, Ci). In contrast, using the ToM network SVC, no clusters survived correction for multiple comparison, and only rTPJ showed a compatibility effect at more lenient threshold (p < .001, uncorrected; see Supplementary Tables S2.1 and S2.2). Anterior mPFC did not show the general compatibility effect even at this more lenient threshold.

The Sex × Compatibility interaction revealed clusters in left SPL extending into postcentral gyrus and a further cluster in the cerebellum (Figure 4B; Table 2B). No clusters emerged following an SVC analysis using the MD and ToM network masks, which demonstrate that the clusters emerging from the Sex × Compatibility interaction do not overlap with the MD or ToM networks (see Supplementary Tables S2.1 and S2.2; Figure 4Cii and Dii).

Power analysis.

We set up Experiment 1 to estimate the appropriate sample size for our critical experiment (Experiment 2). To this end, a power analysis was performed using the fMRIpower software package (fMRIpower.org; Mumford & Nichols, 2008). We performed the power analysis as follows: First, a whole-brain map of the imitation task general compatibility effect (incompatible > compatible) from Experiment 1 was entered into fMRIpower. Next, two ROIs were identified: the MD network (Duncan, 2010) and the ToM network (Saxe & Kanwisher, 2003). The MD and ToM network masks used were the same as in Experiment 2 (see Methods). As recommended, we corrected the alpha value by the number of ROIs (0.05/2 = 0.025) before performing power analyses (Mumford, 2012).

Results from these power analyses showed that testing 50 participants in Experiment 2 would provide 80% power to detect effects as large as (or larger than) the

average effect size that was observed across all nodes in the MD network in Experiment 1 (Cohen's *d* = 0.4, mean signal change = 0.23, *SD* = 0.58). We did not have the same level of power to detect smaller effects than these, such as those observed in the ToM network in Experiment 1. Indeed, the effects in the ToM network in Experiment 1 were so small that we would have needed an impractically large sample size to achieve 80% power. As such, in Experiment 2, we decided to test participants until we had 50 usable data sets.

Design differences between Experiments 1 and 2 are worth considering when interpreting these power calculations because we may be underestimating the power of our design in Experiment 2. The toolbox used to run power calculations (fmripower.org) can only estimate power for a future experiment with the same design as the current data set (Mumford & Nichols, 2008). However, the designs of Experiments 1 and 2 differed in two ways. First, Experiment 1 measured a general compatibility effect, whereas in Experiment 2, we broke this effect down into spatial and imitative compatibility effects. Second, Experiment 2 had more than double the amount of trials per condition as Experiment 1. Therefore, the primary contrast used to determine power was not identical to the contrast used in Experiment 2, but due to a greater number of trials per condition to estimate the effects of interest, we may underestimate power in Experiment 2. Given the lack of sex differences in Experiment 1 in our ROIs, we did not have sufficient power to convincingly investigate neural differences between male and female individuals in Experiment 2. However, given our a priori predictions regarding sex, we continue to report sex difference analyses throughout the article.

Table 2.

General Compatibility Effect and Sex × Compatibility Interaction for the Imitation Inhibition Task (Experiment 1).

Region	Cluster Size	p FWE Corr	t	MNI Coordinates		
				x	y	z
(A) General Compatibility Effect (Incompatible > Compatible)						
L IPL extending into SPL and superior frontal gyrus	**986**	**<.001**	**8.40**	**−39**	**−40**	**43**
			6.50	**−36**	**−37**	**70**
			6.38	**−27**	**−7**	**70**
L cerebellum	**150**	**.001**	**5.79**	**−21**	**−55**	**−41**
			4.95	**−30**	**−55**	**−35**
			4.72	**−9**	**−70**	**−44**
R cerebellum	**198**	**<.001**	**5.71**	**21**	**−58**	**−44**
			5.12	**45**	**−46**	**−32**
			4.32	**39**	**−55**	**−23**
R PrecG extending across superior frontal gyrus and MFG	**183**	**<.001**	**5.12**	**27**	**−1**	**70**
			5.04	**42**	**2**	**58**
			4.39	**39**	**−10**	**61**
R postcentral gyrus extending into SPL and IPL	**481**	**<.001**	**5.18**	**33**	**−40**	**73**
			4.55	**42**	**−40**	**67**
			4.50	**48**	**−34**	**37**
R posterior middle temporal gyrus	41	.179	4.61	66	−46	1
			4.13	57	−40	−8
			3.63	60	−43	13
L insula	24	.458	4.67	−36	17	−2
R posterior medial frontal cortex	20	.564	4.79	3	−4	73
L posterior medial frontal cortex	55	.083	4.40	−3	−1	52
			3.82	−6	11	52
R pallidum extending into thalamus	11	.834	4.14	21	−7	−2
			3.80	15	−6	7
L paracentral lobule	11	.834	3.89	−12	−19	79
R middle cingulate cortex	20	.564	3.85	9	14	43
			3.78	6	8	49
			3.67	6	17	52
(B) Sex × Compatibility [Female (Incompatible > Compatible) > Male (Incompatible > Compatible)]						
L SPL extending into postcentral gyrus	**93**	**.011**	**4.98**	**−21**	**−37**	**70**
			4.80	**−30**	**−19**	**46**
			4.60	**−24**	**−31**	**52**
L cerebellum	16	.679	4.34	−24	−55	−38
			4.04	−21	−55	−29

N.B. Regions surviving a voxel-level threshold of $p < .001$, and 10 voxels are reported for the (A) general compatibility effect and (B) Sex × Compatibility interaction for the imitation inhibition task. Subclusters at least 8 mm from the main peak are listed. **Bold** font indicates clusters that survive correction for multiple corrections using a family-wise error (FWE) correction ($p < .05$). MNI = Montreal Neurological Institute; SPL = superior parietal lobule; IPL = inferior parietal lobule; PrecG = precentral gyrus; MFG = middle frontal gyrus; L = left hemisphere; R = right hemisphere.

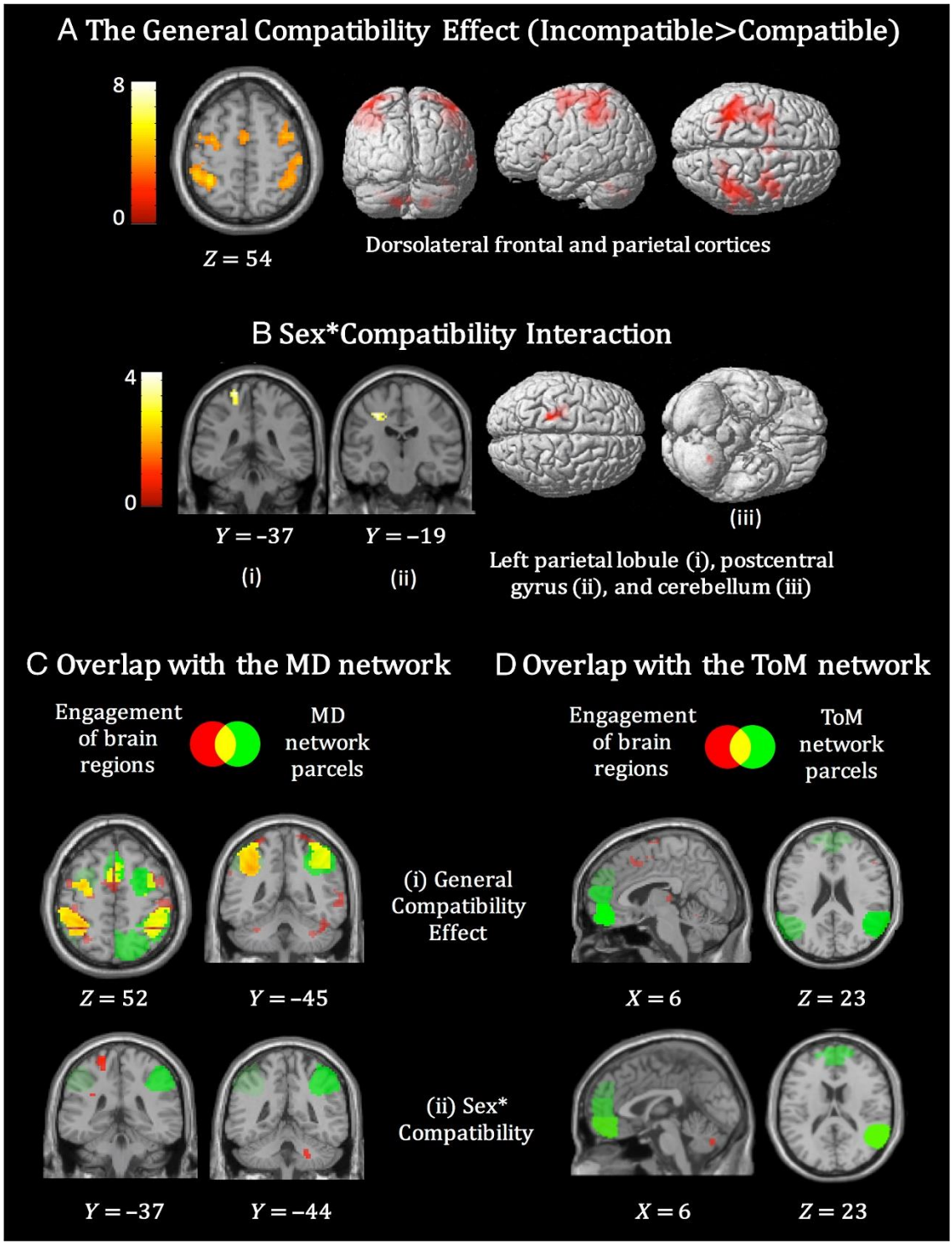


Figure 4. General compatibility effect and Sex × Compatibility interaction in the imitation inhibition task (Experiment 1). (A) Results for the general compatibility effect (incompatible > compatible). Clusters emerged in the dorsal frontoparietal cortices. (B) Results for the Sex × Compatibility interaction (defined as [female (compatibility effect) > male (compatibility effect)]. Clusters emerged in the left superior parietal cortex extending into the postcentral gyrus. The MD network parcels were overlapped with the general compatibility effect (C(i)) and the Sex × Compatibility interaction (C(ii)). An overlap was found between MD parcels and regions engaged by the general compatibility effect. The ToM network parcels were also overlapped with the general compatibility effect (D(i)) and Sex × Compatibility interaction (D(ii)). No overlapping regions were found for the ToM network and regions engaged for the general compatibility effect or for the Sex × Compatibility interaction. Voxel-wise threshold used for all images was $p < .001$, $k = 10$. For a complete set of results, see Table 2 and Supplementary Tables S2.1 and S2.2.

Experiment 2

Behavioural results.

The hand stimulus used in Experiment 2 for the imitation inhibition task was perceived as "neutral" by most participants ($mean_{rating}$ = 5.20, SD_{rating} = 2.04; rated on a scale of 1–9, where 1 = *most masculine*, 5 = *neutral*, and 9 = *most feminine*). To ensure participants were engaging correctly with the task, runs on which participants had less than 80% accuracy (two runs of one participant) were removed. In addition, RTs more than 3 *SD*s away from the mean (two runs of one participant and one run of another participant) were excluded from the analyses. Furthermore, trials on which participants made an incorrect response (1.52%), did not make a response or responded after 2000 msec (0.61%), or pressed an invalid key (0.22%) were also excluded from the analyses. Figure 5 shows the imitative and spatial compatibility effects for both the sexes. For RT data, see Supplementary Table S3.

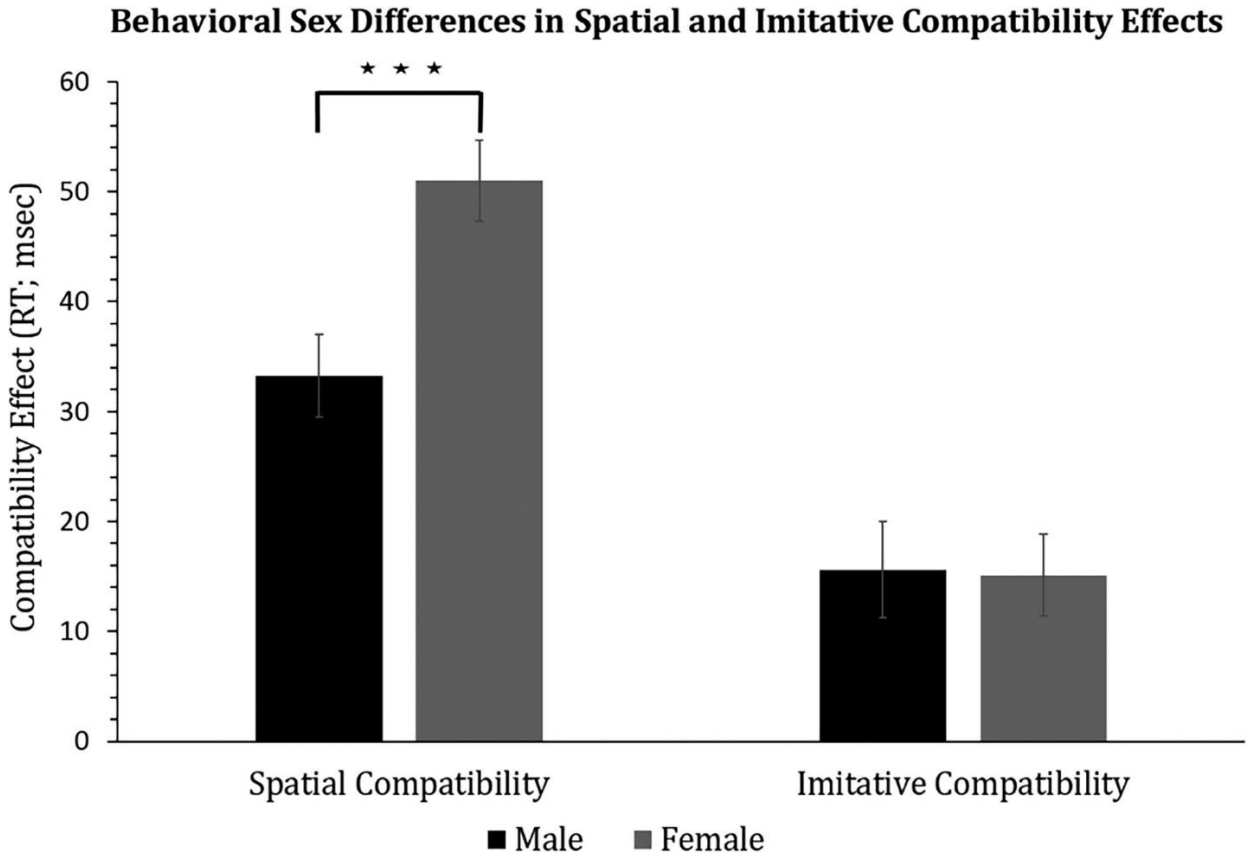


Figure 5. Behavioural sex differences in imitative and spatial compatibility effects. The spatial and imitative compatibility effects (RTs) in male and female individuals displayed in milliseconds. Error bars denote standard error of mean.

Spatial compatibility. A one-sample *t* test confirmed a spatial compatibility effect (mean = 41.94, *SE* = 2.87), $t(54) = 14.618$, $p \leq .001$, 95% CI (36.19, 47.69), Cohen's d_z = 1.97. A one-tailed independent samples *t*test evidenced a greater spatial interference effect for female individuals (mean = 50.98, *SE* = 3.67) as compared with male individuals (mean = 33.20, *SE* = 3.75), $t(53) = -3.38$, $p < .001$, mean difference = 17.76, 95% CI (8.91); Cohen's $d = 0.91$.

Imitative compatibility. A one-sample *t* test showed a significant imitative compatibility effect (mean = 15.37, *SE* = 2.86), $t(54) = 5.37$, $p < .001$, 95% CI (9.63, 21.11), Cohen's $d_z = 0.72$. There was no significant difference between male individuals (mean = 15.62, *SE* = 4.39) and female individuals (mean = 15.11, *SE* = 3.73), $t(53) = 0.09$, $p = .465$, mean difference = −0.51, 95% CI (−10.18), Cohen's $d = 0.02$.

fMRI results.

Five participants were excluded from the fMRI analyses due to lower than 80% accuracy in two runs of the imitation task and the MD network localiser task ($n = 1$) and excessive head motion ($n = 4$; displacement > 4 mm) in all runs of the imitation task and/or all runs of either of the localiser tasks. Thus, the final sample consisted of 50 participants (M_{age} = 22.26, SD_{age} = 3.71; 24 female). From these 50 participants, two sessions of the imitation task were also excluded for one participant due to excessive head motion and one participant's data for one session of the imitation task could not be used because the data file was corrupted.

Localiser tasks.

All fROIs showed the predicted responses to the localiser contrasts (as estimated using data not used for defining ROIs; see Methods). All the MD network fROIs showed a robust hard > easy effect (*t*s > 9.13, *p*s < .0001), and ToM network fROIs showed a robust belief > photo effect (*t*s > 5.70, *p*s < .0001). For responses for each individual fROI separately, see Supplementary Tables S4.1 (MD) and S4.2 (ToM).

The automatic imitation task.

GSS analyses. Figure 6 shows the mean percent signal change for each fROI in the MD and ToM networks for spatial (spatial incompatible > spatial compatible) and imitative compatibility (imitative incompatible > imitative compatible) effects.

MD network fROIs.

Spatial compatibility. All 16 fROIs of the MD network showed a spatial compatibility effect (*t*s > 1.8, *p*s < .04; Figure 6A, Table 3), which survived correction for

multiple comparisons (p < .05, FDR-corrected). The mean percent signal change across the MD network for spatial compatibility was 0.70, SD = 1.66, Cohen's d = 0.42. No significant differences were found between male and female individuals in percent signal change values in any of the fROIs (ts < 1.6, ps > .1), except right SPL which approached significance (p = .062; Figure 7A).

Imitative compatibility. None of the 16 MD network fROIs showed an imitative compatibility effect, which survived correction for multiple comparisons (all ps > .05, FDR-corrected). Five MD network fROIs showed an imitative compatibility effect at an uncorrected threshold (ts > 1.95, ps < .05). These fROIs include bilateral IPL, bilateral IPS, and the right IFG (Figure 6A, Table 3). Four further fROIs showed an imitative compatibility effect that approached significance, which included left IFG (p = .07), right SPL, right MFG, and right PrecG (p = .06). The mean percent signal change across the MD network for imitative compatibility was 0.54, SD = 2.06, Cohen's d = 0.26. There was no significant difference between male and female individuals in any of these fROIs (ts < 1.5, ps > .08; see Figure 7B).

ToM network fROIs.

None of the ToM network fROIs showed imitative (ts < 1.3, ps > .50) or spatial (ts < 1.6, ps > .06) compatibility effects, even at an uncorrected significance threshold (Figure 6B, Table 4). rTPJ showed a spatial compatibility effect that approached significance (p = .065). The mean percent signal change across the ToM network for spatial compatibility was −0.16, SD = 1.88, Cohen's d = −0.08, and the mean percent signal change across the ToM network for imitative compatibility was −0.32, SD = 2.02, Cohen's d = −0.16.

Whole-brain analyses.

For completeness and for use in future meta-analyses, we also computed group-level whole-brain analyses separately for general, spatial, and imitative compatibility effects, as well as for Sex × Compatibility interactions (see Supplementary Table S4).

Open Science

Data for Experiments 1 and 2 are freely available online including behavioural and fROI data (osf.io/45x6z), as well as whole-brain t maps (https://neurovault.org/collections/3218).

Table 3.
Responses in each MD network fROI for Spatial and Imitative Compatibility.

ROI	***ROI Size***	***Intersubject Overlap***	***Average ROI Mask Size (Voxels)***	***Spatial Compatibility***			***Imitative Compatibility***		
				t	*p*	*p-FDR*	*t*	*p*	*p-FDR*
L_SPL	1173	1	117	2.00	.026	**.028**	1.13	.131	.191
L_IPS	287	1	28	2.00	.026	**.028**	1.96	.028	.089
L_IPL	641	1	64	2.72	.005	**.019**	2.05	.023	.089
L_MFG	536	1	53	2.16	.018	**.028**	0.53	.301	.324
L_PrecG	338	1	33	2.17	.018	**.028**	0.91	.184	.227
L_IFG	181	1	18	1.83	.040	**.037**	1.53	.066	.118
L_Insula	197	1	19	2.78	.004	**.019**	0.52	.304	.324
L_SMA	294	1	29	2.52	.008	**.020**	0.39	.349	.349
R_SPL	1181	1	118	2.30	.013	**.026**	1.56	.062	.118
R_IPS	227	1	22	2.03	.024	**.028**	2.30	.013	.069
R_IPL	599	1	59	2.65	.005	**.019**	2.50	.008	.063
R_MFG	535	1	53	3.57	<.001	**.006**	1.55	.064	.118
R_PrecG	269	1	26	2.43	.009	**.021**	1.59	.060	.118
R_IFG	265	1	26	2.61	.006	**.019**	2.53	.007	.063
R_Insula	184	1	18	2.09	.021	**.028**	1.24	.120	.175
R_SMA	328	1	32	2.30	.022	**.028**	1.06	.148	.198

N.B. For each individual, for the MD network mask, the hard > easy contrast was used, and the top 10% of voxels (based on *t*values) within each parcel were defined as that individual's fROI. Uncorrected *p* values as well as FDR-corrected *p* values are reported. Cells in **bold** are fROIs that survive correction for multiple comparisons ($p < .05$, FDR-corrected).

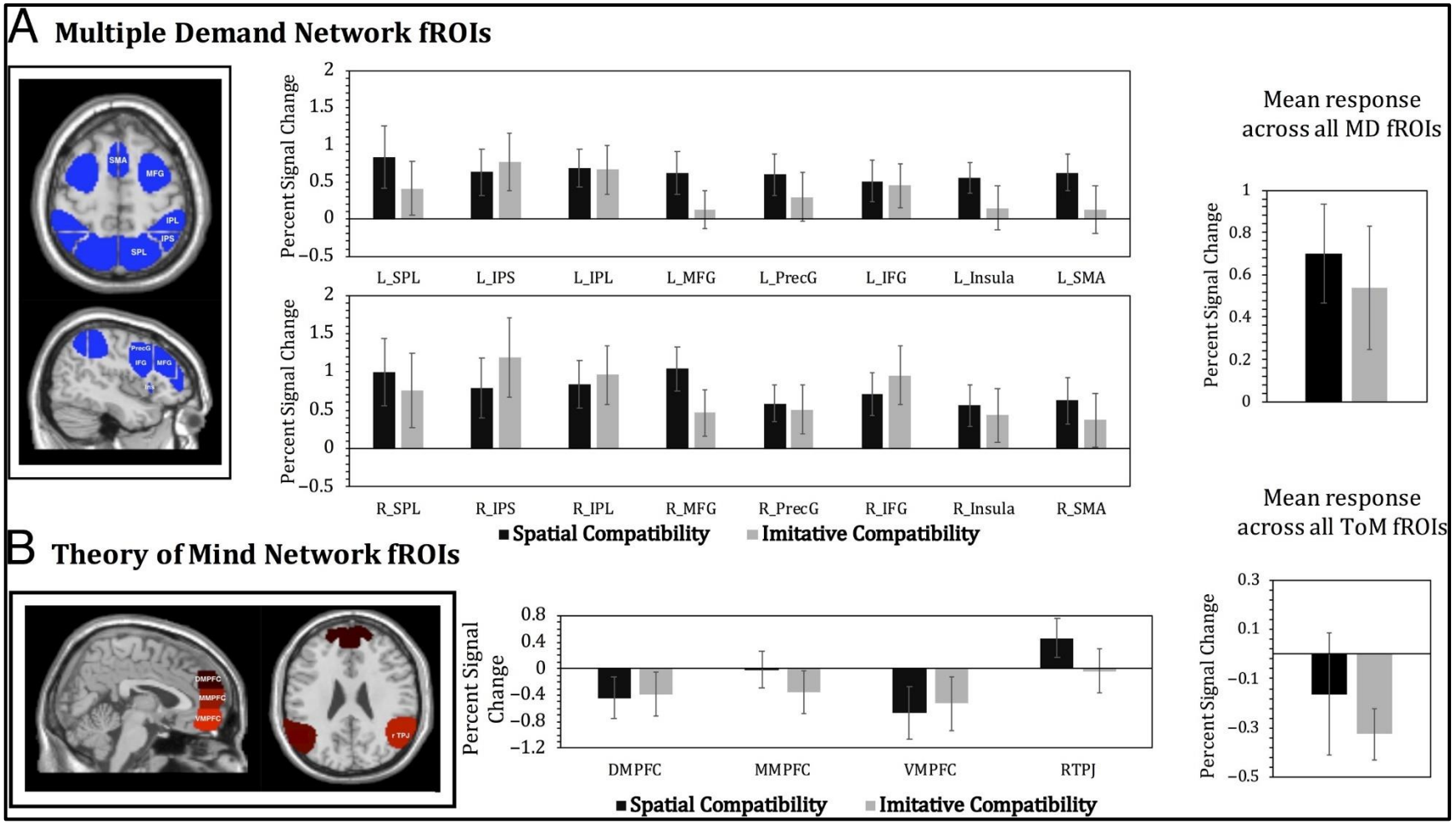


Figure 6. The parcels used to define individual fROIs and the responses to spatial and imitative compatibility effects in the MD (A) and ToM (B) network fROIs are shown. Error bars denote standard error of mean. All MD network fROIs were sensitive to spatial compatibility effects (FDR-corrected, $p < .05$). Bilateral IPL, bilateral IPS, and the right IFG showed a significant response for imitative compatibility effects, but at an uncorrected threshold of $p < .001$. No ToM network fROIs showed engagement for either spatial or imitative compatibility effects.

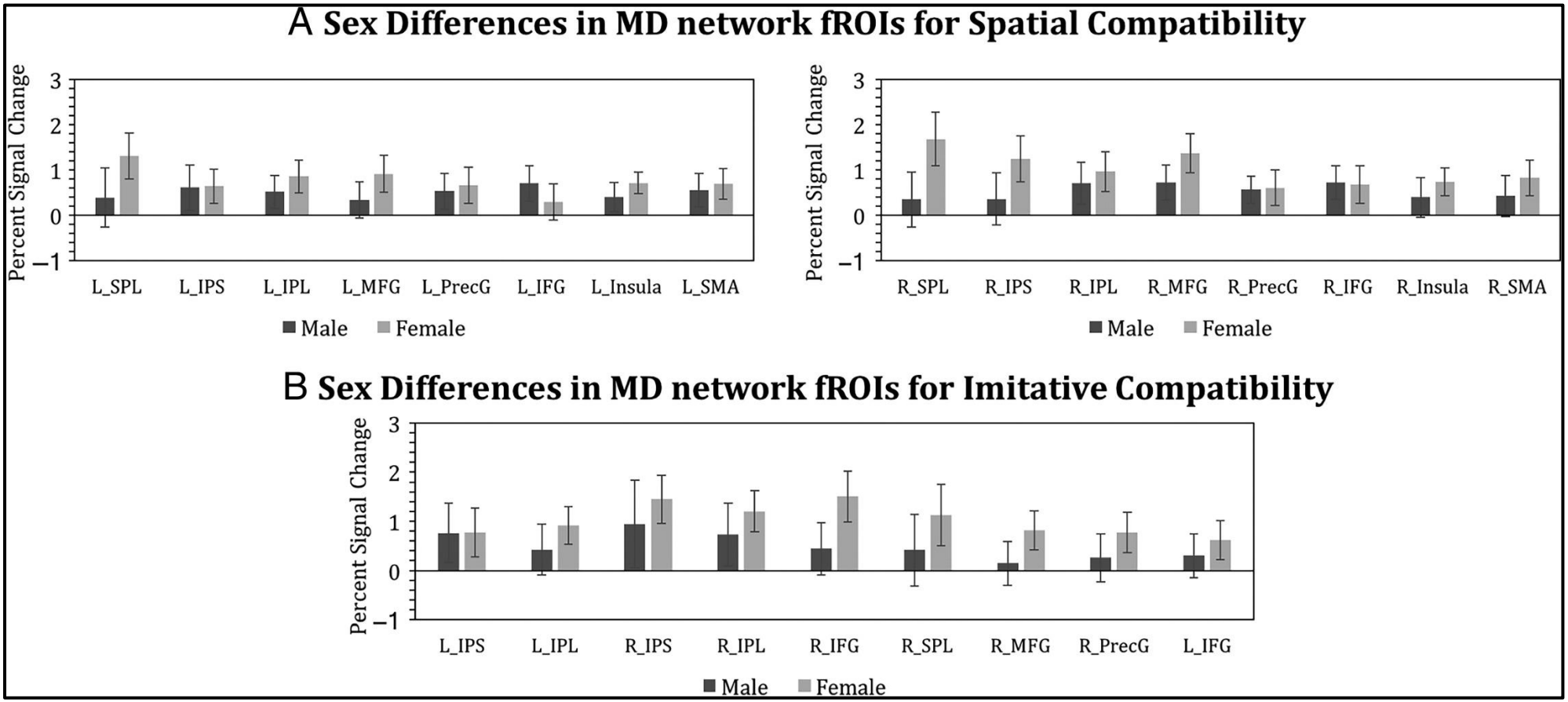


Figure 7. Responses to spatial (A) and imitative compatibility (B) effects separately for male and female participants in the MD network. Error bars denote standard error of mean. None of the fROIs showed a sex difference either in imitative or spatial compatibility that survived correction for multiple comparisons.

Discussion

The current study provides the most robust neuroimaging evidence to date for a lack of functional specificity in the neural circuits supporting the inhibition of automatic imitation. With higher statistical power and functional sensitivity than prior studies, across two experiments the results demonstrate that imitation inhibition engages a domain-general neural network as opposed to a brain network that supports social cognition. As such, models of imitation control need updating to include an increased role for domain-general processes and a reduced or altered role for domain-specific processes. Furthermore, in terms of behaviour, female individuals showed a higher spatial but not imitative compatibility effect than male individuals. However, there was no sex difference in the neural mechanisms underlying spatial or imitation control, which suggests that further exploration of sex differences in inhibitory control is required.

Functional Specificity in Imitation Inhibition

Our findings show that brain regions that are engaged in a verbal WM task, which are associated with the operation of the MD network (Fedorenko et al., 2013; Duncan, 2010), are also engaged during spatial and imitative conflict resolution. These results support the involvement of a domain-general cognitive and neural system during the control of imitation. By contrast, brain regions that are engaged in a belief reasoning task, which are associated with the operation of the ToM network (Van Overwalle, 2009; Saxe & Kanwisher, 2003; Frith & Frith, 1999), show no engagement during the inhibition of imitation. As such, we provide no evidence for domain specificity in cognitive and neural systems that control imitation.

Brass and colleagues (2009) proposed that, in the context of imitation control, rTPJ is involved in self–other distinction, and mPFC enforces the self-generated action over the observed action. Our findings are inconsistent with the hypothesis that a specific neural system related to social cognition is engaged in the inhibition of automatic imitative tendencies. mPFC and rTPJ have both been implicated in imitation inhibition by some studies (Wang et al., 2011; Brass et al., 2009; Spengler et al., 2009; Brass, Derrfuss, & von Cramon, 2005). In contrast, other studies found engagement of mPFC only (Cross et al., 2013; Brass et al., 2001) or of domain-general regions rather than mPFC and rTPJ (Marsh et al., 2016; Cross & Iacoboni, 2013; Crescentini et al., 2011; Bien et al., 2009). In

both experiments in the current study, we had larger sample sizes than prior experiments, and in Experiment 2, we had sufficient statistical power to be confident in detecting effects as large as previously observed in mPFC and rTPJ, should they exist. Taken together with prior findings (Table 1), we suggest that, during the inhibition of imitation, the consistency of mPFC and rTPJ engagement across individuals is relatively low, whereas the consistency of MD network engagement across individuals is relatively high.

These results have potential implications for self–other control theories of social cognition more generally. Mostly based on imitation research, which previously suggested that mPFC and rTPJ are engaged in imitation inhibition, self–other control is thought to be a candidate mechanism for a diverse set of social functions (de Guzman, Bird, Banissy, & Catmur, 2016; Sowden & Shah, 2014; Brass et al., 2009). For example, self–other control processes have been linked to autism, empathy, and theory of mind (de Guzman et al., 2016; Sowden & Shah, 2014; Spengler et al., 2009). However, recent behavioural findings, which used larger sample sizes than prior work and meta-analytical approaches, do not support the view that the control of imitation varies as a function of social disposition as indexed by autistic-like traits and empathy (Cracco et al., 2018; Genschow et al., 2017; Butler et al., 2015). In light of these recent behavioural results, the lack of engagement of mPFC and rTPJ in the current study raises an important question about the reliance of imitation inhibition on a self–other distinction. One possibility is that, instead of a distinctly social mechanism (Bertenthal & Scheutz, 2013; Boyer et al., 2012), inhibiting imitative tendencies may involve the same cognitive processes that are used when inhibiting other nonsocial external influences (Cooper et al., 2012; Heyes, 2011).

Alternatively, the engagement of mPFC and rTPJ during self–other control processes may be more complicated than current models of social cognition suggest. Indeed, a small number of neurostimulation studies have shown that modulation to rTPJ can influence performance on RT measures of imitation (Sowden & Catmur, 2015; Hogeveen et al., 2014). In addition, mPFC and rTPJ have been found to be involved in the modulation of automatic imitation. For example, Klapper et al. (2014) found a higher response in rTPJ when an interaction partner looked human and was believed to be human compared with when neither of these animacy cues was present. Wang and colleagues (2011) demonstrated that mPFC had a top–down influence on other brain

circuits during social modulation of imitation via direct gaze. These studies suggest that mPFC and/or rTPJ may have a regulatory role, be sensitive to social context, and be functionally connected to other regions during the inhibition of automatic imitation. Indeed, regions that do not show direct engagement in a cognitive process of interest have been known to have a regulatory influence on other regions that are directly engaged (Burnett & Blakemore, 2009). In line with this proposal, Cross and colleagues (2013) suggested that imitation control involves top–down regulation between a domain-general cognitive control network and a domain-specific network relevant for imitation. More generally, research from other domains of social cognition shows growing evidence for higher complexity and functional interplay within and between so-called domain-specific and domain-general networks (Spunt & Adolphs, 2015; Baetens, Ma, Steen, & Van Overwalle, 2014; Quadflieg et al., 2011; Zaki, Hennigan, Weber, & Ochsner, 2010). These studies suggest that models including neat divisions between these networks may be an overly simplistic characterization of mental function (Michael & D'Ausilio, 2015; Barrett, 2012). Much like social cognition in general, therefore, imitation control may be best explained by interactions between component functional circuits, which themselves need not be domain-specific (Spunt & Adolphs, 2017). A crucial direction for future research is testing for more complex models of imitation, which may involve connectivity in and between regions of the MD and ToM networks.

An important point to note, however, is that any conclusions made regarding possible domain specificity of mPFC and rTPJ are based on the assumption that mPFC and rTPJ are at least partly specialized for social cognition (Brass et al., 2009). Recent evidence suggests that mPFC and rTPJ may be functionally versatile in the sense that they show general cognitive properties, which may not be specific to social cognition (Dugué, Merriam, Heeger, & Carrasco, 2017; Schurz, Tholen, Perner, Mars, Sallet, 2017; Schuwerk, Schurz, Müller, Rupprecht, & Sommer, 2017; de la Vega, Chang, Banich, Wager, & Yarkoni, 2016; Carter & Huettel, 2013; Alexander & Brown, 2011; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). Thus, the argument that the engagement of mPFC and rTPJ in imitation inhibition may be specific to social cognition might need further validation. In addition, social cognition itself has been broken down in “bottom–up” and “top–down” domains (Zaki & Ochsner, 2012). The bottom–up domain refers to pre-reflective processes that are fast and stimulus driven, whereas the top–down domain maps on reflective, cognitively laborious, and flexible processes (Bohl & van den

Bos, 2012). When extended to imitation control, prior research has consistently implicated regions involved in top–down control for automatic imitation (Brass et al., 2009). However, recent studies suggest that imitation control (and social cognition more broadly) relies on interactions between bottom–up and top–down processes (Christov-Moore, Conway, & Iacoboni, 2017; Cross & Iacoboni, 2014; Bohl & van den Bos, 2012). Thus, another important avenue for future research would be to investigate imitation control based on bottom–up and top–down processes and their interactions, rather than considering these processes as mutually exclusive.

Nonetheless, results from the current study remain clear: The basic imitation inhibition mechanism engages the MD network, which has been consistently associated with domain-general processes (Duncan, 2010). Given the mixed findings in prior imitation studies (Table 1) as well as in psychology and neuroscience more generally (Open Science Framework, 2015; Button et al., 2013), future fMRI research may also consider reliability and reproducibility as key concerns in imitation research and consider the possible use of fROI approaches as a means to quantify consistency across individuals.

Sex Differences in Imitation Inhibition

This study is the first to investigate sex differences in the neural mechanisms that inhibit imitation. The behavioural data demonstrated that female individuals show a greater spatial but not imitative compatibility effect than male individuals. This result extends prior behavioural research on sex differences, which did not separate spatial (or orthogonal spatial) from imitative responses in imitation control (Genschow et al., 2017; Butler et al., 2015). The result is also consistent with reports in a wide range of nonsocial inhibitory control tasks, which show similar sex differences (Figure 2; Stoet, 2010, 2017; Clayson et al., 2011; Rubia et al., 2010; Bayliss et al., 2005). All these tasks share a common feature—they require the inhibition of a response to a task-irrelevant spatial feature to enforce a task-relevant response. Taken together, this pattern of results suggests that response inhibition relating to spatial conflict differs between the sexes, rather than a process that is tied to the control of imitation. An alternative possibility is that the difference between the sexes for spatial compatibility is larger than for imitative

compatibility, and we were unable to detect the imitative effect behaviourally. Future research will have to probe these possibilities further.

Given the proposed role of MD and ToM networks in imitation control, we anticipated sex differences in one or both of these networks. The neuroimaging data, however, demonstrated no sex differences in the ToM or MD networks in either experiment. Furthermore, even though regions outside our ROIs mediated the sex difference in Experiment 1, these regions were not consistently engaged differently for male and female individuals in Experiment 2. Thus, based on data across both experiments, our best estimate is that univariate analyses, which assess the magnitude of BOLD response, do not show large effects of sex in MD or ToM neural networks. This being said, there does seem to be a trend for greater engagement in the MD network for female individuals compared with male individuals for both spatial and imitative effects, but this does not survive our statistical thresholding (Figure 7). As a consequence, we are cautious to interpret this null result as we did not have the same level of statistical power to detect sex differences as we did to detect simple compatibility effects. Indeed, it remains a possibility that small univariate effects exist or that the sex difference is underpinned by more complex neural organization. Future studies that use connectivity measures (Sporns, Tononi, & Kötter, 2005) or multivoxel pattern analysis (Kriegeskorte, Mur, & Bandettini, 2008; Norman, Polyn, Detre, & Haxby, 2006) may show increased sensitivity and be better able to capture the complexity of neural organization that we are aiming to measure.

Limitations

The primary limitation of the current work is that we studied a relatively simple model of brain organization based on univariate measures. Given the mixed evidence from prior studies regarding imitation control (Table 1), we felt it was an important step to first establish the extent to which general and specific systems were engaged in a univariate manner. By doing so, we aimed to build an appropriate foundation for future work to build upon. Moreover, as we only identified the MD and ToM networks, it is possible that neural regions outside our key networks may play a role in imitation inhibition or mediate the sex difference in spatial response inhibition or imitation control. Even though our whole-brain analyses showed no consistent effects outside our fROIs, this only shows that there was no univariate engagement of extended brain

regions. We thus acknowledge that we have tested a relatively simple model of brain organization that is likely to underestimate the complexity of neural processes associated with social and cognitive mechanisms such as imitation control. As mentioned before, future work may consider interactions between general and specific systems and more complex, multivariate measures of brain organization.

A second limitation regards the functional localization approach used to identify the ToM and MD networks in Experiment 2. The validity of the fROI approach is based on assumptions about the functional processes that are engaged by the localisers used to identify fROIs. For example, different ToM localisers may engage partly nonoverlapping aspects of the ToM network (Schaafsma, Pfaff, Spunt, & Adolphs, 2015; Spunt & Adolphs, 2014). Therefore, our conclusions about the role of ToM and MD networks are limited to the type of localiser paradigms that we used in the current study. Future research that uses different functional partitions of these networks would be instructive.

A third point to consider is that we looked at the compatibility effect which was computed by response on incompatible trials minus the response on compatible trials. Incompatible and compatible conditions are independent and can have their own variance. The raw ROI data, however, suggests that the MD network showed a positive percent signal change for both compatible and incompatible conditions over baseline (spatial and imitative) with a higher PSC for the incompatible condition. Thus, engagement in the MD network regions was driven by task conditions. For the ToM network, mPFC fROIs for both spatial and imitative compatibility effects showed a negative PSC. It is unclear from this result what the individual response profile looks like for each condition. For instance, the negative PSC can reflect positive activation for both compatible and incompatible trials with a higher PSC for compatible trials compared to incompatible, or it can reflect de-activation in both conditions. Both response profiles are possible and may lead to differing interpretations of the results. The raw data suggests that both conditions showed a de-activation for incompatible and compatible conditions compared to baseline, with a higher de-activation for incompatible conditions (although not significant). The rTPJ showed a similar profile for imitative incompatible and compatible conditions. In contrast, rTPJ showed a positive PSC for both spatial incompatible and compatible conditions compared to baseline, with a higher PSC for the incompatible condition. Thus, the response in the rTPJ for spatial compatibility was also driven by task demands. Therefore, even though we report only the compatibility effects,

we are confident that the raw data support the conclusions we make. We provide all raw data online for other researchers to test alternative hypotheses. A final potential limitation is that the order of tasks in Experiment 2 could have influenced our results. We ordered the tasks such that the ToM localiser was always performed at the end, but the MD task was interspersed between imitations runs in order to offset boredom. We arranged blocks in this manner because we were primarily concerned that asking people to perform a belief reasoning task would introduce a social bias to treat the person (hand image) in an artificially more social/belief reasoning manner during the imitation inhibition task. We did not share the same level of concern that performing a memory task, which we used to localise the domain-general system, would introduce a memory or "cognitive control" bias to the imitation inhibition task. However, we cannot rule out the possibility in the current experiment that the MD task influenced the way the imitation task was performed. This being said, we did get the same results in Experiment 1, when the MD task was not performed before the imitation task. As such, although possible, we find it unlikely that task order had a meaningful impact on our results in Experiment 2.

CHAPTER 3

The Inhibition of Automatic Imitation – A Meta-Analysis and Synthesis of fMRI Studies

This chapter investigates the consistency of the neural networks that underlie the control of automatic imitation using a multi-level kernel density analysis (MKDA). This is the first study to date that quantitatively meta-analyses existing neuroimaging literature investigating neural mechanisms of automatic imitation. We used MKDA in order to provide a combined quantitative estimate of many individual studies, and avoid problems associated with interpreting individual studies.

This chapter is published as it is from:

Darda, K. M. & Ramsey, R. (2019). The inhibition of automatic imitation – a meta-analysis and synthesis of fMRI studies. *Neuroimage*, 170, 320-329.

Abstract

Humans copy other people without their conscious awareness, a behaviour known as automatic imitation. Although automatic imitation forms a key part of daily social interactions, we do not copy other people indiscriminately. Instead, we control imitative tendencies by prioritising some actions and inhibiting others. To date, neuroimaging studies investigating the control of automatic imitation have produced inconsistent findings. Some studies suggest that imitation control relies on a domain-specific neural circuit related to social cognition (the theory-of-mind network). In contrast, other studies show engagement of a domain-general neural circuit that is engaged during a diverse range of cognitive control tasks (the multiple demand network). Given the inconsistency of prior findings, in the current paper we avoided problems associated with interpreting individual studies by performing a meta-analysis. To do so, we used a multi-level kernel density analysis to quantitatively identify consistent patterns of activation across functional magnetic resonance imaging studies investigating the control of imitation. Our results show clear and consistent evidence across studies that the control of automatic imitation is guided by brain regions in the multiple demand network including dorsolateral frontoparietal cortex. In contrast, there was only limited evidence that regions in the theory of mind network were engaged. Indeed, medial prefrontal cortex showed no consistent engagement and right temporoparietal junction engagement may reflect spatial rather than imitative control. As such, the current meta-analysis reinforces the role of domain-general control mechanisms and provides limited evidence in support of the role of domain-specific processes in regulating imitative tendencies. Consequently, neurocognitive models of imitation need updating to place more emphasis on domain-general control mechanisms, as well as to consider more complex organisational structures of control, which may involve contributions from multiple cognitive systems.

Introduction

The involuntary tendency of human beings to imitate others' gestures, speech patterns, and postures, is known as automatic imitation (Heyes, 2011). It has been suggested that such automatic imitative behaviour functions as a "social glue" as it increases pro-social behaviour, positive rapport, feelings of affiliation and liking between interacting partners (Kavanagh & Winkielman, 2016; van Baaren, et al., 2009; Lakin & Chartrand, 2003; Chartrand & van Baaren, 2009; van Baaren, et al., 2003). Given the influence of imitation on strengthening social bonds, researchers have started to investigate the psychological and biological mechanisms that underpin imitation. For example, over the last 20 years, researchers have used functional magnetic resonance imaging (fMRI) in order to better understand the neural underpinnings of the control of automatic imitative tendencies. However, these studies have provided mixed findings regarding the contributions of domain-general or domain-specific neural networks in imitation control. The current paper, therefore, meta-analyses fMRI studies to date on the control of automatic imitation in order to provide a combined quantitative estimate of the extant evidence of many individual studies (Lipsey and Wilson, 2001).

In the last two decades, automatic imitation has been widely studied with an attempt to interconnect different disciplines like cognitive science, social psychology, evolutionary biology, and cognitive neuroscience (Prinz and Meltzoff, 2002; Bargh and Chartrand, 1999; Byrne and Russon, 1998). This convergence across multiple disciplines has allowed for a range of perspectives on imitation to emerge in which theory and empirical data can strengthen each other. In social psychology, automatic imitation has been studied in naturalistic social interactions (Chartrand and Lakin, 2013). Along with functioning as a "social glue," research performed in more naturalist settings suggests that imitation behaviour is also moderated by other variables including, but not limited to, personality variables, self-construal, goal to affiliate or disaffiliate, cultural and social contexts, as well as the similarity, familiarity, and status of the person being imitated (Chartrand and Lakin, 2013; Caspers et al., 2010; Duffy and Chartrand, 2015).

Even though automatic imitation seems to be an important behaviour that facilitates social interactions, we do not always copy others' behaviours. In many situations, imitation can be maladaptive, and it is essential to circumvent the tendency to automatically imitate (Cross and Iacoboni, 2014; Cross et al., 2013; van Schie et al.,

2008; Newman- Norlund et al., 2007). This need to regulate imitative tendencies indicates the existence of a selection mechanism that inhibits unwanted actions, and prioritises alternatives (Brass et al., 2009). Thus, imitation control can be divided into at least two component processes – action representation and action selection. We observe an interaction partner and their actions (representation), and then select the action that needs to be executed (selection).

In contrast to social psychology approaches, researchers in the field of cognitive psychology and neuroscience have generally used computer-based reaction-time (RT) measures of the inhibition of automatic imitation (Brass et al., 2000; Stürmer et al., 2000). One of the most commonly used tasks in this field is a stimulus response compatibility (SRC) paradigm which consists in making finger movements while simultaneously observing a compatible or incompatible finger movement (Brass et al., 2000). For example, participants may be asked to make a finger movement in response to an imperative cue i.e. they are instructed to lift their index finger when they see a number '1' on screen, and their middle finger when they see a number '2.' Simultaneously, participants also observe a task-irrelevant index or middle finger movement, which is compatible or incompatible with their own response. Other variants of this task include hand opening and closing movements instead of finger movements (Press et al., 2005; Wang et al., 2011) or pre-specifying the participant's response before the imperative cue (i.e. participants are asked to always lift their index finger when they see a finger movement; Brass et al., 2001; Heyes et al., 2005). In these variants as well, participants observe a hand or finger movement which is compatible or incompatible with their own response. Irrespective of the task used, greater cognitive resources are required when inhibiting movements incompatible to one's own responses, thus leading to greater RTs (Heyes, 2011; Brass and Heyes, 2005). The difference between the incompatible and compatible conditions (referred to as the general compatibility effect) is said to be a measure of imitation control (Heyes et al., 2005; Heyes, 2011).

To date, a number of neuroimaging studies have investigated the neural mechanisms of imitation control using RT paradigms. However, the evidence demonstrating the extent to which RT paradigms of imitation control engage domain-general or domain-specific neural networks is mixed. Domain-specific processes operate on particular types of stimuli or aspects of cognition, while domain-general processes operate across a range of stimuli and tasks (Barrett, 2012; Spunt and Adolphs, 2017). One

of the prevailing theories of automatic imitation proposes that imitation control relies on a domain-specific neural circuit related to social cognition (Brass et al., 2009). This "specialist" theory has gained traction with evidence from patient and neuroimaging data pointing to the engagement of two key candidate regions – the anterior medial prefrontal cortex (mPFC) and the right temporoparietal junction (rTPJ) (Brass and Heyes, 2005; Brass et al., 2009). For example, mPFC and rTPJ have been engaged in human brain imaging investigations of imitation inhibition (Brass et al., 2001, 2005; 2009; Spengler et al., 2009; Wang et al., 2011). Brass and colleagues further proposed a dissociation of roles for the mPFC and rTPJ during imitation control - the rTPJ distinguishes between self- and other-generated actions, and the mPFC enforces the self-generated action when faced with conflict from an action representation generated by another agent (Brass et al., 2009). In addition, patients with frontal lobe lesions show disrupted imitation inhibition behaviour (Brass et al., 2003; Spengler et al., 2010) and an increased tendency to automatically imitate even when they are clearly instructed to not do so (Lhermitte et al., 1986). More evidence for the involvement of rTPJ comes from neuro-stimulation studies: inhibiting the activity in the rTPJ by transcranial magnetic stimulation (TMS) interfered with imitative responses impairing imitation inhibition (Hogeveen et al., 2014; Sowden & Catmur, 2015). Irrespective of the method used, it is worth noting that, to date, there have only been a small number of studies implicating mPFC and rTPJ in the control of imitation. Moreover, these studies have used relatively small sample sizes between 10 and 25 participants and there have been few, if any, direct replications. Therefore, the sum total of evidence for mPFC and rTPJ engagement during imitation control is suggestive rather than compelling.

Along with imitative control, neuroimaging findings suggest mPFC and rTPJ are also engaged in a variety of socio-cognitive tasks that are associated with theory of mind, including distinguishing between self from other, perspective taking, as well as attributing beliefs, desires and attitudes to others (ToM; Gallagher et al., 2000; Amodio and Frith, 2006; Ruby and Decety, 2001; Aichhorn et al., 2006; Decety et al., 2002; Santiesteban et al., 2012; Brass et al., 2009; Spengler et al., 2010). Based on these findings, self-other control processes have thus been proposed as a candidate mechanism for a range of socio-cognitive functions. For example, it is important to inhibit one's own perspective or mental state and enhance that of the other when empathising with others, taking their perspective, or engaging a successful theory-of-mind (de Guzman et al.,

2016; Sowden and Shah, 2014). Further, atypical self-other control has been linked to disorders characterised by social dysfunction including autism and schizophrenia (Cook and Bird, 2012; Ferri et al., 2012). Overall, this evidence suggests that in imitation control, it is crucial to inhibit the representation of the other's action, and enforce your own, and this mechanism is guided by a domain-specific neural circuit unique to social cognition (Brass et al., 2009).

In contrast to this "specialist" view of imitation control, however, "generalist" theories of imitation suggest that the inhibition of automatic imitation does not differ from any other pre-potent tendencies or general cognitive functions (Heyes, 2011; Cooper et al., 2012). Multiple cognitive control tasks like the Flanker, Stroop, and Simon tasks, which require the inhibition of automatic overlearned response tendencies, have been found to engage a domain-general control network identified in the dorsolateral fronto-parietal cortices (Aron et al., 2014; Bunge et al., 2002; Hazeltine et al., 2007; Nee et al., 2007; Wager et al., 2005). This network is also called the multiple demand (MD) network as it is engaged across a diversity of mental operations (Duncan, 2010). Across studies that investigate imitation inhibition, some have found engagement of the mPFC and rTPJ (Brass et al., 2001, 2005; 2009; Spengler et al., 2009), whereas others show engagement of the MD network (Bien et al., 2009; Crescentini et al., 2011; Cross and Iacoboni, 2013; Mengotti et al., 2012; Marsh et al., 2016). However, most previous fMRI studies have been limited by low statistical power and small sample sizes. More recently, a multi-experiment study using larger sample sizes (N = 28, N = 50) and a functional region of interest (fROI) approach that bolsters statistical power and functional sensitivity has shown that imitation control engages only the MD network, and not mPFC or rTPJ (Darda et al., 2018). Indeed, even with an *a priori* power analysis ensuring 80% power to detect medium effect sizes, Darda et al. (2018) did not even find a directional trend to suggest that the ToM network was directly engaged during imitation control.

As mentioned before, imitation control can be divided into at least two component processes – action representation and action selection. The above review of literature suggests two possible neural mechanisms as being key to action selection during imitation control. On one hand, during imitation control, the neural representation generated by the observed person's action is inhibited, and the self-generated action is selected and enforced and this selection mechanism engages a domain-specific neural

network i.e. the mPFC and rTPJ. On the other hand, the selection mechanism may be guided by a domain-general neural network i.e. the MD network. In both possible mechanisms, the input is the same i.e. the observed person and action may engage domain-specific socio-perceptual neural circuits. However, the difference lies in the selection or control mechanism that underlies the inhibition of automatic imitative tendencies which finally leads to consequent behaviour (see graphical representation in Fig. 1).

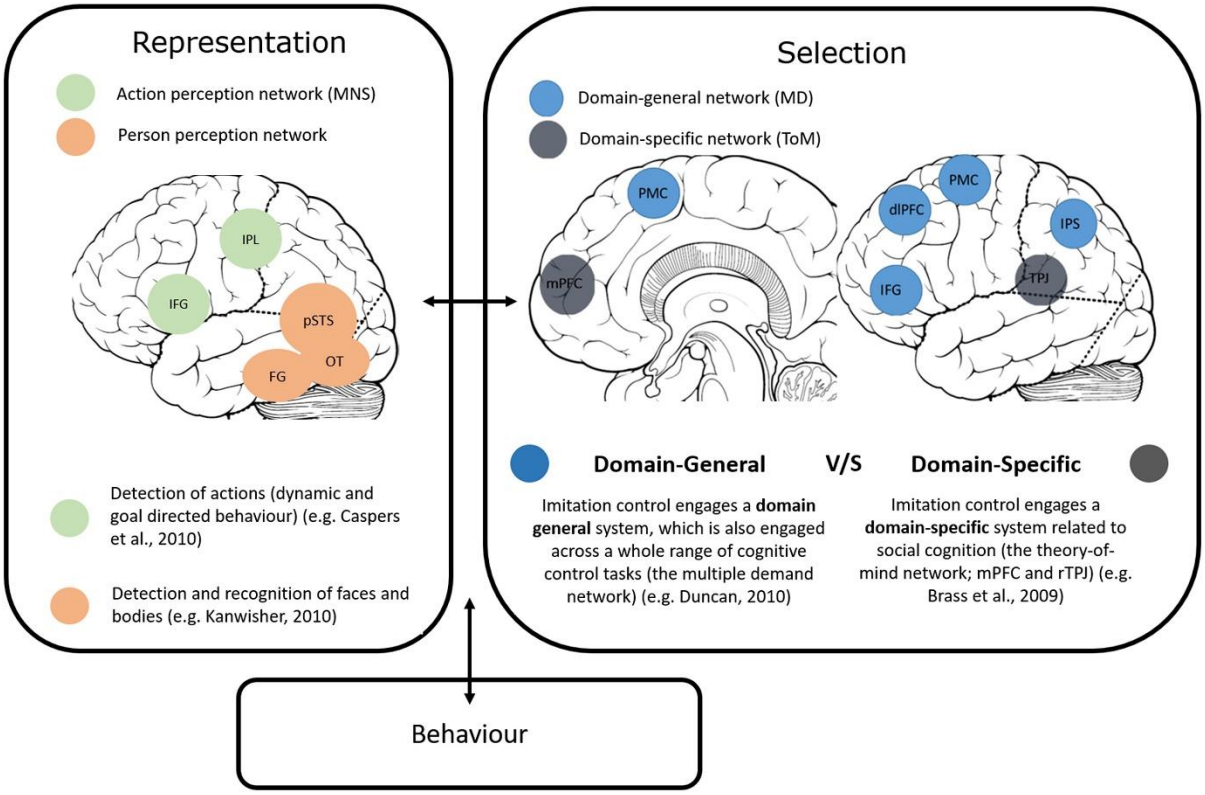


Figure 1. Brain networks associated with the control of automatic imitation**.** This graphical representation divides imitation control into two constituent processes – representation of the person and their action, and the selection (control) of the right action to be executed. In the context of automatic imitation, the representation system consists in face, body, biological motion, and action perception. The neural substrates for person and action perception span the fusiform gyrus, occipitotemporal cortex, and posterior superior temporal sulcus, as well as the mirror neuron system (Kanwisher, 2010; Caspers et al., 2010). The control or selection system consists in a brain network that is either domain-general (i.e. the multiple demand network) or domain-specific (i.e. the theory-of-mind network). N.B. Abbreviations: MNS = mirror neuron system; IPL = inferior parietal lobule, IFG = inferior frontal gyrus; pSTS = posterior superior temporal sulcus; OT = occipito-temporal cortex; FG = fusiform gyrus, MD = multiple demand network; ToM = theory-of-mindnetwork; mPFC = medial prefrontal cortex; PMC = primary motor cortex; dlPFC = dorsolateral prefrontal cortex; TPJ = temporo-parietal junction. The bidirectional arrow "↔" indicates links between the different nodes of imitation control.

The question of interest for the current meta-analysis, therefore, lies at the selection stage of imitation control with the evidence to date for engagement of domain-specific and domain-general neural networks being inconsistent. Even though the most statistically powerful fMRI study to date only shows the engagement of the MD network (Darda et al., 2018), the interpretation of individual studies remains limited in scope for several reasons. First, many single studies are likely to be underpowered leading to missed or spurious results (Button et al., 2013). Second, empirical work involves design choices that strongly influence results, making it harder to generalise effects across analysis pipelines and differing experimental procedures (Carp, 2012). Given the inconsistency of prior findings and the absence of a quantitative synthesis of evidence, taking a meta-analytical approach to further investigate the neural basis of imitation has many benefits (Cumming, 2014). As such, by means of a meta-analysis, the current paper enables the detection of consistent patterns of activation across studies.

In order to quantify the consistency and specificity of regional activation for imitation control across studies, we performed a multi-level kernel density analysis (MKDA; see Methods and Materials for details). We included all fMRI studies (N = 12) investigating imitation control using the RT measure of imitation inhibition (see Table 1). Our primary measure aimed to quantify the consistency of region engagement across studies with particular focus on the engagement of the ToM network and the MD network. The dependent variable was the blood oxygen level dependent (BOLD) response measured in the included fMRI studies. Given the prior mixed findings across studies, this meta-analysis aimed to quantify the extent to which ToM, MD or both neural networks may be engaged during the inhibition of automatic imitation.

We also ran two more exploratory analyses, which were based on a small subset of the total studies. The most common measure of imitation inhibition, the general compatibility effect, also includes a spatial component (Heyes, 2011). In order to measure imitative compatibility more specifically, therefore, imitative and spatial effects need to be dissociated (Gowen et al., 2016; Boyer et al., 2012; Catmur and Heyes, 2011). However, only a few fMRI studies have measured the imitative compatibility effect independent of the spatial component (Darda et al., 2018; Marsh et al., 2016; Cross et al., 2013). This makes it difficult to interpret the roles of the ToM (mPFC and rTPJ) and MD networks in imitation control – their engagement could reflect both social (imitative) and/or non-social (spatial) control. Indeed, the rTPJ has been previously associated with

orienting to both social and non-social stimuli (Corbetta et al., 2008; Thiel et al., 2004). Thus, given that only a few studies have dissociated between imitative (N = 3) and spatial compatibility (N = 4) effects, we also ran two further exploratory MKDAs in order to quantify consistency of patterns across studies for both imitative and spatial compatibility effects. Indeed, given the low number of studies included in the secondary analyses, these results provide only suggestive, and not compelling, evidence regarding the role of the MD and ToM networks in imitative and spatial control.

Table 1.

Data extracted from the studies included in the meta-analysis.

Authors	Year	Sample	Contrasts			Fixed/Random Effects Model	MNI or Talaraich coordinates
			GC	SC	IC		
Brass et al.	2001	10	x			Fixed	Talaraich
Brass et al.	2005	20	x			Fixed	Talaraich
Spengler et al.	2009	20	x			Random	Talaraich
Crescentini et al.	2011	19	x			Random	MNI
Wang et al.	2011	20	x			Random	MNI
Mengotti et al.	2012	22	x	x		Random	MNI
Cross & Iacoboni	2013	24	x	x		Random	MNI
Cross et al.	2013	20	x		x		
Klapper et al.	2014	19	x			Random	MNI
Marsh et al.	2016	24	x	x	x	Random	MNI
Darda et al. (Exp1)	2018	28	x			Random	MNI
Darda et al. (Exp2)	2018	50	x	x	x	Random	MNI
Campbell et al.	2018	24	x			Random	MNI
TOTAL = 12		300					

N.B. GC = General Compatibility, SC = Spatial Compatibility, IC = Imitative Compatibility.

Methods and Materials

Literature search and data collection

In the current paper, we follow recent guidelines put forward for meta-analysing neuroimaging studies (Müller et al., 2018). FMRI studies exploring the inhibition of automatic imitative tendencies were searched for on the online database PubMed, as well as the article search engine Google Scholar. Combinations of keywords including 'imitation inhibition,' 'fMRI,' 'imitation,' 'automatic imitation,' and 'imitation control' were used to identify relevant literature (prior to January 2019). A total of 15 studies were found. We rejected studies if the primary method of investigation was not fMRI, if the study did not report results in stereotactic coordinate space (either Montreal neurological Institute (MNI) or Talaraich coordinates) (N = 1; Bien et al., 2009), if reported results were based on region-of-interest (ROI) analyses, and the study did not report whole-brain analysis coordinates either in the main article or in supplementary materials (or we could not obtain them from the authors) (N = 1; Brass et al., 2009), and if the study involved children or atypical populations (and the coordinates for controls were not reported separately) (N = 1; Spengler et al., 2010).

A wide variety of contrasts are used in studies that investigate the inhibition of automatic imitation. However, in order to minimise heterogeneity, studies that used a paradigm that was not based on or was not conceptually similar to the Brass et al. (2000) paradigm for measuring inhibition of automatic imitation were also excluded. Thus, 12 studies with a total of 300 participants were included in the meta-analysis (see Table 1).

Even though our main analysis was on the general compatibility effect, we also ran two separate meta-analyses for spatial and imitative compatibility. Table 2 summarises the contrasts used in the current meta-analysis for general, spatial, and imitative compatibility effects. A total of 13 contrasts across 12 studies with 142 foci were used for general compatibility, 4 contrasts across 4 studies with 42 foci were used for spatial compatibility, and a total of 3 contrasts across 3 studies with 20 foci were used for imitative compatibility.

Table 2.

Contrasts used in the meta-analysis for general, spatial, and imitative compatibility.

Authors	Year	General Compatibility	Spatial Compatibility	Imitative Compatibility
Brass et al.	2001	General Incompatible > General Compatible		
Brass et al.	2005	General Incompatible > General Compatible		
Spengler et al.	2009	General Incompatible > General Compatible		
Crescentini et al.	2011	General Incompatible > General Compatible		
Wang et al.	2011	General Incompatible > General Compatible*		
Mengotti et al.	2012	Non-specular > Specularˆ	Spatially Incompatible > Spatially Compatible	
Cross & Iacoboni	2013	General Incompatible > General Compatible	Spatially Incompatible > Spatially Compatible	
Cross et al.	2013	General Incompatible > General Compatible		General Compatibility > Spatial Compatibility
Klapper et al.	2014	General Incompatible > General Compatible*		
Marsh et al.	2016	General Incompatible > General Compatible	Spatially Incompatible > Spatially Compatible	Imitatively Incompatible > Imitatively Compatible
Darda et al. (Exp1)	2018	General Incompatible > General Compatible		
Darda et al. (Exp2)	2018	General Incompatible > General Compatible	Spatially Incompatible > Spatially Compatible	Imitatively Incompatible > Imitatively Compatible
Campbell et al.	2018	General Incompatible > General Compatible		
No. of studies		12	4	3
No. of contrasts		13	4	3
No. of foci		142	42	20

Table 2 shows the contrasts used in the current meta-analysis, and the number of contrasts, foci, and studies for each compatibility type (general, spatial, and imitative). *Collapsed across conditions; for Wang et al. (2011): collapsed across direct and averted gaze, for Klapper et al., 2014: collapsed across belief (motion-capture, computer animation) and form (human, non-human). ˆNon-specular > Specular i.e. {(spatially incompatible and imitatively compatible) + (imitatively incompatible and spatially compatible) > general compatible)}].

Data analysis

All analyses in the current paper were performed in MatlabR2015b (Mathworks, Naticks, MA) using the MKDA toolbox developed by Wager et al. (2007); http://wagerlab.colorado.edu). MKDA is an analysis technique that uses a random effects model to assess convergence across studies. This allows for assessing convergence across studies as opposed to between individual foci (as implemented in classical meta-analysis techniques that use fixed effects analyses). Thus, results are not biased by a small number of individual studies. Further, each contrast is weighted by the sample size and study quality (i.e. whether the study used a fixed or random effects model; Wager et al., 2007; Kober and Wager, 2010).

MKDA was performed on all three compatibility types separately. Before performing the analyses, we extracted the following information from each study and included it in our database: authors, year of publication, sample size, task contrasts, fixed or random effects model, and MNI or Talaraich co-ordinates. Co-ordinates reported in Talairach space were converted to MNI stereotactic space using Lancaster transformation (tal2icbm transform; Lancaster et al., 2007). Peak coordinates from each contrast map were then convolved with a 10 mm spherical kernel in order to create a contrast indicator map (CIM). The resulting voxels within 10 mm of the peak were deemed "significant" and given a value of one; other voxels were given a value of zero which indicated no significant effect. A density map was then created by averaging the indicator maps, weighted by sample size, and whether the study used a fixed or random effects model. More specifically, as recommended by Wager and colleagues (Wager et al., 2007), this density map was weighted by the square root of the sample size of the study, and then multiplied by an adjustment factor of 1 for random effects analysis, and 0.75 for a fixed effects analysis.

Each voxel of the density map was given a density statistic P. P stands for the proportion of contrasts included in the analysis that show activity within 10 mm of the peak. A Monte Carlo simulation (with 5000 iterations) was then carried out in order to identify voxels that had a P-statistic that was higher than the frequency predicted by chance. This was tested against the null hypothesis that activated regions in the resulting pairwise contrast maps (from the 5000 iterations) were randomly distributed across the brain. To test for the significance of the cluster size, a similar procedure was used. This allowed for the identification of a threshold for cluster size at which a specific number of

voxels needed to be activated contiguously so that the cluster could be deemed significant.

In order to maximise sensitivity in testing our hypotheses, we report results using two thresholding techniques. One thresholding technique is based on height and the other is based on cluster size. For the weighted P-statistic (height-based threshold), the family wise error (FWE) corrected threshold is the proportion of studies which yielded activity within 10 mm of a voxel that showed a higher P-statistic than the maximum P-statistic across 95% of the Monte Carlo maps. For the cluster size threshold, the FWE corrected threshold is the contiguous voxels observed at two different thresholds ($p < .001$ and $p < .01$) whose cluster size is more than the extent of clusters found across 95% of the Monte Carlo maps. We use two cluster-based thresholds in order to also detect regions that show a lower response in magnitude over a larger cluster size both at more stringent ($p < .001$) and less stringent ($p < .01$) thresholds. Voxels that exceed the height-based threshold in our analysis appear on the resulting maps in Fig. 2 in yellow, and those that exceed the cluster extent-based threshold appear in orange ($p < .001$) and red ($p < .01$).

In Table 3, peak activation foci that pass the height-based threshold are reported. If activations do not pass the height-based threshold, foci of the cluster-extent-based thresholding are reported. The number of voxels in each cluster that survived height-based and/or extent-based thresholding is also reported. Resulting coordinates were localised using the SPM Anatomy Toolbox (Eickhoff et al., 2005). The database of co-ordinates, and code used to perform the meta-analysis are available online (https://osf.io/dbuwr/).

Results

For the general compatibility effect, across 13 contrasts from 12 studies, consistent activation was found in right inferior parietal lobule, right supramarginal gyrus, right superior temporal gyrus, and right temporo-parietal junction (see Table 3; Fig. 2A). These clusters survived both height-based and the more stringent extent-based thresholding ($p < .001$). Activation was also found in right superior frontal gyrus, and right middle frontal gyrus, which survived both height and the less stringent extent-based thresholding ($p < .01$). Activation in the left and right insula survived the more stringent extent-based threshold ($p < .001$), but not the height-based threshold.

Activation in the right IFG survived the less stringent extent-based threshold ($p < .01$) but not the height-based threshold.

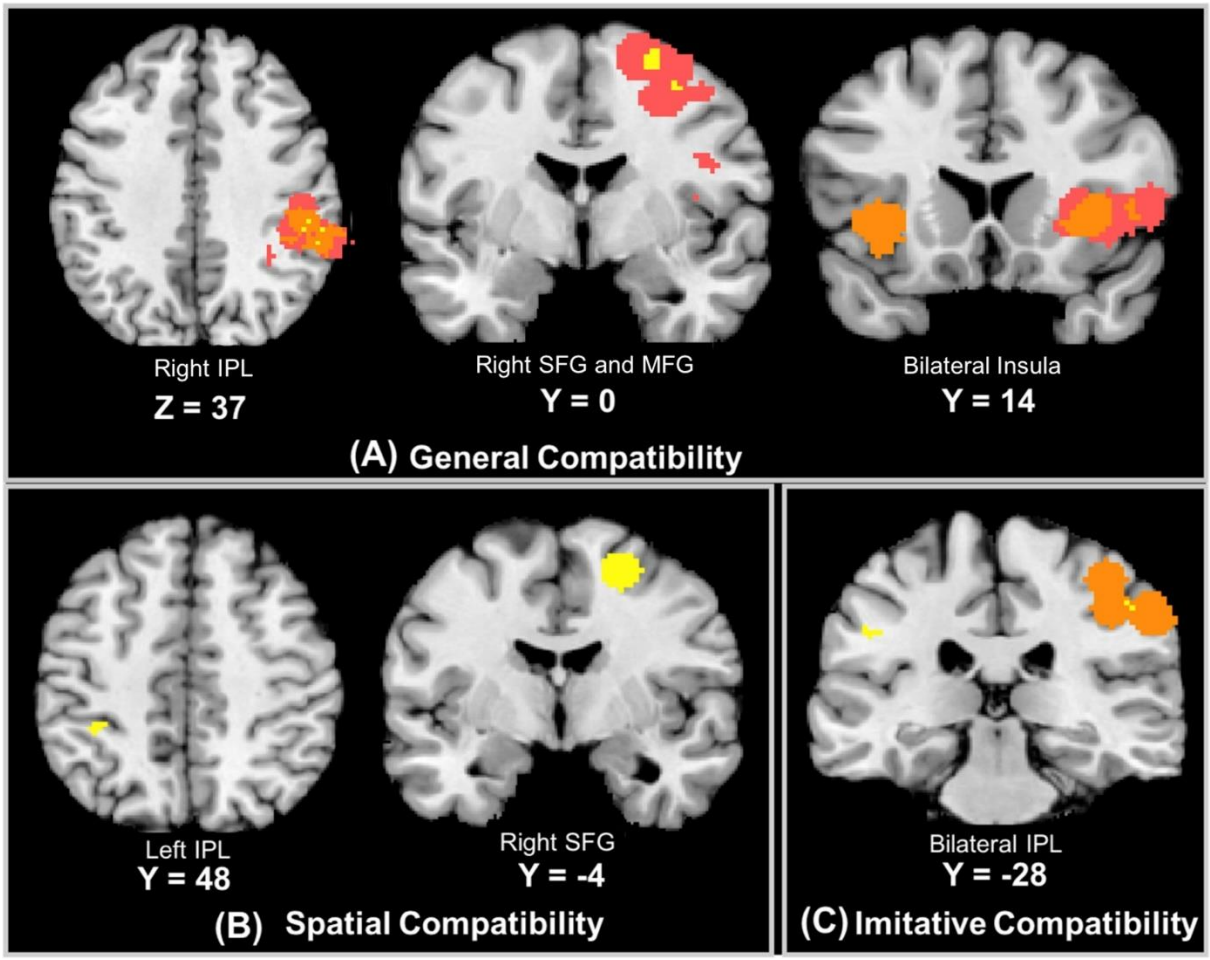


Figure 2. Consistency of brain activation from the MKDA Analyses. Brain areas that are consistently engaged for general compatibility (A), spatial compatibility (B), and imitative compatibility (C). Voxels that exceed the height-based threshold ($p < .05$, FDR corrected) in our analysis appear in yellow, and those that exceed the cluster extent-based threshold appear in orange ($p < .001$) and red ($p < .01$).

We ran two further MKDAs separately for spatial and imitative compatibility. For spatial compatibility, across 4 contrasts from 4 studies, we found consistent activation that withstood the height-based thresholding in the left IPL and the right SFG (see Table 3; Fig. 2B). No regions withstood cluster-based thresholding. For imitative compatibility, across 3 contrasts from 3 studies, we found consistent activation in the left IPL that survived the height-based threshold (see Table 3; Fig. 2C). Activation was also found in the right IPL, which withstood height-based as well as the less stringent extent-based thresholding ($p < .01$).

Table 3.

Areas consistently activated for general compatibility, spatial compatibility, and imitative compatibility.

GENERAL COMPATIBILITY						
Region	**MNI**			**Maximum P**	**No. of Voxels**	**Threshold**
	x	**y**	**z**			
Right TPJ	60	−46	22	0.42	1	^**
Right TPJ	56	−46	32	0.41	1	^**
Right supramarginal gyrus	56	−36	36	0.44	9	^**
Right IPL	60	−34	34	0.43	2	^**
Right IPL	52	−30	38	0.37	5	^**
Right MFG	34	0	54	0.45	15	^*
Right SFG	26	−2	64	0.37	46	^*
Left Insula	−36	14	0	0.32	453	**
	−34	12	−2		213	
	−36	18	2		240	
Right Insula	38	16	4	0.36	405	**
	34	18	0		172	
	46	12	2		75	
	38	16	6		158	
Right IFG	46	14	10	0.28	1269	*
	44	16	−4		74	
	28	24	−4		44	
	32	12	0		41	
	46	22	2		113	
	56	10	2		116	
	28	20	6		105	
	36	26	6		66	
	56	16	6		128	
	52	12	14		219	
	48	2	22		92	
	40	8	22		132	
	50	8	28		139	
SPATIAL COMPATIBILITY						
Left IPL	−36	−40	48	0.78	8	^
Right superior frontal gyrus	24	−4	58	0.78	190	^
	24	−6	54		56	^
	24	−4	60		134	^
IMITATIVE COMPATIBILITY						
Left supramarginal gyrus/IPL	−48	−28	34	0.73	11	^
Right supramarginal gyrus/IPL	48	−26	44	0.73	18	*^
	52	−30	42		7	
	46	−26	46		11	

N.B. Table 3 shows areas consistently activated for general compatibility, spatial compatibility, and imitative compatibility. Maximum P stands for the maximum proportion of studies exhibiting the effect at the peak density weighted by sample size. MNI = Montreal Neurological Institute (MNI) standard stereotaxic space coordinates. The voxel size is 2 × 2 × 2 mm^3. *Clusters withstanding $p < .01$ cluster extent-based threshold. **Clusters withstanding $p < .001$ cluster extent-based threshold. ^Clusters withstanding the height-based threshold.

These density maps showing regions that withstood both height and/or cluster-extent thresholding for each compatibility type were then overlaid with the ToM and MD network masks separately. The ToM network mask consisted of four parcels including the dorsal, medial, and ventral medial prefrontal cortex (DMPFC, MMPFC, VMPFC), and the right temporo-parietal junction (rTPJ), which have previously been implicated in mentalising or theory-of-mind (Dufour et al., 2013). For the MD network mask, 16 parcels were used which included areas in bilateral superior and inferior parietal lobules (SPL, IPL), intraparietal sulcus (IPS), inferior and middle frontal gyrus (IFG, MFG), precentral gyrus (PrecG), insula (Ins), and the supplementary motor area (SMA) (available at: https://evlab.mit.edu/funcloc/download-parcels). Overlay of the density maps with the ToM and MD network masks allowed for identification of overlap between regions that were consistently activated in the MKDA and the ToM and MD networks (Fig. 3). For all compatibility types (general, imitative and spatial), all regions that passed height or extent-based thresholding overlapped with regions in the MD network (Fig. 3A). Additionally, one cluster, which showed consistent activation for general compatibility, also overlapped with the right TPJ in the ToM network (Fig. 3B). There was no overlap with the mPFC node of the ToM network for any compatibility type.

In order to break down the role of the right TPJ in general compatibility, we performed a further, more exploratory analysis. We compared peak coordinates from prior studies with a right TPJ mask, which has been previously implicated in theory-of-mind (Dufour et al., 2013). To do so, the ToM network mask for rTPJ was overlaid with the contrast indicator maps of all studies used for general (N = 12), imitative (N = 3) and spatial (N = 4) compatibility. The contrast indicator maps include 10 mm spherical kernels around peak coordinates of each contrast. This allows coordinates from prior general, imitative and spatial compatibility contrasts to be displayed without any thresholding restrictions and overlaid with the rTPJ node of the ToM network. Fig. 4 shows overlap between contrast indicator maps for general compatibility and spatial compatibility with the right TPJ node of the ToM network mask. By contrast, there is no overlap between contrast indicator maps for imitative compatibility and the same right TPJ mask.

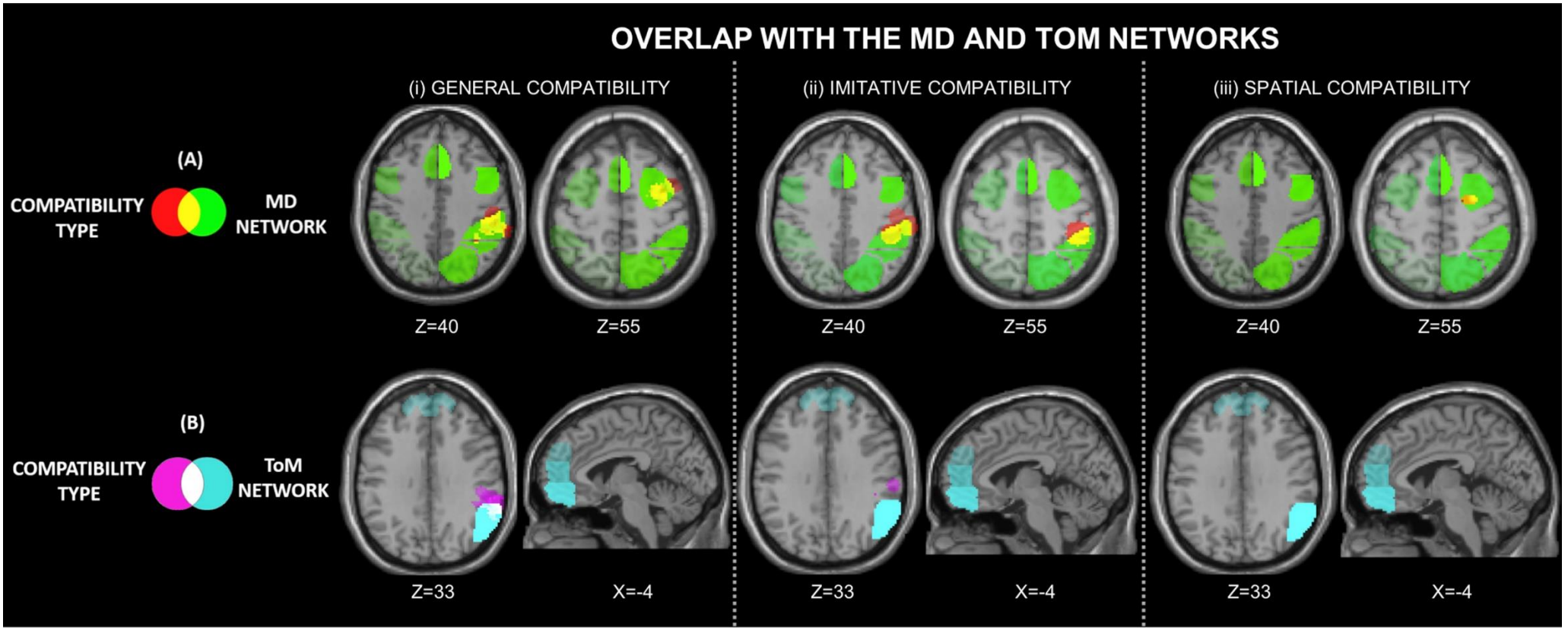


Figure 3. Overlay of the MKDA maps with the ToM and MD network masks**.** Overlay of the density maps with the ToM and MD network masks allowed for identification of overlap between regions that were consistently activated in the MKDA and the ToM and MD networks. For all compatibility types (general, imitative and spatial), all regions that passed height or extent-based thresholding overlapped with regions in the MD network (A). Additionally, one cluster, which showed consistent activation for general compatibility, also overlapped with the right TPJ in the ToM network (B). There was no overlap with the mPFC node of the ToM network for any compatibility type.

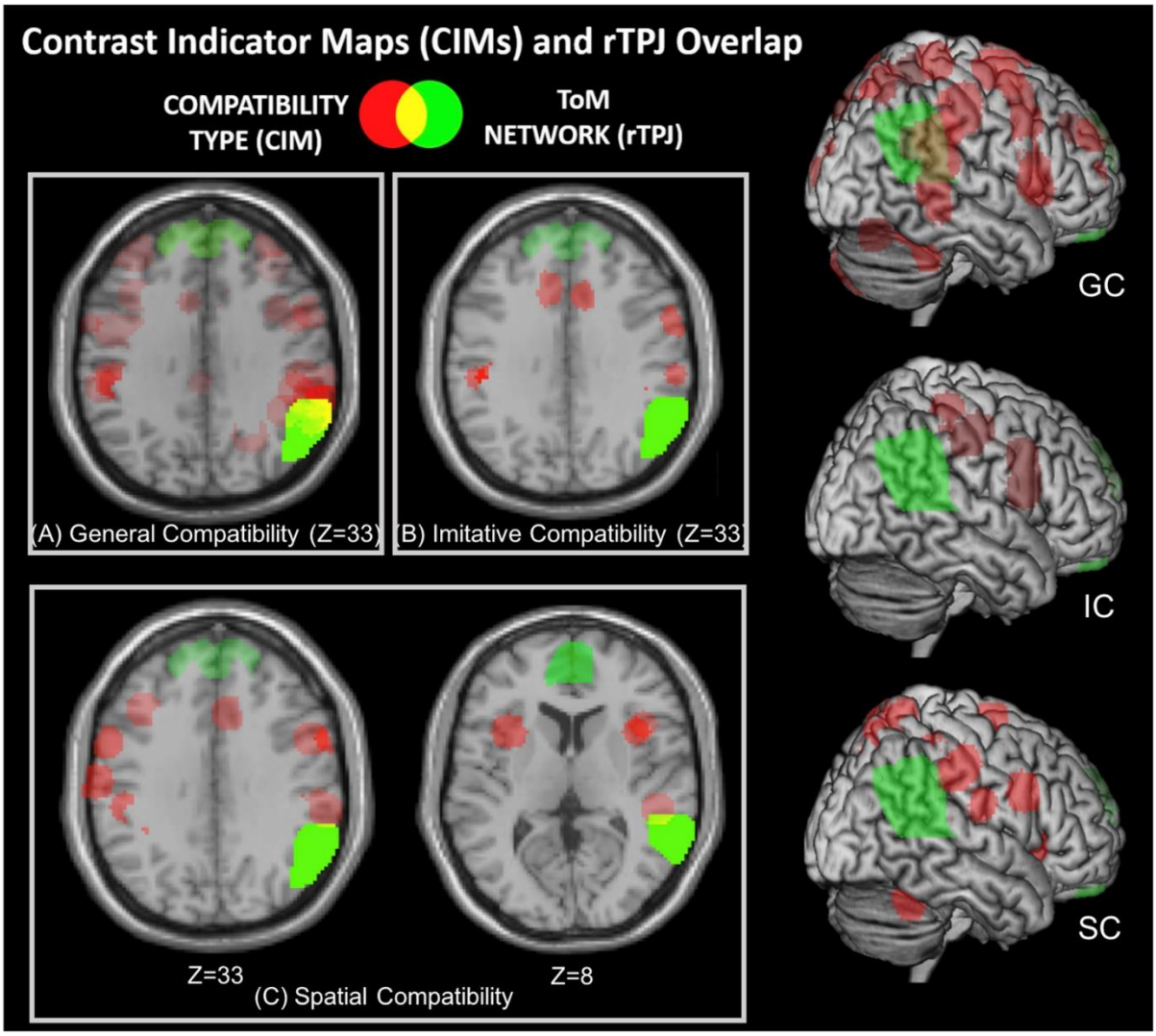


Figure 4. Overlay of Contrast Indicator Maps with rTPJ. The ToM network mask for rTPJ overlaid with the contrast indicator maps of all studies used for general (N = 12; A) imitative (N = 3; B) and spatial (N = 4; C) compatibility. There was overlap between contrast indicator maps for general compatibility and spatial compatibility with right TPJ. There was no overlap between contrast indicator maps for imitative compatibility and the same right TPJ mask. Abbreviations: IC = Imitative Compatibility; SC = Spatial Compatibility, GC = General Compatibility.

Discussion

In the current paper, we performed a meta-analysis of fMRI studies in order to quantify the consistency and specificity of regional activation during the inhibition of automatic imitation. Our results supported a "generalist" view of imitation control – we found clear engagement of dorsolateral frontoparietal cortices when observing an action that conflicted with a current motor intention. These regions overlapped with regions associated with the MD network. We found less evidence for a "specialist" view of imitation control, which relies on the ToM network. Indeed, there was no engagement of mPFC across studies and there was no clear evidence regarding the engagement of rTPJ; there was only suggestive evidence that it may reflect spatial rather than social control. Thus, our results provide unambiguous support for the engagement of a domain-general neural network during the control of imitation, and only limited evidence for the engagement of a domain-specific neural network that is tied to social cognition.

Studies investigating the neural correlates of imitation control have to date shown mixed evidence for the engagement of domain-general and domain-specific neural networks in imitation inhibition. While some studies have found engagement of the mPFC and rTPJ (Brass et al., 2001, 2005; 2009; Spengler et al., 2009), others show engagement of the MD network (Bien et al., 2009; Marsh et al., 2016; Darda et al., 2018). The current MKDA demonstrated that brain regions in the multiple demand network are reliably and consistently engaged across studies that investigate imitation inhibition using the general compatibility effect. Brain regions in the MD network are also engaged for imitative (bilateral IPL) and spatial compatibility effects (left IPL, right SFG). Thus, our findings suggest that brain regions that are engaged across a range of cognitive control tasks are also reliably engaged when controlling the automatic tendency to imitate others, as measured by general and imitative compatibility effects.

Evidence supporting the engagement of a domain-specific neural circuit that is central to social cognition and includes mPFC and rTPJ was less consistent in the current meta-analysis. Brass et al. (2009) proposed that the rTPJ was involved in distinguishing between self- and other-generated actions, whereas the mPFC was engaged when enforcing the correct action. However, the current MKDA did not find any evidence of anterior mPFC engagement for either general, spatial, or imitative compatibility effects. An absence of mPFC engagement for imitation control across studies is thus inconsistent

with the hypothesis that a specific neural system related to social cognition is also engaged during the inhibition of automatic imitation (Brass et al., 2009).

In contrast to the results reported in mPFC, across 12 studies investigating imitation inhibition as measured by the general compatibility effect, the current meta-analysis found engagement of rTPJ. However, it is difficult to interpret the role of rTPJ in imitation control for at least two reasons. First, the general compatibility effect is a product of both spatial and imitative effects, which makes it hard to interpret in a straightforward manner. Second, rTPJ is involved in both social and non-social processes, which makes it a functionally heterogenous region (Corbetta et al., 2008; Krall et al., 2015, 2016; Lee and McCarthy, 2014; Schuwerk et al., 2017).

To further consider our findings, it is important to distinguish between a synthesis of evidence based on a descriptive approach, and a quantitative meta-analysis (Gigerenzer, 2018). To date, 14 fMRI studies have investigated imitation control by measuring the general compatibility effect in typical populations (we excluded Bien et al., 2009 and Brass et al., 2009 in the meta-analysis, see Methods). Out of the 14 studies, only 4 studies report the engagement of rTPJ for the general compatibility effect (see Darda et al., 2018, Table 1 for more details). Thus, we find that only 28.6% of fMRI studies on imitation control to date (4/14) show any evidence in support of a role of rTPJ in imitation control, and these studies do not dissociate between spatial and imitative effects. Both the descriptive and quantitative approaches, therefore, produce similar findings.

It is also informative to compare our results with the largest and most sensitive fMRI study of imitation inhibition to date. Darda et al. (2018) showed no engagement of rTPJ for the imitative compatibility effect, but engagement of rTPJ for general and spatial compatibility effects. Similarly, in the current meta-analysis, when we explored the unthresholded spatial and imitative compatibility effect maps separately, there was partial overlap between the spatial compatibility effect and rTPJ, but no overlap between the imitative compatibility effect and rTPJ. Given that the small number of fMRI studies investigating spatial and imitative compatibility effects separately (N = 4 and N = 3, respectively), the current findings need to be interpreted with caution. However, when taken together with prior findings, the results provide consistently limited evidence for the univariate engagement of rTPJ in the control of imitative tendencies. In contrast, current fMRI findings provide more evidence that rTPJ is involved in resolving spatial

conflict, which is in keeping with patient work (Vallar and Perani, 1987; Vallar, 1993), as well as evidence using spatial cueing tasks like the Posner paradigm (Posner and Cohen, 1984; Thiel et al., 2004; Corbetta et al., 2008). More recent work also suggests that rTPJ may play a more domain-general role in the process of contextual updating, acting on changing expectations after unexpected events (Geng and Vossel, 2013; Mengotti et al., 2017). Assuming that on incompatible trials expectations are violated, rTPJ may play a more generalised role of context updating in imitation and spatial control. However, irrespective of whether it plays a domain-specific or domain-general role, in the current meta-analysis, we find limited evidence for the univariate engagement of rTPJ in the control of automatic imitative tendencies.

Limitations and alternative interpretations

Before moving on to the wider theoretical implications of these results, we first acknowledge possible limitations to the current meta-analytical approach. The current meta-analysis did not include work by Brass et al. (2009), which implicated rTPJ and mPFC in imitation control, due to the whole-brain data being unavailable. Nonetheless, as mentioned before, only 28.6% (4/14) of fMRI studies, which have investigated imitation control, found engagement of mPFC and rTPJ, and they all had small sample sizes (between 10 and 20 participants). It is, therefore, unlikely that the inclusion of an additional study with a relatively small sample size (Brass et al., 2009) would change the results of the meta-analysis, given that they are weighted by sample size.

A further consideration is the relative size of the MD and ToM networks that we used in our analyses. Given that the MD network spans a much larger area than the ToM network, our analysis may be biased toward finding results in the MD network over the ToM network. Although this is true in a relative sense, we do not feel that it hinders our interpretation of the results in the ToM network for several reasons. First, regions of interest in the ToM network were not particularly small areas. The mPFC regions included several portions of the dorsal, middle, and ventral mPFC, and the rTPJ covered a relatively large area of cortex. Second, both networks were defined accurately based on prior work, which used large samples of participants. Thus, even though ToM areas were comparatively smaller than the MD network, they still covered a swath of cortex in regions functionally and precisely defined as the ToM network. Consequently, we feel confident that had these regions been consistently engaged across studies, we would have been able to detect them. Third, even if we only use CIMs across the whole brain,

which report activation peaks from prior studies, thus avoiding issues with thresholding or choice of masks, we still do not find evidence for engagement near rTPJ and mPFC for the imitative compatibility effect (Fig. 4).

An additional possibility to consider is that the difference between the results in terms of domain-specific and domain-general network engagement could be due to the differences in stimuli used in the studies included in the meta-analysis. However, the tasks are all conceptually, visually and cognitively similar to each other with only minor differences across all studies. For example, in Darda et al. (2018; Exp1 and Exp2), the stimuli consist of index and middle finger movements, whereas in Wang et al. (2011), hand opening and closing movements are used. Moreover, a recent meta-analysis also showed that behavioural performance is consistent across a range of studies that cover a range of minor methodological differences (Cracco et al., 2018). Given the lack of substantial differences between the studies and the consistent pattern of behavioural data, it seems unlikely that small differences could be responsible for these effects.

Finally, we acknowledge that fMRI is only one form of measurement, and it is important to consider how these findings mesh with results from other neuroscience techniques. For instance, neurostimulation studies have implicated rTPJ in imitation control (Santiesteban et al., 2015; Bardi et al., 2017). Using repetitive transcranial magnetic stimulation (TMS), dampening of activity in the rTPJ interfered with imitative, but not spatial responses (Hogeveen et al., 2014; Sowden & Catmur, 2015), whereas excitatory stimulation of the rTPJ by anodal transcranial direct current stimulation (tDCS) caused increased performance on the imitation task (Santiesteban et al., 2012). Further, in patients with lesions in the temporoparietal junction area, imitation inhibition deficits have been found to correlate with deficits in visual and cognitive perspective taking tasks, further supporting the role of rTPJ in imitation control (Spengler et al., 2010). Thus, there seems to be a discrepancy between neurostimulation and patient studies, and results from the current meta-analysis of fMRI studies. The evidence from neurostimulation and patient studies for the engagement of rTPJ in imitation control is, however, limited to a few studies with small sample sizes. Under any yardstick, therefore, the sum total of evidence from neurostimulation and patient studies can only be judged to be suggestive at present. It is based on a few studies with small sample sizes that lack formal power analyses and replications. Therefore, for more confirmatory evidence, future investigations with pre-registered and adequately

powered replications are essential (Munafo et al., 2017; Zwaan et al., 2017; Nelson et al., 2018). In addition, it is also possible that the role of rTPJ in imitation control cannot be captured by univariate measurements and a more complex neural organisation is at play during imitation control.

Theoretical implications

The lack of consistent activation in mPFC in the current meta-analysis and a difficulty in interpreting the role of rTPJ have implications for "specialist" theories of imitation. "Specialist" theories suggest that based on a dedicated neural circuit for social cognition, self-other control is crucial for the regulation of imitation, empathy, autism, and theory-of-mind (Brass et al., 2009; de Guzman et al., 2016; Sowden and Shah, 2014). However, more recent behavioural evidence suggests that imitation may not vary as a function of autistic-like traits or empathy, thus questioning the reliance of imitation inhibition on a distinctly social mechanism (Butler et al., 2015; Cracco et al., 2018; Genschow et al., 2017). Instead, imitation control may involve domain-general cognitive control mechanisms, which are also engaged during the control of other non-social pre-potent response tendencies (Heyes, 2011; Cooper et al., 2012). Indeed, the dual-route model of automatic imitation proposed by Heyes (2011) can explain the control of automatic imitative tendencies without assuming a reliance on a self-other distinction. The model suggests that like other stimulus-response compatibility tasks, imitation control is mediated by long-term stimulus-response associations which are a product of learning. In line with this, the computational model put forth by Cooper et al. (2012) further substantiates this notion by demonstrating that spatial and imitative compatibility effects depend on similar cognitive processes, and any behavioural differences are accounted for by different sets of input nodes for spatial and imitative effects in a general dual-route framework (but see Bertenthal and Scheutz (2013) for a critique of this model).

Even though it is possible that imitation and spatial compatibility rely on a partly shared set of cognitive processes, this does not address the question of whether these processes also rely on similar or distinct neurobiological mechanisms. The current meta-analysis suggests that the selection mechanism in imitation inhibition is guided by a domain-general multiple demand system, which is also engaged during the inhibition of other non-social external influences. However, a lack of engagement of mPFC (and possibly rTPJ) in imitation control does not imply that they do not also play a regulatory

role in imitation control. For example, mPFC has been demonstrated to exert a top-down influence during modulation of imitation via direct gaze (Wang et al., 2011). In addition, rTPJ showed a higher response when an interaction partner was believed to be human and looked human compared to when these animacy cues were absent (Klapper et al., 2014). These findings suggest that mPFC and rTPJ may play a regulatory role in imitation control and may be functionally connected to other networks without being directly engaged (Burnett & Blakemore, 2009). The current findings suggest that future work should postulate and test more complex models of imitation control, which extend beyond the operations of the theory of mind network.

In a similar manner, other socio-perceptual circuits, which extend beyond the MD network, may also be involved when inhibiting automatic imitative tendencies. In this regard, it is important to note the distinction between input- and mechanism-specificity. Of course, the input in the imitation inhibition task can be readily identified as emanating from a social entity i.e. a human hand. Thus, the observed input is clearly social in the sense that the observed agent offers opportunity for social interaction. Although the perceptual input is social, a domain-general selection mechanism may still operate in imitation control. Indeed, it is possible that the same selection mechanism operates across both social and non-social contexts. In the context of imitation, therefore, domain-specific action observation and person perception networks may functionally interact with domain-general control mechanisms in the MD network (see Fig. 1). Thus, similar to other domains of social information processing, an interplay between domain-general and domain-specific networks may result in the control of automatic imitative tendencies (Baldauf & Desimone, 2014; Spunt and Adolphs, 2017; Zaki et al., 2010). Thus, the engagement of domain specific and domain general neural networks in imitation control may be more complicated that what current models of imitation suggest. Consequently, theories that move beyond a neat division and posit links between domain-general and domain-specific systems in imitation control need to be given greater emphasis in future work (Barrett, 2012; Spunt and Adolphs, 2017; Michael & D'Ausilio, 2015; Binney and Ramsey, 2019).

In conclusion, the current meta-analysis provides evidence that the selection mechanism when inhibiting automatic imitative tendencies is guided by the regions of the domain-general multiple demand network rather than a domain-specific system related to social cognition. Our meta-analysis questions the role of mPFC and right TPJ in

imitation control and suggests that current neurocognitive models of imitation control need further revision in order to account for the more complex nature of functional interplay between domain-general and domain-specific systems.

Data and code availability statement

Database of co-ordinates and code used to perform the meta-analysis are available online at: https://osf.io/dbuwr/.

4. INDIVIDUAL DIFFERENCES

CHAPTER 4

Individual Differences in Social and Non-Social Cognitive Control

This chapter investigates individual differences in automatic imitation using multiple large-sample behavioural studies (N>600 across 3 experiments). This study integrates experimental and differential psychology approaches in order to investigate the extent to which cognitive systems related to social (imitative) and non-social control differ between individuals, and whether these differences rely on domain-general or specialised control mechanisms.

This chapter is published as it is from:

Darda, K. M., Butler, E. E., & Ramsey, R. (2019, May 23). Individual differences in social and non-social cognitive control. Preprint: https://psyarxiv.com/8b4cz/.

N.B. Emily Butler is thanked for input on design and data collection for Experiments 1 and 2.

Acknowledgements

The authors thank masters and undergraduate students for assistance with data collection on all three experiments.

Abstract

Cognitive control refers to the ability of human beings to adapt flexibly and quickly to continuously changing environments. Several decades of research have identified a diverse range of mental processes that are associated with cognitive control but the extent to which shared systems underlie cognitive control in social and non-social contexts, as well as how these systems may vary across individuals, remains largely unexplored. By integrating methodological approaches from experimental and differential psychology, the current study is able to shine new light on the relationships between stable features of individuals, such as personality and sex, and the architecture of cognitive control systems. Across three large-sample experiments (>600 participants in total), we demonstrate that cognitive control systems are largely invariant to stable aspects of personality, but exhibit a sex difference, such that females show greater task-interference than males. Moreover, we further qualified this sex difference in two ways. First, we showed that the sex difference was unrelated to the sex of the interaction partner and therefore did not reflect an in-group bias based on sex. Second, we showed that the sex difference was tied to a form of spatial interference control rather than social (imitative) control and therefore it does not reflect a specialised mechanism for guiding social interactions exclusively. Instead, our findings suggest that a robust sex difference exists in the system (or set of subsystems) that operate in resolving a form of spatial interference control, and that such systems are unaffected by social factors such as the sex of the interaction partner. The results highlight the value of integrating approaches from experimental and differential psychology by providing a deeper understanding of the structure of cognitive control systems, whilst also providing new dimensions to incorporate into theories and models of social and non-social control.

General Introduction

A remarkable feature of the human cognitive system is its ability to quickly and flexibly adapt behaviour to guide interactions with people and objects in the environment. The mental processes behind such adaptability are collectively referred to as cognitive control and have been the focus of growing research in cognitive psychology (Botvinick et al., 2001; Inzlicht, Bartholow, & Hirsch, 2015). Several decades of research have identified a diverse range of mental processes that are associated with cognitive control, as well as the psychological, neural, and computational mechanisms that support it (Banich, 2009; Botvinick & Cohen, 2014; O'Reilly et al., 2010). However, the extent to which shared systems underlie cognitive control in social and non-social contexts, as well as how these systems may vary across individuals, remains largely unexplored.

A typical assumption in cognitive psychology is that all individuals rely on a common set of cognitive mechanisms, despite obvious individual differences (de Schotten & Shallice, 2017). Therefore, understanding how general cognitive mechanisms operate across all individuals, as well as how different individuals vary from these general patterns, is essential for understanding the structure of social and non-social cognition (Fischer-Baum et al., 2018). Thus, in the current paper, across three experiments, we integrate methodological approaches from experimental and differential psychology to investigate the extent to which cognitive systems relating to social (imitative) and non-social control differ between individuals, and whether such individual differences rely on domain-general or specialised control mechanisms.

Cognitive control is multi-faceted, with a core function being the ability to inhibit unwanted but dominant responses, in order to prioritise alternative, more context-appropriate responses (Chaiken & Trope, 1999; Miyake et al., 2000; Payne, 2005). For example, non-social cognitive control may involve inhibiting automatic reading responses in a Stroop task (MacLeod, 1991), whereas social cognitive control may involve controlling automatic social biases based on race, sex or other social groupings (Amodio et al., 2004). The study of cognitive processes during social interactions has received much attention in the last couple of decades across a range of methodologies (Adolphs, 2009; Frith, 2008; Frith & Frith, 2012; Ochsner & Lieberman, 2001). For example, researchers in the domain of social cognition have used methodologies from cognitive psychology to provide experimental control over phenomena of interest to social

psychologists (Lambert & Scherer, 2013). One such example is that of methodologies used to study automatic imitation.

Humans imitate a wide range of behaviours from their interaction partners, including speech patterns, body postures, gestures and facial expressions (Bernieri, 1988; Brass et al., 2000; Dimberg, 1982; Hansen et al., 2016; Webb, 1972). This behaviour is usually not intended, often occurs without the conscious awareness of the imitator, and is termed as automatic imitation (Heyes, 2009; 2011). Although unintended, automatic imitation has been argued to play a central role in nonverbal social behaviour. Indeed, research has suggested that automatic imitation strengthens human bonds by functioning as a "social glue," powers cognitive and social development, enhances emotional reciprocity, and increases feelings of affiliation, positive rapport and pro-social behaviour (Cacioppo et al., 2000; Chartrand & Bargh, 1999; Kavanagh & Winkielman, 2016; Lakin & Chartrand, 2003; van Baaren et al., 2003; 2009).

Given the influential and wide-ranging role that automatic imitation plays in our social world, researchers from different disciplines have assessed automatic imitative behaviours as well as the antecedents to automatic imitation in order to better understand the role of imitation in social cognition. For instance, in social psychology, studies of automatic imitation (known as motor mimicry) have typically involved measuring overt copying behaviours during live social interactions (Chartrand & Bargh, 1999; Ray & Heyes, 2011). In one such study, Chartrand & Bargh (1999) asked participants to interact with a confederate and perform a card sorting task. During the task, the confederate either waggled their foot or touched their face. Behaviours of the participant were recorded, and it was found that participants noticeably copied the confederate i.e. they touched their face more than waggled their foot when the confederate also touched their face, and vice versa.

By contrast, in cognitive psychology, stimulus-response compatibility (SRC) paradigms have been used in order to measure the automatic tendency to imitate (Brass et al., 2000; Kilner et al., 2003; Stürmer et al., 2000). SRC paradigms have been commonly used in social cognition research to investigate processes associated with the perception of eye gaze (Schillbach et al., 2011), as well as joint attention (Sebanz & Knoblich, 2009). In a typical SRC task measuring automatic imitation, participants are required to lift their index or middle finger in response to a number cue ('1' for index finger, '2' for middle finger). Simultaneously, they also see either the same finger movement (compatible

condition) or a different finger movement (incompatible condition). Participants respond slower in the incompatible condition as the observed movement interferes with their response. This difference in reaction time between compatible and incompatible conditions is referred to as the compatibility effect and is considered to be a measure of the control of automatic imitation. That is, it has been argued to index the cognitive resources required to inhibit the automatic tendency to copy an observed (incorrect) action and instead prioritise the alternative (correct) action (Brass & Heyes, 2005).

After establishing the basic SRC paradigm in order to measure a form of social (imitative) control, subsequent research on automatic imitation in cognitive psychology has provided insight into the factors that influence automatic imitation (Heyes, 2011; Cracco et al., 2018). To do so, these studies have typically used an experimental method, which measures the average influence of a manipulation across a group of participants, rather than a differential approach that measures differences across individuals. For example, previous research has found that factors such as eye gaze and facial expressions of the interacting partner modulate the tendency to automatically imitate (Butler et al., 2016; Crescentini et al., 2011; Grerucci et al., 2013; Rauchbauer et al., 2015; Wang et al., 2011; Wang & Hamilton, 2014). These findings suggest that social and contextual factors serve as antecedents to automatic imitative behaviours.

A focus on experimental over differential approaches is consistent with research practice in psychology more generally, which over the last 100 years has seen a division between two historic streams of method or thought (Cronbach, 1957). One stream – experimental psychology - manipulates conditions in order to observe effects on average across a group of participants. The other stream – differential psychology - is focussed on variation across individuals. Although a confluence of both these methods has been suggested for a unification and progress of psychological science as a whole (Cronbach, 1975; Eysenck & Eysenck, 1985; Eysenck, 1997), these two streams of thought have remained largely autonomous (Cramer et al., 2010; Cronbach, 1957). Indeed, by focussing on the experimental method, the contribution of individual differences tends to be neglected (Eysenck, 1997). For example, in the context of social information processing, a recent study found that tasks measuring mental state reasoning may reflect socioeconomic characteristics of the sample as much as socio-cognitive processes (Dodell-Feder et al., 2019). Thus, it is essential to embrace both experimental and differential methods (including but not limited to sex, age, social class, culture, and

personality traits) in order to fully understand the complex underpinnings of social interactions.

To aid cross-pollination between experimental and differential approaches, more recent imitation research has started to take an individual differences approach by investigating how characteristics of the imitator also influence automatic imitation. For example, imitative behaviour has been argued to vary across stable personality traits such as empathy (Chartrand & Bargh, 1999), narcissism (Hogeveen & Obhi, 2013; Obhi et al., 2014), interoceptive awareness (Ainley et al., 2014) as well as alexithymia (Sowden et al., 2016). Such claims, however, are limited due to the small number of studies reported to date, together with the use of relatively small sample sizes and a lack of powerful replications. Moreover, further studies, which used considerably larger sample sizes, have not been able to replicate the moderating influence of personality variables on automatic imitation (Butler et al., 2015; Cracco et al., 2018). For example, in the study by Butler and colleagues (2015), personality variables such as empathy, narcissism, extraversion, agreeableness, as well as autistic-like and schizotypal traits, did not predict automatic imitation as measured on the SRC task. Interestingly, however, Butler and colleagues (2015) showed that the sex of the participant modulated the compatibility effect such that females showed a greater compatibility effect compared to males. Therefore, it is possible that biological sex is a factor to consider further when attempting to understand how cognitive mechanisms supporting imitation vary across individuals.

Sex is an important individual difference that influences a wide range of cognitive abilities and skills in social as well as non-social contexts (Geary, 2010; Hall, 1978). However, few studies have investigated how socio-cognitive abilities vary as a function of one's biological sex, and the ones that have studied sex differences have typically focused on mental reasoning or emotion perception (Campbell et al., 2002; Krach et al., 2009; Rahman et al., 2004; Russel et al., 2007). Further, such prior studies have often produced mixed results based on relatively small sample sizes (Hyde, 2014; Miller & Halpern, 2014). Therefore, the potential influence of sex on complex cognitive mechanisms that control non-verbal interactions, remains largely unknown.

The extent to which sex differences operate in imitative behaviour has also received minimal attention to date. For example, no sex differences have been found on the automatic imitation of actions or gestures (Chartrand & Bargh, 1999; Larsen et al., 2010). By contrast, studies on facial mimicry have shown that females automatically

imitate facial expressions more than males (Dimberg, 1990; Sonnyby-Borgstrom et al., 2008). Although there is no clear and consistent empirical evidence to date, which speaks to sex differences in imitation, there is theoretical reason to think that sex differences may exist in imitative behaviour. Indeed, there is robust evidence for females to be more empathetic than males (Baron-Cohen & Wheelwright, 2004; Christov-Moore et al., 2014). Further, empathy has been associated with a variety of paradigms investigating imitation (Chartrand & Bargh, 1999; Müller et al., 2013; Sonnby-Borgstrom 2002). Females could thus be more pro-social than males in general, which may lead to more imitative tendencies in particular social contexts.

Further, a core question pertains to whether the sex difference seen in the SRC task is a genuine difference between males and females, or reflects an in-group or own-sex bias. The stimuli used in the task by Butler and colleagues (2015) were of a female hand. It may be that automatic imitation may increase when the sex of the participant and the interacting partner are matched. In other words, female participants observing a female hand may show an in-group bias. Group biases (typically in-group favouritism and out-group dislike) are prevalent in day-to-day social interactions (Allport, 1954; Cameron et al., 2001). Individuals show in-group favouritism for members of their own race and ethnicity (Aboud 1988; Ito & Bartholow, 2009; Kubota et al., 2012; Malpass & Kravitz, 1969; Milner, 1983; van Bavel & Cunningham, 2009), sex (Brown, 1995; Fishbein, 1996; Powlishta, 1995; Rudman & Goodwin, 2004; Yee & Brown, 1994), as well as when groups are arbitrarily assigned (Bernstein et al., 2007; Tajfel et al., 1971). Therefore, ingroup biases seem like a powerful candidate mechanism, which may guide imitative behaviour based on the sex of the interaction partner.

In imitation research more specifically, children have been known to imitate same-sex models more than others (Shutts et al., 2010). Further, facial imitation and SRC measures of automatic imitation have both been found to increase when the interacting partner is an in-group member compared to an out-group member based on race, ethnicity and arbitrary group assignment (Gleibs et al., 2016; Mondillon et al., 2007; Rauchbauer et al., 2015). Thus, the sex difference on the SRC task may be explained by females being more sensitive to female stimuli, thus showing a higher compatibility effect compared to males. In line with this possibility, recent work provides suggestive evidence for a sex difference and/or in-group bias in the automatic imitation task (Cracco et al., 2018; Genschow et al., 2017). For example, a meta-analysis found a higher reaction time

compatibility effect when the sex of the stimuli matched the sex of most participants in the sample[2]. In addition, a study using a female model also found a higher compatibility effect in females compared to males (Genschow et al., 2017). However, the extent to which this sex difference reflects an in-group bias remains unclear because no existing study has manipulated the sex of the stimuli across male and female participants.

Another possible explanation of the sex difference on the automatic imitation SRC task is that females tend to automatically imitate more than males, and therefore require more cognitive resources to inhibit the tendency to automatically imitate, leading to a greater compatibility effect. If so, the sex difference would be tied to a process related to imitation specifically. Alternatively, the sex difference may be more domain-general in nature i.e. it may reflect a basic difference in the cognitive systems that underlie performance on SRC tasks more generally.

Consistent with a domain-general explanation, sex differences have been found on a whole range of non-social inhibitory control tasks which, like the imitation task, require the inhibition of task-irrelevant automatic response tendencies in order to enforce a task-relevant response. For example, on a typical flanker task, participants are instructed to respond to the central stimulus, and inhibit their automatic response tendencies to the flankers or task-irrelevant stimuli that are also presented (Eriksen & Eriksen, 1974). Females were found to have a higher compatibility effect than males i.e. they were slower to respond when flankers were incompatible to the task-relevant response (Stoet, 2010; Judge & Taylor, 2012; Clayson et al., 2011). A similar profile of sex difference has also been found on other inhibitory control tasks like the oddball, gaze- and arrow-cueing, and Simon tasks (Rubia et al., 2010; Bayliss et al., 2005; Merritt et al., 2007; Alwall et al., 2010). As such, sex differences in SRC tasks may reflect differences in cognitive systems that operate across these tasks such as selective attention (Clayson et al., 2011) and/or spatial processing (Stoet, 2017).

These findings suggest that it is as yet unclear whether the sex difference on SRC measures of automatic imitation reflect more domain-general processes or processes solely tied to imitative control (Butler et al., 2015; Cracco et al., 2018; Darda et al., 2018;

[2] However, the authors of the meta-analysis categorised a sample as "female" if more than half the population was female. Thus, even samples with 51% females would be classified as a female sample, biasing the interpretation of the consequent analysis and making clear conclusions difficult to reach.

Darda & Ramsey, 2019; Genschow et al., 2017). Indeed, the SRC task used by previous researchers to demonstrate the existence of a sex difference was a composite of both spatial and imitative components (Butler et al., 2015; Genschow et al., 2017). A sex difference solely tied to imitative control might suggest that a distinct mechanism, or a partially distinct set of mechanisms, may underpin performance on the automatic imitation task compared to other inhibitory control tasks. Thus, in order to understand the cognitive architecture of social interactions, it is critical to unpack the relative contributions of both general and specific components in socio-cognitive processes. Such a viewpoint is consistent with recent proposals, which suggest that it may be fruitful to view social cognition as a composite of both domain-general and domain-specific processes, rather than overly focus on domain-specific contributions (Michael & D'Ausilio, 2015; Spunt & Adolphs, 2017; Binney & Ramsey, 2019; Ramsey, 2018). Therefore, in the current study, we investigate sex differences on the automatic imitation task as well as a non-social control task in order to investigate whether the sex difference relies on domain-general and/or domain-specific mechanisms.

In the current paper, across three large sample experiments, we integrate approaches from experimental and differential psychology approaches to investigate three critical questions pertaining to individual differences in a form of social (imitative) and non-social cognitive control. First, consistent with recent suggestions to make replication a common and foundational practice in psychology (Zwaan et al., 2018), we aim to confirm the sex difference found previously (on both social and non-social cognitive control tasks) and provide a more precise estimate of the effect size. Further, we aimed to replicate the lack of variation in automatic imitation as a function of personality traits that has been reported previously in large sample research designs (Butler et al., 2015; Cracco et al., 2018). Second, we aim to investigate whether the sex difference on the imitation task reflects an actual difference between males and females, or an in-group or own-sex bias. Third, we aim to uncover whether mechanisms underlying the sex difference are domain-general or domain-specific (or a combination of both).

Experiment 1

Introduction

In the first experiment, we aim to replicate the sex difference on the general compatibility effect as found previously (Butler et al., 2015; Genschow et al., 2017). We further extend this research by investigating whether performance on a non-social inhibitory control task (i.e. a flanker task) also varies between the sexes. A similar sex difference on both tasks would indicate that the sex difference is supported by differences in a basic domain-general control system that underpin performance across social and non-social tasks. Alternatively, a sex difference on only one task would indicate at least partially distinct mechanisms as a function of sex.

Further, we also investigate the extent to which stable dimensions of personality influence the control of automatic imitation as measured on the SRC task. Prior work has provided mixed evidence regarding this question. Some studies have found a link between automatic imitation and empathy (Chartrand & Bargh, 1999) and narcissism (Hogeveen & Obhi, 2013; Obhi et al., 2013). There are theoretical grounds to also posit a link between automatic imitation and two of the Big Five personality factors. Agreeableness and extraversion have been previously linked to empathy, altruism, and sociability (Ashton et al., 1998; Barrio et al., 2004; McCrae & Costa, 1999), and are thus considered as contributors to prosocial behaviour (Graziano & Eisenberg, 1997). Thus, individuals who are more agreeable and more extraverted may be more prosocial and could thus imitate their interacting partners more than others. In addition, although controversial and debated (Hamilton, 2013; Southgate & Hamilton, 2008), imitation abilities have been argued to vary in atypical populations including autism spectrum disorders and schizophrenia (Oberman & Ramachandran, 2007; Thakkar et al., 2014; Williams et al., 2001), indicating that a relationship may exist between autistic-like and schizotypal traits and automatic imitation.

The largest datasets to date, however, show that performance on the SRC task is invariant to stable personality variables (Butler et al., 2015; Cracco et al., 2018). One concern with such null effects of personality is that they may reflect the impoverished social context of the SRC task. That is, effects of interest may only operate in more socially meaningful contexts. Therefore, in Experiment 1, we make the social context more meaningful by including emotional facial expressions within our design and investigate the extent to which automatic imitation continues to remain invariant as a function of

stable personality traits. We included five face images depicting five different emotional expressions (fearful, angry, happy, sad, neutral), and seven personality variables such as extraversion, agreeableness, autistic-like and schizotypal traits, narcissism (including grandiose narcissism and vulnerability narcissism), empathy (including empathic concern and perspective taking), and alexithymia (for detailed information about measures used, see supplementary material). Following Butler and colleagues (2015), although there is theoretical reason to expect pro-social dimensions of personality to be related to automatic imitation, we would expect automatic imitation to be invariant to stable dimensions of personality.

Method

Across all experiments, we report how the sample size was determined, all data exclusions, and all measures in the study (Simmons et al., 2011; 2012). Following open science initiatives, all raw data are available online for other researchers to pursue alternative questions of interest. For all three experiments, data pre-processing, statistical analyses, and data visualisations were performed using R (R Core Team, 2018), unless otherwise specified. All raw data and code used for analyses are available online (https://osf.io/fsh9b/).

We determined the sample size for our experiments as follows. For experiment 1, we aimed to collect as many participants as possible over a two-day data collecting session. Therefore, the stopping rule was to terminate data collection after day 2 of data collection. For Experiments 2 and 3, in order to focus our design on the primary research question, which concerned sex differences, we set a minimum sample size of 100 male and 100 female participants. Sensitivity analyses revealed that given a sample size of two hundred participants (100 per sex), we would have 80% power to detect an effect size of Cohen's d > 0.35 for the mean difference between the two sexes, and an effect size of η_p^2 > 0.04 for a 2 x 2 mixed ANOVA. Such a design, therefore, provides reasonable confidence (80%) to detect effect sizes of interest that are conventionally considered small-to-medium.

Participants.

Two hundred and three participants took part in this experiment for monetary compensation (£6) or course credit. All participants provided informed consent and had normal or corrected-to-normal vision. Approval was obtained from the Research Ethics

and Governance Committee of the School of Psychology at Bangor University. Participants were excluded if performance was three standard deviations away from the group mean average performance per condition in terms of accuracy or reaction time (N=14 for the imitation task, N=7 for the flanker task). A further 14 participants were excluded as demographic information (age and sex of the participant) was not recorded. For the imitation task, the final sample included 175 participants (59 males, $Mean_{age}$ = 20.9, SD_{age} = 4.23). For the flanker task, the final sample included 182 participants (59 males, $Mean_{age}$ = 20.9, SD_{age} = 3.73).

Stimuli, tasks, and procedure.

Automatic imitation task. The automatic imitation task was based on the stimulus response compatibility (SRC) paradigm developed by Brass and colleagues (2000), which consisted of the observation and execution of finger lifting movements (Figure 1). In order to explore whether facial cues signalling emotional states influenced automatic imitation, and to make the social context more meaningful, five face images depicting five different emotional states were also presented along with the hand stimuli of the imitation task. The face stimuli were images of 5 individuals from the NimStim data set with five different expressions (neutral, sad, happy, fearful, and angry) (Tottenham et al., 2009). Closed mouth variations of sad and neutral, and open mouth versions of smiling, frowning, and fearful expressions were chosen as these stimuli were most often correctly identified (see validation data provided by Tottenham and colleagues: http://www.macbrain.org/faq.htm). To avoid any effects of race, models were chosen from the largest ethnic group represented in the NimStim set (European-American). From this group, five male and five female models whose expressions were identified with the highest accuracy across the five relevant expressions were chosen. Thus, there were a total of 50 different face images.

The hand stimuli comprised five images of a female hand positioned in the centre of the screen and viewed from a third person perspective such that the fingers extended towards the participants. The first image was of the hand in a neutral position, while the remaining four images showed either an index or middle finger lift with a number ‘1’ or ‘2’ presented between the index and middle finger. Participants were asked to hold down the “m” and “n” keys on the keyboard with their index and middle fingers of the right hand, respectively. They were instructed to lift their index finger when they saw a number “1” and their middle finger when they saw the number “2”. Thus, there were four

possible trial types, two of which were compatible, and two of which were incompatible. In the compatible condition, participants were cued to perform the same finger-lifting movement that they observed (i.e. an index finger movement with a '1' or a middle finger movement with a '2'). In the incompatible condition, the executed and observed movements were different (i.e. an index finger movement with a '2' or a middle finger movement with a '1').

Each trial began with the presentation of a fixation cross for 500 milliseconds (ms). A face image was presented after the fixation cross for 500 ms, followed by the neutral hand image. The face image remained on the screen above the neutral hand and target hand image for the remainder of the trial. The neutral hand was presented for a random inter-stimulus interval (ISI) of 500, 700, or 1000 ms, followed by the target hand image. The succession of neutral and target hand images was such that it produced apparent motion of either an index or middle finger lift simultaneously with the presentation of the number cue. The target hand image remained on the screen until the participant made a response (but no longer than 2000 ms). The total trial length varied depending on the ISI, but was never longer than 3500 ms. Trials were pseudo-randomised in such a way that no more than 4 identical trials were presented consecutively. There were four blocks of 50 trials each which included 25 compatible trials and 25 incompatible trials with equal number of trials per face image.

Flanker task. The flanker task was based on the paradigm developed by Eriksen and Eriksen (1974; Figure 1). The stimuli consisted of five equally sized and spaced white arrows on a black background. Participants were instructed to respond to the direction of the central arrow – they were asked to press key 'm' with their right index finger if the central arrow pointed to the right, and press key 'n' with their left index finger if the central arrow pointed to the left. The direction of the flanker arrows was either compatible (<<<<< OR >>>>>) or incompatible (<<><< OR >><>>) to the central arrow direction. This produced four trial types and two conditions (compatible and incompatible).

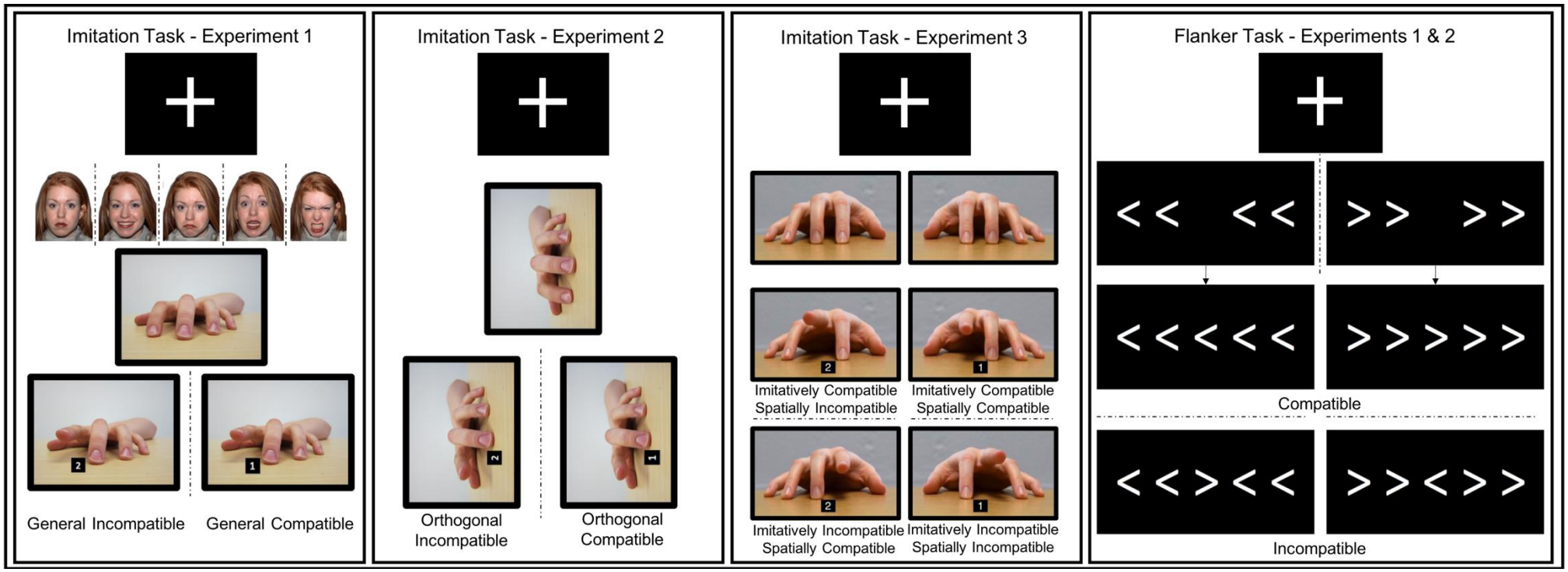


Figure 1. Imitation and Flanker Tasks. Stimuli and trial design for the imitation and flanker tasks in Experiments 1, 2, and 3. Flanker tasks were similar for Experiment 1 and 2. In Experiment 1, in the imitation task, hand stimuli were preceded by a face depicting either a neutral, happy, sad, fearful, or angry image. In Experiment 2, hand stimuli were presented orthogonal to the participant's response hand, and in Experiment 3, both left- and right-hand images were used in order to measure imitative and spatial effects independent of each other.

Each trial started with a fixation cross for 800 ms, 1000 ms, or 1200 ms. The flanker arrows then appeared on the screen for 100 ms, followed by the central arrow in between the flankers. The five arrows remained on the screen till the participant responded (but for no longer than 1600 ms). Participants were first presented with the fixation cross for 800 ms, 1000 ms, or 1200 ms, followed by the presentation of the four flanker arrows for 100 ms. Total trial length was never longer than 2900 ms. Trials were pseudo-randomised in such a way that no more than 4 identical trials were presented consecutively. Each participant did one block of 64 trials, with 32 compatible, and 32 incompatible trials. Further, in this experiment, we addressed an additional unrelated question – in half of the compatible and incompatible trials, flanker arrows flipped arrow direction during the trial between their initial presentation on the screen and the appearance of the central arrow. However, as we were interested in the basic compatibility effect, we collapsed trials across conditions irrespective of whether they changed direction mid trial or not.

Participants first completed the automatic imitation task, followed by the flanker task. Before starting each task, they completed a 10-trial practice block.

Questionnaires. Participants also completed a range of self-report questionnaires which included the Mini International Personality Item Pool (mini IPIP; Donnellan et al., 2005; the Short Autism Spectrum Quotient (AQ-10 Adult; Baron-Cohen et al., 2001; Allison et al., 2012), the Brief Schizotypal Personality Questionnaire (SPQ-B; Raine & Benishay, 2005), the Narcissistic Personality Inventory (NPI-16; Ames et al., 2006), the Hypersensitivity Narcissism Scale (HSNS; Hendin & Cheek, 1997), the Interpersonal Reactivity Index (IRI; Davis, 1980), and the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994). For more details on the measures used, see the Supplementary Material.

Data analysis.

Accuracy on the imitation task was recorded as the proportion of trials that were correct i.e. when participants lifted the correct finger in response to the number cue. Reaction time (RT) was recorded as time taken from target onset to participant's response. Only correct trials were used to calculate RT. Trials on which participants responded incorrectly, i.e. lifted the wrong finger, responded after 2000 ms, or before target onset were all excluded from the analysis (5.64%).

Accuracy on the flanker task was recorded as the proportion of trials that were correct i.e. when participants pressed the correct button in response to the central arrow

direction. RT was calculated as the time taken from target onset (i.e. presentation of the arrow) to when the participant made a response. Only correct trials were used to calculate RT. Trials on which participants responded incorrectly, i.e. lifted the wrong finger, responded after 1600 ms, or before target onset were all excluded from the analysis (4.77%). Compatibility effects were calculated by subtracting reaction times on compatible trials from reaction times on incompatible trials.

Data was analysed as follows: first, for both the RT and accuracy data on the imitation task, a 2 (compatibility: incompatible, compatible) x 5 (emotion: neutral, sad, happy, fearful, angry) repeated measures ANOVA was performed to investigate whether facial cues signalling emotional states modulated the compatibility effect on the imitation task. Second, on both RT and accuracy data, for the flanker and imitation tasks separately, a 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA was performed in order to investigate whether the compatibility effect on the imitation and flanker tasks varies as a function of sex.

Based on prior research (Heyes, 2011; Brass et al., 2000; Eriksen & Eriksen, 1974), we expected a main effect of compatibility such that RT would be higher, and accuracy would be lower on incompatible trials compared to compatible trials. In support of our hypothesis, we also expected a Compatibility*Sex interaction such that the compatibility effect would be higher for females as compared to males. The interaction effect was central to testing our primary hypothesis, and thus, we calculated compatibility effects for male and female participants separately by computing the mean difference and 95% confidence intervals between compatible and incompatible conditions. In order to directly estimate the size of the difference in compatibility effects between males and females, we then again computed the mean difference and 95% confidence interval. We used one-tailed 95% confidence intervals as we had a directional hypothesis that females would have a higher compatibility effect than males on both the imitation and flanker tasks. Third, in order to investigate whether the flanker and imitation compatibility effects were correlated, a one-tailed Pearson's correlation was performed. A positive correlation would suggest that the two compatibility effects were related to each other.

We also report standardised effect sizes for ANOVA using partial eta-squared (η_p^2) for independent samples t-tests using Cohen's *d* and for paired samples t-tests using Cohen's d_z (Cohen, 1992; Lakens, 2013). We also report and interpret the point and interval estimate using 95% CIs for effect sizes of interest in line with recent suggestions (Cumming, 2012; Amrhein et al., 2019). In order to quantify the evidence for a null

hypothesis over the experimental hypothesis (where a null result was found using null hypothesis significance testing; NHST), we calculated the Bayes factor (BF_{01}) by performing a Bayesian independent samples t-test to investigate the sex difference between the sexes. The Bayes factor was interpreted using benchmark criteria from Jeffereys (1961). Bayesian analyses, Cohen's d and d_z, as well as 95% CIs were calculated using JASP (JASP Team, 2018).

Further, as previous research demonstrated that the compatibility effect (as measured on the SRC imitation task) is invariant to stable traits of personality (Butler et al., 2015), we also investigated whether personality variables influenced automatic imitation by using multiple regression analyses. We introduced a more social context to the task by introducing facial cues signalling emotional expressions simultaneously with the hand images. Based on prior work, we predicted that facial cues signalling positive emotions would increase automatic imitation compared to neutral and negative emotional expressions (Rauchbauer et al., 2015; Butler et al., 2016). Following Butler and colleagues (2015), we set up a base model comprising mean RT (collapsed across all conditions), participant sex, and the mean RT * sex interaction, as these factors have been shown to explain variance in automatic imitation previously (Butler et al., 2015). We then individually tested the contribution of each of the personality measures by adding them to the base model in separate hierarchical multiple regression analyses. By doing so, we are able to address the extent to which personality measures predict variance in the SRC imitation task above and beyond the base model. To transparently visualise and report the data, we also include zero-order correlations between personality measures and performance on the SRC imitation task. As sex differences have been previously found on personality measures (Schmitt et al., 2008), we computed sex*trait interaction terms for all personality variables, and evaluated them in separate multiple regression models.

Results

Automatic imitation task.

Accuracy. Average accuracy on the imitation task was above 90% for both males and females on both compatible and incompatible conditions (Figure 2, Supplementary Table 1). A 2 (compatibility: compatible, incompatible) x 5 (emotion: neutral, sad, happy, fearful, angry) ANOVA showed no main effect of emotion ($F(4, 696) = 0.50$, $p=0.729$, $\eta_p^2 = 0.003$) and no significant Compatibility*Emotion interaction ($F(4, 656) = 1.20$, $p=0.310$, $\eta_p^2 = 0.007$). Thus, for all further analyses of accuracy, trials are collapsed across all emotion conditions.

The 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA showed a main effect of compatibility such that participants were more accurate on compatible trials than incompatible trials ($F(1, 173) = 258.09$, $p<.001$, $\eta_p^2 = 0.60$; Figure 2). The effect size for the main effect of compatibility is conventionally considered to be large. There was no significant main effect of sex ($F(1, 173) = 0.22$, $p = 0.64$, $\eta_p^2 = 0.001$) and no significant Compatibility*Sex interaction ($F(1, 173) = 0.60$, $p = 0.44$, $\eta_p^2 = 0.003$) and the effect sizes were close to zero (Supplementary Table 2).

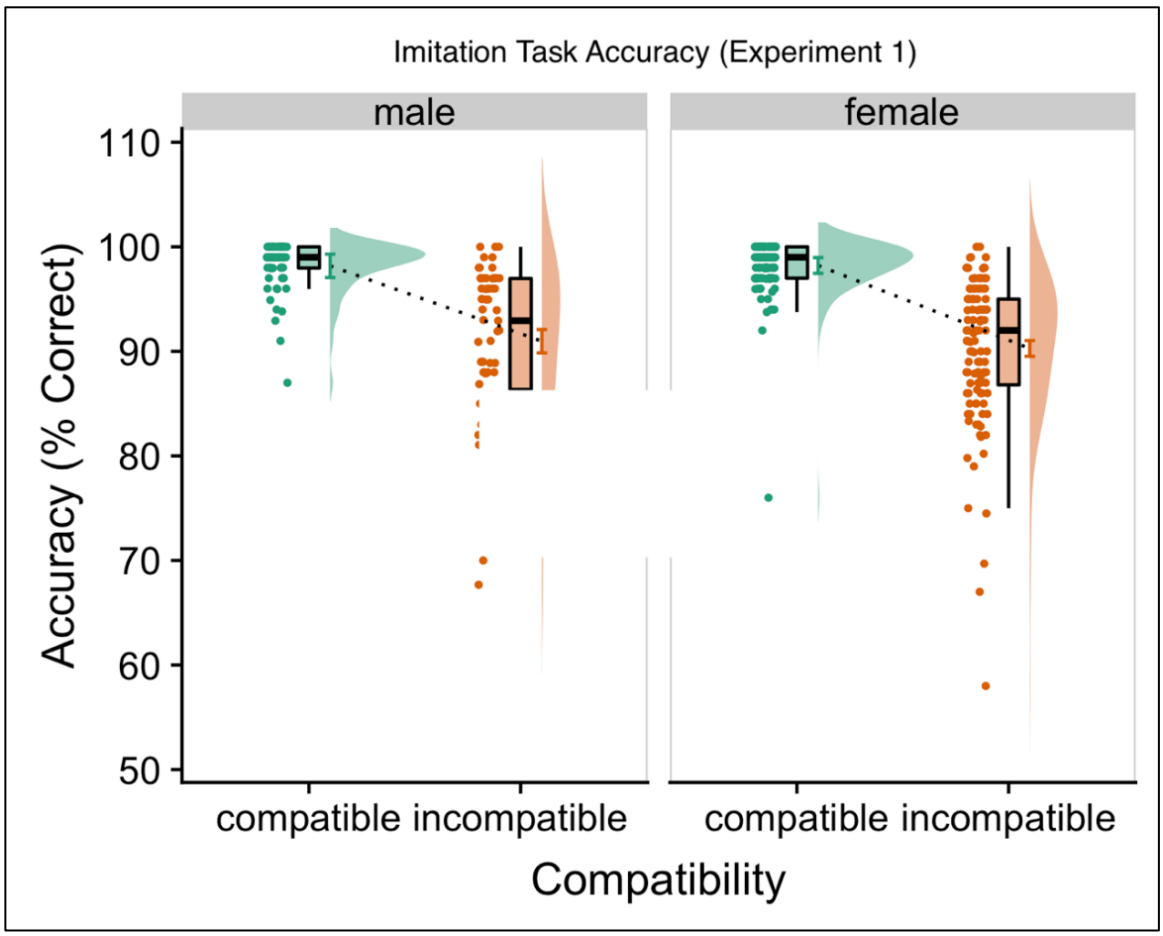


Figure 2. Experiment 1 – Imitation Task Accuracy. Raincloud plot (Allen et al., 2018) illustrating imitation task accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females.

Reaction time. Average reaction times on the imitation task for both males and females on both compatible and incompatible conditions were between 485 and 585 milliseconds (Figure 3, Supplementary Table 1). A 2 (compatibility: compatible, incompatible) x 5 (emotion: neutral, sad, happy, fearful, angry) ANOVA showed no main effect of emotion ($F (4, 696) = 1.81$, $p=0.127$, $\eta_p^2 = 0.004$). Importantly, there was no Compatibility*Emotion interaction and the effect size was close to zero ($F (4, 696) = 0.40$,

p=0.796, η_p^2 = 0.002, see Supplementary Figure 1). Thus, for all further analyses of RT, trials are collapsed across all emotion conditions.

The 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA showed a main effect of compatibility such that participants were slower to respond on incompatible trials than compatible trials ($F(1, 173) = 669.77$, $p<.001$, η_p^2 = 0.80; Figure 3). The effect size for the main effect of compatibility is conventionally considered to be large. There was no significant main effect of sex and the effect size was close to zero ($F(1, 173) = 0.26$, $p = 0.61$, $\eta_p^2 = 0.001$). There was a Compatibility*Sex interaction approaching significance and the effect size is generally considered to be a small effect ($F(1, 173) = 3.16$, $p = 0.08$, $\eta_p^2 = 0.018$; Supplementary Table 2).

To further explore our primary research question regarding sex differences in the imitation task, compatibility effects were computed separately for males and females, and then compared to each other. For both males and females, compatibility effects had a large standardised effect size (Cohen's $d_z > 2.07$) with the lower bound of the 95% confidence interval at 1.68 or higher. When compatibility effects for males and females were directly compared to each other, we found a mean difference in the direction that was predicted (females > males). Indeed, the compatibility effect for females was 12.40ms higher than males and the lower bound of the 95% confidence interval was 0.87ms (Mean Difference = 12.40 ms, 95% CI[0.87, ∞], Cohen's d = 0.28, 95% CI [0.02 , ∞]; Figure 3, Table 1A). The standardised effect of d = 0.28 is conventionally considered a small-to-medium effect, and the lower bound of the 95% confidence interval was just above zero (0.02). Thus, these findings suggest that performance on the imitation task differs as a function of sex in a manner that is consistent with our predictions, such that females had a greater compatibility effect than males.

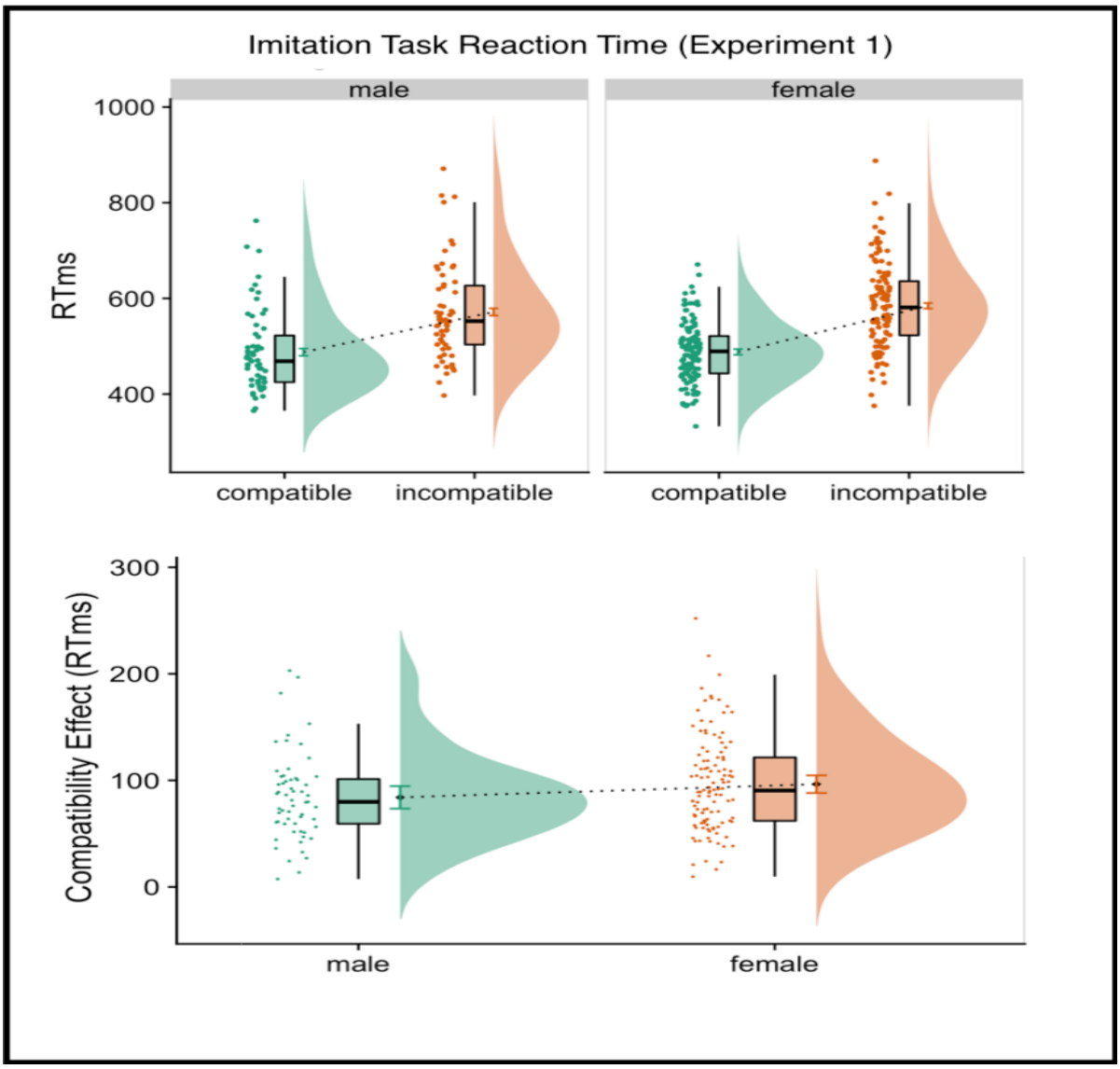


Figure 3. Experiment 1 – Imitation Task Reaction Time. Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females. The lower panel shows the compatibility effect for both males and females. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds.

Table 1.
Compatibility effects for the imitation and flanker tasks across Experiments 1, 2, and 3.

A) Experiment 1	**Compatibility effect for the imitation and flanker task**			
	Mean Difference (ms)	95% CI	Cohen's d_z/d	BF_{01}
Imitation task (General Compatibility Effect)				
Males	84.03	(75.19, ∞)	2.07 [1.68, ∞]	
Females	96.43	(89.49, ∞)	2.14 [1.86, ∞]	
Females - Males	12.40	(0.87, ∞)	0.28 [0.02, ∞]	
Flanker Task (Flanker Compatibility Effect)				
Males	51.17	(42.46, ∞)	1.28 [0.98, ∞]	
Females	56.79	(51.42, ∞)	1.58 [1.36, ∞]	
Females - Males	5.62	(-4.14, ∞)	0.15 [-0.11, ∞]	3.85
B) Experiment 2	**Compatibility effect for the imitation and flanker task**			
	Mean Difference (ms)	95% CI	Cohen's d_z/d	BF_{01}
Imitation task (Orthogonal Compatibility Effect)				
Males	25.79	(21.72, ∞)	1.02 [0.82, ∞]	
Females	32.77	(28.82, ∞)	1.28 [1.07, ∞]	
Females - Males	6.98	(1.34, ∞)	0.27 [0.05, ∞]	
Flanker Task (Flanker Compatibility Effect)				
Males	93.88	88.89	3.11 [2.71, ∞]	
Females	94.87	89.96	2.98 [2.62, ∞]	
Females - Males	0.98	(-6.01, ∞)	0.03 [-0.19, ∞]	6.58
C) Experiment 3	**Spatial and imitative compatibility effects (collapsed across all levels of stimulus sex)**			
	Mean Difference (ms)	95% CI	Cohen's d_z/d	BF_{01}
Spatial Compatibility (Spatially Incompatible – Spatially Compatible)				
Males	30.74	(27.64, ∞)	1.18 [1.02, ∞]	
Females	37.28	(33.97, ∞)	1.37 [1.20, ∞]	
Females - Males	6.53	(2.02, ∞)	0.24 [0.07, ∞]	
Imitative Compatibility (Imitatively Incompatible – Imitatively Compatible)				
Males	7.02	(4.50, ∞)	0.33 [0.21. ∞]	
Females	8.35	(5.95, ∞)	0.42 [0.30, ∞]	
Females - Males	1.33	(-2.15, ∞)	0.06 [-0.10, ∞]	4.95

N.B. Compatibility effects for males and females, as well as the difference between males and females, for the imitation and flanker tasks are reported for Experiments 1, 2, and 3, along with 95% CIs, effect sizes and BF_{01}. Abbreviations: ms = milliseconds, CI = confidence intervals, BF = Bayes Factor.

Flanker task.

Accuracy. Average accuracy on the flanker task was above 94% for both males and females on both compatible and incompatible conditions (Figure 4, Supplementary Table 1). A 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 4) showed no main effect of compatibility ($F(1, 180) = 2.24$, $p = 0.136$, $\eta_p^2 = 0.01$). The effect size for the main effect of compatibility was close to zero. There was no significant main effect of sex ($F(1, 180) = 0.04$, $p = 0.85$, $\eta_p^2 = <0.001$) and no significant Compatibility*Sex interaction ($F(1, 180) = 0.09$, $p = 0.759$, $\eta_p^2 < 0.001$) and the effect sizes were almost zero (Supplementary Table 2).

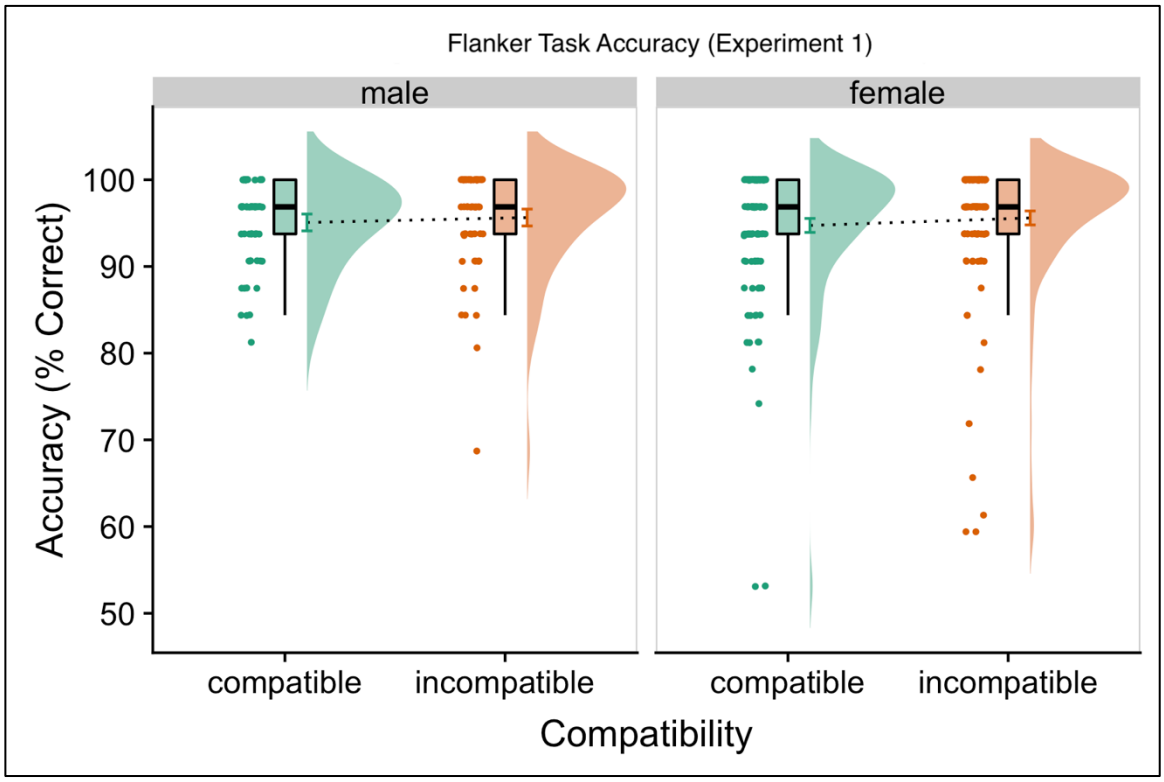


Figure 4. Experiment 1 – Flanker Task Accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females.

Reaction time. Mean reaction time on the flanker task for both males and females on both compatible and incompatible conditions was between 420 and 495 milliseconds (Figure 5, Supplementary Table 1). A 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 5) showed a main effect of compatibility such that participants were slower to respond on incompatible trials than compatible trials (F(1,

180) = 334.15, p<.001, η_p^2 = 0.65). The effect size for the main effect of compatibility is conventionally considered to be large. There was a main effect of sex approaching significance and the effect size was a small effect (F(1, 180) = 3.40, p = 0.08, η_p^2 = 0.02). There was no significant Compatibility*Sex interaction and the effect was close to zero (F(1, 180) = 0.90, p = 0.34, η_p^2 = 0.005). The main effect of sex showed that females were overall slower than males on the flanker task (see Supplementary Table 2).

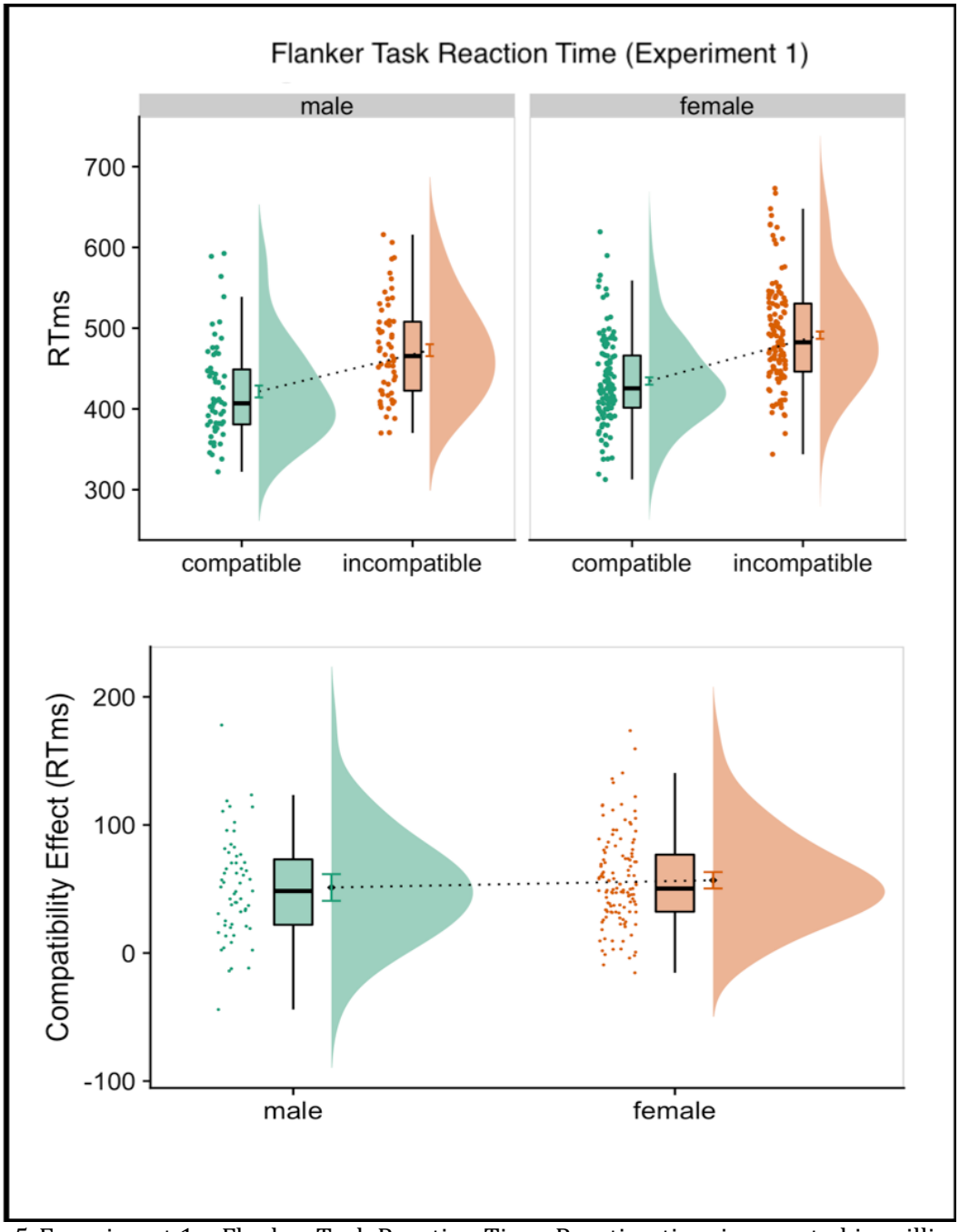


Figure 5. Experiment 1 – Flanker Task Reaction Time. Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females. The lower panel shows the compatibility effect for both males and females. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds.

To further compare with the automatic imitation task, compatibility effects in the flanker task were computed separately for males and females, and then compared to each other. For both males and females, compatibility effects had a large effect size (Cohen's $d_z > 1.2$) with the lower bound of the 95% confidence interval at 0.98 or higher. When compatibility effects for males and females were directly compared to each other, there was a trend toward females showing a higher compatibility effect than males by 5.62ms, with the lower bound of the 95% confidence interval at -4.14 ms below zero (Mean Difference = 5.62 ms, 95% CI[-4.14, ∞], Cohen's d = 0.15, 95% CI [-0.11]; see Figure 5, Table 1A).The effect size was small, with the lower bound of the 95% confidence interval at -0.11. A Bayesian independent samples t-test showed that the null was 3 to 4 times more likely than the alternative hypothesis ($BF_{01} = 3.85$). Thus, a reasonable estimate for the mean difference between males and females on the flanker compatibility effect ranges from -4.14ms to 5.62 ms, with our best estimate being a small difference between females and males, such that females may show a higher compatibility effect than males.

Correlational analysis.

In order to investigate whether the flanker and imitation compatibility effects were correlated, a one-tailed skipped correlation was performed. To do so, only those participants who performed both the tasks were included in the analysis (N=165). The skipped correlation analyses were performed using a Matlab-based toolbox (Mathworks Inc., MA; http://sourceforge.net/projects/robustcorrtool/, Pernet, Wilcox, & Rousselet, 2013). Skipped correlation takes into consideration the overall structure of the data, and protects against bivariate outliers. In order to perform a skipped correlation analysis, we first tested the assumptions of normality and homogeneity of variance. The Henze-Zirkler's multivariate normality test (Trujillo-Ortiz et al., 2007) indicated that the data were close to normally distributed, and the test for heterogeneity indicated that the data have the same variance. Next, we estimated the robust centre of the data using the minimum covariance determinant (MCD) estimator. The MCD estimator is considered to be a robust estimator of the scatter and location of multivariate data (Rousseeuw, 1984; Rousseeuw & van Drissen, 1999; Verboten & Hubert, 2005). Bivariate outliers were then identified by using a projection technique – data points were orthogonally projected by lines joining each data point to the robust centre of the data cloud. Five bivariate outliers were removed using the box-plot rule relying on the interquartile range (Carling, 2000), and skipped correlation was computed on the remaining data. Following guidelines put

forward previously (Pernet et al., 2013), as the data were close to not being normally distributed, we used a skipped Spearman correlation analysis. A one-tailed skipped Spearman correlation analysis showed a small positive correlation between the imitation and flanker compatibility effects, which did not pass our statistical threshold, with the lower bound of the 95% CI extended below zero (r (160) = 0.04, 95% CI [-0.12, ∞]; see Figure 6).

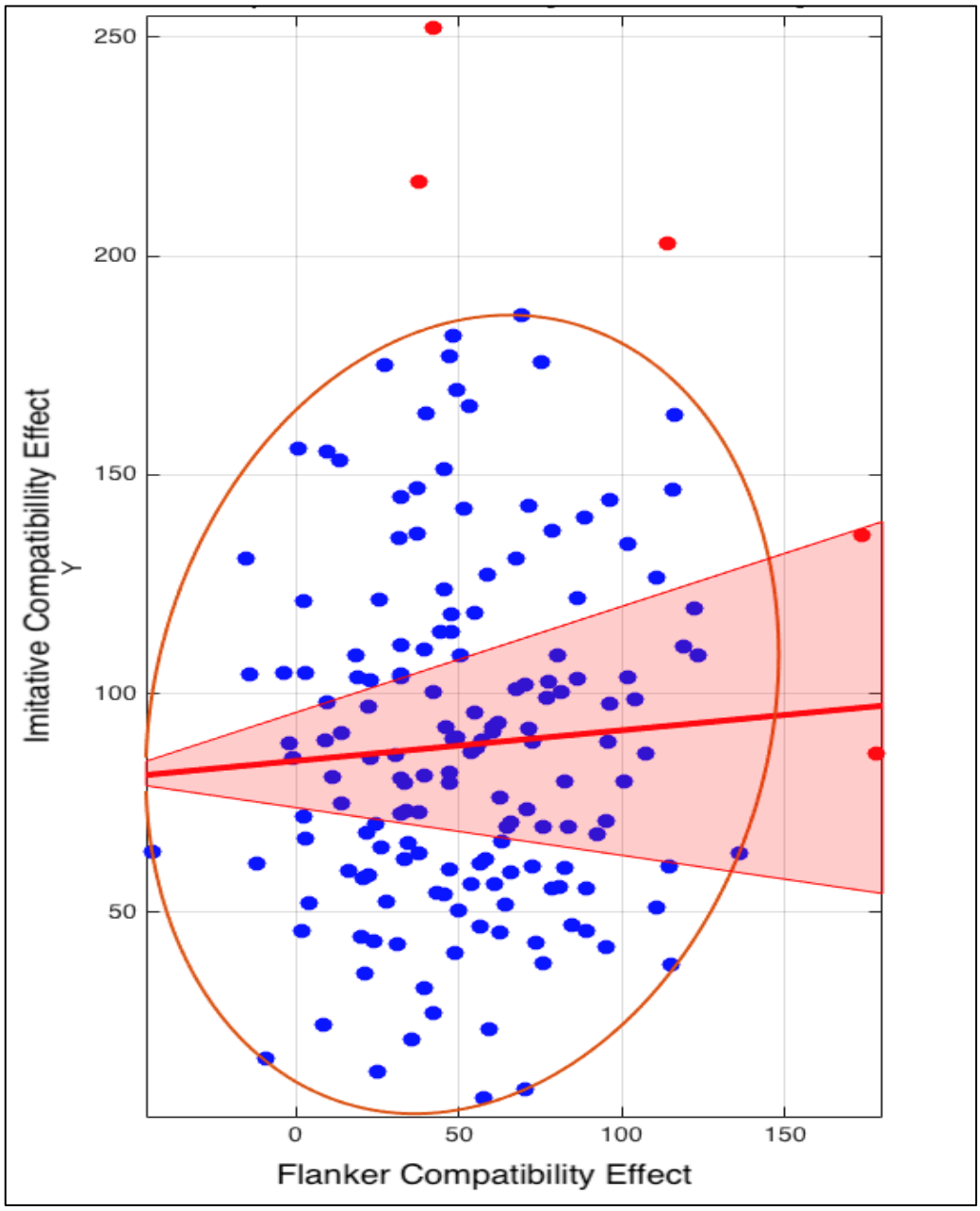


Figure 6. Experiment 1 - Correlation Analysis. A skipped Spearman correlation shows a small positive correlation between the flanker and imitative compatibility effects that does not pass our statistical thresholding. Abbreviations: RTms = reaction time in milliseconds). Dots in red are the bivariate outliers.

Our results thus suggest that the compatibility effects across these tasks were largely unrelated, and participants with greater interference in one task did not experience a greater interference in the other task.

Multiple regression analyses.

We also investigated the relationship between stable personality measures and the general compatibility effect as measured on the SRC task. Tests for multicollinearity indicated that a very low level of multicollinearity was present (*VIF* for all predictor variables < 2). The base model explained 33.6 % of the variance in the congruency effect ($F(3,171)=28.88$, $p<.001$, $R^2=.34$, $f^2=0.51$) and indicated a medium effect size (Cohen, 1992). Mean RT predicted the compatibility effect, with increasing CE as mean RT increases (B=0.27, SEB=0.03, t(171)=8.02, p<.001, 95% CI [0.20; 0.34]). In addition to mean RT, sex marginally predicted the compatibility effect (B=5.40, SEB=2.88, t(171)=1.87, p=.06, [-0.29; 11.10]) with a higher compatibility effect for females than males. The mean RT * sex interaction was also a significant predictor (B=0.10, SEB=0.03, t(171)=3.01, p=.003, [0.03; 0.17]), suggesting that increases in mean RT predicted larger increases in the compatibility effect for females (B=0.37, SEB=0.04, t(171)=8.43, p=.001) compared to males (B=0.17, SEB=0.05, t(171)=3.32, p=.001). Results from the base model are very similar to the results of prior work using a same SRC task and analytical approach (Butler et al., 2015).

Agreeableness, extraversion, grandiose and vulnerability narcissism, empathy, autistic-like and schizotypal trials, and alexithymia did not predict the general compatibility effect above and beyond the base model (all p's > .11, all CIs overlapping with zero; see Figure 7). Effect sizes attributable to the addition of the personality variables (beyond the base model) indicated very small effects (Cohen's f^2 for all models <0.01; Cohen, 1992). The multiple regression models are summarized in Supplementary Table 3. Zero-order correlations are also consistent with the findings from the multiple regression analyses, such that there are no relationships between stable personality measures and CE (see Supplementary Table 4, Supplementary Figure 2).

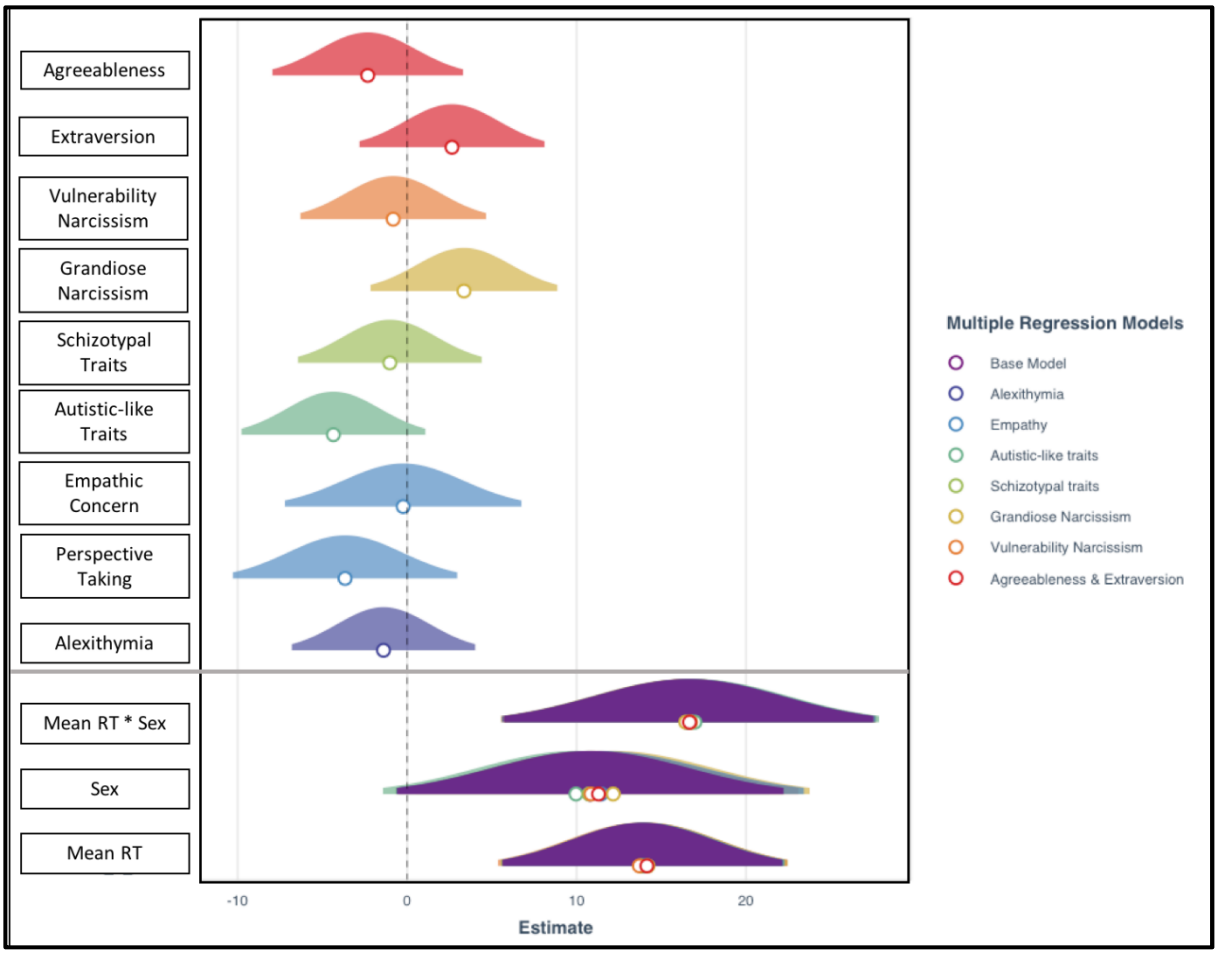


Figure 7. Experiment 1 – Multiple Regression Analyses. Values of standardised coefficients are plotted for each predictor variable (personality trait) along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

To evaluate the sex*trait interaction terms, we computed additional models – each model consisted of the base model, one trait predictor, and the sex*trait interaction term. For alexithymia, when the sex*trait term was included in the model, the model explained 34.9% of the variance. The sex*alexithymia interaction term marginally predicted the compatibility effect (B=0.48, SEB=0.27, t(169)=-1.71, p=0.09, 95% CI [-0.07, 1.03]), and explained an additional 1.3 % of the variance (ΔR2= .013, F(5,169)=18.09,p<.001). A decrease in alexithymia marginally predicted an increase in the compatibility effect only for males (B=-0.80, SEB=0.47, t(169)=-1.72, p=0.09) and not for females (B=0.15, SEB=0.30, t(169)=0.49, p=0.62; see Figure 8A). The effect size attributable to the addition of alexithymia and the sex*trait interaction term was very small (Cohen's f^2 = 0.02).

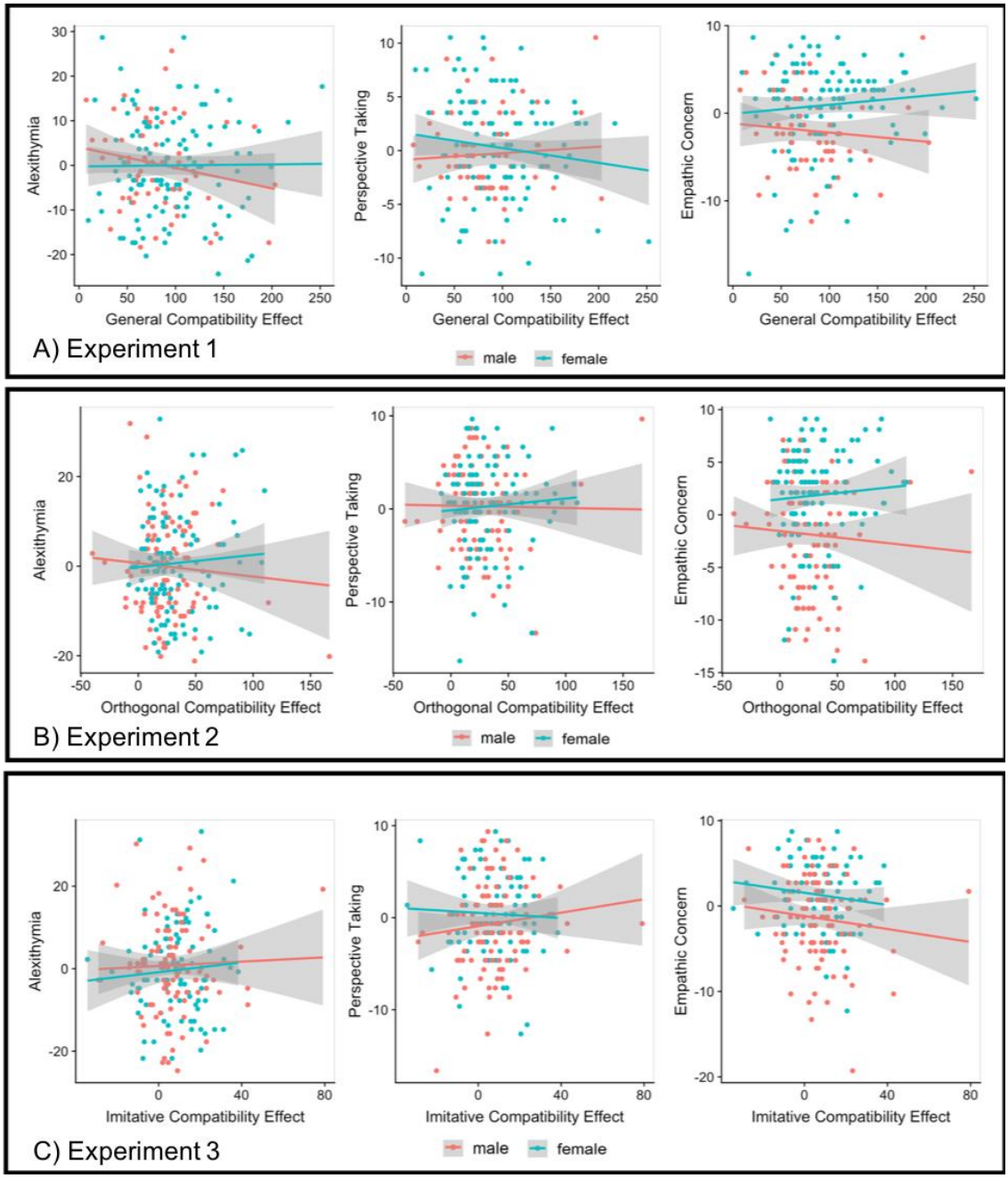


Figure 8. Sex by trait interactions for Experiments 1, 2, 3. Sex by trait interactions for alexithymia, perspective taking, and empathy personality traits for Experiments 1 (A), 2 (B), and 3 (C). X axis denotes the imitative compatibility effect in milliseconds, and Y axis denotes mean centred scores on the personality traits.

For empathy, when the sex*trait term was included in the model, the model explained 36.9% of the variance. Both sex*empathic concern (B=1.62, SEB=0.74, t(167)=2.17, p=0.03, 95% CI [0.15, 3.10]) and sex*perspective taking (B=-1.91, SEB=0.81, t(167)=-2.34, p=0.02, 95% CI [-3.52, -0.30]) predicted the compatibility effect above and beyond the base model and explained an additional 3.3 % of the variance (ΔR2= .03, F(7,167)=13. 95, p<.001). An increase in empathic concern marginally predicted a

decrease in the compatibility effect for males (B=-2.24, SEB=1.23, t(167)=-1.82, p=.07) whereas in females, there was a trend for an increase in empathic concern predicting an increase in the compatibility effect (B=1.01, SEB=0.85, t(169)=1.19, p=.24). An increase in perspective taking predicted a decrease in the compatibility effect in females (B=-1.90, SEB=0.85, t(167)=-2.24, p=.026). In males, there was a trend for an increase in perspective taking predicting an increase in the compatibility effect (B=1.92, SEB=1.39, t(167)=1.38, p=0.169; see Figure 8A). The effect size attributable to the addition of empathy and the sex*trait interaction term was small (Cohen's f^2 = 0.05). None of the other sex*trait interaction terms predicted the compatibility effect above and beyond the base model (Cohen's f^2 for all models <0.02, Supplementary Figure 3, Supplementary Table 5).

Discussion

The results demonstrate a sex difference in the general compatibility effect on the imitation task such that females showed a higher general compatibility effect than males, thus replicating the direction of results found previously (Butler et al., 2015; Genschow et al., 2017). The sex difference observed on the imitation task was not seen to the same degree on the flanker task. Moreover, flanker and general compatibility effects were largely unrelated to each other. At first glance, therefore, this suggests that the sex difference may be tied to a form of cognitive control that is not shared between the two tasks, such as social (imitative) control.

Before we can make firm conclusions regarding the type of cognitive structure supporting the sex difference, however, we first consider some limitations of these results. First, the general compatibility effect is a sum of both spatial and imitative features. Participants respond with their right hand to a number cue – they are asked to lift their index or middle finger (for '1' or '2' respectively) while simultaneously observing a left hand making either the same or different finger movements. However, in this task, the observed and executed movements are not just imitatively compatible or incompatible, but also on the same or different side of space i.e. spatially compatible or incompatible. Thus, the task measures a general compatibility effect i.e. it does not measure the control of automatic imitation or the imitative compatibility effect independent of spatial compatibility effects (Catmur & Heyes, 2011). Therefore, the sex difference may reflect a difference in spatial compatibility with respect to a finger location in space, as opposed to specifically in imitation control.

A second limitation to these initial conclusions is that the flanker task used in the current experiment employed fewer trials than those used in previous studies where a sex difference was found (e.g. Clayson et al., 2011; Stoet, 2011). Therefore, a lack of sex difference might reflect a lack of precision in measuring the effect. This might also explain why we did not find a main effect of compatibility on the accuracy data in the flanker task. Thus, although the current experiment employed a larger sample size than previous studies using the flanker task, we are still cautious to interpret the lack of evidence for the sex difference in the first experiment.

Further, in the current experiment, we did not find any effect of the type of emotional expression on automatic imitation. These findings add to previous research that shows mixed evidence for a link between the emotional expression of the interacting partner and the tendency to automatically imitate (Crescentini et al., 2011; Grerucci et al., 2013; Rauchbauer et al., 2015). Finally, if we turn to consider the effects of stable personality measures, a clear picture begins to emerge. Even in a more socially meaningful context where emotional expressions are signalled, we further support the claim that imitative control in general (across the entire group of participants), shows a general invariance to stable dimensions of personality like narcissism, agreeableness, extraversion, autistic-like and schizotypal traits (Butler et al., 2015; Cracco et al., 2018b). Of course, it is possible that the emotional expressions failed to add to the social context of the task – the face image signalling the emotional expression was presented simultaneous to the hand image, however they were two separate images. Further, it has been recently suggested that individuals cannot readily infer a person's emotional state from their facial movements or expressions (Barrett et al., 2019). It is possible, therefore, that participants ignored the face image, did not think of the hand as connected to the face image, and did not infer the emotional state of the stimuli. Even if this were true, however, we add a further large dataset to the prior work (Butler et al., 2015; Cracco et al., 2018b), which all show that personality variables have little relationship to performance on the imitation task in general. We suggest, therefore, that studies purporting alternative patterns of relationship between imitation and personality measures in general across the population (Chartrand & Bargh, 1999; Hogeveen & Obhi, 2013; Obhi et al., 2013) perform powerful replications to enable a cumulative science to develop (Munafo et al., 2017; Zwaan et al., 2018).

Although there were no clear main effects of personality across the entire group, there was some suggestive evidence that the effect of personality on imitation differed by

sex. Given prior evidence linking automatic imitation and alexithymia, we expected that an increase in automatic imitation would be predicted by a decrease in alexithymia (Sowden et al., 2016). In the current experiment, this was true only for males, and not for females. We further predicted that an increase in empathic concern and perspective taking would predict an increase in automatic imitation (Chartrand & Bargh, 1999). However, the current findings suggest that a decrease in empathic concern predicts the compatibility effect in males, but not females, and a decrease in perspective taking predicts the compatibility effect in females, but not males. It has been suggested that males score higher on measures of alexithymia as compared to females (Levant et al., 2009), and females score higher on empathic concern and perspective taking as compared to males (Christov-Moore et al., 2014; Van der Graaff et al., 2013). We had no *a priori* hypotheses however as to whether females and males would show a link between personality and imitation in different directions. In addition, these sex*trait interactions were small effects and contributed to only an additional 1.3% (alexithymia) and 3.3 % (empathy) of the model. Thus, before making any firm conclusions, these results require replication in order to confirm that they do not reflect false positives as a result of sampling error.

Overall, however, these initial results from Experiment 1 demonstrate that cognitive control systems may operate differently on some (sex) but not other (personality) stable dimensions of individuals.

Experiment 2

Introduction

In the second experiment, we extend findings from Experiment 1 and address its limitations by making the following changes. First, in the automatic imitation task, stimuli were displayed orthogonal to the response hand in order to minimise the effect of spatial compatibility. Thus, instead of the general compatibility effect, we now investigate the sex difference on the orthogonal compatibility effect. The orthogonal compatibility effect allows us the measure automatic imitation dissociated from right-left spatial compatibility effects, thus allowing for a more precise measure of the imitative effect. Second, we again compare between males and females on the flanker task but increase the number of trials such that both the imitation and flanker tasks are equal. Similar to Experiment 1, we performed a correlational analysis to see whether flanker and orthogonal compatibility effects were related to each other or not.

In Experiment 1, three sex*trait interactions, which covered empathic concern, perspective taking, and alexithymia, predicted the general compatibility effect. Thus, in order to further confirm these findings, we included empathy (empathic concern and perspective taking) and alexithymia measures in Experiment 2 to investigate whether these traits modulated the orthogonal compatibility effect.

Method

Participants.

Two hundred and thirty-eight participants took part in this experiment for monetary compensation (£6) or course credit. All participants provided informed consent and had normal or corrected-to-normal vision. Approval was obtained from the Research Ethics and Governance Committee of the School of Psychology at Bangor University. One participant was excluded because data on only half the trials was recorded on the flanker task.

Participants were excluded if performance was three standard deviations away from the group mean average performance per condition in terms of accuracy or reaction time (N=15 for the imitation task, N=21 for the flanker task). For the imitation task, the final sample included 223 participants (107 males, $Mean_{age}$ = 20.0, SD_{age} = 4.33; $Mean_{age}$ and SD_{age} are based on 203 participants as some participants did not enter their age in the demographic questionnaire). For the flanker task, the final sample included 217 participants (101 males, $Mean_{age}$ = 20.7, SD_{age} = 4.31; $Mean_{age}$ and SD_{age} are based on 198 participants).

Stimuli, tasks, and procedure.

Automatic imitation task. The automatic imitation task was similar to the one used in Experiment 1, with the following changes: one, no face image was presented during the task (Figure 1). Two, the hand stimuli were presented orthogonal to the response (Figure 1). Three, there were 360 trials in total, which comprised six blocks of 60 trials, each of which included 30 compatible and 30 incompatible trials.

Flanker task. The flanker task was the same as Experiment 1 with only one change – participants completed 360 trials in total, with 6 blocks of 60 trials each (30 compatible and 30 incompatible trials; Figure 1).

The order of the tasks was counterbalanced such that half the participants did the flanker task first, whereas the remaining half did the imitation task first.

Questionnaires. Participants also completed two self-report questionnaires which included the Interpersonal Reactivity Index (IRI; Davis, 1980), and the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994). For more details on the measures used, see the Supplementary Material.

Data analysis.

Accuracy and RT on the imitation and flanker tasks were recorded in the same way as Experiment 1 and only correct trials were used to calculate RT. Trials on which participants responded incorrectly, i.e. lifted the wrong finger, responded after 2000 ms, or before target onset (imitation = 5.59%; flanker = 5.97%) were all excluded from the analysis.

Data were analysed in the same way as Experiment 1. For the imitation task, a Sex*Compatibility interaction showing a higher compatibility effect for females compared to males would indicate that the sex difference on the imitation task persists even when stimuli are presented orthogonally to the response. Alternatively, similarly sized compatibility effects between the sexes would suggest that reducing the spatial component of the task largely removes the sex difference.

Results

Automatic imitation task.

Accuracy. Average accuracy for both males and females for both compatible and incompatible trials was over 92% (see Figure 9, Supplementary Table 6). The 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 9) showed a main effect of compatibility such that participants were more accurate on compatible trials than incompatible trials ($F(1, 221) = 96.22$, $p<.001$, $\eta_p^2 = 0.30$). The effect size for the main effect of compatibility is conventionally considered to be large. There was no significant main effect of sex ($F(1, 221) = 1.87$, $p = 0.17$, $\eta_p^2 = 0.008$) and no significant Compatibility*Sex interaction ($F(1, 221) = 0.14$, $p = 0.71$, $\eta_p^2 <0.001$) and the effect sizes were close to zero (see Supplementary Table 7).

Reaction time. Mean reaction times were between 435 and 485 milliseconds for both males and females on both compatible and incompatible trials (see Figure 10, Supplementary Table 6). The 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 10) showed a main effect of compatibility such that participants were slower to respond on incompatible trials than compatible trials (F(1,

221) = 293.18, p<.001, η_p^2 = 0.56). The effect size for the main effect of compatibility is conventionally considered to be large. The main effect of sex was also significant such that females were generally slower than males, but this was a relatively small effect size (F(1, 221) = 4.23, p = 0.040, η_p^2 = 0.02). There was a significant Compatibility*Sex interaction and the effect size is conventionally considered to be a small effect (F(1, 221) = 4.17, p = 0.042, η_p^2 = 0.02; Supplementary Table 7).

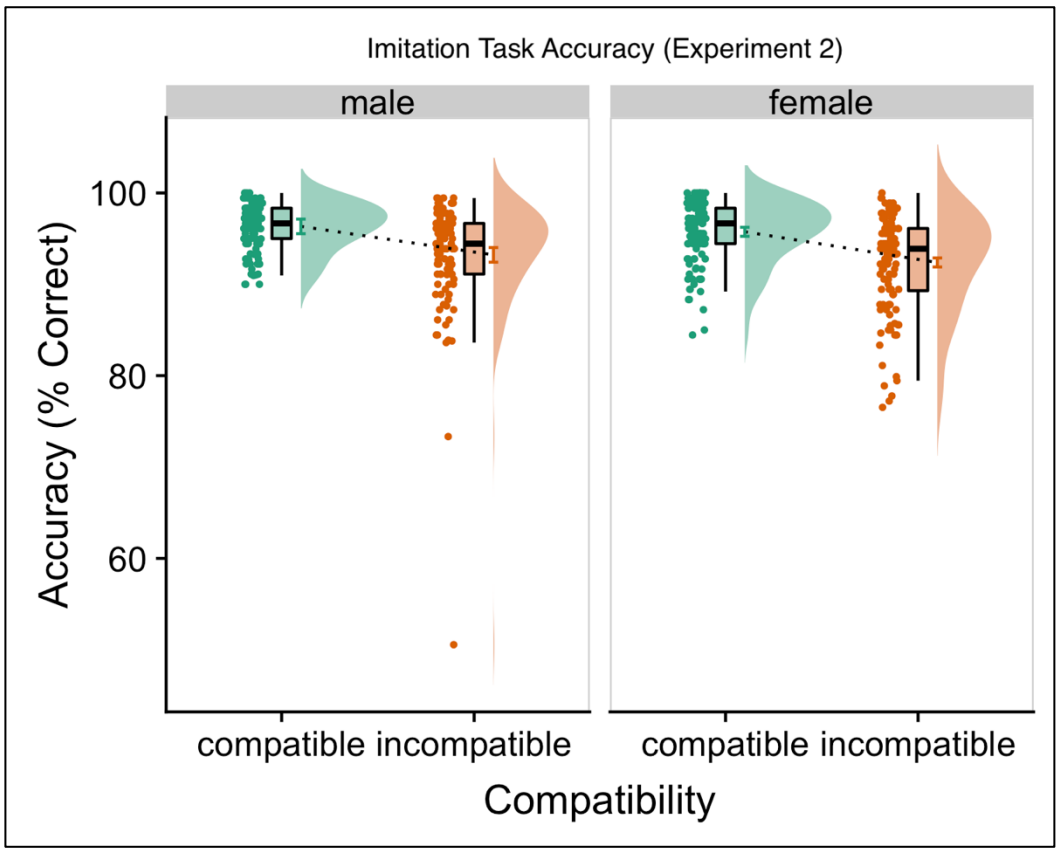


Figure 9. Experiment 2 – Imitation Task Accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females.

In order to interrogate our primary hypothesis regarding sex differences in the imitation task, we computed compatibility effects separately for males and females, and then compared them to each other. For both males and females, compatibility effects had a large effect size (Cohen's d_z > 1.0) and the lower bound of the 95% confidence interval was at least 0.82. When compatibility effects for males and females were directly compared to each other, we found a mean difference of 6.98 ms in the direction that was

predicted i.e. the compatibility effect for females was greater than the compatibility effect for males with the lower bound of the 95% confidence interval above zero (Mean Difference = 6.98 ms, 95% CI[1.34, ∞], Cohen's d = 0.27, 95% CI[0.05, ∞]; Figure 10, Table 1B). The effect size was a small-to-medium effect, with the lower bound of the 95% CI at 0.05.

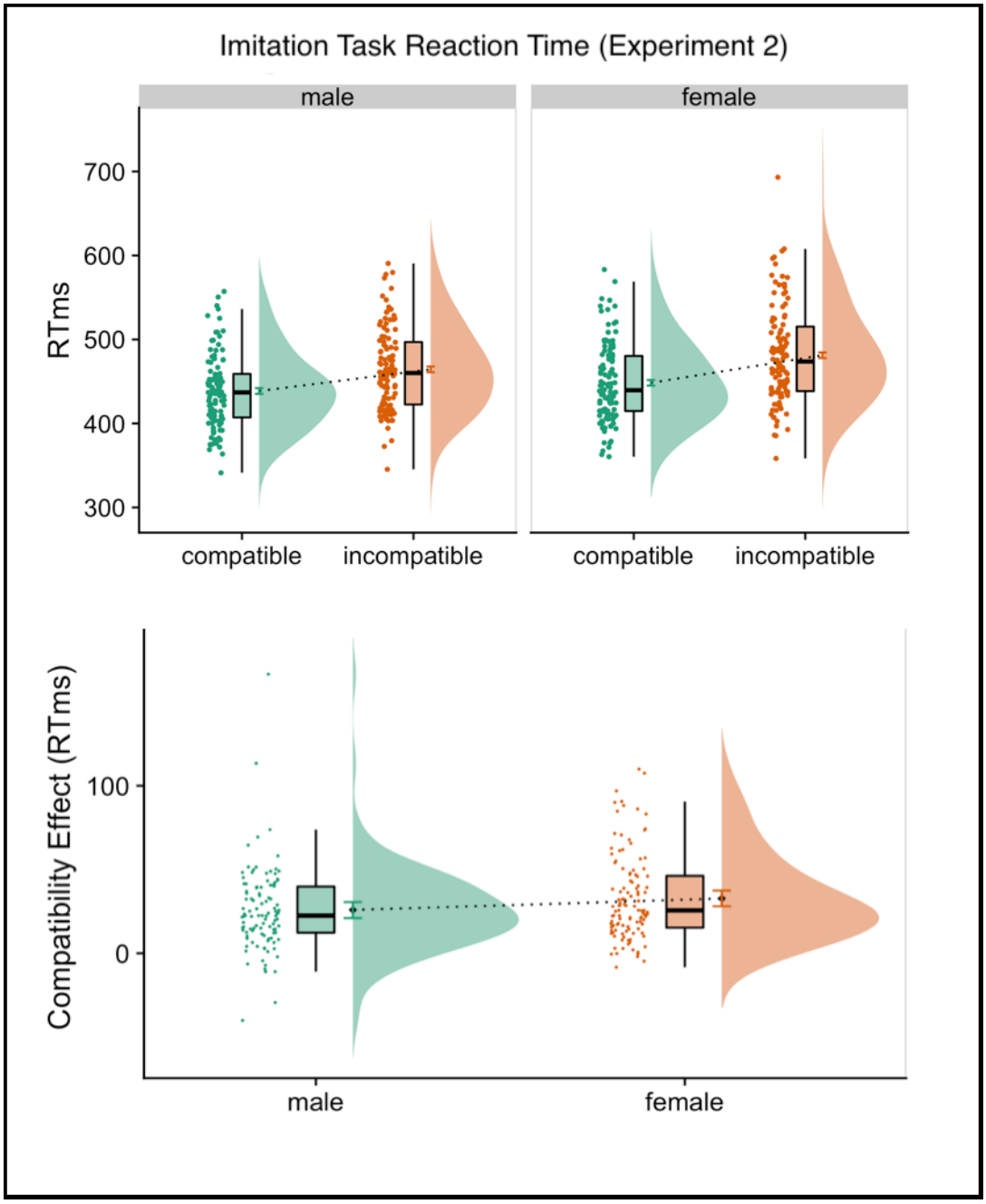


Figure 10. *Experiment 2 – Imitation Task Reaction Time.* Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females. The lower panel shows the compatibility effect for both males and females. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds.

The absolute size of the difference between the sexes as measured in original units (i.e., ms) is smaller than Experiment 1, as the orthogonal compatibility effect is smaller than the general compatibility effect measured in Experiment 1. Indeed, when measured in original units, the compatibility effect in Experiment 2 is approximately half the size of Experiment 1 and the same is true for the sex difference in compatibility effect between the two experiments. However, the standardised effect size for the sex difference is nearly identical across the two experiments (Exp. 1 = 0.28; Exp. 2 = 0.27). Therefore, when measured in comparable units, which account for differences in absolute values, these results suggest that the sex difference measured is quite consistent across experiments. In sum, the orthogonal compatibility effect on the imitation task differed as a function of sex in the same manner and to a similar degree as Experiment 1, such that females had a greater orthogonal compatibility effect than males.

Flanker task.

Accuracy. Average accuracy was over 88% for both males and females on both compatible and incompatible trials (see Figure 11, Supplementary Table 6). A 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 11) showed a main effect of compatibility ($F(1, 215) = 151.33$, $p<.001$, $\eta_p^2 = 0.41$). The main effect of compatibility showed that participants were more accurate on compatible trials compared to incompatible trials. The main effect of sex was also significant such that females showed a lower accuracy overall compared to males ($F(1, 215) = 5.78$, $p = 0.017$, $\eta_p^2 = 0.03$). The Compatibility*Sex interaction approached significance ($F(1, 215) = 3.17$, $p = 0.076$, $\eta_p^2 = 0.01$) with the difference in accuracy between compatible and incompatible trials being greater for females compared to males. The effect sizes for both the main effect of sex, and the interaction were relatively small (see Supplementary Table 7).

Reaction time. Mean reaction times for both males and females for both compatible and incompatible conditions was between 400 and 500 milliseconds (see Figure 12, Supplementary Table 6). A 2 (compatibility: compatible, incompatible) x 2 (sex: male, female) mixed ANOVA (Figure 12) showed a main effect of compatibility such that participants were slower to respond on incompatible trials than compatible trials ($F(1, 215) = 1986.89$, $p<.001$, $\eta_p^2 = 0.90$). The effect size for the main effect of compatibility is conventionally considered to be large. There was no significant main effect of sex and

the effect size was zero (F(1, 215) = 0.03, p = 0.854, η_p^2 < 0.001). There was no significant Compatibility*Sex interaction (F(1, 215) = 0.05, p = 0.816, η_p^2 < 0.001) and the effect size was close to zero (see Supplementary Table 7).

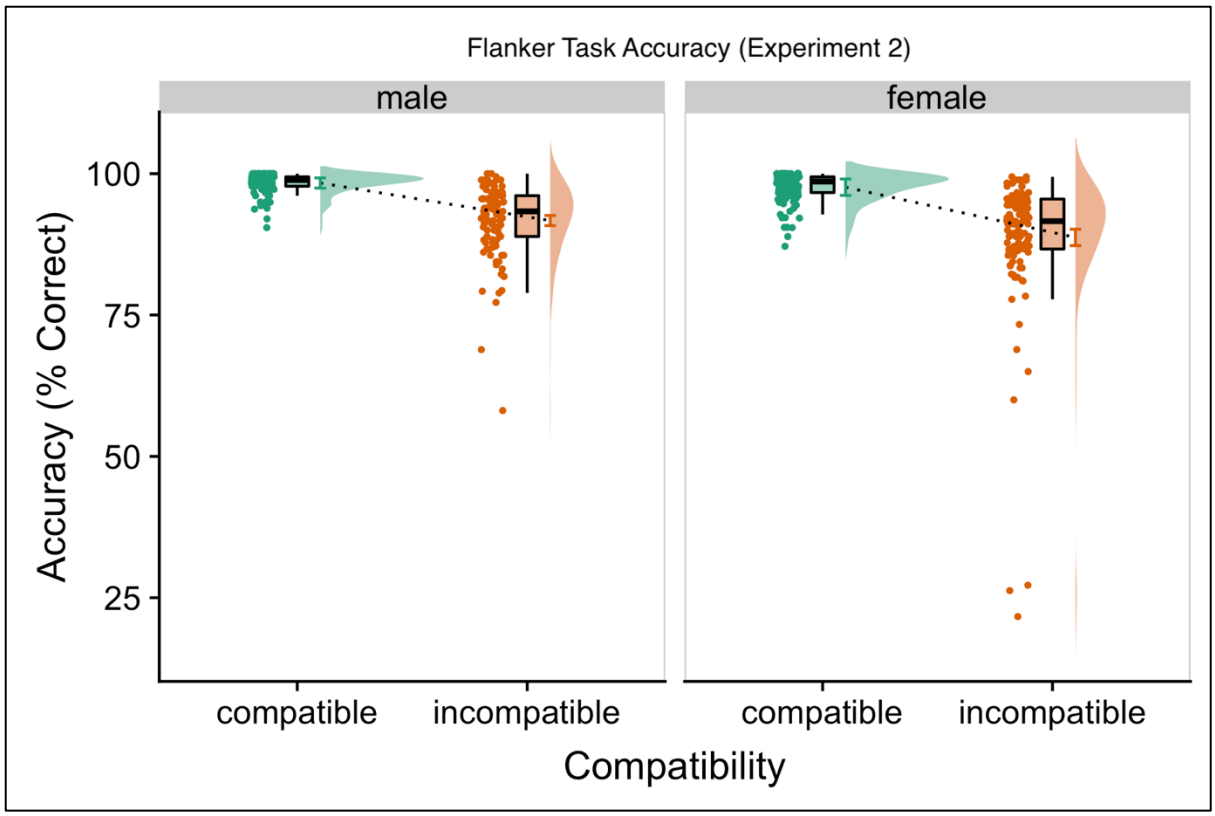


Figure 11. Experiment 2 – Flanker Task Accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females.

To explore sex differences in the flanker task further, compatibility effects were computed separately for males and females, and then compared to each other. For both males and females, compatibility effects had a large effect size (Cohen's d_z > 2.9) with the lower bound of the 95% confidence interval at least 2.62. When compatibility effects for males and females were directly compared to each other, females showed a higher compatibility effect than males, but the effect size was very small, with the lower bound of the 95% confidence interval reaching -6.01ms (Mean Difference = 0.98 ms, 95% CI[-6.01, ∞], Cohen's d = 0.03, 95%CI[-0.19, ∞]; Figure 12, Table 1B). The effect size was close to zero with the lower bound of the confidence interval at -0.19 (below zero). A Bayesian independent samples t-test showed that the null was 6 to 7 times more likely than the alternative hypothesis (BF_{01} = 6.58). Thus, although both males and females separately

showed a compatibility effect, there was a negligible difference between males and females on the flanker compatibility effect.

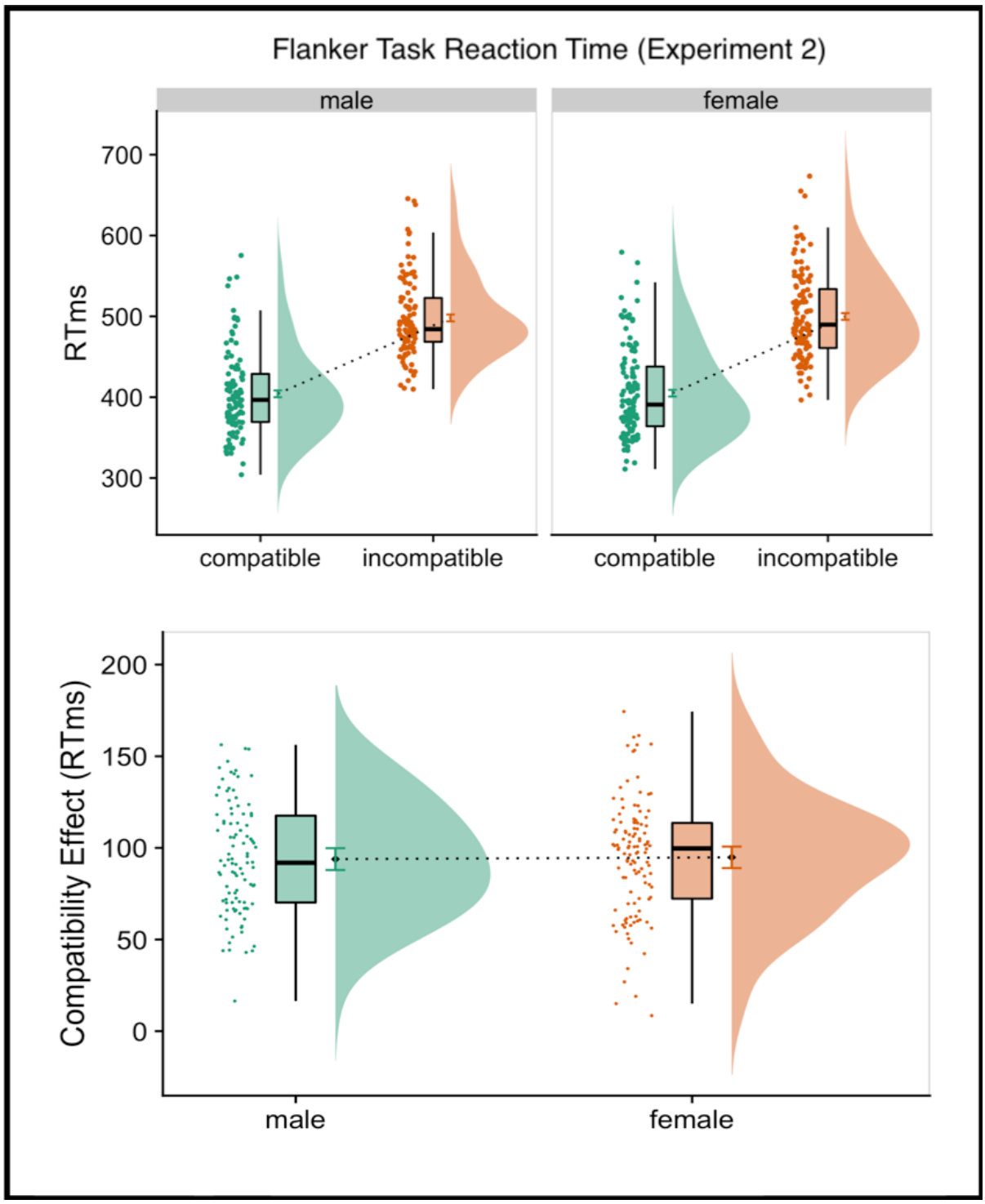


Figure 12. Experiment 2 – Flanker Task Reaction Time. Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females. The lower panel shows the compatibility effect for both males and females. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds.

Correlational analysis.

In order to investigate whether the flanker and imitation compatibility effects were correlated, a one-tailed skipped correlation was performed. For the correlational analysis, only those participants who performed both the tasks were included in the analysis (N=205). As in Experiment 1, we also performed a more robust correlation analysis. The data was not normally distributed, but was homoscedastic. Thus, we performed a skipped Spearman correlation analysis on 191 participants as 14 bivariate outliers were detected. Results indicated that flanker and imitation compatibility effects showed a weak positive correlation that did not pass our statistical threshold (Spearman r(191)=0.07, 95% CI [-0.07, ∞]; Figure 13). Our findings thus suggest that flanker and imitative compatibility effects are largely unrelated, and interference on one task did not predict interference on the other.

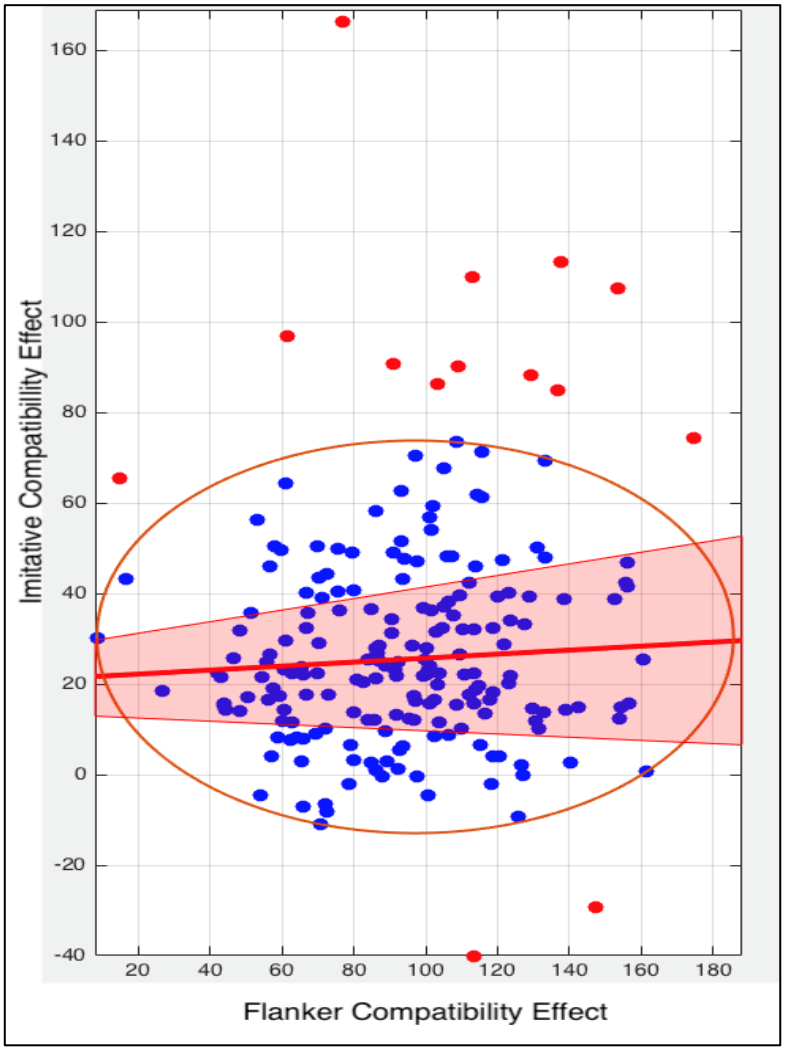


Figure 13. Experiment 2 - Correlation Analysis. A skipped Spearman correlation shows a small positive correlation between the flanker and imitative compatibility effects that does not pass our statistical thresholding. Abbreviations: RTms (reaction time in milliseconds). Dots in red are the bivariate outliers.

Multiple regression analyses.

We also investigated the relationship between personality variables (empathy and alexithymia) and the orthogonal compatibility effect as measured on the SRC task. Tests for multicollinearity indicated that a very low level of multicollinearity was present (*VIF* for all predictor variables < 2). The base model (which included mean RT, sex, and the mean RT*Sex interaction) explained 16.2% of the variance in the congruency effect ($F(3,204)=13.13$, $p<.001$, $R2=.16$, $f^2 = 0.19$) and indicated a medium effect. Mean RT predicted the orthogonal compatibility effect ($B=0.19$, $SEB=0.03$, $t(204)=5.36$, $p<.001$, [0.12; 0.26]) with increasing CE as mean RT increased. Sex did not predict the orthogonal compatibility effect ($B=2.14$, $SEB=1.67$, $t(204)=1.27$, $p=.21$, [-1.19; 5.47]). The mean RT * sex interaction was a marginally significant predictor ($B=0.07$, $SEB=0.03$, $t(204)=1.90$, $p=.06$, [-0.002; 0.13]), suggesting that increases in mean RT predicted larger increases in the compatibility effect for females ($B=0.25$, $SEB=0.05$, $t(204)=5.45$, $p<.001$) compared to males ($B=0.12$, $SEB=0.05$, $t(204)=2.33$, $p<.001$).

Alexithymia and empathy (empathic concern and perspective taking) did not predict the orthogonal compatibility effect above and beyond the base model (all ps>0.3, all CIs overlapping with zero; see Figure 14). Effect sizes attributable to the addition of the personality variables (beyond the base model) indicated extremely small effects (Cohen's $f^2 = <.001$ for alexithymia and Cohen's $f^2 = .005$ for empathy). The multiple regression models are summarized in Supplementary Table 8. Zero-order correlations are also consistent with the findings from the multiple regression analyses (see Supplementary Table 9, Supplementary Figure 4).

To evaluate the sex*trait interaction terms, we computed additional models – each model consisted of the base model, one trait predictor, and the sex*trait interaction term. None of the sex*trait interaction terms predicted the orthogonal compatibility effect above and beyond the base model (all ps>0.3, all CIs overlapping with zero; Figure 8B, Supplementary Figure 5). Effect sizes attributable to the addition of the sex*trait interaction terms (beyond the base model) indicated extremely small effects (Cohen's $f^2 = 0.01$ for both alexithymia and empathy). The multiple regression models are summarized in Supplementary Table 10.

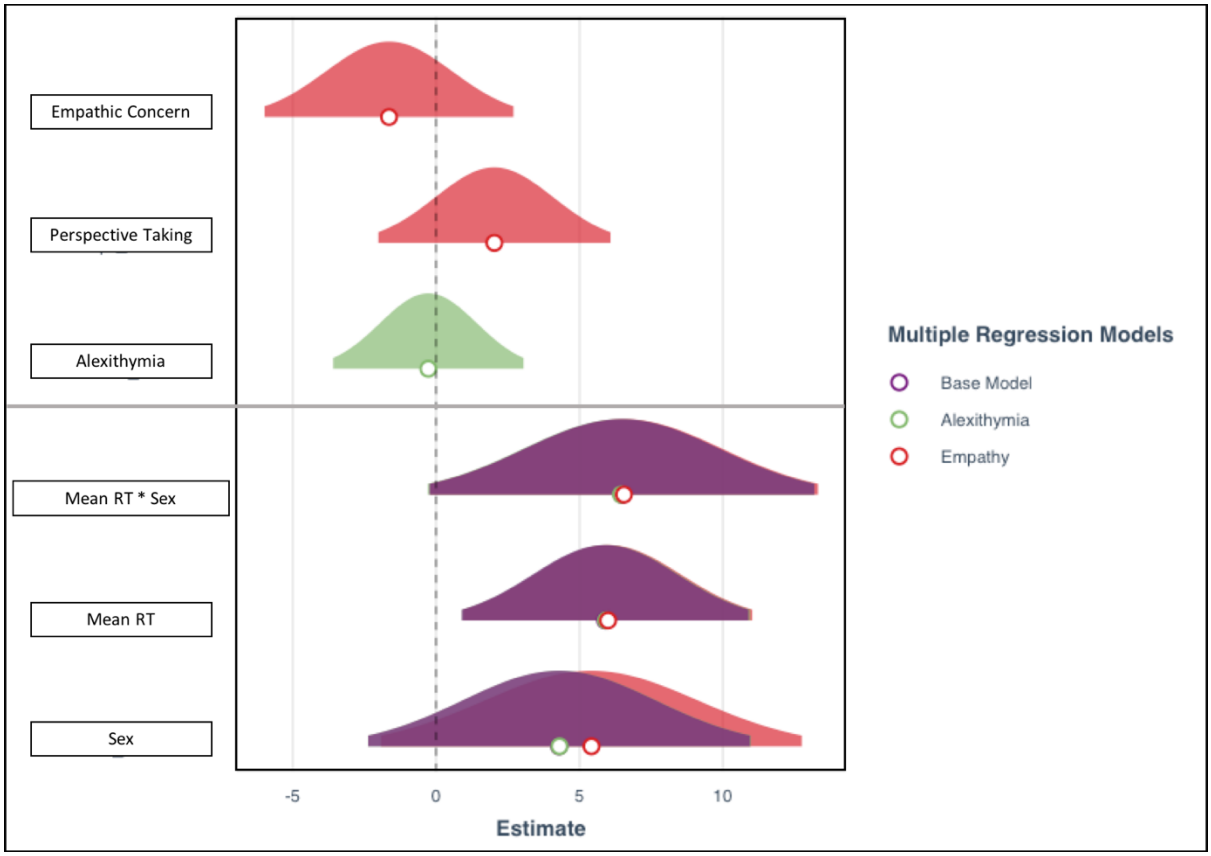


Figure 14. Experiment 2 – Multiple Regression Analyses. Values of standardised coefficients are plotted for each predictor variable (personality trait) along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

Discussion

As in Experiment 1, results indicated a clear sex difference in the orthogonal compatibility effect such that females showed a higher orthogonal compatibility effect compared to males on the automatic imitation task. The sex difference persisted on the imitation task in spite of presenting stimuli orthogonal to the response. However, this sex difference on the RT compatibility effect was not found on the flanker task even after increasing the number of trials. Further, the correlational analysis suggested that flanker and orthogonal compatibility effects were only marginally correlated with each other and explained only 0.8% of the variance. Thus, greater interference on one task is able to predict only a very small amount of interference on the other task.

Thus, across Experiments 1 and 2, we show a lack of consistent evidence for a sex difference in the flanker task. However, the interpretation of the sex difference on the imitation task still has two potential limitations. One, while the presentation of

orthogonal stimuli reduces spatial compatibility effects on the left-right axis, they do not rule out the possibility of orthogonal spatial compatibility effects i.e. the propensity of participants to show an advantage for an up-right and down-left pairing (Cho & Proctor, 2003; Weeks & Proctor, 1999; Weeks, Proctor, & Beyak, 1994). For instance, when stimuli were presented orthogonal to the response hand (see Figure 1), the index finger was always below the middle finger, and the participant's index finger was to the left side of space. Thus, a preference for responding to "up" stimuli with a right response and "down" stimuli with a left response may be observed along with imitative effects on the automatic imitation task used in the current experiment. The sex difference may therefore reflect a difference in orthogonal spatial effects as opposed to purely imitative effects.

Two, the stimuli used in both Experiments 1 and 2 were those of a female model. We did not manipulate the sex of the stimulus, and therefore, the sex difference can either reflect a genuine difference between males and females, or an in-group bias. A difference between male and female participants (irrespective of the sex of the stimulus) would reflect distinct (or partially distinct) cognitive mechanisms underlying imitative or spatial control as a function of sex. On the contrary, an in-group bias or own-sex bias would suggest that sex differences as evidenced previously on the automatic imitation task (Butler et al., 2015; Genschow et al., 2017) do not actually reflect a sex difference - females show a higher compatibility effect because they favour members of the in-group i.e. of their own sex compared to members of the out-group i.e. of the opposite sex (Brown, 1995; Gleibs et al., 2016; Rauchbauer et al., 2015; Rudman & Goodwin, 2004).

Finally, in terms of personality measures, empathy and alexithymia (and sex*trait interactions) did not modulate the orthogonal compatibility effect. Although we found suggestive evidence in Experiment 1 for a small link between personality traits (alexithymia and empathy) and imitation that differed between the sexes, the current experiment did not replicate these findings. Therefore, overall, these results provide limited support for a link between personality traits and automatic imitation and confirm and replicate findings from previous large sample studies (Butler et al., 2015; Cracco et al., 2018) that suggest automatic imitation is largely invariant to stable traits of personality.

Experiment 3

Introduction

Experiment 3 addressed two remaining issues. First, we measured the imitative compatibility effect independently from the spatial compatibility effect, in order to estimate whether the sex difference reflects a spatial or more specialised (social) aspect of cognitive control. Second, we assessed the extent to which the sex difference reflects a basic difference between males and females and/or an in-group bias based on sex.

To separate imitative and spatial components of the task, we used a modified version of the SRC task of automatic imitation that allowed us to manipulate imitative and spatial effects separately (Bertenthal et al., 2006; Boyer et al., 2012; Catmur & Heyes, 2011). A sex difference on spatial compatibility alone would indicate that the sex difference observed in Experiments 1 and 2 can be explained by differences associated with processing spatial information. Alternatively, a sex difference on imitative compatibility alone, would suggest that a greater compatibility effect for females reflects a difference in the control of automatic imitation specifically.

To compare a sex difference account with an in-group bias account of our findings so far, we manipulated the sex of the stimuli used in the SRC task and again tested male and female participants. A greater compatibility effect for females for female stimuli compared to male stimuli would indicate that an own-sex bias contributes to the sex difference observed on the automatic imitation task. Alternatively, a sex difference on the task and relative invariance to the sex of the stimuli would suggest that there is a basic control mechanism that differs between males and females that seems resistant to possible contextual factors, such as group biases.

In order to investigate whether personality variables influence automatic imitation, in Experiment 3, we included all personality variables included in Experiment 1 (alexithymia, empathy, autistic-like and schizotypal traits, narcissism, extraversion, and agreeableness). In Experiment 1 and 2, the compatibility effect measured on the imitation task was a composite of spatial and imitative effects. Therefore, the invariance of the compatibility effect may be related to spatial effects as opposed to imitative effects. Therefore, we included all the personality measures in order to investigate whether imitative compatibility when measured independently of spatial effects is also invariant to stable personality traits.

Method

Participants.

Two hundred and one participants took part in this experiment for monetary compensation (£6) or course credit. All participants provided informed consent and had normal or corrected-to-normal vision. Approval was obtained from the Research Ethics and Governance Committee of the School of Psychology at Bangor University. Participants were excluded if performance was 3 standard deviations away from the group mean average performance per condition in terms of accuracy or reaction time on the imitation task (N = 12). The final sample included 189 participants (97 males, $Mean_{age}$ =21.4, SD_{age} = 4.08, age range = 18 to 42) ($Mean_{age}$ and SD_{age} are based on 182 participants as 7 participants did not enter their age in the demographics questionnaire).

Stimuli, tasks, and procedure.

Automatic imitation task. The automatic imitation task was similar to the one used in Experiment 2, with the following changes: one, stimuli were not presented orthogonally to the response. Two, we calculated an imitative compatibility effect independent of the spatial compatibility effect (Catmur & Heyes, 2011). For this, both left- and right-hand images were used as stimuli, but participants always responded with their right hand. This resulted in eight trial types and four conditions of interest (Figure 1):

1. imitatively and spatially compatible (for example, when participants are cued to lift their index finger, and watch an index finger lift of the left hand, the observed finger movement is both spatially and imitatively compatible to the executed movement),
2. imitatively and spatially incompatible (for example, when participants are cued to lift their index finger, and watch a middle finger lift of the left hand, the observed finger movement is both spatially and imitatively incompatible to the executed movement),
3. imitatively compatible and spatially incompatible (for example, when participants are cued to lift their middle finger, and watch a middle finger lift of the right hand, the observed finger movement is imitatively compatible, but spatially incompatible to the executed movement),
4. imitatively incompatible and spatially compatible (for example, when participants are cued to lift their middle finger, and watch an index finger lift of the right hand, the observed finger movement is imitatively incompatible, but spatially compatible to the executed movement).

Thus, participants performed the same (imitatively compatible) or different (imitatively incompatible) movement on the same (spatially compatible) or different (spatially incompatible) side of space.

A third change in comparison to Experiment 2, is that in order to investigate whether the sex difference was due to an own-sex bias, the hand stimuli presented included 4 female and 4 male hands. The hand stimuli were chosen based on a pilot study. In the pilot study (see Supplementary Material), eighteen hand stimuli were rated by 51 participants on a scale of 1 to 9 with one being most masculine, 5 being neutral, and 9 being most feminine. Four hand stimuli rated as most masculine, and four hand stimuli rated as most feminine were chosen for the current experiment. There were 360 total trials, with 90 trials per condition. Timing information and pseudo-randomisation was the same as in Experiment 1 and 2.

Questionnaires. Participants also completed a range of self-report questionnaires which included the Mini International Personality Item Pool (mini IPIP; Donnellan et al., 2005; the Short Autism Spectrum Quotient (AQ-10 Adult; Baron-Cohen et al., 2001), the Brief Schizotypal Personality Questionnaire (SPQ-B; Raine & Benishay, 2005), the Narcissistic Personality Inventory (NPI-16; Ames et al., 2006), the Hypersensitivity Narcissism Scale (HSNS; Hendin & Cheek, 1997), the Interpersonal Reactivity Index (IRI; Davis, 1980), and the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994). For more details on the measures used, see the Supplementary Material. In order to confirm that participants perceived male and female stimuli differently, participants also rated the hand stimuli used in the experiment after they completed the task. Participants were asked to rate the stimuli on a scale of 1 to 9, with 1 being extremely masculine and 9 being extremely feminine.

Data analysis.

Accuracy and RT on the imitation task were recorded in the same way as Experiment 1 and 2 and only correct trials were used to calculate RT. Trials on which participants responded incorrectly, i.e. lifted the wrong finger, responded after 2000 ms, or before target onset (7.41 %) were all excluded from the analysis. As the primary aim was to investigate the sex difference and own-sex bias in imitative and spatial compatibility (and not compare between the two), we performed analyses separately on the spatial and imitative compatibility effects. For each compatibility effect separately, we performed a 2 (compatibility: incompatible, compatible) x 2 (stimulus sex: male hand, female hand) x 2 (participant sex: male, female) mixed ANOVA on the RT and accuracy

data. Based on prior research (Catmur & Heyes, 2011; Gowen et al., 2016; Marsh et al., 2016; Darda et al., 2018), we expected a main effect of spatial and imitative compatibility such that RT would be higher, and accuracy would be lower on spatially incompatible trials compared to spatially compatible trials, and on imitatively incompatible trials compared to imitatively compatible trials.

In addition, a Sex*Compatibility interaction for spatial compatibility (such that females show a higher spatial compatibility effect that males) would be expected if the sex difference observed in Experiments 1 and 2 was largely driven by the spatial component of the task. In contrast, a Sex*Compatibility interaction for imitative compatibility (such that females show a higher imitative compatibility effect than males) would suggest that the sex difference is largely a reflection of the imitative component of the task.

Alternatively, if the sex difference in the spatial or imitative compatibility effect is because of an own-sex bias, we would expect a three-way interaction (Sex*Compatibility*Stimulus Sex) such that females would be more interfered by a female stimulus, and males would be more interfered by a male stimulus i.e. females would show a higher compatibility effect than males for female stimuli compared to male stimuli.

As in Experiment 1, the interaction effect was central to testing our primary hypotheses, and thus, we calculated compatibility effects for male and female hand stimuli separately and independently for both male and female participants. To do so, we computed the mean difference and 95% confidence intervals between compatible and incompatible conditions across the levels of stimulus sex and participant sex. Spatial compatibility was calculated as RT on spatially incompatible trials minus RT on spatially compatible trials. Imitative compatibility was calculated as RT on imitatively incompatible trials minus imitatively compatible trials. In order to directly estimate the size of the difference in spatial and imitative compatibility effects between males and females, we then again computed the mean differences between the sexes and 95% confidence intervals. We used one-tailed 95% confidence intervals as we had a directional hypothesis that females would have a higher spatial or imitative compatibility effect than males. For the secondary analyses, multiple regression analyses were performed in the same way as Experiments 1 and 2 in order to investigate whether personality variables and sub-clinical traits modulate automatic imitation when measured independent of spatial effects.

Results

Spatial compatibility.

Accuracy. Average accuracy was over 92% for both males and females for all conditions of compatibility and stimulus sex (see Figure 15, Supplementary Table 11). A 2 (compatibility: incompatible, compatible) x 2 (stimulus sex: male hand, female hand) x 2 (participant sex: male, female) mixed ANOVA showed a main effect of compatibility such that participants were more accurate on compatible trials than incompatible trials ($F(1,187) = 563.35$, $p<.001$, $\eta_p^2 = 0.75$; Figure 15). The effect size of the main effect of compatibility was large. There was a main effect of stimulus sex such that participants were more accurate when observing male hand stimuli as compared to female hand stimuli ($F(1,187) = 335.47$, $p<.001$, $\eta_p^2 = 0.64$). There was a significant Compatibility*Stimulus Sex interaction such that the difference in accuracy between incompatible and compatible trials was overall bigger for female stimuli compared to male stimuli ($(F(1,187) = 202.31$, $p<.001$, $\eta_p^2 = 0.52$), Figure 15). All other main effects and interactions were not significant and effect sizes were relatively small (see Supplementary Table 12).

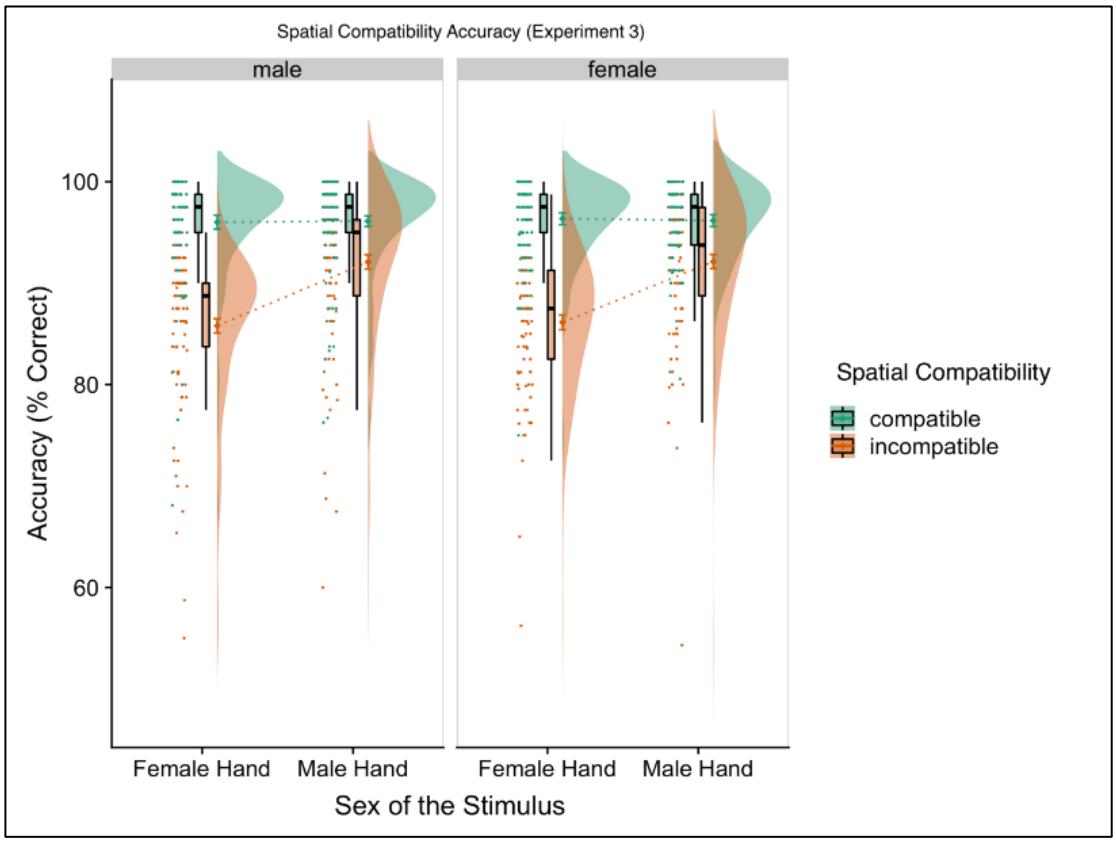


Figure 15. Experiment 3 – Spatial Compatibility Accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females when responding to both male and female hand stimuli.

Reaction time. Mean reaction times for both males and females on all conditions of compatibility and stimulus sex were between 415 to 475 milliseconds (see Figure 16, Supplementary Table 11). A 2 (compatibility: incompatible, compatible) x 2 (stimulus sex: male hand, female hand) x 2 (participant sex: male, female) mixed ANOVA (Figure 16) showed a main effect of compatibility such that participants were slower to respond on spatially incompatible trials compared to spatially compatible ($F(1,187) = 459.71$, $p<.001$, $\eta_p^2 = 0.71$). The effect size of the main effect of compatibility was large. There was a main effect of stimulus sex, which had a medium effect size, and suggested that overall participants responded slower to female hand stimuli than male hand stimuli ($F(1,187) = 5.63$, $p = 0.019$, $\eta_p^2 = 0.03$). There was a significant Sex*Compatibility interaction and the effect size was a small-to-medium effect ($F(1,187) = 4.24$, $p = 0.041$, $\eta_p^2 = 0.02$). To interrogate the sex difference in spatial compatibility, we computed the difference in compatibility effects between males and females, collapsed across all conditions of stimulus sex. Females showed a higher compatibility effect than males by 6.53 ms, and the lower bound of the 95% CI was over zero at 2.02ms (Mean difference = 6.53ms, 95% CI [2.02, ∞], Cohen's d = 0.24, 95% CI [0.07, ∞]; Figure 19A, Table 1C). The effect size was a small-to-medium effect, and the lower bound of the 95% CI was above zero at 0.07.

The three-way (Compatibility*Stimulus Sex*Sex) interaction was not significant ($F(1,187) = 1.77$, $p=0.185$, $\eta_p^2 = 0.01$). However, there was a trend for females showing a higher compatibility effect for female hand stimuli compared to male hand stimuli, and males showing a higher compatibility effect for male hand stimuli compared to female hand stimuli, although the effect size was close to zero (Figure 16). No other main effects or interactions were significant (see Supplementary Table 12). In order to investigate whether an in-group bias explains the sex difference in spatial compatibility, we computed compatibility effects on all levels of participant sex and stimulus sex. For both males and females, spatial compatibility effects were present when observing both male (Cohen's $d_z > 1.25$) as well as female stimuli (Cohen's $d_z > 1.10$). There was a trend for females showing a higher compatibility effect for female stimuli compared to male stimuli (Mean difference = 2.74 ms, 95% CI [-2.40, ∞], Cohen's $d_z = 0.09$, 95% CI[-0.08, ∞]), and for males showing a higher compatibility effect for male stimuli compared to female stimuli (Mean difference = 3.19, 95% CI [-2.12, ∞], Cohen's $d_z = 0.10$; 95% CI[-0.07, ∞]), but these were relatively small effect sizes (see Figure 16).

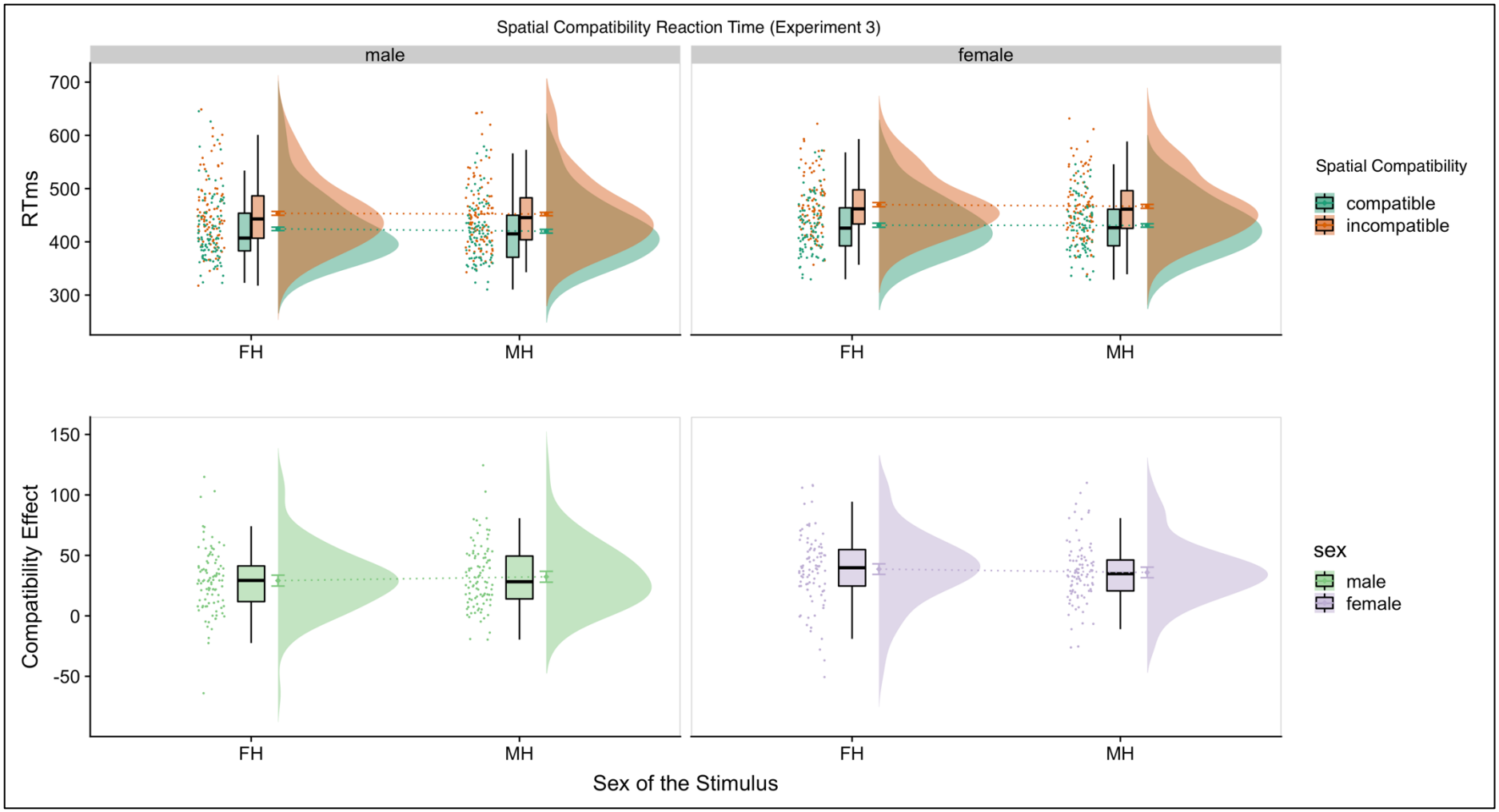


Figure 16. Experiment 3 – Spatial Compatibility Reaction Time. Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females, when responding to both male and female hand stimuli. The lower panel shows the compatibility effect for both males and females when responding to both male and female hand stimuli. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds, MH=male hand stimuli, FH=female hand stimuli.

Imitative compatibility.

Accuracy. Average accuracy for both males and females for all conditions of stimulus sex and compatibility was above 87% (see Figure 17, Supplementary Table 11). A 2 (compatibility: incompatible, compatible) x 2 (stimulus sex: male hand, female hand) x 2 (participant sex: male, female) mixed ANOVA (Figure 17) showed a main effect of compatibility ($F(1,187) = 205.65$, $p<.001$, $\eta_p^2 = 0.52$) such that participants were more accurate on compatible trials than incompatible trials. The effect size of the main effect of compatibility was large. There was a main effect of stimulus sex ($F(1,187) = 335.47$, $p<.001$, $\eta_p^2 = 0.64$) such that participants were more accurate when observing male hand stimuli as compared to female hand stimuli. There was a significant Compatibility*Stimulus Sex interaction ($F(1,187) = 162.98$, $p<.001$, $\eta_p^2 = 0.46$) such that the difference in accuracy between compatible and incompatible trials was bigger for female stimuli compared to male stimuli (see Figure 17). All other main effects and interactions were relatively small or close to zero (see Supplementary Table 12).

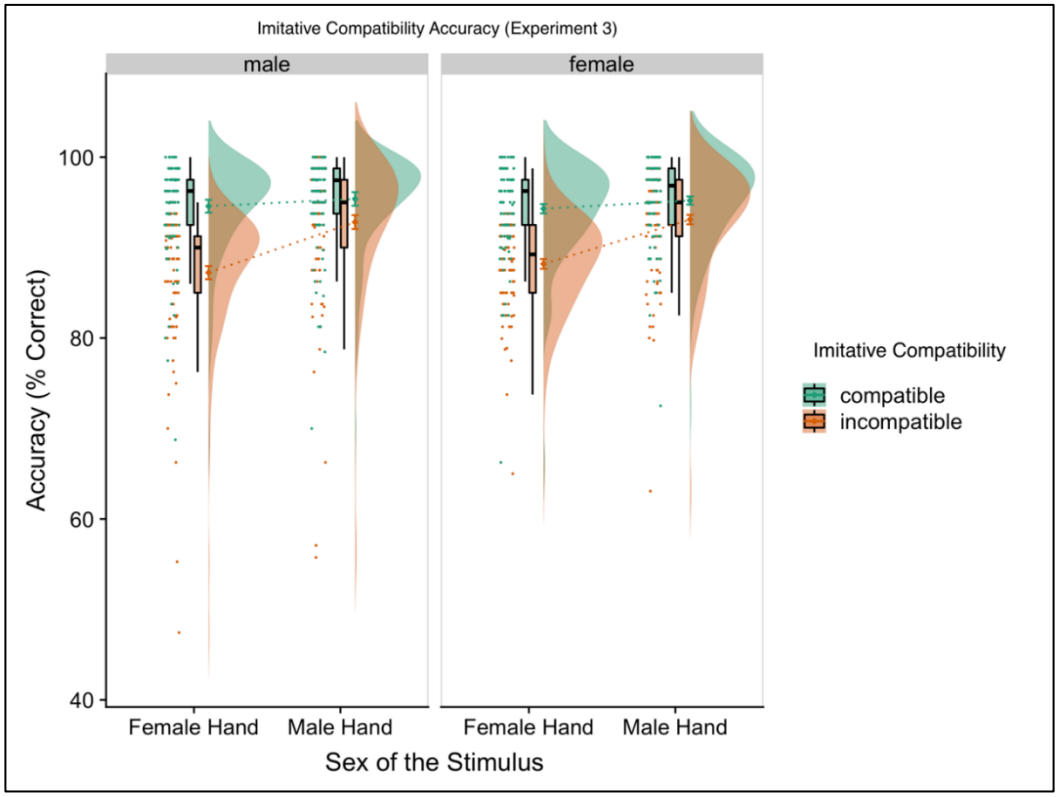


Figure 17. Experiment 3 – Imitative Compatibility Accuracy. Accuracy is reported in percentage of correct responses. Error bars show 95% confidence intervals. Green and orange markers show individual data points for compatible and incompatible conditions for males and females when responding to both male and female hand stimuli.

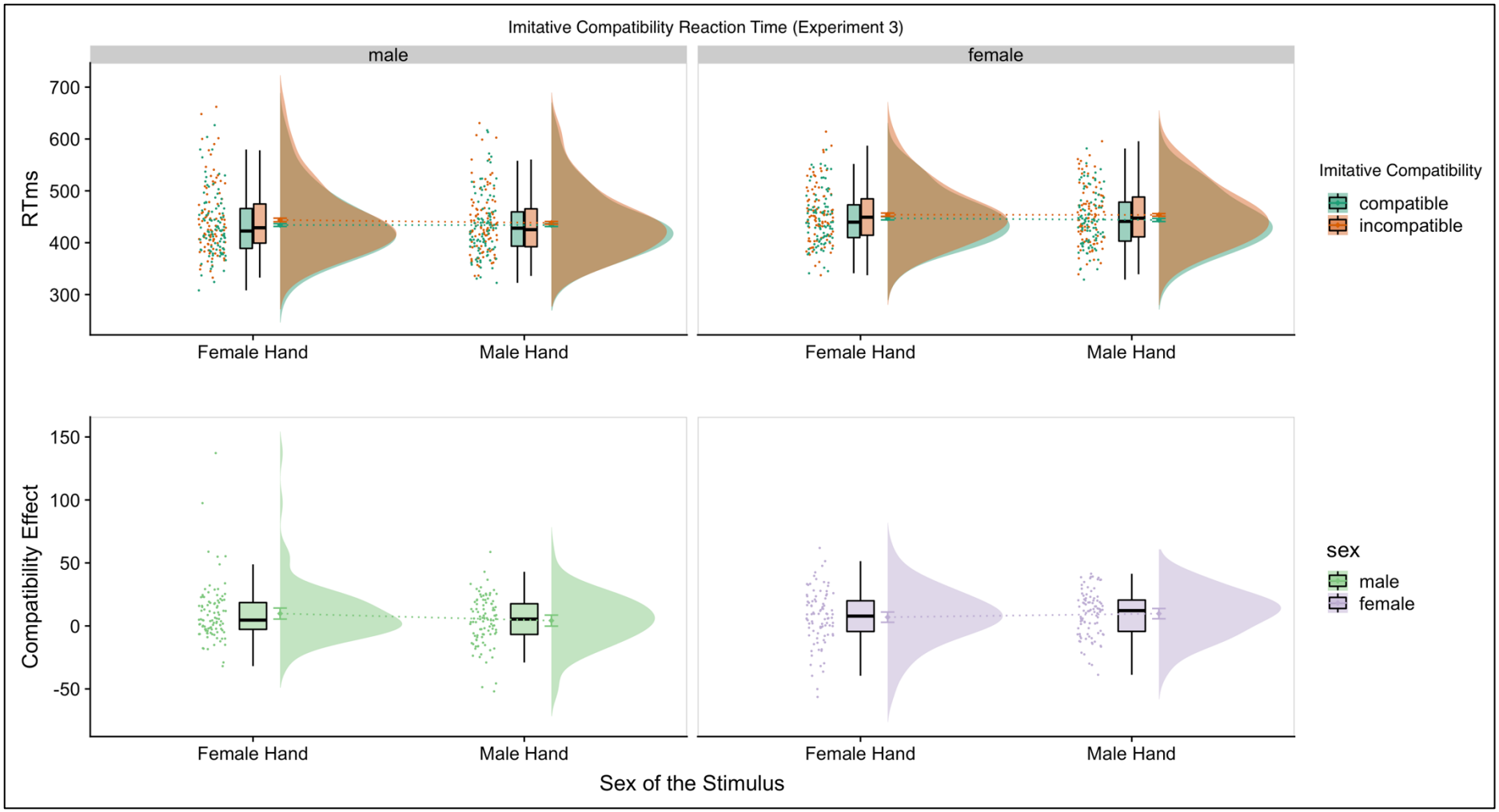


Figure 18. Experiment 3 – Imitative Compatibility Reaction Time. Reaction time is reported in milliseconds (ms). The upper panel shows mean reaction times for compatible and incompatible conditions for both males and females when responding to both male and female hand stimuli. The lower panel shows the compatibility effect for both males and females when responding to both male and female hand stimuli. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials. Error bars represent 95% confidence intervals. Abbreviations: RTms = reaction time in milliseconds.

Reaction time. Mean reaction times were between 430 and 460 milliseconds for both males and females on all conditions of compatibility and stimulus sex (see Figure 18, Supplementary Table 11). A 2 (compatibility: incompatible, compatible) x 2 (stimulus sex: male hand, female hand) x 2 (participant sex: male, female) mixed ANOVA showed a main effect of compatibility ($F(1,187) = 54.96$, $p<.001$, $\eta_p^2 = 0.23$) such that participants were slower to respond on imitatively incompatible trials compared to imitatively compatible. The effect size of the main effect of compatibility was large. There was a main effect of stimulus sex which was a small-to-medium effect and suggested that participants responded slower to female hand stimuli than male hand stimuli ($F(1,187) = 5.70$, $p = 0.018$, $\eta_p^2 = 0.03$).

There was no significant Sex*Compatibility interaction ($F(1,187) = 0.41$, $p = 0.52$, $\eta_p^2 = 0.002$) and the effect size was close to zero. Given the importance to our primary research question regarding sex differences in the compatibility effect, we interrogated the RT data further by computing the difference in compatibility effects between males and females, collapsed across all conditions of stimulus sex. Although females showed a marginally higher compatibility effect than males by 1.33 ms, the lower bound of the 95% CI was below zero at -2.15 ms. The effect size was small with the lower bound of the 95% CI below zero at -0.10 (Mean difference = 1.33, 95% CI [-2.15, ∞], Cohen's d = 0.06, 95 % CI[-0.10, ∞]; see Figure 19B, Table 1C).

The three-way (Compatibility*Stimulus Sex*Participant Sex) interaction approached significance ($F(1,187) = 3.86$, $p=0.051$, $\eta_p^2 = 0.02$) and was a relatively small effect. No other main effects or interactions were significant (see Figure 18, Supplementary Table 12).

In order to investigate the three-way interaction and explore whether the sex difference can be explained by an in-group bias, we computed compatibility effects on all levels of participant sex and stimulus sex. For both males and females, imitative compatibility effects were present when observing both male (Cohen's $d_z > 0.2$) as well as female stimuli (Cohen's $d_z > 0.3$). However, there was not even a trend in the direction we predicted i.e. females did not show a higher compatibility effect for female stimuli compared to male stimuli (Cohen's $d_z = -0.10$), and for males showing a higher compatibility effect for male stimuli compared to female stimuli (Cohen's $d_z = -0.18$). On the contrary, the direction of the interaction was contrary to our hypothesis i.e. females

showed a higher compatibility effect for male stimuli compared to female stimuli, and males showed a higher compatibility effect for female stimuli compared to male stimuli, but these effects were small (see Figure 18). As such, not only are these effects relatively small, they are also inconsistent with the sex difference being a result of an ingroup bias based on the sex of the interaction partner.

In sum, our results indicated a sex difference in spatial compatibility, but not imitative compatibility. An in-group bias/own-sex bias did not explain the sex difference found in the spatial compatibility effect.

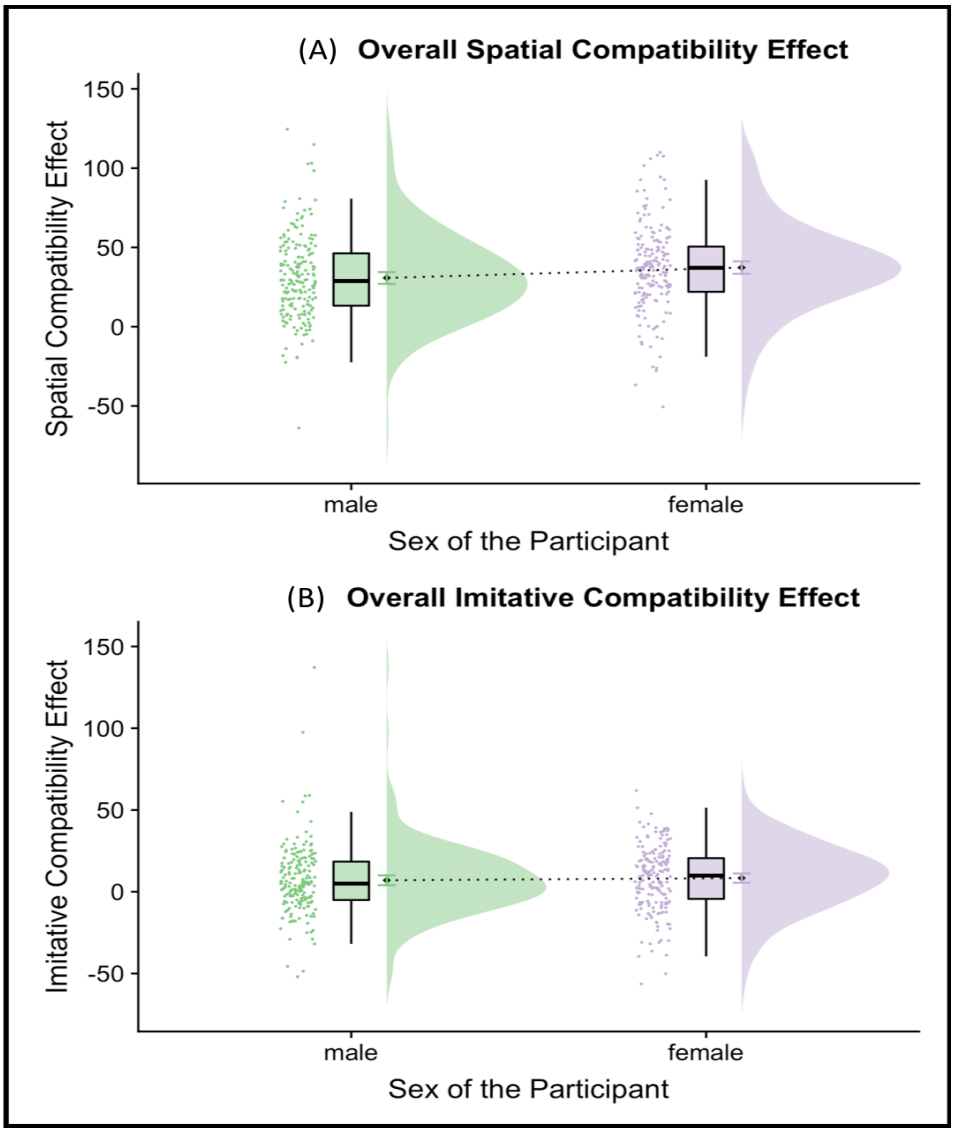


Figure 19. Experiment 3 – Overall Compatibility Effects. The upper panel (A) shows the spatial compatibility effect collapsed across sex of the stimulus for both males and females. The lower panel (B) shows the imitative compatibility effect collapsed across sex of the stimulus for both males and females. The compatibility effect is calculated by subtracting reaction times on compatible trials from incompatible trials and is measured in milliseconds. Error bars represent 95% confidence intervals.

Multiple regression analyses.

We also investigated the relationship between stable personality measures and the imitative compatibility effect as measured on the SRC task independent of spatial effects. Tests for multicollinearity indicated that a very low level of multicollinearity was present (*VIF* for all predictor variables < 2). The base model (including sex, mean RT and the sex*mean RT interaction) explained 4.59 % of the variance in the imitative compatibility effect (F(3,181)=2.90, p=0.036, R2=.04, f^2=0.05) indicating a small effect. Mean RT was a significant predictor (B=0.05, SEB=0.02, t(181)=2.79,p=.006, [0.01; 0.09]), but both sex (B=0.23, SEB=1.04, t(181)=0.22,p=.82, [-1.82; 2.29]) and the sex*mean RT interaction (B=-0.003, SEB=0.02, t(181)=-0.21,p=.83, [-0.04; 0.03]) did not predict the imitative compatibility effect (see Figure 20).

When the model included empathy, the model predicted 7.04 % of the variance. Empathic concern predicted the imitative compatibility effect above and beyond the base model (B=-0.51, SEB=0.24, t(179)=-2.08, p=.04, [-0.99; -0.03]), and explained an additional 2.45 % of the variance (ΔR2= .02, F(5,179)=2.71, p=.022; Figure 21). A decrease in empathic concern predicted a higher imitative compatibility effect. When agreeableness and extraversion were included in the model, the model predicted 7.09% of the variance. Both agreeableness (B=-2.68, SEB=1.53, t(179)=-1.75,p=.081, [-5.67; 0.33]) and extraversion (B=1.81, SEB=1.09, t(179)=1.67,p=.096, [-0.32; 3.96]) marginally predicted the imitative compatibility effect above and beyond the base model, predicting an additional 2.5% of the variance (ΔR2= .02, F(5,179)=2.73, p=.021). Higher extraversion predicted higher imitative compatibility, whereas higher agreeableness predicted a lower imitative compatibility effect (see Figure 21). Effect sizes attributable to the addition of empathy (Cohen's f^2 = 0.03), and agreeableness and extraversion (Cohen's f^2 = 0.03) (beyond the base model) indicated very small effects.

Grandiose and vulnerability narcissism, autistic-like and schizotypal trials, and alexithymia did not predict the imitative compatibility effect above and beyond the base model (all p's > 0.3, all CIs overlapping with zero; see Figure 20). The multiple regression models are summarized in Supplementary Table 13. Zero-order correlations are also consistent with the findings from the multiple regression analyses (see Supplementary Table 14, Supplementary Figure 6).

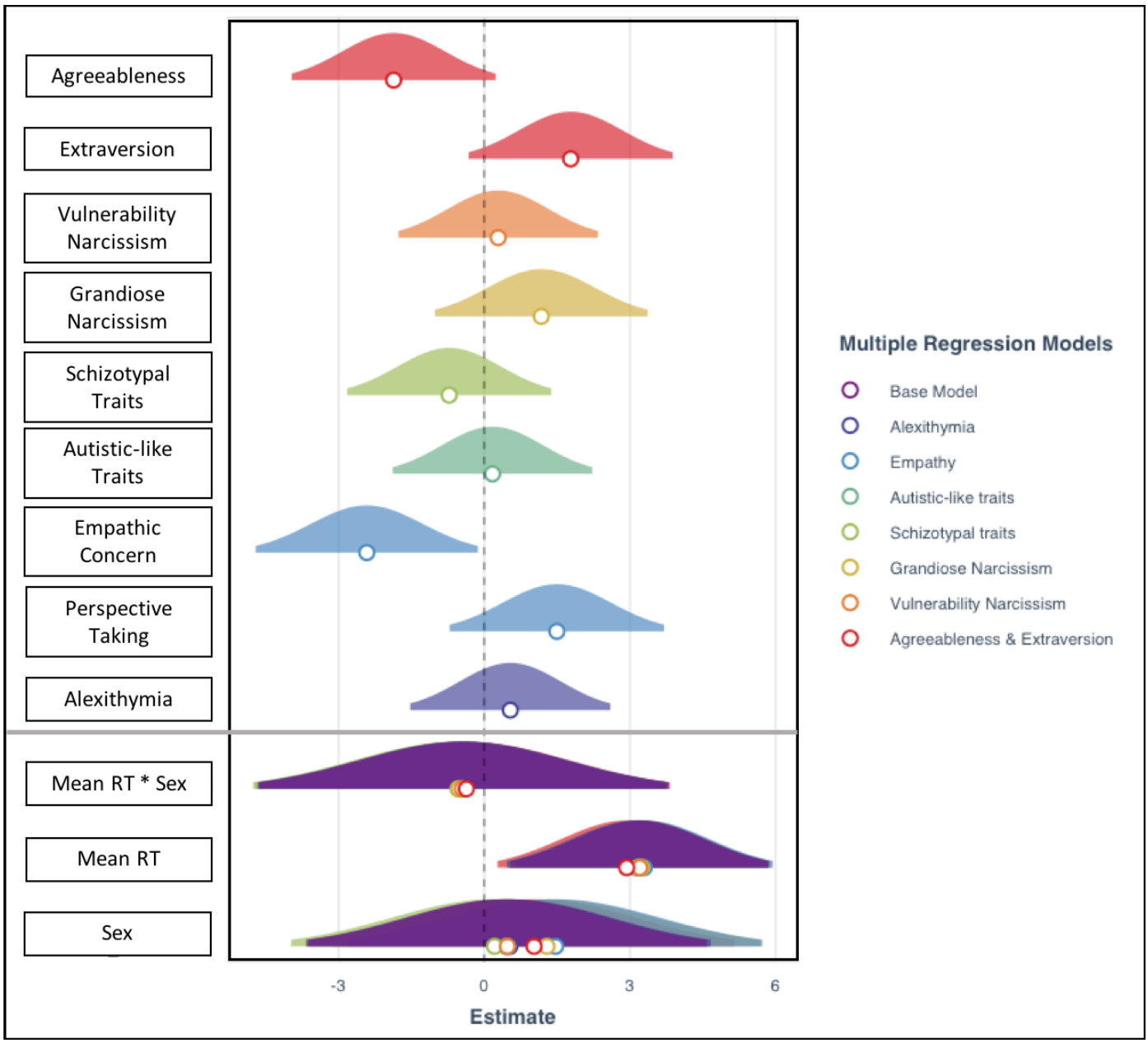


Figure 20. Experiment 3 – Multiple Regression Analyses. Values of standardised coefficients are plotted for each predictor variable (personality trait) along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

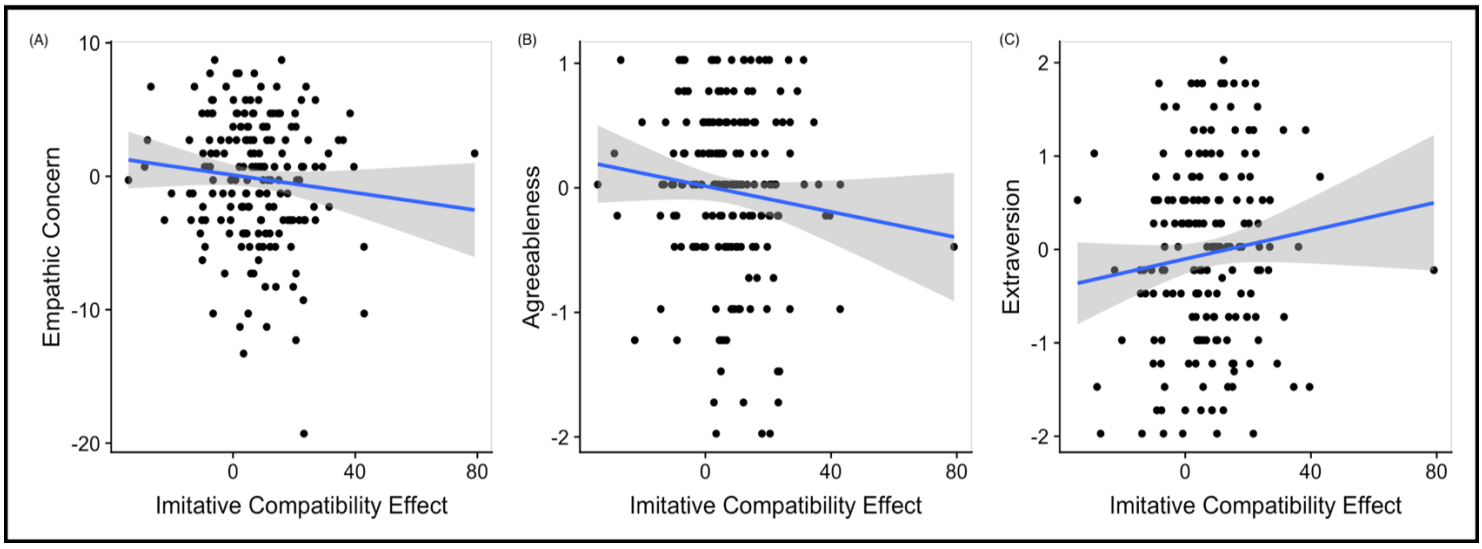


Figure 21. Experiment 3 – Scatterplots. Scatterplots depicting the relationship between imitative compatibility effect and personality traits – empathic concern (A), agreeableness (B), and extraversion (C). X axis denotes the imitative compatibility effect in milliseconds, and Y axis denotes mean centred scores on the personality traits.

To evaluate the sex*trait interaction terms, we computed additional models – each model consisted of the base model, one trait predictor (subscales were included in the same model), and the sex*trait interaction term. None of the sex*trait interaction terms predicted the compatibility effect above and beyond the base model (Supplementary Figure 7). Multiple regression models are summarised in Supplementary Table 15. Effect sizes attributable to the addition of the sex*trait interaction terms (beyond the base model) indicated very small effects (Cohen's f^2 = <0.04 for all models). The pattern of results seen in Experiment 1 for the empathy*sex and alexithymia*sex models did not replicate in Experiment 3 (Figure 8C).

Although our main question of interest was the link between personality traits and automatic imitation, for completeness, we also report results from the multiple regression analyses for spatial compatibility in the supplementary material (see Supplementary Tables 16 and 17, Supplementary Figures 8 and 9).

Stimuli rating.

All participants also rated the male and female hand stimuli on a scale of 1 to 9, with 1 being most masculine, and 9 being most feminine. All male hand stimuli were rated as masculine (Mean rating = 2.93, SD = 0.30). All female stimuli were rated as relatively feminine (Mean rating = 5.68, SD = 0.65). Although the female stimuli were not rated as

strongly feminine, the ratings suggest that both male and female stimuli were perceived differently on average by the participants. The stimuli rating data is also available online.

Discussion

Results from Experiment 3 clearly show that a sex difference exists on the spatial compatibility effect such that females show a higher spatial compatibility effect than males. This difference did not persist when imitative compatibility was measured independently. This suggests that females and males do not differ in the control of automatic imitation as measured by the imitative compatibility effect.

Furthermore, for the first time to date, we manipulated the sex of the stimuli across both male and female participants. Results indicated that there was no own-sex bias in the imitative compatibility effect. For the spatial compatibility effect, although the findings showed a trend toward an own-sex bias such that females showed a greater compatibility effect on female stimuli than male stimuli, this was a relatively small effect size, and thus does not explain much of the sex difference observed in the spatial compatibility effect.

The findings from Experiment 3 thus suggest that it is unlikely that there is a sex difference in the imitative compatibility effect. Instead, our findings suggest that there is a sex difference in the spatial compatibility effect, which may reflect a difference in spatial control between males and females that in the case of this experiment is triggered by the location of a finger in space.

The multiple regression analyses suggest that the imitative compatibility effect is invariant to stable traits of personality including grandiose and vulnerability narcissism, autistic-like and schizotypal traits, as well as alexithymia.

Given prior evidence, we predicted that individuals who report higher empathy, extraversion, and agreeableness would be more prosocial, and would therefore imitate more than those who scored lower on these measures. In the current experiment, although higher extraversion predicted higher imitation, we found the opposite pattern for empathy and agreeableness. An increase in empathic concern and agreeableness predicted a decrease in the imitative compatibility effect. The effects, however, were small and predicted only an additional 2.45% (empathy) and 2.5% (extraversion and agreeableness) of the variance. Before making any firm conclusions, these results would need to be replicated using large sample sizes to ensure that these findings do not reflect false positives. In addition, none of the sex*trait interactions predicted the imitative

compatibility effect, and the pattern of results from Experiment 1 for the sex*empathy and sex*alexithymia interactions did not replicate over Experiment 2 and 3 (see Figure 8). Overall, therefore, these results provide support for the suggestion that automatic imitation is largely invariant to stable traits of personality (Butler et al., 2015; Cracco et al., 2018).

For all three experiments, we performed all the analyses again by further excluding participants who were three standard deviations away from the group mean on the compatibility effect on either of the tasks. For Experiment 1, no additional participants were excluded. For both Experiment 2 and 3, one additional participant was excluded. Obtained results were very similar to those reported above.

General Discussion

By integrating methodological approaches from experimental and differential psychology, the current study shines new light on the relationships between stable features of individuals, such as personality and sex, and the architecture of cognitive control systems. Across three experiments, we consistently showed that cognitive control systems are largely invariant to stable aspects of personality, but exhibit a sex difference, such that females show greater interference than males. Moreover, we further qualified this sex difference in two ways. First, we showed that the sex difference was unrelated to the sex of the interaction partner and therefore did not reflect an in-group bias based on sex. Second, we showed that the sex difference was tied to a form of spatial interference control rather than imitative control and therefore it is unlikely to reflect a specialised mechanism for guiding social interactions exclusively. Instead, our findings suggest that a robust sex difference exists in the system (or set of subsystems) that operate in resolving a form of spatial interference control. The implications of these findings for understanding cognitive control systems in social and non-social contexts are discussed.

Are individual differences in interference control robust and replicable?

In recent years, a key question in psychology and neuroscience has concerned the credibility of reported findings (Button et al., 2013; Open Science, 2015; Munafo et al., 2017; Pashler et al., 2012; Vazire, 2018) with estimates of replicability ranging between 25 and 75% (Camerer et al., 2018; Marsman et al., 2017; Matzke et al., 2015; Nosek & Lakens, 2014). Studies that integrate experimental and differential approaches are rare in general, and in the context of imitation control, prior studies have typically used small

sample sizes (Ainley et al., 2014; Chartrand & Bargh, 1999; Hogeveen & Obhi, 2013; Obhi et al., 2013; Santiesteban et al., 2015). As such, one important contribution from the current study is a more robust and precise estimate of the size and replicability of sex differences in cognitive control. To do so, we used relatively large sample sizes, which could detect small-to-medium effect sizes with a high degree of confidence and ran three separate experiments using designs that combined approaches from experimental and differential psychology.

In Experiments 1 and 2, we replicated the sex difference found previously both when the SRC task measured automatic imitation as a composite of imitative and left-right spatial effects (Butler et al., 2015), as well as orthogonal spatial compatibility effects (Genschow et al., 2017). Furthermore, in Experiment 3, we measured imitative compatibility effects independent of spatial compatibility effects (Berthental et al., 2006; Boyer et al., 2012; Catmur & Heyes, 2011; Jimenez et al., 2012). In Experiment 3, females showed a greater spatial compatibility effect than males, but there was no difference between the sexes on imitative compatibility. Thus, it is clear that the sex difference on the SRC task reflects a difference in spatial control between males and females, rather than a difference in a specialised system that is dedicated to social control.

According to Cohen's benchmarks for interpreting effect sizes (Cohen, 1992), the difference between the sexes was a small-to-medium effect size (Cohen's d = 0.28) and was relatively consistent across the three experiments, with the lower bound of the 95% CI > 0.02 (see Figure 22). Considering the sensitivity of our design, it is important to note that these effect sizes were below the 80% power mark, which our power analysis identified, as we had 80% power to detect effects greater than Cohen's d 0.36. Each individual experiment, therefore, has less than 80% power. This said, all three experiments showed results similar to Butler and colleagues (2015), in that they were in the same (predicted) direction and of a consistent magnitude. Further, by replicating the effects in separate large sample designs, it makes it less likely that these results represent sampling error (Zwaan et al., 2018). Moreover, if we interpret the length of the confidence interval (Armhein et al., 2019; Cumming, 2012), then our best estimate is a small to medium effect, with all likely effects being in the predicted direction (i.e., greater than zero). Therefore, building on prior work (Butler et al., 2015), across three large-sample experiments, we have provided a robust and relatively precise estimate of the size of the sex difference and shown that it reflects spatial rather than social control mechanisms.

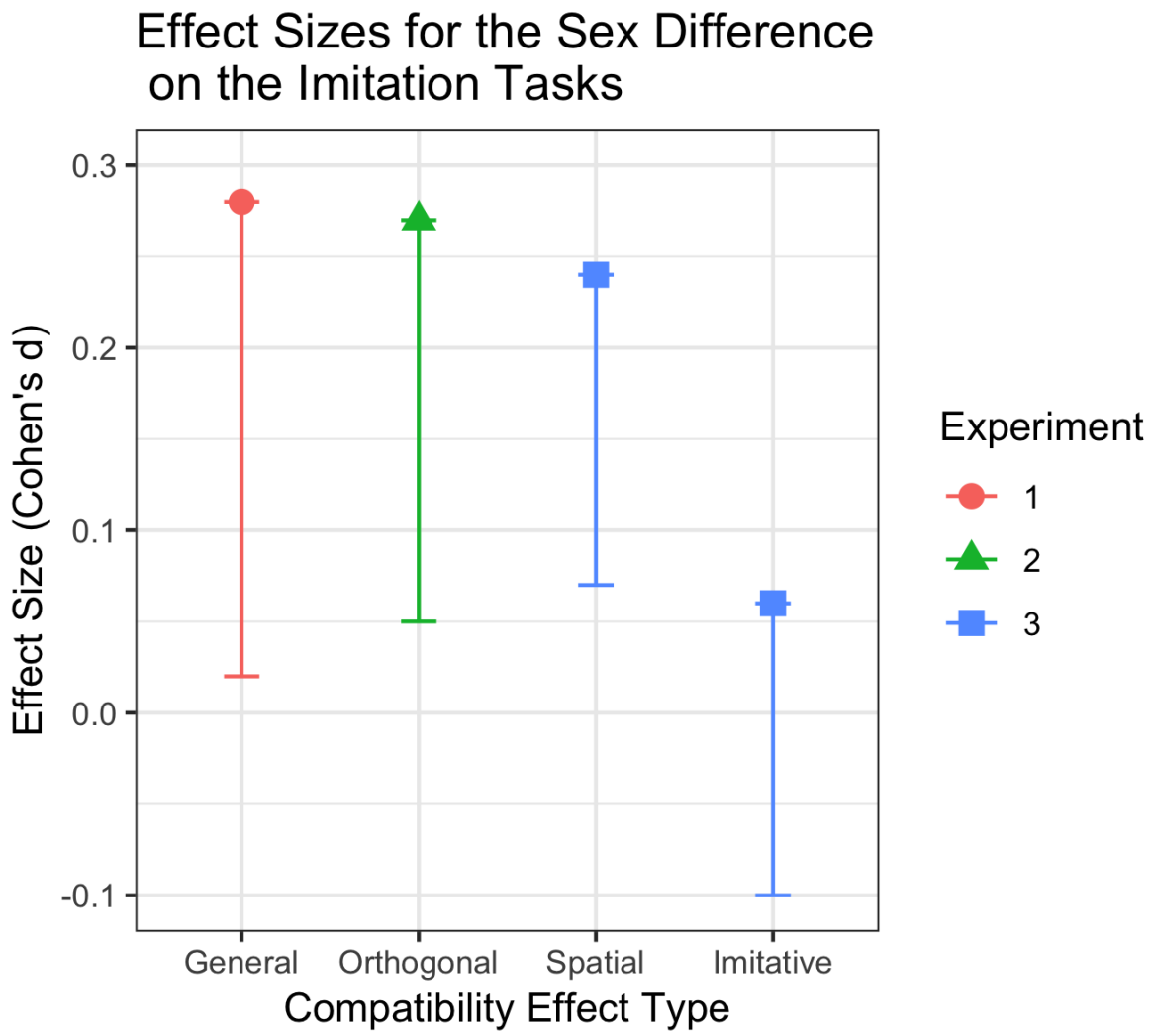


Figure 22. Effect Sizes of the Sex Difference. Cohen's d effect sizes of the sex difference on the imitation task across Experiments 1, 2, and 3. Error bar denotes one-tailed 95% confidence interval.

Moreover, across all three experiments, we consistently found that the control of automatic imitative tendencies, as measured by the SRC task, is invariant to differences in personality traits across individuals. Compared to prior studies (Chartrand & Bargh, 1999; Obhi et al., 2013; Obhi & Hogeveen, 2013), we provide a more robust test of hypotheses regarding individual differences as we used larger sample sizes, which produce higher statistical power, and we looked for consistent patterns of data across multiple experiments. By doing so, a more stable picture is emerging with regard to personality and SRC measures of automatic imitation, which suggests that mechanisms of imitative control are largely invariant to dimensions of personality (Butler et al., 2015;

Cracco et al., 2018b; Genschow et al., 2017), even they are operating in more socially rich contexts (Exp. 1) and when spatial and imitative effects are more clearly separated (Exps. 2 and 3). In short, any effects of personality were small and inconsistent across experiments. Of course, our design did not have sufficient power to detect small effects with reasonable confidence (> 80%), and such effects would require considerably larger sample sizes to be able to confidently confirm that they exist. As such, at present, our best estimate is that the effects of personality on SRC measures of automatic imitation are negligible or small.

Do differences in cognitive control reflect a sex difference or an in-group bias?

In Experiments 1 and 2, as well as in prior studies that have observed sex differences in the SRC imitation task (Butler et al., 2015; Genschow et al., 2017), the stimuli used were of a female hand. Thus, it was possible that the sex difference reflected an in-group bias leading to higher compatibility effects for females compared to males (Cracco et al., 2018). Indeed, there is already suggestive evidence (from studies with relatively small sample sizes), that both facial imitation and SRC measures of imitation have been found to increase when the interacting partner is an in-group member compared to an out-group member based on race, ethnicity, and arbitrary group assignment (Gleibs et al., 2016; Mondillon et al., 2007; Rauchbauer et al., 2015).

In the current study, based on the sex of the interaction partner, we show no clear evidence for an own-sex bias for either spatial or imitative compatibility. Moreover, in terms of sensitivity, the use of a larger sample size than is typical means that our study had 80% power to detect effect sizes at or above Cohen's d = 0.36, which means that we can be reasonably confident that effect sizes of this magnitude or larger are unlikely. Taken together, although ingroup biases are potent in everyday life and relate to sex, race and ethnicity (Brown, 1995; Fischbein, 1996; Kubota et al., 2012; Malpass & Kravitz, 1969; Powlishta, 1995; Rudman & Goodwin, 2004; Yee & Brown, 1995), the difference in interference control reported here reflects the sex of the participant, rather than an in-group bias based on the sex of the interaction partner. As such, these results are contrary to proposals put forward by Cracco and colleagues (2018), and highlight a stable individual difference in interference control, rather than an effect of the social context (i.e., the sex of the interaction partner).

What type of cognitive system underpins sex differences in interference control?

Three broad structures of cognitive system were candidates to underpin the sex difference in interference control: 1) a sex difference specific to social imitative control; 2) a sex difference generalised across all types of control; 3) a sex difference specific to a form of non-social control. If the sex difference was solely tied to imitative control and reflected the workings of a specialised and domain-specific cognitive structure, we would have observed a sex difference only on the imitative compatibility component of the task. Likewise, if the sex difference reflected the operation of a straightforwardly domain-general system, we would have expected a difference between males and females on the flanker task, as well as both the spatial and imitative components of the automatic imitation SRC task. As such, these findings demonstrate that the sex difference is neither completely domain-general i.e., it does not generalise across all types of compatibility effects nor is it domain-specific i.e., it is not solely tied to the control of automatic imitation.

Our findings show more support for the third type of cognitive system outlined above, which suggests that the sex difference reflects a particular type of non-social interference, which is not shared across all SRC tasks. Indeed, across our experiments, the sex difference was tied to a type of spatial interference observed in the spatial component of the automatic imitation SRC task, but not the imitative component of the same task or the non-social flanker task. A sex difference on spatial control, but not on imitative control, when measured on the same task, suggests that although general cognitive control systems are engaged for both tasks to some extent, they may not be engaged in an identical manner across both the compatibility effects. Moreover, it is unlikely that the sex difference on spatial compatibility reflects a difference in the perceptual processing of the social stimulus (i.e. the hand on the screen) as the stimuli are the same across both compatibility effects, but no sex difference emerges on the imitative compatibility effect. For both imitative and spatial compatibility, therefore, the input to the control mechanism that resolves conflict is the same i.e. a finger. However, the way conflict is resolved for spatial and imitative effects might involve mechanisms that operate differently as a function of sex.

In addition to the sex difference not being tied to social or imitative control, it also reflects a component that is not shared with the flanker task. A lack of sex difference on the flanker task, and little or no correlation between the compatibility effects on the two

tasks, has at least two possible interpretations, which are not mutually exclusive. First, it could reflect a lack of sensitivity. The differences between females and males on behavioural indices (such as RT) on the flanker task may be small (Fischer et al., 2015; Clayson et al., 2011; Stoet, 2011). In the current experiment, our sensitivity analysis suggests that we could detect effect sizes of Cohen's d > 0.36 with reasonable confidence (80%), but the effects of sex on the flanker were smaller than this in Experiment 1 and 2 (Cohen's d = 0.15 and 0.03, respectively). Moreover, a large sample study with 895 participants found a small sex difference in the predicted direction on the flanker task using arrows such that females showed a greater compatibility effects than males (Fischer et al., 2015). Thus, there could also be a non-zero sex difference on the flanker task, but even if this turns out to be the case, it is clear that the size of the sex difference varies across different types of non-social cognitive control tasks.

A second possible reason for the lack of sex difference on the flanker task is that the sex difference is underpinned by a particular type of non-social control. Previous studies help contextualise this finding by showing that females differ from males across a wide range of cognitive control tasks, especially those involving spatial processing (Bayliss & Tipper, 2005; Clayson et al., 2011; Stoet, 2011; Stoet et al., 2017). One possibility, therefore, is that the sex difference may reflect a difference in the two types of spatial conflict measured by the flanker and spatial compatibility effect. For example, in the SRC task measuring spatial compatibility, the conflict arises because a stimulus feature is inconsistent to the response, whereas on the flanker task, a stimulus feature is inconsistent to another stimulus feature (Kornblum, 1994; Kornblum et al., 1990). Further, it has been proposed that stimulus-response (S-R) conflicts and stimulus-stimulus (S-S) conflicts are underpinned by different processing patterns (Frühholz et al., 2011; Kornblum & Lee, 1995; Kornblum et al., 1990; 1999; Li et al., 2014; Zhang et al., 1999).

Based on the distinction between stimulus-stimulus and stimulus-response mappings, it may be that sex differences may reflect spatial processing differences in different types of conflicts, such that females show a higher compatibility effect on spatial S-R conflicts, but not S-S conflicts. In line with this, prior evidence on Stroop tasks (Stroop, 1935; which is a type of S-S conflict) suggests that females and males do not differ on the Stroop task (MacLeod, 1991; Daniel et al., 2000) or show a sex difference in the opposite direction such that females show a lower compatibility effect than males (Van der Elst et

al., 2006). However, no prior study has looked at whether sex differences manifest in different ways on S-R and S-S conflicts, and future research would be needed to investigate this interpretation.

More generally, other sex differences, which do not rely on SRC paradigms, can further contextualise our findings. Indeed, prior research suggests that females differ from males on a range of social processes (Baron-Cohen, 2002). For example, females show greater empathy than males, which may lead to more pro-social behaviour, thus suggesting that females may imitate more than males (Baron-Cohen & Wheelwright, 2004; Christov-Moore et al., 2014; Schulte-Rüther et al., 2008). However, although empathy has been associated with a variety of paradigms investigating automatic imitation (Chartrand & Bargh, 1999; Müller et al., 2013; Sonnby-Borgstrom 2002), there does not seem to be a clear link between empathy and automatic imitation as measured on the SRC task (Butler et al., 2015; Genschow et al., 2017). Moreover, while females show higher facial mimicry than males (Dimberg, 1990; Hess & Bourgeois, 2010; Korb et al., 2015; Lundqvist, 1995; Sonnby-Borgstrom 2002; 2008), studies investigating imitation of other behaviours, such as nose-scratching, have not found any reliable sex differences, although such studies have been limited by small sample sizes (Chartrand & Bargh, 1999). Inconsistent and equivocal results across imitation tasks might suggest that these tasks engage different cognitive mechanisms. Indeed, although different measures of automatic imitation have been previously assumed to rely on the same underlying mechanisms, there is accumulating empirical and theoretical reason to question such an assumption (Genschow et al., 2017). Thus, divergent sex differences across different measures of automatic imitative behaviour may reflect differences in cognitive mechanisms that underpin these tasks.

Conclusion

The current findings provide a general insight into the relationship between individual differences and cognitive control systems in social and non-social contexts. Integrating experimental and differential psychology approaches, across three large sample experiments, we show that there is negligible or no evidence for a link between social control and stable personality traits. However, cognitive control systems vary as a function of biological sex, such that females show a greater interference than males. Further, this sex difference does not reflect an in-group bias based on the sex of the

interacting partner and is not tied specifically to social control but reflects differences in the cognitive systems that operate in resolving a form of spatial interference. Therefore, we show that the sex difference exists in the system (or set of subsystems) that operate in resolving a form of spatial interference control, and that such systems are unaffected by social factors such as facial expression or the sex of the interaction partner. More generally, the results highlight the value of integrating approaches from experimental and differential psychology, as well as using large sample sizes, in order to investigate the relationship between cognitive control architectures and stable traits of individuals, which few studies have achieved to date.

4. INDIVIDUAL DIFFERENCES

CHAPTER 5

GENERAL DISCUSSION

In the last few decades, there has been burgeoning interest in automatic imitation given its importance for understanding non-verbal communication and social interactions. The empirical chapters of this thesis aimed at investigating critical questions pertaining to individual differences in automatic imitation, as well as the underlying specialised and/or generalised cognitive and neural mechanisms. The results from the present empirical work addresses important questions and gaps in existing knowledge about automatic imitation and enhances understanding of the nature of automatic imitative processes as well as social cognition more generally. In the current chapter, I first provide a summary of the empirical findings from Chapters 2, 3, and 4. Next, I discuss what these findings mean for research on automatic imitation as well as the broader implications for the field of social cognitive neuroscience as a whole. These implications centre around three main topics – domain-generality v/s domain-specificity models, individual differences, and wider issues in the imitation literature. Within each of these sub-sections, I also outline limitations of the current work as well as interesting and exciting avenues for future research that the current work can lead to.

5.1. Summary of findings from this thesis.

The empirical chapters comprising this thesis aimed at elucidating the neural and cognitive mechanisms underlying automatic imitation, and how these mechanisms differ as a function of individual differences. In Chapter 2, I investigated functional specificity and sex differences in the neural circuits underlying the control of automatic imitation. This was the first fMRI study to date investigating sex differences in the neural correlates of automatic imitation as well as the first study using a functional region-of-interest (fROI) approach. With higher statistical power and functional sensitivity than prior studies, across two experiments, the results demonstrated that the control of automatic imitation engaged a domain-general multiple demand network, as opposed to a unique brain network that supports social cognition. Further, I also investigated whether females and males differed in the control of automatic imitation. Behaviourally, I found a

sex difference on the spatial compatibility effect such that females showed a higher interference effect that males. However, there was no difference between females and males on the imitative compatibility effect. Neural mechanisms underlying spatial and imitative control did not vary as a function of biological sex, pointing toward more similarities than differences between the sexes. It does remain a possibility that small univariate effects may exist, or that the sex difference is underpinned by a more complex neural organisation that was not tested in the present study.

In Chapter 3, I performed a meta-analysis of fMRI studies investigating automatic imitation in order to quantify the consistency and specificity of regional activation during the control of automatic imitation. Given the inconsistency of prior findings, I used multi-level kernel density analysis (MKDA) in order to overcome limitations of interpreting individual studies. Similar to Chapter 1, our results from the meta-analysis also provided unambiguous support for the engagement of a domain-general multiple demand network spanning the dorsolateral frontoparietal cortex during the control of automatic imitation. In contrast, I found less evidence for the specialist view of imitation control which relies on the theory-of-mind (ToM) network. Indeed, mPFC showed no consistent engagement, and engagement in the rTPJ may reflect spatial rather than imitative control. As such, findings from Chapters 2 and 3 suggest that models of imitation control need updating in order to include an increased role for domain-general processes and a reduced or altered role for domain-specific processes.

In Chapter 4, I used three large-sample behavioural experiments in order to investigate individual differences in social (imitative) and non-social control. Integrating methodological approaches from experimental and differential psychology, I demonstrated that cognitive control systems are largely invariant to personality traits but show a sex difference such that females show a greater interference than males. I further qualified this sex difference in two ways. First, I showed that the sex difference was unrelated to the sex of the interaction partner and therefore did not reflect an in-group bias based on sex. Second, I showed that the sex difference was tied to a form of spatial interference control rather than social (imitative) control and therefore it does not reflect a specialised mechanism for guiding social interactions exclusively.

Taken together, the current findings unambiguously support the role of domain-general neural processes in the control of automatic imitation. The results also demonstrate that automatic imitation is not modulated by individual differences such as

stable dimensions of personality including empathy, narcissism, extraversion, agreeableness, alexithymia, and autistic-like and schizotypal traits. In addition, automatic imitation does not differ as a function of biological sex. Instead, the sex difference exists in the system (or set of subsystems) that operates in resolving a form of spatial interference control, and that such systems are unaffected by social factors such as the sex of the interaction partner. Theoretical and methodological implications of these findings are discussed in the following sections of this chapter.

5.2. Implications for imitation and social cognition

5.2.1. Domain-specificity v/s domain-generality.

Before the cognitive revolution in the middle of the 20th century, the brain was looked at as a black box, and any investigations regarding what the black box contained and how it worked were considered as unscientific. The cognitive revolution, however, changed this perspective and many researchers across different domains started investigating the "parts" of this black box, and the links between these parts and corresponding behavioural outputs. A central tenet of these investigations has been that of modularity: are mental phenomena the product of many distinct, specialised processes, or a single generalised one?

It is generally accepted that the mind has *some* internal structure, and different kinds of information processing might engage different neural structures (Barrett & Kurzban, 2006). There is some agreement that there are neural networks that are specialised for perceptual and motor processes such as for processing faces, bodies, objects, biological motion, and representing mental states of other conspecifics (Downing et al., 2001; Grossman et al., 2000; Kanwisher et al., 1997; Saxe et al., 2006). However, specialisation for higher-level cognitive abilities such as reasoning and language is a much more controversial topic of research (Kanwisher, 2010). Researchers across multiple disciplines including evolutionary biology, psychology, cognitive sciences, and social and cognitive neuroscience have considered these questions as key scientific objectives, and postulated and theorised many accounts of domain-specificity across different disciplines which, however, do not see wide consensus (Spunt & Adolphs, 2017).

In its broadest sense, domain-specificity refers to the link between specialised domains and the brain and cognitive systems that are engaged by these domains (Spunt

& Adolphs, 2017). As such, domain specific processes can be described as processes that are tailored to particular types of stimuli or task features. In contrast, domain-general processes operate across different types of stimuli or task features (Barrett, 2012). Thus, there seems to be a clear divide between domain specific and domain general processes (Hirschfield & Gelman, 1994; Kanwisher, 2010).

In Chapters 2 and 3 of my thesis, I investigated whether the control of automatic imitation relied on domain-specific neural architectures that are unique to social cognition as proposed by some researchers (Brass et al., 2001; 2003; 2009; Brass & Heyes, 2005; Santiesteban et al., 2012; Spengler et al., 2009; 2010). However, the results showed robust evidence for the engagement of a domain-general neural architecture which spans dorsolateral fronto-parietal cortices and is engaged for not just the control of automatic imitation, but a wide range of cognitive control tasks (Duncan, 2010; Fedorenko et al., 2013). Thus, the control of automatic imitation does not rely on a specialised neural circuitry, but on a domain-general multiple demand network.

Much like neuroscience in general, social cognitive neuroscience has focused on domain-specificity – by taking a functional segregation approach, the brain has been divided into distinct processing "modules" that perform specific functions (Park & Friston, 2013). A critical focus of research within the field of social cognitive neuroscience has been that of whether there exists a "social brain" i.e. whether social cognition is domain-specific or just one instance of general cognition (Adolphs, 2001). There has been an over-reliance on domain-specificity in the last few decades, and researchers have largely neglected the role of domain-general processes (Barrett, 2012; Barrett & Satpute, 2013; Frith & Frith, 2012; Spunt & Adolphs, 2015; 2017). However, results from the current thesis as well as evidence from other domains of social cognition suggest that domain-general neural networks are also engaged during social interactions and social information processing (Baetens et al., 2014; Baldauf & Desimone, 2014; Quadflieg et al., 2011; Zaki et al., 2010).

In line with this, more recent models of social information processing have suggested that domain-general and domain-specific systems play complementary roles in social cognition. Social interactions are likely brought about by a combination of general and specific processes, as well as interactions and interplay between these two types of processes (Binney & Ramsey, 2019; Barrett, 2012; Michael & D'Ausilio, 2015; Spunt & Adolphs, 2017). These models reduce the burden of explanatory power

traditionally placed on domain-specific systems. For example, Binney and Ramsey's model draws from the semantic cognition literature and suggests that social information processing relies on two primary systems – that of representation and control (Binney & Ramsey, 2019; Ralph et al., 2017). The representation system stores information related to other people and the control system uses this information, reorients attention and prioritises responses depending on task and context. The representation system relies on functionally domain-specific networks, whereas the control system is domain-general. Thus, social information processing and cognition can occur by means of interplay and interactions between control systems and representational systems depending on the type of input.

Of course, characterising the roles of domain-general and domain-specific processes in social cognition is not as easy as one may think given various interpretational difficulties. For instance, evidence suggests that although some brain regions have clearly specific functions, there may be fewer of these regions than previously suggested (Downing et al., 2006; Kanwisher, 2010). So-called domain-specific regions also show engagement across different domains, are functionally heterogenous, and do not work alone but in concert with other brain regions (Baetens et al., 2014; Baldauf & Desimone, 2014; Quadflieg et al., 2011; Spunt & Adolphs, 2015; Zaki et al., 2010;). The temporo-parietal junction (TPJ) is one of many examples of a functionally heterogenous region – it has been implicated in disparate cognitive processes including social cognition, attentional reorienting, memory, and even language (Cabeza et al., 2012; Corbetta et al., 2008, Lee & McCarthy, 2016; Schuwerk et al., 2017). In addition, different groups of neurons and cells within the same brain regions have been known to have different specific and/or general properties (Swanson & Petrovich, 1998). It is difficult, therefore, to parse out the different roles and functions that TPJ performs in a task resulting in misleading interpretations and characterisation of the brain region.

An additional difficulty in interpreting the roles of domain-general and domain-specific networks in social cognition exists because different networks show overlap in the brain. Consider the debate on whether or not high-level language selective regions share computational demands with the mirror neuron system (Rizzolatti & Airbib, 1998) – for instance, the inferior frontal gyrus (IFG) shows activation during both language and action observation tasks. However, in a study by Pritchett and colleagues, functionally defined language regions showed little or no activation during action observation and

imitation (Pritchett et al., 2018). In a similar vein, some of the regions of the fronto-parietal domain-general MD network overlap with key nodes of the mirror-neuron system (Rizzolatti & Sinigaglia, 2016). This makes it difficult to interpret their functional role in the phenomenon of interest. In Chapter 2 of this thesis, I used a functional localiser approach and found that regions which were engaged for a working memory task were also engaged when controlling the tendency to automatically imitate, thus characterising the functional contribution of the MD network in automatic imitation. Of course, this does not mean the MNS was not engaged – both networks, although overlapping, can be simultaneously engaged when performing the task, but future research needs to test this proposal. A further possibility is that non-overlapping networks in the brain could in fact also share some core features and computational demands (Douglas et al., 1989; Harris & Shepherd, 2015). For instance, principles of learning or Bayesian updating which are considered as signatures of domain-general processing can be shared by different domain-specific networks (Barrett, 2012).

I believe that the focus should not be on *whether* domain-general processes contribute to social cognition, but on unpacking the division of labour between domain-general and domain-specific processes and their integration and interactions. After all, integration and segregation are not two opposing ends of a continuum – several functionally segregated units can integrate to bring about behaviour (Fox & Friston, 2012; Friston, 2011; Tononi et al., 1998). Newer theories and models of social cognition need to consider that when asking questions about functional segregation and integration, we need to move beyond exclusivity (brain region A only performs function A), sufficiency (function A is performed only by brain region A), uniformity (the *entire* brain region A performs function A), and independence (brain region A performs function A without interacting with other brain regions), and focus on a more flexible, multi-conceptual, and multi-dimensional characterisation and organisation of the human brain in social cognition.

Future directions. The results from the current thesis provide interesting insight into the construction and testing of models which emphasize the roles of both domain-general and specific systems. Indeed, in some other domains of social neuroscience, these links between domain-specific representational and domain-general control systems have already been initially tested (Quadflieg et al., 2011; Baldauf & Desimone, 2014). For instance, in the study by Quadflieg et al. (2011), judgements inconsistent with

stereotypes recruited areas of the brain associated with domain-specific person perception and domain-general conflict resolution, and these areas functionally interacted with each other. Therefore, it would be of particular interest in the realm of imitation control to test whether representational systems such as those involved in person perception and action perception functionally interact with the control system i.e. the multiple demand network when observed actions are incompatible with the ones that need to be executed and automatic imitation needs to be controlled. Such models can also be extrapolated to other domains of social cognition, allowing for more precise predictions and hypotheses within this framework that can be empirically tested.

In addition, future work that can distinguish between the functional contributions and computational features of overlapping (and non-overlapping) regions would be valuable in better characterising the role of domain-general and specific neural networks in social cognition. This approach would also help in understanding whether domain-general computational principles and features are shared between domain-specific networks. While the current work using univariate fMRI measurements cannot address these questions directly, more nuanced and refined measures of brain function such as multi-voxel pattern analyses (MVPA), repetition suppression analyses (RSA), encoding models, and functional connectivity analyses can better elucidate how and why functional segregation and integration in the brain leads to behaviour (Norman, Polyn, Detre, & Haxby, 2006; Kriegeskorte, Mur, & Bandettini, 2008).

One framework within which interactions between domain-specific and domain-general neural structures in a broader cognitive architecture can be explained is by conceptualising the brain as a Bayesian machine (Friston, 2012). Within such a framework, the Bayesian brain processes information by computing probabilities through reciprocal and recurrent interactions that minimise prediction error (Friston, 2012; Genon et al., 2018; Kilner, et al., 2007; Knill & Pouget, 2004). In such a framework, the brain represents the world as "predictions" and not as "precepts." For instance, people may use prior knowledge and experience to predict action outcomes during social interactions. With new incoming sensory inputs, generated prediction errors are used to update prior predictions and form new ones (Tamir & Thornton, 2018). Therefore, when we see a ball in flight, or a person fly into a rage, we can predict the trajectory of the ball or the person's emotional trajectory (e.g. they might punch the wall) in order for us to react (catch the ball or stop the person from punching the wall).

In a similar vein, if an imitatively compatible action is "expected", or a complementary (instead of imitative) action is "expected" depending on situation and context, the predicted reaction would benefit from facilitated processing (Yon et al., 2018; Yon & Press, 2018). One possibility is that the domain-general control networks works independently and flexibly with errors, predictions, and updating, and may be particularly efficient for social stimuli. Of course, even if I had found the engagement of a domain-specific social control system for the control of automatic imitation, the results could still fit within this framework – the brain may choose between control systems depending on its predictions. Thus, a Bayesian brain may not care whether modules in the brain are domain-general or domain-specific. A second related (and not necessarily antagonistic) possibility is that there is constant information exchange and dynamic switching between domain-general and domain-specific systems in a highly context-dependent manner. That is, the representational and control systems in the brain can bring about response behaviour within this predictive framework by "talking" to each other (Ploran et al., 2007; 2011). The predictive mind is thus the result of a flexible Bayesian computational framework which assumes predictive "message-passing" as a ubiquitous feature of cortical function (Friston, 2005; Kilner et al., 2007; Tamir & Thornton, 2018). Of course, although this is not the only framework that explains integration between domain-general and domain-specific functional networks in the brain, it shows much potential (e.g. Apps & Tsakiris, 2013; Rao & Ballard, 1999; Tamir & Thornton, 2018; Yon & Press, 2018). Thus, using such frameworks can allow for more nuanced and systematic investigations in elucidating the underlying mechanisms of social cognition.

5.2.2. Individual differences.

Just a quick glance around our social world tells us that individuals are different – our friends and colleagues differ in the ways they see, think and act. The discipline of psychology has traditionally seen a divide between experimental and differential psychology approaches (Cronbach, 1957). The experimental approach focuses on characterising a cognitive mechanism based on the average or group response to environmental variables. Alternatively, the differential approach focuses on inter-individual differences i.e. it distinguishes between individuals within a population instead of focusing on group effects. Although there has been a call for integrating these

two approaches, they have remained largely isolated (Cronbach, 1957; Cramer et al., 2010).

In Chapters 2 and 4 of the current thesis, in order to aid cross-pollination between experimental and differential psychology, I integrated approaches from both disciplines to investigate individual differences in the control of automatic imitation. In Chapter 2, I investigated sex differences in automatic imitation using both behavioural and fMRI measures (N=50). Behaviourally, females showed a higher spatial compatibility effect than males, but there was no sex difference on the imitative compatibility effect. No differences were found between the sexes on the neural correlates of either spatial or imitative compatibility effects. The behavioural sex difference on spatial compatibility was further supported by findings from Chapter 4 where across three large sample experiments (with N= ~200 each), a sex difference was found on spatial compatibility, but not on flanker or imitative compatibility effects. This finding suggests that females and males differ on a particular type of spatial conflict which is not shared by the flanker task. Further, I found no evidence of a relationship between automatic imitation and stable personality traits of empathy, alexithymia, autistic-like and schizotypal traits, narcissism, agreeableness, and extraversion.

The current findings contrast with previous studies that have implicated a relationship between stable traits of personality and automatic imitation, and those that have found sex differences on facial mimicry (Ainley et al., 2014; Chartrand & Bargh, 1999; Dimberg, 1990; Hogeveen & Obhi, 2013; Obhi et al., 2014; Sonnby-Borgstrom et al., 2008; Sowden et al., 2016). Such claims, however, are limited due to the small number of studies reported to date, relatively small sample sizes, lack of powerful replications, and a difficulty in generalising across different measures of automatic imitation. Almost all of these studies have used approaches that do not control for other confounding variables such as age, sex, and mean reaction times.

The findings from the current thesis (Chapter 4) as well as recent large-sample studies (Butler et al., 2015; Cracco et al., 2018b; Genschow et al., 2017) strongly suggest that there is no sex difference and personality traits do not have a moderating influence on the SRC measure of automatic imitation. I am confident about these findings because in the current work, I used three large sample behavioural experiments and multiple regression analyses (Chapter 4), as opposed to correlational analyses and ANOVAs used in prior studies (e.g. Hogeveen & Obhi, 2013; Obhi et al., 2014; Sowden et al., 2016). The

current approach provided a more rigorous test of individual differences by controlling for the effects of potentially confounding variables, and by avoiding false positives that may be a result of small sample sizes (Cumming, 2014; Maxwell et al., 2008).

Does this mean that imitative tendencies (and other socio-cognitive processes more generally) do not show individual differences at all? Human beings are extremely adept in their social world - we encounter a myriad of social contexts every day and navigate our way effortlessly with only the occasional faux pas (Adolphs, 2001). Thus, social cognition is an inherently complex process and it is not impossible to assume that individuals differ in the ways in which they interact with their social world (Baron-Cohen et al., 2009; Conway, Catmur, & Bird, 2019; Hamann & Canli, 2004). In addition, atypical social cognition has been observed in clinical samples including individuals with personality disorders (Herpertz & Bertsch, 2014; Hengartner et al., 2014; Semerari et al., 2014). These studies point toward investigations in individual differences that have suggested that socio-cognitive functions vary across individuals.

Integrating experimental and differential psychology approaches, however, comes with its own obstacles. The investigation of an experimental effect at the group level, and individual differences within that effect are questions that can often be at odds with each other. This is because group effects need low variability within the sample whereas differential psychology questions, especially those using correlational methods, are dependent on high variability among individuals (Rogosa, 1988). We cannot assume that robust experimental paradigms such as the SRC tasks frequently used in social cognition research will lend themselves well to approaches used by individual differences researchers. In fact, these paradigms may not be optimal for differential approaches because they are optimal for experimental psychology approaches (De Schryver et al., 2016; Hahn et al., 2011; Ross, Richler, & Gauthier, 2015; Hedge, Powell, & Sumner, 2017). For instance, Hedge and colleagues demonstrated that the reliabilities of many cognitive control tasks such as the Stroop and stop-signal tasks do not meet outlined standards (e.g. Barch et al., 2008) in individual differences contexts (Hedge et al., 2017). Their results suggested that although a difference between groups can be detected if the group means are sufficiently far away from each other to be detectable, the tasks used do not distinguish between individuals in the population consistently.

Further, most prior studies investigating individual differences in socio-cognitive processes have used small sample sizes. A recent meta-analysis in individual differences

research using 708 correlations found that 75% of the effect sizes reported are less that .29 Cohen's d (which by Cohen's standards is a small effect size; Cohen, 1988; Gignac & Szodorai, 2016). Sample sizes required to detect small effect sizes while also taking reliability of the measures used in consideration greatly exceed those that are typically used in most research investigating individual differences in social cognition (Hedge et al., 2017).

Of course, this does not mean that such paradigms should not be used, but understanding their limitations improves their use and interpretation in an individual differences' context. Given these obstacles, it is all the more important that studies investigating individual differences using such tasks use powerful statistical approaches that control for confounding variables, perform powerful replications, use large sample sizes, and ensure adequate statistical power in order to detect the effect of interest (Hedge et al., 2017). In light of these concerns, the current work with multiple experiments, large sample sizes, and higher statistical power goes a step further in establishing a reliable effect of the invariance of automatic imitation to individual differences such as stable traits of personality and biological sex.

These concerns are also especially important and have been raised when investigating individual differences in neuroimaging studies (Bennett & Miller, 2010; Cooper et al., 2019; Dubois & Adolphs, 2016; Vul et al., 2009; Yarkoni & Braver, 2010). Traditionally, cognitive neuroscience has focused on experimental methods as well – however, it is necessary to investigate how individuals vary in brain responses while performing various cognitive tasks, especially if interventions to address cognitive impairments need to be effective at the individual level (Cooper et al., 2019). Increasing sample sizes in neuroimaging, however, is more difficult as these methods are expensive and resource intensive. It has been suggested that in order to establish brain-behaviour links, a sample size of more than a hundred participants is required, but such sample sizes are rarely to be found (Boy et al., 2010; Hedge et al., 2017; Yarkoni & Braver, 2010). In Chapter 2, I investigated sex differences in automatic imitation using fMRI methods. However, I did not find any sex differences in the functional regions of interest (fROIs) for either spatial or imitative compatibility, although a behavioural sex difference was found on spatial compatibility. This suggests that even with a higher sample size than most previous studies using fMRI (N=50), I possibly did not have enough power to detect a sex difference in neural correlates.

Future directions. Of course, this does not mean that the tasks and measures used in experimental psychology are not useful – understanding limitations and statistical concerns of measures used would significantly aid how robust experimental tasks should be used and evaluated in individual differences research.

In the current work, I provided a more rigorous test of individual differences as compared to previous studies investigating individual differences in automatic imitation. Whereas findings from other domains of social cognition suggest that there might be individual differences in socio-cognitive functioning, this does not seem to be the case for the control of automatic imitation (as measured by the SRC task). It is possible that control systems that resolve imitative conflict do not vary as a function of individual differences whereas other systems underlying the processing or representation of social information may show a difference. Future research needs to investigate in what contexts and for what processes or systems does social cognition vary as a function of individual differences. It has recently been suggested that investigations on individual differences would benefit more by using alternative statistical analyses such as linear and generalised linear mixed-effects models as these provide greater flexibility when dealing with confounding variables (Nakagawa & Schielzeth, 2010). Further, latent variable models have also been suggested as an alternative in order to address psychometric concerns in individual differences research in both behavioural and neuroimaging methods (Cooper et al., 2019). Future research using these approaches would provide greater insight into the cognitive mechanisms underlying individual differences in social cognition.

Researchers have previously suggested that it may not be of theoretical or practical importance to study small effect sizes that require large sample sizes especially using neuroimaging methods (Friston, 2012). However, I think that the size of the effect that can be considered meaningful depends on the context. The investigation of individual differences using both behavioural and neuroimaging methods is beneficial to understanding the structure of social cognition, constrain cognitive theories, validate neurophysiological measures, and characterise brain function at an individual level in health and disease (Dubois & Adolphs, 2016; Vogel & Awh, 2008). Using alternative methods to correlational analyses and increasing sample sizes can greatly enhance our understanding and improve reliability of measures that are used in investigations of individual differences using both behavioural and neuroimaging methods.

A further exciting and interesting avenue for future research is to move beyond univariate measures and existing models to investigate individual differences using fMRI. In Chapter 2, I did not find any differences between males and females in the response profile of our functional regions of interest for the control of automatic imitation. It would be interesting to explore whether differences between individuals in social cognition lie in the representational system, the control system, and/or their interaction. For instance, divergent effects of age have been observed on the representational and control systems in semantic cognition (Hoffman, 2018).

In the context of the current work, it is possible that individual differences exist in the representational systems (either in person perception or action perception systems), the control system, or in the interactions between the representational and control systems, and these differences may manifest in varying forms for imitative and spatial compatibility. For instance, there is some evidence suggesting that females encode actions of others as well as social stimuli more saliently than males, however these findings are mixed (Cheng et al., 2007; 2008; Halpern, 2012; Miller & Halpern, 2014; Rahman et al., 2004; Russel et al., 2007). Similarly, the pattern of activation in the control system i.e. the MD network may be different for imitative and spatial compatibility effects, and these patterns may show a varying sex difference depending on the type of conflict being resolved. Multi-voxel pattern analyses may prove to be useful when investigating such subtle effects in representational and control systems. Further, individual differences have been found in resting-state, functional, and structural connectivity in the brain suggesting a more complex neural organisation (Dubois & Adolphs, 2016; Ingalharikar et al., 2014; Zhang, Dougherty, Baum, White, & Michael, 2018). Future research using more nuanced and refined methods is needed to elucidate how individual differences manifest in domain-general and domain-specific systems, as well as their interactions.

5.2.3. Wider issues within the imitation literature.

Cognitive systems underlying imitation. Often, but not always, the discussion of generality v/s specificity of cognitive mechanisms has also been housed under a broader nature/nurture debate – i.e. whether the ability to imitate is innate or brought about by domain-general associative learning. Some researchers argue that human beings have an

innate, genetically inherited module for imitation (Decety & Meltzoff, 2011; Meltzoff, 1988; Meltzoff & Moore, 1997). In contrast, the empiricist account argues that sensorimotor links support imitation, and these are formed by associative learning (Heyes, 2010; 2011; Cook, Bird, Catmur, Press, & Heyes, 2014; Catmur et al., 2009).

At the level of description of the cognitive systems at play, whether specialised or generalised cognitive mechanisms underlie automatic imitation as measured by the SRC task is a topic of much interest. Findings from Chapters 2 and 4 substantiate the claim that automatic imitation is not reducible to spatial compatibility effects as also evidenced in prior research (Heyes, 2011; Catmur & Heyes, 2011, Marsh et al., 2016). On one hand, this has been considered by some as evidence that imitative and spatial effects rely on distinct cognitive processes (Bertenthal et al., 2006), with some researchers suggesting that automatic imitation relies on either a special direct matching mechanism (Bertenthal & Scheutz, 2012) and/or a self-other distinction mechanism (Brass et al., 2009) while spatial compatibility does not.

Self/other distinction has been considered as a candidate mechanism for automatic imitation, based primarily on the engagement of the mPFC and rTPJ (nodes of the domain-specific ToM network; Brass et al., 2003; 2009). Further, based on research on automatic imitation, researchers have further suggested that self/other distinction forms the basis for not just imitation, but a whole range of social processes (e.g. Brass et al., 2009; Sowden & Shah, 2014). However, one issue here is that of reverse inference – claims made about the cognitive functions at work based on the anatomical localisation of its neural correlates (e.g. activation in TPJ, therefore self-other distinction is different from self-other distinction, therefore activation in TPJ; Poldrack 2006). When discussing the mental processes underlying automatic imitation (and social cognition), I think it is essential and necessary to keep the different levels of description in mind (cognitive and/or neural; Morton, 2008; Morton & Frith, 1995).

On the other hand, other researchers believe that spatial and imitative compatibility effects are mediated by the same cognitive processes of domain-general associative learning (Cooper et al., 2012; Heyes, 2011; 2017). Thus, sensorimotor contingencies for imitation are formed in the same way as other non-social sensorimotor contingencies – in this sense, imitation is brought about by domain-general cognitive mechanisms (Catmur et al., 2009; Cook et al., 2014; Heyes, 2010; 2011). In contrast, ideomotor theorists suggest that although sensorimotor contingencies for imitation and

spatial control are brought about by learning mechanisms, there is a crucial difference in how this learning occurs (Brass & Muhle-Karbe, 2014). For instance, ideomotor representations are formed when the relationship between actions and their sensory consequences is learnt (R-E learning). This type of learning occurs when we perform an action with our hand, and we see our hand moving. In contrast, associative learning involves learning the relationship between responses and the preceding stimuli (S-R learning; Greenwald, 1970; Prinz, 2005).

Our findings from Chapters 2 and 3 unambiguously suggest that the MD network is engaged for imitation and spatial control. Does this say anything about whether specialised or generalised *cognitive* mechanisms are at play when resolving spatial or imitative control? While it is perhaps simpler (and more intuitive) to assume that a generalised neural mechanism might underlie a generalised cognitive mechanism, we cannot deny the possibility that a specialised mechanism (of self/other distinction or ideomotor representation, for example) may not necessarily recruit a domain-specific neural architecture. For instance, imitative control can be brought about by multiple functional circuits, that themselves do not need to be domain specific (Spunt & Adolphs, 2017).

Whatever the cognitive structure underlying automatic imitation, the current findings demonstrate that a domain-general neural system is engaged for the control of automatic imitation. Therefore, "specialist" or "generalist" models at any level of description need to take into account that generalised neural systems are involved in imitation control. Any model that makes "specialist" claims about imitation control (or social cognition more generally) needs to provide compelling evidence linking both cognitive and neural levels of description in order to substantiate the claim that imitation control is underpinned by specialised cognitive and/or neural mechanisms.

Future directions. The current work therefore suggests that although both spatial and imitative effects recruit similar regions of the MD network, they may or may not do so in an identical manner. An important focus for future research would be to elucidate the specialised and/or generalised neural as well as cognitive mechanisms underlying spatial and imitative control. For instance, dissociating between ideomotor theory and associative learning by investigating the anticipatory nature of ideomotor representations, as well as elucidating the differential working of the MD system for non-imitative and imitative control are exciting avenues for future investigation.

In some sense, debates of domain-generality and specificity for the "control" of automatic imitation are similar to debates on "representation" seen in the field of social cognition in the last few decades. For any function or process to qualify as "domain-specific" it needs to fulfil certain criteria – for instance, the domain-specificity of face processing has been demonstrated by showing that the fusiform face area (FFA) is most used for face processing, is active more for faces than other visual stimuli, contains neurons that are face-selective, is necessary (and not simply involved) for face processing, and evolved to be face-specific (Kanwisher, Mcdermott, & Chun, 1997; Kaniwsher, 2010; Tsao, Freiwald, Tootell, & Livingstone, 2006). Similarly, in order to establish whether the control of automatic imitation is underpinned by domain-specific neural systems, compelling evidence on these criteria needs to be provided by proponents of "specialist" models.

Function(s) of imitation. It is also essential to distinguish between the origin and functions of both imitative and non-imitative (or social and non-social) sensorimotor contingencies (Cook et al., 2007). For instance, irrespective of how sensorimotor contingencies are learnt, the functions that imitation performs may be different to functions performed by non-imitative sensorimotor contingencies. In contrast, a nihilist account would argue that imitation performs no "special" function and is indistinguishable from other non-imitative sensorimotor contingencies (Farmer et al., 2018). However, a popular theory is that automatic imitation functions as a "social glue" by increasing positive rapport and affiliation, and thus has a strategic social function (Chartrand & Bargh, 1999; Chartrand & Lakin, 2013; Dijksterhuis, 2005; Lakin et al., 2003; van Baaren et al., 2009; Wang & Hamilton, 2012). Fundamental assumptions of the social glue hypothesis are that automatic imitation is a pro-social signal and the production of automatic imitation is influenced by factors and in contexts where we would expect more pro-social signalling, and that automatic imitation would generate positive consequences (Farmer, Ciaunica, & Hamilton, 2018).

In line with this, SRC measures of automatic imitation have found an increase in automatic imitation after pro-social priming (Cook & Bird, 2011, 2012; De Coster et al., 2014; Leighton et al., 2010). Empathy is linked to pro-sociality, and it has been suggested that more empathic individuals tend to show greater automatic imitation (Chartrand & Bargh, 1999). Similarly, individuals lacking in emotional and cognitive empathy show decreased automatic imitative tendencies (Hogeveen & Obhi, 2013; Obhi et al., 2013).

Automatic imitation also has positive consequences such as increasing trust and rapport between interacting partners (Chartrand & Bargh, 1999; Chartrand & Lakin, 2013).

However, these findings have recently been challenged. A study by Newey and colleagues did not find a link between pro-sociality and automatic imitation (Newey, et al., 2019). Findings from the current thesis (Chapter 4) also corroborate the suggestion that automatic imitation is invariant to stable traits of personality such as empathy and narcissism (Butler et al., 2015; Cracco et al., 2018b; Genschow et al., 2017). In addition, evidence also suggests that being mimicked does not lead to increased rapport or trust (Hale & Hamilton, 2016a; 2016b). These findings demonstrate that the effects of prosocial priming, stable traits of personality like empathy and narcissism, as well as positive consequences such as increased trust and rapport on and due to automatic imitative behaviours may be small (or negligible) and less robust than what prior research has proposed.

In addition, even though imitation may have many positive consequences, it may not always be the sincerest form of flattery – in some circumstances, over-imitation may lead to negative consequences. For instance, some previous work suggests that high levels of mimicry and similarity lead to negative reactions toward the imitator (Lynn & Snyder, 2002; White & Argo, 2011). These studies, however, have looked at intentional imitation, and it is yet unclear whether over-imitation without the conscious awareness of either of the interacting partners may or may not lead to negative outcomes.

Future directions. Along with functioning as a social glue, automatic imitation is also thought to aid in action understanding (although this is controversial; Gallese & Goldman, 1998) and serve as a tool for social and cultural learning (Heyes, 2017). These functions of imitation are not independent and can co-exist (Farmer et al., 2018). Future research with larger sample sizes and high statistical power is needed to understand and identify the contexts and conditions under which automatic imitation functions as a social glue, and has positive or negative consequences.

Are all the tasks measuring the same thing? In the automatic imitation literature, it is often accepted that there are consistent effects across studies. However, the evidence presented above and in Chapter 4 of this thesis suggests that the effect of stable personality traits and other social modulators such as pro-sociality, trust and rapport are not consistently replicated. However, all these studies have used different tasks. For instance, in the current thesis, I used SRC measures of automatic imitation (also

in Butler et al., 2015; Genschow et al., 2017). Hale and Hamilton (2016b) measured mimicry using a virtual reality paradigm whereas social psychology researchers have used naturalistic paradigms (e.g. Chartrand & Bargh, 1999). Therefore, it is difficult to estimate how reliable these effects are for the phenomenon of automatic imitation as a whole (Lakens, Hilgard, & Staaks, 2016).

Further, although automatic imitation is considered to be a laboratory equivalent of mimicry, and the two phenomena are thought to rely on similar mechanisms (Heyes, 2011), this has been recently questioned (Genschow et al., 2017; Ramsey, 2018). Genschow and colleagues found no correlation between mimicry and automatic imitation, suggesting that these tasks may not be measuring the exact same thing and are not related to each other (Genschow et al., 2017). One possibility is that one or both tasks lack validity and reflect measurement error i.e. they do not measure what they are supposed to measure (i.e. automatic imitation). For instance, SRC measures of automatic imitation may not be measuring anything related to the overt copying of behaviours but something unrelated to imitation (Ramsey, 2018). A second possibility is that SRC measures and naturalistic paradigms measure different dimensions or flavours of imitation (Cracco & Brass, 2019; Ramsey, 2018).

A similar discussion can be seen on tasks measuring theory-of-mind or an ability to understand others' mental states (Apperly, 2012; Schaafsma et al., 2015; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014). Researchers have suggested that theory-of-mind may not be a unitary construct but a diverse, multi-dimensional construct that intersects with different social and cognitive abilities depending on the task and dimension it is measuring (Warnell & Redcay, 2019). Extrapolating these findings to imitation research, it may be that different tasks measuring imitation tap into different dimensions of a diverse construct, and thus intersect with different social and cognitive variables. This can perhaps explain why, for example, mimicry paradigms find a link between empathy and mimicry, whereas SRC measures of automatic imitation do not as these tasks may be enlisting different cognitive mechanisms (Butler et al., 2015; Chartrand & Bargh, 1999; Genschow et al., 2017; Mueller et al., 2013; Sonnyby-Borgstrom 2002).

Future directions. A fundamental issue that, however, lies unresolved is whether the SRC task actually links to social cognition more broadly. The domain of social cognition research has seen the use of cognitive psychology paradigms and

methodologies to answer questions that are of interest to social psychologists (Lambert & Scherer, 2013). An outstanding question remains – are these paradigms truly measuring what we think they are measuring? Throughout this thesis, I have used the SRC task to measure automatic imitation. However, more and more recent evidence suggests that social cognition is fundamentally different when we are involved in live social interactions with each other as compared to when we are doing tasks in a controlled environment (Schillbach et al., 2013; Redcay & Schillbach, 2019). For instance, researchers have suggested that eye-gaze behaviours when measured using screen-based tasks cannot be validly generalised to and used as a proxy for understanding gaze behaviours in live social interaction settings (Grossman, Zane, Mertens, & Mitchell, 2019). Thus, I believe that more empirical evidence is needed to know whether the SRC task of automatic imitation is actually measuring a social cognitive process (or even one component of a multi-dimensional construct like imitation).

Therefore, as a field, it might be beneficial to assess the validity of tasks before assuming that inferences hold at the level of latent constructs – we should use operational terminology (e.g. we have a Stroop effect) rather than use it as a substitute for a theoretical construct (e.g. we have a response inhibition effect; Dennett, 1991; Peters & Crutzen, 2016). Elucidating similar or different cognitive mechanisms underlying different tasks measuring imitation, the relationship between overt and covert measures of copying behaviours, as well as assessing the validity of automatic imitation tasks (as well as other tasks that tap into social processes) are important directions for future research.

5.3. Conclusion

For centuries, thinkers, philosophers, and scientists have been interested in and questioned the processes underlying the tendency of humans to automatically imitate others. These explorations and dialogues have spanned many domains and have tried to answer fundamental questions about humankind including the freedom of will, the underlying unity of all sentience, and the nature, functions, and intricacies of social cognition. While processes such as the control of automatic imitation are only the microcosm representing the intricacy of the human mind and brain, insight into the cognitive and neural mechanisms underlying such processes, and how these mechanisms

differ between individuals, are important pursuits for solving the puzzle of human existence and behaviour in an inherently complex social world.

The findings from the current thesis are instrumental in providing a better understanding of the methodological and theoretical debates surrounding automatic imitation and social cognition. In the empirical chapters comprising this thesis, I have used multi-experiment and meta-analytical approaches as well as large-sample experiments to investigate the neural mechanisms and individual differences in the control of automatic imitation. The results from the current thesis provide novel insight into models of social cognition - newer models need to place greater emphasis on the role of domain-general processes, and the interactions between domain-specific and domain-general processes, instead of focusing only on domain-specificity. Further, the current findings also highlight the importance (and obstacles) of integrating experimental and differential psychology approaches to explore individual differences in social (and non-social) cognition. While the current results suggest that automatic imitation is invariant to stable personality traits and biological sex, the cognitive and neural organisation of individual differences in social and non-social cognition seems to be more complex and intricate than what has been previously conceived. In conclusion, the current thesis has important implications for research on automatic imitation, and points to many exciting and interesting avenues for future research that would further enhance our knowledge and understanding of social cognition.

REFERENCES

Aboud, F. E. (1988). *Children and prejudice*. Oxford: Basil Blackwell.

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, *3*(12), 11.

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology, 11 (2),* 231–239.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual review of psychology, 60,* 693-716.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind?. *Neuroimage, 30*(3), 1059-1068.

Aicken, M. D., Wilson, A. D., Williams, J. H. G., & Mon-Williams, M. (2007). Methodological issues in measures of imitative reaction times. *Brain and Cognition, 63*(3), 304–308.

Ainley, V., Brass, M. & Tsakiris, M. (2014). Heartfelt imitation: High interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia, 60,* 21–28.

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14,* 1338–1344.

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). Raincloud plots: a multiplatform tool for robust data visualization. *PeerJ Preprints, 6*, e27137v1.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202-212.

Allport, G. W. (1954). *The nature of prejudice*. Cambridge/Reading, MA: Addison-Wesley.

Alwall, N., Johansson, D., & Hansen, S. (2010). The gender difference in gaze-cueing: Associations with empathizing and systemizing. *Personality and Individual Differences, 49*(7), 729-732.

Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality, 40*, 440-350. doi: 10.1016/j.jrp.2005.03.002

Amodio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci. 7*:268-277.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15*(2), 88-93.

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature, 567,* 305-307

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology, 65*(5), 825-839.

Apps, M. A., & Tsakiris, M. (2013). Predictive codes of familiarity and context during the perceptual learning of facial identities. *Nature communications, 4*, 2698.

Aron AR, Robbins TW, Poldrack RA. 2014. Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci. 18*, 177-185.

Ashton, M. C., Paunonen, S. V., Helmes, E., & Jackson, D. N. (1998). Kin altruism, reciprocal altruism, and the big five personality factors. *Evolution and Human Behaviour, 19,* 243-255

Ashton–James, C., Van Baaren, R. B., Chartrand, T. L., Decety, J., & Karremans, J. (2007). Mimicry and me: The impact of mimicry on self–construal. *Social cognition*, *25*(4), 518-535.

Baetens, K., Ma, N., Steen, J., & Van Overwalle, F. (2014). Involvement of the mentalizing network in social and non-social high construal. *Social Cognitive Affective Neuroscience, 9*(6), 817-824, doi: 10.1093/scan/nst048

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of psychosomatic research, 38*(1), 23-32.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16(10),* 814-819.

Bailey, P. E., & Henry, J. D. (2009). Subconscious facial expression mimicry is preserved in older adulthood. *Psychology and Aging, 24,* 995–1000. doi:10.1037/a0015789

Baldauf, D., & Desimone, R. (2014). Neural Mechanisms of Object-Based Attention. *Science, 344,* 424–427.

Banich MT. 2009. Executive function: the search for an integrated account. *Curr. Dir. Psychol. Sci,. 18,* 89–94

Barch, D. M., Carter, C. S., Comm, C. E., & 4. (2008). Measurement issues in the use of cognitive neuroscience tasks in drug development for impaired cognition in schizophrenia: A report of the second consensus building conference of the CNTRICS initiative. *Schizophrenia Bulletin, 34*, 613–618. doi: 10.1093/schbul/sbn037

Bardi, L., Bundt, C., Notebaert, W., & Brass, M. (2015). Eliminating mirror responses by instructions. *Cortex, 70,* 128e136. doi: 10.1016/j.cortex.2015.04.018

Bardi, L., Gheza, D., & Brass, M. 2017. TPJ-M1 interaction in the control of shared representations: new insights from tDCS and TMS combined. *Neuroimage, 146*, 734-740.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6,* 248–254

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*, 163–175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.

Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences*, *109*(Supplement 1), 10733–10740. doi: 10.1073/pnas.1201898109.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. Psychological Science in the Public Interest, 20(1), 1-68.Barrett, L. F. & Satpute, A. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology, 23*(3), 361–372. doi:10.1016/j.conb.2012.12.012

Barrio, V. D., Aluja, A., & García, L. F. (2004). Relationship between empathy and the Big Five personality traits in a sample of Spanish adolescents. *Social Behaviour and*

Personality: An International Journal, 32, 677–681. doi:10.2224/sbp.2004.32.7.677

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1987). Motor mimicry as primitive empathy. In *Cambridge Studies in Social and Emotional Development*. *Empathy and its development* (pp. 317–338). New York, NY, US: Cambridge University Press.

Bayliss, A. P., di Pellegrino, G., & Tipper, S. P.(2005). Sex differences in eye gaze and symbolic cueing of attention. *Quarterly Journal of Experimental Psychology, 58A,* 631–650.

Begley, C. G. (2013). Reproducibility: Six red flags for suspect work. *Nature*, *497*(7450), 433.

Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Year in Cognitive Neuroscience, 2010*(1191), 133–155. doi: 10.1111/j.1749-6632.2010.05446.x

Bernieri FJ (1988) Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behaviour, 12,* 120–138

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*(8), 706-712.

Berry, C. S. (1906). The imitative tendency of white rats. *Journal of Comparative Neurology, 16*, 333-361.

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 210–225.

Bertenthal, B. I., & Scheutz, M. (2013). In praise of a model but not its conclusions: Commentary on Cooper, Catmur, and Heyes (2012). *Cognitive science*, *37*(4), 631-641.

Bien N, Roebroeck A, Goebel R, Sack AT. 2009. The brain's intention to imitate: the neurobiology of intentional versus automatic imitation. *Cereb Cortex, 19,* 2338-51.

Binney, R. J., & Ramsey, R. (2019, March 18). Social Semantics: The role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. doi: 10.31234/osf.io/36tm5

Bird, G., Leighton, J., Press, C., & Heyes, C. (2007). Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society B, 274(1628),* 3027-3031. doi: 10.1098/rspb.2007.1019

Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*(2), 260-267.

Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: where are we heading?. *Trends in cognitive sciences*, *8*(5), 216-222.

Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology, 112,* 1105–1118.

Bock, J. K. (1986). Syntactic persistence in language production. Cognitive Psychology, 18(3), 355–387. doi: 10.1016/0010-0285(86)90004-6

Bohl, V., & van den Bos, W. (2012). Toward an integrative account of social cognition: Marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience, 6,* 274.

Bortoletto, M., Mattingley, J. B., & Cunnington, R. (2013). Effects of Context on Visuomotor Interference Depends on the Perspective of Observed Actions. *Plos One, 8*(1), e53248– e53248.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review, 108*(3), 624.

Botvinick, M. M., & Cohen, J. D. (2014). The computational and neural basis of cognitive control: charted territory and new frontiers. *Cognitive science, 38*(6), 1249-1285.

Bouquet, C. A., Gaurier, V., Shipley, T., Toussaint, L., & Blandin, Y. (2007). Influence of the perception of biological or non-biological motion on movement execution. *Journal of Sports Sciences, 25*(5), 519–530.

Boy, F., Evans, C. J., Edden, R. A., Singh, K. D., Husain, M., & Sumner, P. (2010). Individual differences in subconscious motor control predicted by GABA concentration in SMA. *Current Biology, 20*(19), 1779–1785. doi: 10.1016/j.cub.2010.09.003

Boyd, R., & Richerson, P. J. (1996, January). Why culture is common, but cultural evolution is rare. In Proceedings-British Academy (88), 77-94. Oxford University Press Inc.

Boyer, T. W., Longo, M. R., & Bertenthal, B. I. (2012). Is automatic imitation a specialized form of stimulus-response compatibility? Dissociating imitative and spatial compatibilities. *Acta Psychologica, 139*(3), 440–448.

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44(2),* 124-143. doi: 10.1006/brcg.2000.1225

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia, 43(1),* 89-98.

Brass, M., Derrfuss, J., Matthes-von Cramon, G., & von Cramon, D. Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology, 17(2),* 265-271.

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences, 9(10),* 489-495. doi:10.1016/j.tics.2005.08.007

Brass, M., & Muhle-Karbe, P. S. (2014). More than associations: an ideomotor perspective on mirror neurons. *Behavioural and Brain Sciences, 37*(2), 195-196.

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1528),* 2359-2367. doi: 10.1098/rstb.2009.0066

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage, 14(6),* 1416-1423.

Brown, R. (1995). *Prejudice in social psychology*. Blackwell, Oxford, U.K.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., & Rizzolatti, G. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron*, *42*(2), 323-334.

Bunge SA, Hazeltine E. Scanlon MD, Rosen AC, Gabrieli JDE. 2002. Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage. 17,* 1562-1571.

Burnett, S., & Blakemore, S. J. (2009). Functional connectivity during a social emotion task in adolescents and in adults. *European Journal of Neuroscience, 29*, 1294–1301.

Butler, E. E., Ward, R., & Ramsey, R. (2015). Investigating the Relationship between Stable Personality Characteristics and Automatic Imitation. *PLoS ONE, 10(6),* e0129651. doi:10.1371/journal. pone.012965

Butler, E. E., Ward, R., & Ramsey, R. (2016). The influence of facial signals on the automatic imitation of hand actions. Frontiers in psychology, 7, 1653.Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioural and brain sciences, 21*(5), 667-684.

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. *Trends in cognitive sciences, 16*(6), 338-352.

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist*, *47*(8), 1019–1028. doi: 10.1037/0003-066X.47.8.1019

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., and McClintock, M. K. (2000). Multi-level integrative analyses of human behaviour: social neuroscience and the complementing nature of social and biological approaches. *Psychol. Bull. 126,* 829–843.

Camerer, C. F., Dreber, A., Holzmeister, H., H, T., Huber, J., Johannesson, M., et al. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour., 2,* 637–644.

Cameron, J. A., Alvarez, J. M., Ruble, D. N., & Fuligni, A. J. (2001). Children's lay theories about ingroups and outgroups: Reconceptualizing research on prejudice. *Personality and Social Psychology Review, 5*(2), 118-128.

Campbell, R., Elgar, K., Kuntsi, J., Akers, R., Terstegge, J., Coleman, M., (2002). The classification of “fear” from faces is associated with face recognition skill in females. *Neuropsychologia, 40,* 575–584.

Carling, K. (2000). Resistant outlier rules and the non-Gaussian case. *Stat. Data Anal. 33*, 249–258.

Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *Neuroimage, 63*(1), 289-300.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal–parietal junction. *Trends in Cognitive Sciences, 17*, 328–336.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage, 50*(3), 1148-1167.

Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance, 37(2),* 409-421

Catmur, C., Walsh, V., Heyes, C.M., 2007. Sensorimotor learning configures the human mirror system. *Current Biology, 17 (17),* 1527–1531. doi: 10.1016/j.cub.2007.08.006

Catmur C., Walsh V.& Heyes C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Phil. Trans. R. Soc. B364*, 2369–2380. doi:10.1098/rstb.2009.0048.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology.* Guilford Press.

Chakrabarti, A., & Weber, R. (Eds.). (2015). *Comparative philosophy without borders*. Bloomsbury Publishing.

Chaminade, T., Franklin, D. W., Oztop, E., & Cheng, G. (2005). Motor interference between humans and humanoid robots: Effect of biological and artificial motion. In Proceedings of 2005. 4th IEEE International Conference On Development and Learning (pp. 96–101).

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behaviour link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. doi: 10.1037/0022-3514.76.6.893

Chartrand, T. L., & Lakin, J. L. (2013). The Antecedents and Consequences of Human Behavioural Mimicry. *Annual Review of Psychology, 64(1),* 285-308. doi:10.1146/annurev-psych-113011-143754

Chartrand, T. L., & Van Baaren, R. (2009). Human mimicry. *Advances in Experimental Social Psychology, 41*, 219–274. doi:10.1016/S0065-2601(08)00405-X

Cheng, C. M., & Chartrand, T. L. (2003). Self-monitoring without awareness: using mimicry as a nonconscious affiliation strategy. *Journal of personality and social psychology*, *85*(6), 1170.

Cho, Y. S., & Proctor, R. W. (2003). Stimulus and response representations underlying orthogonal stimulus-response compatibility effects. *Psychonomic Bulletin & Review*, *10*(1), 45-73.

Christov-Moore, L., Conway, P., & Iacoboni, M.(2017). Deontological dilemma response tendencies and sensorimotor representations of harm to others. *Frontiers in Integrative Neuroscience, 11,* 34.

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behaviour. *Neuroscience & Biobehavioural Reviews, 46*, 604–627.

Clayson, P. E., Clawson, A., & Larson, M. J. (2011). Sex differences in electrophysiological indices of conflict monitoring. *Biological Psychology, 87*, 282–289.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology, 112*, 155–159.

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. *Psychonomic bulletin & review,* 1-15.

Cook, J., & Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Experimental Brain Research, 211,* 601-612. doi:10.1007/s00221-011-2584-4

Cook, J., & Bird, G. (2012). Atypical Social Modulation of Imitation in Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders, 42(6),* 1045-1051. doi: 10.1007/s10803-011-1341-7

Cook, R., Bird, G., Lünser, G., Huck, S., and Heyes, C. (2012a). Automatic imitation in a strategic context: players of rock-paper-scissors imitate opponents' gestures. *Proceedings of the Royal Society B, 279*, 780–786. doi: 10.1098/rspb.2011.1024

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. *Behavioural and Brain Sciences, 37*(2), 177-192.

Cooper, R. P., Catmur, C., & Heyes, C. (2012). Are automatic imitation and spatial compatibility mediated by different processes? *Cognitive science*, *37*(4), 605-630.

Cooper, S., Jackson, J., Barch, D., & Braver, T. (2019). Neuroimaging of individual differences: A latent variable modeling perspective. *Neuroscience and Biobehavioural Reviews, 98*, 29-46.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron, 58*(3), 306-324.

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., ... & Brass, M. (2018b). Automatic imitation: A meta-analysis. *Psychological Bulletin*, *144*(5), 453.

Cracco, E., & Brass, M. (2019). Reaction time indices of automatic imitation measure imitative response tendencies. *Consciousness and cognition, 68*, 115-118.

Cracco, E., Genschow, O., Radkova, I., & Brass, M. (2018a). Automatic imitation of pro- and antisocial gestures: Is implicit social behaviour censored? *Cognition, 170,* 179–189

Cramer, A. O., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioural and brain sciences*, *33*(2-3), 137-150.

Crescentini, C., Mengotti, P., Grecucci, A., & Rumiati, R. I. (2011). The effect of observed biological and non biological movements on action imitation: an fMRI study. *Brain Research, 1420,* 80-92. doi:10.1016/j.brainres.2011.08.077

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American psychologist, 12*(11), 671. 87

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American psychologist, 30*(2), 116.

Cross KA, Iacoboni M. 2013. Optimised neural coding? Control mechanisms in large cortical networks implemented by connectivity changes. *Human Brain Mapping. 34,* 213-25.

Cross, K. A., Torrisi, S., Losin, E. A. R., & Iacoboni, M. 2013. Controlling automatic imitative tendencies: interactions between mirror neuron and cognitive control systems. *Neuroimage, 83,* 493-504.

Cross, K. A., & Iacoboni, M. 2014. To imitate or not: Avoiding imitation involves preparatory inhibition of motor resonance. *Neuroimage, 91,* 228-236.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

Cumming, G. (2014). The new statistics: Why and how. *Psychological science*, *25*(1), 7-29.

Daniel, D. B., Pelotte, M., & Lewis, J. (2000). Lack of sex differences on the Stroop Color-Word Test across three age groups. *Perceptual and motor skills, 90*(2), 483-484.

Darda, K. M., Butler, E. E., & Ramsey, R. (2018). Functional Specificity and Sex Differences in the Neural Circuits Supporting the Inhibition of Automatic Imitation*. Journal of cognitive neuroscience, 30*(6), 914-933.

Darda, K. M., & Ramsey, R. (2019). The inhibition of automatic imitation: a meta-analysis and synthesis of fMRI studies. *NeuroImage. 197,* 320-329.

Davis, M. H. (1980). Interpersonal reactivity index (IRI). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Decety, J., Chaminade, T., Grezes, J., & Meltzoff, A. N. (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation. *Neuroimage, 15*(1), 265-272.

Decety, J., & Meltzoff, A. N. (2011). Empathy, imitation, and the social brain. Empathy: *Philosophical and psychological perspectives*, 58-81.

De Coster, L., Mueller, S. C., Sjoen, G. T., Saedeleer, L. De, & Brass, M. (2014). The influence of Oxytocin on automatic motor simulation. Psychoneuroendocrinology, 50, 220–226. doi: 10.1016/j.psyneuen.2014.08.021

De Guzman M, Bird G, Banissy MJ, Catmur C. 2016. Self–other control processes in social cognition: from imitation to empathy. *Phil. Trans. R. Soc.B. 371*(1686):20150079.

De la Vega, A., Chang, L. J., Banich, M. T., Wager, T. D., & Yarkoni, T. (2016). Large-scale meta-analysis of human medial frontal cortex reveals tripartite functional organization. *Journal of Neuroscience, 36,* 6553–6562.

Dennett, D. C. (1991). Real patterns. *The journal of Philosophy, 88*(1), 27-51.

De Schotten, M. T., and Shallice, T. (2017). Identical, similar or different? is a single brain model sufficient? *PLoS One, 86*, 172–175. doi: 10.1016/j.cortex.2016.12.002

De Schryver, M., Hughes, S., Rosseel, Y., & De Houwer, J. (2016). Unreliable Yet Still Replicable: A Comment on LeBel and Paunonen (2011). *Frontiers in Psychology, 6.* doi: 10.3389/Fpsyg.2015.07039.

Dijksterhuis, A. (2005). Why we are social animals: The high road to imitation as social glue. Perspectives on imitation: *From neuroscience to social science, 2,* 207-220.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology, 19,* 643-647.

Dimberg, U. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology, 30,* 151–159.

Dimberg, U., & Lundquist, L. O. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology, 30*, 151–159.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180. doi: 10.1007/BF00230027

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage, 55,* 705–712.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. Psychological Assessment, 18, 192–203.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*, 2470–2473

Downing, P. E., & Peelen, M. V. (2011). The role of occipitotemporal body-selective regions in person perception. *Cognitive Neuroscience, 2*(3-4), 186-203.

Dubois, J., & Adolphs, R. (2016). Building a science of individual differences from fMRI. *Trends in cognitive sciences, 20*(6), 425-443.

Duffy, K. A., & Chartrand, T. L. (2015). Mimicry: causes and consequences. *Current Opinion in Behavioural Sciences, 3*, 112-116.

Dufour N, Redcay E, Young L, Mavros PL, Moran JM, Triantafyllou C, Gabrieli JD, Saxe R. 2013. Similar brain activation during false belief tasks in a large sample of adults with and without autism. *PloS One. 8*:e75468.

Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2017). Specific visual subregions of TPJ mediate reorienting of spatial attention. *Cerebral Cortex, 13,* 1–16.

Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, *36*(5), 562–572. doi: 10.1080/03014460902960289

Duncan J. 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci. 14,* 172-9.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25,* 1325–1335.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, U.S.A., 113,* 7900–7905.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16,* 143–149.

Eysenck, H. J. (1997). Personality and experimental psychology: the unification of psychology and the possibility of a paradigm. *Journal of Personality and social Psychology, 73*(6), 1224.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: a natural science approach.* Plenum Press. New York.

Farmer, H., Ciaunica, A., & Hamilton, A. F. D. C. (2018). The functions of imitative behaviour in humans. *Mind & Language, 33*(4), 378-396.

Fedorenko, E., Behr, M. K., & Kanswisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 16428–16433.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 110*, 16616–16621.

Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S. & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology, 104,* 1177–1194.

Ferri, F., Frassinetti, F., Mastrangelo, F., Salone, A., Ferro, F. M., and Gallese, V. (2012). Bodily self and schizophrenia: the loss of implicit self-body knowledge. *Conscious. Cogn. 21,* 1365–1374. doi: 10.1016/j.concog.2012.05.001

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender influences on brain responses to errors and post-error adjustments. *Scientific reports, 6,* 24435.

Fischer-Baum, S., Kook, J. H., Lee, Y., Ramos Nuñez, A. I., & Vannucci, M. (2018). Individual Differences in the Neural and Cognitive Mechanisms of Single Word Reading. *Frontiers in human neuroscience, 12,* 271.

Fishbein H. D. 1996. *Peer Prejudice and Discrimination: Evolutionary, Cultural, and Developmental Dynamics.* Westview Press: Boulder, CO

Flynn, E., & Smith, K. (2012). Investigating the Mechanisms of Cultural Acquisition. *Social Psychology*, *43*(4), 185–195. doi: 10.1027/1864-9335/a000119

Fox, P. T., & Friston, K. J. (2012). Distributed processing; distributed functions? *Neuroimage, 61*(2), 407-426.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences, 360*(1456), 815-836.

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain connectivity, 1*(1), 13-36.

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage, 62*(2), 1230-1233.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1,* 210–220.

Frith, C. D. (2008). Social cognition. Philosophical Transactions of the Royal Society B: *Biological Sciences, 363*(1499), 2033-2039.

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science, 286,* 1692–1695.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual review of psychology, 63,* 287-313.

Frühholz, S., Godde, B., Finke, M., & Herrmann, M. (2011). Spatio-temporal brain dynamics in a combined stimulus–stimulus and stimulus–response conflict task. *Neuroimage, 54*(1), 622-634.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind'in verbal and nonverbal tasks. *Neuropsychologia, 38(1*), 11-21.

Galef, B. G. (2013). Imitation and local enhancement: Detrimental effects of consensus definitions on analyses of social learning in animals. *Behavioural Processes*, *100*, 123–130. doi: 10.1016/j.beproc.2013.07.026

Gallese, V. (2001). The 'shared manifold' hypothesis. From mirror neurons to empathy. *Journal of consciousness studies*, *8*(5-6), 33-50.

Gallese, Vittorio, Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. doi: 10.1093/brain/119.2.593

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

Gazzaniga, M. S., & Bizzi, E. (Eds.). (2009). *The cognitive neurosciences* (4. ed). Cambridge, Mass.: MIT Press.

Geary, D. C. (2010). Male, female: The evolution of human sex differences. *American Psychological Association.*

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: contextual updating? *Neuroscience & Biobehavioural Reviews, 37*(10), 2608-2620.

Genon, S., Reid, A., Langner, R., Amunts, K., & Eickhoff, S. B. (2018). How to characterize the function of a brain region. Trends in cognitive sciences, 22(4), 350-364.

Genschow, O., Van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. *PLoS ONE, 12*(9), 1–21. doi: 10.1371/journal.pone.0183784

Gerdes, K. E., Lietz, C. A., & Segal, E. A. (2011). Measuring empathy in the 21st century: Development of an empathy index rooted in social cognitive neuroscience and social justice. Social Work Research, 35(2), 83-93.

Gigerenzer, G. (2018). Statistical Rituals: The Replication Delusion and How We Got There. Advances in Methods and Practices in Psychological Science, 2515245918771329.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. Personality and Individual Differences, 102, 74–78.

Gleibs, I. H., Wilson, N., Reddy, G., & Catmur, C. (2016). Group dynamics in automatic imitation*. PloS one, 11*(9), e0162880.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26–42.

Gowen, E., Bolton, E., & Poliakoff, E. (2016). Believe it or not: Moving non-biological stimuli believed to have human origin can be represented as human movement. *Cognition, 146,* 431–438. doi: 10.1016/j.cognition.2015.10.010

Gowen, E., Stanley, J., & Miall, R. C. (2008). Movement interference in autism-spectrum disorder. *Neuropsychologia, 46*(4), 1060–1068.

Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. Handbook of personality psychology (pp. 795-824). Academic Press.

Grecucci, A., Koch, I., & Rumiati, R. I. (2011). The role of emotional context in facilitating imitative actions. *Acta Psychologica, 138(2),* 311e315. doi: 10.1016/j.actpsy.2011.07.005.

Grecucci, A., Brambilla, P., Siugzdaite, R., Londero, D., Fabbro, F., Rumiati, R. I. (2013). Emotional resonance deficits in autistic children. *Journal of Autism and Developmental Disorders, 43,* 616-628. doi: 10.1007/s10803-012-1603-z

Greenwald, A. G. (1970) Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review, 77*(2), 73–99.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, *21*(2), 744-750.

Haggarty, M. E. (1909). Imitation in monkeys. *Journal of Comparative Neurology, 19,* 337-455.

Hahn, E., Thi, M. T. T., Hahn, C., Kuehl, L. K., Ruehl, C., Neuhaus, A. H., & Dettling, M. (2011). Test retest reliability of Attention Network Test measures in schizophrenia. *Schizophrenia Research, 133*(1–3), 218–222. doi: 10.1016/j.schres.2011.09.026

Hale, J., & Hamilton, A. F. D. C. (2016a). Cognitive mechanisms for responding to mimicry from others. *Neuroscience & Biobehavioural Reviews, 63*, 106-123.

Hale, J., & Hamilton, A. F. D. C. (2016b). Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific reports, 6*, 35295.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological bulletin, 85*(4), 845.

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current opinion in neurobiology, 14*(2), 233-238.

Hamilton, A. F. de C. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience, 3,* 91–105.

Hamilton, A. F. de C. (2014). Cognitive underpinnings of social interaction. *The Quarterly Journal of Experimental Psychology, 68(3),* 417-432 doi: 10.1080/17470218.2014.973424

Hamilton, A. F. D. C., Brindley, R. M., & Frith, U. (2007). Imitation and action understanding in autistic spectrum disorders: how valid is the hypothesis of a deficit in the mirror neuron system?. *Neuropsychologia*, *45*(8), 1859-1868.

Hansen J, Alves H, Trope Y (2016) Psychological distance reduces literal imitation: Evidence from an imitation-learning paradigm. *Journal of Experimental Psychology: Human Perception and Performance,* 320–330. doi: 10.1037/xhp0000150

Hazeltine E, Poldrack R, Gabrieli JDE. 2007. Neural activation during response competition. *J Cogn Neurosci. 12,* 118-129.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology, 13*(3), e1002106.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behaviour Research Methods, 50*(3), 1166-1186.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience*, *17*(5), 1123-1128.

Hendin, H. M., & Cheek, J. M. (1997). Assessing hypersensitive narcissism: A re-examination of Murray's Narcissism Scale. *Journal of research in personality, 31*(4), 588-599.

Hengartner, M. P., De Fruyt, F., Rodgers, S., Müller, M., Rössler, W., & Ajdacic-Gross, V. (2014). An integrative examination of general personality dysfunction in a large community sample. *Personality and mental health, 8*(4), 276-289.

Herpertz, S. C., & Bertsch, K. (2014). The social-cognitive basis of personality disorders. *Current opinion in psychiatry, 27*(1), 73-77.

Hess, U., & Bourgeois, P. (2010). You smile–I smile: Emotion expression in social interaction. *Biological psychology, 84(*3), 514-520

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research, 22,* 233–240. doi: 10.1016/j.cogbrainres.2004.09.009

Heyes, C. (2009). Evolution, development and intentional control of imitation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 364(1528),* 2293–2298. doi:10.1098/rstb. 2009.0049

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioural Reviews, 34*(4), 575-583.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin, 137(3),* 463-483. doi: 10.1037/a0022288

Heyes, C. (2017). When does social learning become cultural learning? *Developmental Science, 20*(2), e12350.

Hickok, G. (2013). Do mirror neurons subserve action understanding? *Neuroscience letters, 540*, 56-58.

Hirschfeld, L. A. & Gelman, S. A. (1994). *Mapping the mind: Domain specificity in cognition and culture.* New York: Cambridge University Press.

Hoffman, P. (2018). An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Scientific reports, 8*(1), 8145.

Hogeveen, J., & Obhi, S. S. (2013). Automatic imitation is automatic, but less so for narcissists. *Experimental Brain Research, 224,* 613-621. doi: 10.1007/s00221-012-3339-6

Hogeveen J, Obhi SS, Banissy MJ, Santiesteban I, Press C, Catmur C, Bird G. 2014. Task-dependent and distinct roles of the temporoparietal junction and inferior frontal cortex in the control of imitation. *Social Cognitive and Affective Neuroscience*, *10,* 1003-1009.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65,* 373–398.

Iacoboni, M. (2009). Imitation, Empathy, and Mirror Neurons. *Annual Review of Psychology*, *60*(1), 653–670. doi: 10.1146/annurev.psych.60.110707.163604

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *science*, *286*(5449), 2526-2528.

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., ... & Verma, R. (2014). Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences, 111*(2), 823-828.

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control*. Trends in cognitive sciences, 19*(3), 126-132.

Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in cognitive sciences, 13*(12), 524-531.

James, W. (1961). *Psychology: The briefer course*. New York: Harper.

Jansson, E., Wilson, A. D., Williams, J. H. G., & Mon-Williams, M. (2007). Methodological problems undermine tests of the ideo-motor conjecture. *Experimental Brain Research, 182*(4), 549–558

JASP Team (2018). JASP (Version 0.9)[Computer software]

Jeffreys, H. (1961). *Theory of Probability*. Oxford: UK: Oxford University Press.

Jimenez, L., Recio, S., Mendez, A., Lorda, M. J., Permuy, B., & Mendez, C. (2012). Automatic imitation and spatial compatibility in a key-pressing task. *Acta Psychologica, 141*(1), 96– 103

Josephs, O., & Henson, R. N. (1999). Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 354*, 1215–1228.

Judge, J., & Taylor, P. J. (2012). Gender differences on the semantic flanker task using transposed-letter target words. *The Quarterly Journal of Experimental Psychology, 65*(10), 2008-2017.

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage, 60,* 2357–2364.

Kanwisher N. 2010. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc Natl Acad Sci. 107,* 11163-70.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*(11), 4302–4311.

Kavanagh, L. C., & Winkielman, P. (2016). The functionality of spontaneous mimicry and its influences on affiliation: An implicit socialization account. *Frontiers in psychology, 7,* 458.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive processing, 8*(3), 159-166.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of Mirror Neurons in Human Inferior Frontal Gyrus. *Journal of Neuroscience*, *29*(32), 10153–10159. doi: 10.1523/JNEUROSCI.2668-09.2009

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*(6), 522–525.

Klapper, A., Ramsey, R., Wigboldus, D., & Cross, E. S. (2014). The Control of Automatic Imitation Based on Bottom-Up and Top-Down Cues to Animacy: Insights from Brain and Behaviour. *Journal of Cognitive Neuroscience, 26(11).* 2503-13. doi: 10.1162/jocn_a_00651.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27*(12), 712-719.

Kober, H., & Wager, T. D. (2010). Meta-analysis of neuroimaging data. Wiley Interdisciplinary Reviews: *Cognitive Science, 1*(2), 293-300.

Kohler, W. (1925). *The mentality of apes*. London: Routledge & Kegan Paul

Korb, S., Malsert, J., Rochas, V., Rihs, T. A., Rieger, S. W., Schwab, S., ... & Grandjean, D. (2015). Gender differences in the neural network of facial mimicry of smiles–An rTMS study. *Cortex, 70,* 101-114.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological review, 97*(2), 253.

Kornblum, S., & Lee, J. W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 855.

Kornblum, S. (1994). The way irrelevant dimensions are processed depends on what they overlap with: The case of Stroop- and Simon-like stimuli. *Psychological Research/Psychologische Forschung, 56*, 130-135.

Kornblum, S., Stevens, G. T., Whipple, A., & Requin, J. (1999). The effects of irrelevant stimuli: 1. The time course of stimulus–stimulus and stimulus–response consistency effects with Stroop-like stimuli, Simon-like tasks, and their factorial combinations. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 688.

Krach, S., Blumel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience, 10*, 9.

Krall, S. C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P. T., Eickhoff, S. B., ... & Konrad, K. (2015). The role of the right temporoparietal junction in attention and social

interaction as revealed by ALE meta-analysis. *Brain Structure and Function, 220(2),* 587-604.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2,* 4.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12,* 535–540.

Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature neuroscience, 15*(7), 940.

Kymissis, E., & Poulson, C. L. (1990). The history of imitation in learning theory: The language acquisition process. *Journal of the Experimental Analysis of Behaviour*, *54*(2), 113–127. doi: 10.1901/jeab.1990.54-113

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863.

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC psychology, 4*(1), 24.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioural mimicry to create affiliation and rapport. *Psychological science*, *14*(4), 334-339.

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The Chameleon Effect as Social Glue: Evidence for the Evolutionary Significance of Nonconscious Mimicry. *Journal of Nonverbal Behaviour*, *27*(3), 145–162. doi: 10.1023/A:1025389814290

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*(1), 42.

Lambert, A. J., & Scherer, L. (2013). Measurement and methodology in social cognition: A historical perspective. *The Oxford Handbook of Social Cognition*, 33.

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... & Fox, P. T. (2007). Bias between MNI and Talairach coordinates analysed using the ICBM-152 brain template. *Human brain mapping, 28*(11), 1194-1205.

Larsen, H., Overbeek, G., Granic, I., & Engels, R. C. (2010). Imitation of alcohol consumption in same-sex and other-sex dyads. *Alcohol and Alcoholism*, *45*(6), 557-562.

Lee, S. M., & McCarthy, G. (2014). Functional heterogeneity and convergence in the right temporoparietal junction. *Cerebral Cortex, 26*(3), 1108-1116.

Leighton, J., & Heyes, C. (2010). Hand to Mouth: Automatic Imitation Across Effector Systems. *Journal of Experimental Psychology-Human Perception and Performance, 36*(5), 1174–1183

Levant, R. F., Hall, R. J., Williams, C. M., & Hasan, N. T. (2009). Gender differences in alexithymia. *Psychology of men & masculinity, 10*(3), 190.

Levelt, W. J., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive psychology*, *14*(1), 78-106.

Levitin, D. J. (2014). *The organized mind: Thinking straight in the age of information overload*. Penguin.

Lhermitte, F., Pillon, B., & Serdaru, M. (1986). Human autonomy and the frontal lobes. Part I: Imitation and utilization behaviour: a neuropsychological study of 75 patients. *Annals of neurology, 19*(4), 326-334.

Li, Q., Nan, W., Wang, K., & Liu, X. (2014). Independent processing of stimulus-stimulus and stimulus-response conflicts. *PLoS One, 9*(2), e89249.

Liepelt, R., Prinz, W., & Brass, M. (2010). When do we simulate non-human agents? Dissociating communicative and non-communicative actions. *Cognition, 115*(3), 426– 434.

Likowski, K. U., Weyers, P., Seibt, B., Stöhr, C., Pauli, P., & Mühlberger, A. (2011). Sad and lonely? Sad mood suppresses facial mimicry. *Journal of Nonverbal Behaviour*, *35*(2), 101-117.

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Sage Publications, Inc.

Longo, M. R., & Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Experimental Brain Research, 192*(4), 739–744

Lundqvist, L. O. (1995). Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian journal of psychology, 36*(2), 130-141.

Lynn, M., & Snyder, C. R. (2002). Uniqueness seeking. In C. R. Snyder & S. J. Lopez (Eds.), Handbook of positive psychology (pp. 395-410). New York, NY, US:

Oxford University Press.MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin, 109*(2), 163.

Malpass, R.S., & Kravitz, J. (1969). Recognition for faces of own and other ‘‘race.’’ *Journal of Personality and Social Psychology, 13*, 330–334

Maran, T. (2017). *Mimicry and Meaning: Structure and Semiotics of Biological Mimicry*. Retrieved from https://www.springer.com/gp/book/9783319503158

Marsh LE, Bird G, Catmur C. 2016. The imitation game: Effects of social cues on ‘imitation’ are domain-general in nature. *NeuroImage. 139,* 368-75.

Marsman, M., Schönbrodt, F. D., Morey, R. D., Yao, Y., Gelman, A., & Wagenmakers, E. J. (2017). A Bayesian bird's eye view of ‘Replications of important results in social psychology’. *Royal Society Open Science, 4*(1), 160426.

Matzke, D., Nieuwenhuis, S., van Rijn, H., Slagter, H. A., van der Molen, M. W. & Wagenmakers, E.J. (2015) The effect of horizontal eye movements on free recall: A preregistered adversarial collaboration. *Journal of Experimental Psychology: General, 144*(1):e1–15.

Maurer, R. E., & Tindall, J. H. (1983). Effect of postural congruence on client's perception of counselor empathy. *Journal of counseling psychology*, *30*(2), 158.

Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology, 59*, 537-563. doi: 10.1146/annurev.psych.59.103006.093735

McCormick, T. H., Salganik, M. J., & Zheng, T. (2010). How Many People Do You Know?: Efficiently Estimating Personal Network Size. *Journal of the American Statistical Association*, *105*(489), 59–70. https://doi.org/10.1198/jasa.2009.ap08518

McCrae, R. R., & Costa Jr, P. T. (1999). A five-factor theory of personality. Handbook of personality: *Theory and research, 2*(1999), 139-153.

McHugo, G. J., Lanzetta, J. T., & Bush, L. K. (1991). The effect of attitudes on emotional reactions to expressive displays of political leaders. *Journal of Nonverbal Behaviour*, *15*(1), 19-41.

Mead, W. R., & Kurzweil, R. (2006). The Singularity Is near: When Humans Transcend Biology. *Foreign Affairs*, *85*(3), 160. doi: 10.2307/20031996

Meltzoff, A. N. (1988). Imitation of televised models by infants. *Child development, 59*(5), 1221.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Infant and child development, 6*(3-4), 179-192.

Mengotti, P., Dombert, P. L., Fink, G. R., & Vossel, S. (2017). Disruption of the right temporoparietal junction impairs probabilistic belief updating. *Journal of Neuroscience, 37*(22), 5419-5428.

Mengotti P, Corradi-Dell'Acqua C, Rumiati RI. 2012. Imitation components in the human brain: an fMRI study. *Neuroimage. 59,* 1622-30.

Merritt, P., Hirshman, E., Wharton, W., Stangl, B., Devlin, J., & Lenz, A. (2007). Evidence for gender differences in visual selective attention. *Personality and individual differences, 43*(3), 597-609.

Michael, J., & D'Ausilio, A. (2015). Domain-specific and domain-general processes in social perception—A complementary approach. *Consciousness and Cognition, 36*, 434–437.

Miller DI, Halpern DF (2014) The new science of cognitive sex differences. *Trends Cogn Sci 18*(1):37–45. doi:10.1016/j.tics.2013.10.011

Milner D. 1983.Children and Race:Ten Years On. Ward Lock Educational: London. Nesdale D. 2000. Developmental changes in children's ethnic preferences and social cognition. *Journal of Applied Developmental Psychology. 20*, 501–519.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology, 41*(1), 49-100.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioural Reviews*, *36*(1), 341–349. doi: 10.1016/j.neubiorev.2011.07.004

Mondillon, L., Niedenthal, P. M., Gil, S., & Droit-Volet, S. (2007). Imitation of in-group versus outgroup members' facial expressions of anger: a test with a time perception task. Social Neuroscience, 2(3-4), 223–37. doi: 10.1080/17470910701376894

Morgan, C. L. (1900). Animal behaviour. London: Arnold.

Morton, J. (2008). *Understanding developmental disorders: A causal modelling approach.* John Wiley & Sons.

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology.

Müller, B. C., Van Leeuwen, M. L., Van Baaren, R. B., Bekkering, H., & Dijksterhuis, A. (2013). Empathy is a beautiful thing: Empathy predicts imitation only for attractive others. *Scandinavian journal of psychology*, *54*(5), 401-406.

Mumford, J. A. (2012). A power calculation guide for fMRI studies. *Social Cognitive and Affective Neuroscience, 7*, 738–742.

Mumford, J. A., & Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *Neuroimage, 39,* 261–268.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., . . . Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1,* 0021. doi: 10.1038/s41562-016-0021

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews, 85*(4), 935–956. doi: 10.1111/j.1469-185X.2010.00141.x

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective & Behavioural Neuroscience, 7,* 1–17.

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annu Rev Psychol, 69*, 511-534. doi: 10.1146/annurev-psych-122216-011836.

Nee DE, Wager TD, Jonides J. 2007. Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cogn Affect Behav Neurosci. 7,* 1-17.

Newman-Norlund RD, van Schie HT, van Zuijlen AM, Bekkering H. 2007. The mirror neuron system is more active during complementary compared with imitative action. *Nat Neurosci, 10,* 817-819.

Newey, R., Koldewyn, K., & Ramsey, R. (2019). The influence of prosocial priming on visual perspective taking and automatic imitation. *PloS one*, *14*(1), e0198867.

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localisers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage, 63,* 1646–1669.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences, 10,* 424–430.

Nosek, B. A., & Lakens, D. (2014). Registered reports.

Oberman, L. M., & Ramachandran, V. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin, 133*(2), 310-327. doi: 10.1037/0033- 2909.133.2.310

Obhi, S. S., Hogeveen, J., Giacomin, M., & Jordan, C. H. (2013). Automatic imitation is reduced in narcissists. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 920-928. doi: 10.1037/a0034056

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, *56*(9), 717–734. doi: 10.1037/0003-066X.56.9.717

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

O'Reilly RC, Herd SA, Pauli WM. 2010. Computational models of cognitive control. *Curr. Opin. Neurobiol. 20,* 257–61

Over, H., & Carpenter, M. (2012). Putting the social into social learning: explaining both selectivity and fidelity in children's copying behaviour. *Journal of Comparative Psychology, 126*(2), 182.

Park, H. J., & Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science, 342*(6158), 1238411.

Park, S., Matthews, N., & Gibson, C. (2008). Imitation, Simulation, and Schizophrenia. Schizophrenia Bulletin, 34(4), 698-707. doi:10.1093/schbul/sbn048

Pashler, H., Coburn, N., Harris, C. R. (2012). Priming of Social Distance? Failure to Replicate Effects on Social and Food Judgments. PLoS ONE, 7(8), e42510. doi: 10.1371/journal.pone.0042510.

Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of personality and social psychology, 89*(4), 488.

Pernet, C.R., Wilcox, R. & Rousselet, G.A. (2013). Robust correlation analyses: false positive and power validation using a new open source Matlab toolbox. *Front. in Psychology, 3*, 606. doi: 10.3389/fpsyg.2012.00606

Peters, G. J. Y., & Crutzen, R. (2017). Pragmatic nihilism: how a Theory of Nothing can help health psychology progress*. Health Psychology Review, 11*(2), 103-121.

Piaget, J. (1952). The Origins of Intelligence in Children. New York, NY: W.W. Norton & Co. doi: 10.1037/11494-000

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage, 56*, 2356–2363.

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience, 27*(44), 11912-11924.

Ploran, E. J., Tremel, J. J., Nelson, S. M., & Wheeler, M. E. (2011). High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex. *Cerebral Cortex, 21*(11), 2650-2662.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. Attention and performance X: *Control of language processes, 32*, 531-556.

Powlishta K. K. 1995. Intergroup processes in childhood: Social categorization and sex role development. *Developmental Psychology, 31,* 781–788.

Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Cognitive Brain Research, 25*(3), 632–640.

Press, C., Bird, G., Walsh, E., & Heyes, C. (2008). Automatic imitation of intransitive actions. *Brain and Cognition, 67*(1), 44–50.

Press, C., Gillmeister, H., & Heyes, C. (2006). Bottom-up, not top-down, modulation of imitation by human and robotic models. *European Journal of Neuroscience, 24*(8), 2415–2419.

Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B-Biological Sciences, 274*(1625), 2509–2514.

Press, C., Richardson, D., & Bird, G. (2010). Intact imitation of emotional facial actions in autism spectrum conditions. *Neuropsychologia*, *48*(11), 3291-3297.

Prinz, W., & Meltzoff, A. N. (2002). An introduction to the imitative mind and brain. *The imitative mind: Development, evolution and brain bases*, 1-15.

Prinz, W. (2005). An ideomotor approach to imitation. *Perspectives on imitation: From neuroscience to social science, 1,* 141-156.

Pritchett, B. L., Hoeflin, C., Koldewyn, K., Dechter, E., & Fedorenko, E. (2018). High-level language processing regions are not engaged in action observation or imitation. Journal of neurophysiology, 120(5), 2555-2570.

Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., Macrae, C. N. (2011). Stereotype-based modulation of person perception. *Neuroimage, 57,* 549–557.

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition, 54,* 179–185.

Raine, A., & Benishay, D. (1995). The SPQ-B: A brief screening instrument for schizotypal personality disorder. *Journal of Personality Disorders, 9*(4), 346-355.

Ramachandran, V. S. (2000). Mirror neurons and imitation learning as the driving force behind “the great leap forward” in human evolution.

Ramsey, R. (2018). What are reaction time indices of automatic imitation measuring? *Consciousness and cognition, 65*, 240-254.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience, 2(1), 79.

Rauchbauer, B., Majdandžić, J., Hummer, A., Windischberger, C., & Lamm, C. (2015). Distinct neural processes are engaged in the modulation of mimicry by social group-membership and emotional expressions*. Cortex, 70,* 49-67, doi:10.1016/j.cortex.2015.03.007

Ray, E., & Heyes, C. (2011). Imitation in infancy: The wealth of the stimulus. *Developmental science*, *14*(1), 92-105.

Reber, P. (2010). What is the memory capacity of the human brain. *Scientific American*, *4*, 2010.

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. Nature Reviews Neuroscience, 1.Rizzolatti G. & Arbib, M.A. (1998). Language within our grasp. *Trends in Neuroscience, 21,* 188–94.

Rizzolatti, G., & Craighero, L. (2004). THE MIRROR-NEURON SYSTEM. *Annual Review of Neuroscience*, *27*(1), 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews neuroscience, 11*(4), 264.

Rogosa, D. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. C. Rawlings (Eds.), Methodological issues in ageing research (pp. 171–210). New York: Springer.

Ross, D. A., Richler, J. J., & Gauthier, I. (2015). Reliability of composite-task measurements of holistic face processing. *Behaviour Research Methods, 47*(3), 736–743. doi: 10.3758/s13428-014-0497-4

Rousseeuw, P.J. (1984), Least Median of Squares Regression. *Journal of the American Statistical Association, 79*, 871-881.

Rousseeuw, P.J. and Van Driessen, K. (1999), A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics, 41*, 212-223.

Rubia, K., Hyde, Z., Halari, R., Giampietro, V., & Smith, A. (2010). Effects of age and sex on developmental neural networks of visual-spatial attention allocation. *Neuroimage, 51,* 817–827.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature neuroscience, 4*(5), 546.

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of personality and social psychology, 87*(4), 494.

Russell, T. A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and Emotion, 21,* 1554–1564.

Santiesteban I, Banissy MJ, Catmur C, Bird G. 2012. Enhancing Social Ability by Stimulating Right Temporoparietal Junction. *Curr Biol. 22,* 2274-2277.

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. 2015a. Functional lateralization of temporoparietal junction–imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience, 42*(8), 2527-2533.

Santiesteban, I., Bird, G., Tew, O., Cioffi, M. C., & Banissy, M. J. (2015b). Mirror-touch synaesthesia: Difficulties inhibiting the other. *Cortex, 71*, 116-121, doi:10.1016/j.cortex.2015.06.019

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localisers. *Neuroimage, 30,* 1088–1096.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *Neuroimage, 19,* 1835–1842.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19*, 65–72.

Scheflen, A. E. (1964). The significance of posture in communication systems. *Psychiatry*, *27*(4), 316-331.

Schilbach L., Eickhoff S. B., Cieslik E. C., Kuzmanovic B., Vogeley K. (2011). Shall we do this together Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. Autism 16, 151–16210.1177/1362361311409258

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. Behavioral and brain sciences, 36(4), 393-414.Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in big five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*(1), 168-182. doi: 10.1037/0022-3514.94.1.168

Schulte-Rüther, M., Markowitsch, H. J., Shah, N. J., Fink, G. R., & Piefke, M. (2008). Gender differences in brain networks supporting empathy. *Neuroimage, 42*(1), 393-403.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioural Reviews, 42*, 9-34.

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping, 38,* 4788–4805.

Schuwerk T, Schurz M, Müller F, Rupprecht R, Sommer M. 2017. The rTPJ's overarching cognitive function in networks for attention and theory of mind. *Social cognitive and affective neuroscience. 12*(1):157-168.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1*(2), 353-367.

Semerari, A., Colle, L., Pellecchia, G., Buccione, I., Carcione, A., Dimaggio, G., ... & Pedone, R. (2014). Metacognitive dysfunctions in personality disorders: Correlations with disorder severity and personality styles. *Journal of Personality Disorders, 28*(6), 751-766.

Shea, N. (2009). Imitation as an inheritance system. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2429-2443.

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.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science, 22*(11), 1359-1366.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*, 174.

Singer, T. (2012). The past, present and future of social neuroscience: A European perspective. *NeuroImage*, *61*(2), 437–449. doi: 10.1016/j.neuroimage.2012.01.109

Smith, A. (1759). *The theory of moral sentiments*. London: A. Miller.

Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology, 43,* 433-443.

Sonnby-Borgström, M., Jönsson, P., & Svensson, O. (2003). Emotional empathy as related to mimicry reactions at different levels of information processing. *Journal of Nonverbal Behaviour, 27*(1), 3-23.

Sonnby-Borgström, M., Jönsson, P., & Svensson, O. (2008). Gender differences in facial imitation and verbally reported emotional contagion from spontaneous to emotionally regulated processing levels. *Scandinavian Journal of Psychology, 49*, 111–122. doi: 10.1111/j.1467-9450.2008.00626.x

Southgate, V., & Hamilton, A. F. de C. (2008). Unbroken mirrors: challenging a theory of Autism. *Trends in Cognitive Sciences, 12*(6), 225-229. doi: 10.1016/j.tics.2008.03.005

Sowden, S., Brewer, R., Catmur, C., & Bird, G. (2016). The specificity of the link between alexithymia, interoception, and imitation. *Journal of Experimental Psychology: Human Perception and Performance, 42*(11), 1687.

Sowden, S., & Catmur, C. (2015). The role of the right temporoparietal junction in the control of imitation*. Cerebral Cortex, 25*, 1107–1113.

Sowden, S., Koehne, S., Catmur, C., & Dziobek, I. (2015). Intact automatic imitation and typical spatial compatibility in autism spectrum disorder: Challenging the broken mirror theory. *Autism Research.*

Sowden, S., & Shah, P. (2014). Self-other control: a candidate mechanism for social cognitive function. *Frontiers in human neuroscience*, *8*, 789.

Spengler, S., Bird, G., & Brass, M. (2010a). Hyperimitation of Actions Is Related to Reduced Understanding of Others' Minds in Autism Spectrum Conditions. *Biological Psychiatry, 68*(12), 1148–1155

Spengler S, von Cramon DY, Brass M. 2009. Control of shared representations relies on key processes in mental state reasoning. *Human Brain Mapping. 30,* 3704-3718.

Spengler S, von Cramon DY, Brass M. 2010b. Resisting motor mimicry: Control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience, 5,* 401-416.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology, 1,* e42.

Spunt, R. P., & Adolphs, R. (2014). Validating the why/how contrast for functional MRI studies of theory of mind. *Neuroimage, 99,* 301–311.

Spunt, R. P., & Adolphs, R. (2015). Folk Explanations of Behaviour: A Specialized Use of a Domain-General Mechanism. *Psychological Science, 26*(6), 724–736. doi: 10.1177/0956797615569002

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nat Rev Neurosci, 18*(9), 559-567. doi: 10.1038/nrn.2017.76

Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology Human Perception and Performance, 33*(4), 915–926.

Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology*, *101*(2), 311-323.

Stoet, G. (2010). Sex differences in the processing of flankers. *The Quarterly Journal of Experimental Psychology, 63,* 633–638.

Stoet, G. (2017). Sex differences in the Simon task help to interpret sex differences in selective attention. *Psychological Research, 81,* 571–581.

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1746–1759. doi:10.1037/0096-1523.26.6.1746

Swanson, L. W., & Petrovich, G. D. (1998). What is the amygdala? *Trends in neurosciences, 21*(8), 323-331.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European journal of social psychology, 1*(2), 149-178.

Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. Trends in cognitive sciences, 22(3), 201-212.

Thakkar, K. N., Peterman, J. S., & Park, S. (2014). Altered brain activation during action imitation and observation in schizophrenia: A translational approach to investigating social dysfunction in schizophrenia. *American Journal of Psychiatry, 171*(5), 539-548

Thayer, J. F., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology, 41,* 243–246.

Thiel, C. M., Zilles, K., & Fink, G. R. (2004). Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: an event-related fMRI study. *Neuroimage, 21*(1), 318-328.

Thorndike, E. 1.. (1911). *Animal intelligence*. New York: Macmillan

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA, US: Harvard University Press.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioural and Brain Sciences*, *16*(3), 495–511. doi: 10.1017/S0140525X0003123X

Tononi, G., Edelman, G. M., & Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in cognitive sciences, 2*(12), 474-484.

Trujillo-Ortiz, A., R. Hernandez-Walls, K. Barba-Rojo and L. Cupul-Magana. (2007). HZmvntest:Henze-Zirkler's Multivariate Normality Test. A MATLAB file.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670-674.

Uzgiris, I. C. (1981). Two functions of imitation during infancy. *International Journal of Behavioural Development*, *4*(1), 1-12.

Vallar, G. (1993). The anatomical basis of spatial hemineglect in humans. In *Unilateral neglect* (pp. 37-69). Psychology Press.

Vallar, G., & Perani, D. (1987). The anatomy of spatial neglect in humans. In *Advances in psychology,* 45, 235-258). North-Holland.

van Baaren, R. B., Fockenberg, D. A., Holland, R. W., Janssen, L., & van Knippenberg, A. (2006). The moody chameleon: The effect of mood on non–conscious mimicry. *Social Cognition*, *24*(4), 426-437.

Van Baaren, R. B., Holland, R. W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioural consequences of imitation. *Journal of Experimental Social Psychology*, *39*(4), 393-398.

Van Baaren, R. B., Holland, R. W., Kawakami, K., & Van Knippenberg, A. (2004). Mimicry and prosocial behaviour. *Psychological science*, *15*(1), 71-74.

van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1528),* 2381-2389. doi:10.1098/rstb.2009.0057

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin, 35*(3), 321-335.

Van der Elst, W., Van Boxtel, M. P. J., Van Breakelen, G. J. P., Jolles, J. (2006). The Stroop Color-Word Test: Influence of Age, Sex, and Education; and Normative Data for a Large Sample Across the Adult Age Range. Assessment, 13(1), 62-79, doi: 10.1177/1073191105283427

Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental psychology, 50*(3), 881.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30,* 829–858.

van Schie HT, van Waterschoot BM, Bekkering H. 2008. Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J Exp Psychol Hum Percept Perform*. *34*, 1493.

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science, 13*(4), 411-417.

Verboten, S., & Hubert, M. (2005). LIBRA: a MATLAB Library for Robust Analysis, *Chemometrics and Intelligent Laboratory Systems*(75), 127–136

Vital, A. (2013, April 29). Counting The People Your Life Impacts - Infographic. Retrieved 22 July 2019, from: https://blog.adioma.com/counting-the-people-you-impact-infographic/

Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science, 17*(2), 171-176.

Vul, E., Harris, C., Wimkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality and social cognition. Perspectives on *Psychological Science, 4*(3), 274–290.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18*, 293–309.

Wager, T. D., Lindquist, M., & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: current and future directions*. Social cognitive and affective neuroscience, 2*(2), 150-158.

Wager TD, Sylvester CYC, Lacey SC, Nee DE, Franklin, M, Jonides J. 2005. Common and unique components of response inhibition revealed by fMRI. *NeuroImage. 27,* 323-340.

Wang, Y., & Hamilton, A. F. de C. (2014a). Anterior medial prefrontal cortex implements social priming of mimicry. *Social Cognitive and Affective Neuroscience, 10(4),* 486e493. doi:10.1093/scan/nsu076

Wang, Y., & Hamilton, A. F. de C. (2014b). Why does gaze enhance mimicry? Placing gaze-mimicry effects in relation to other gaze phenomena*. Quarterly Journal of Experimental Psychology, 67(4),* 747-762. doi:10.1080/17470218.2013.828316

Wang, Y., Newport, R., & Hamilton, A. F. de C. (2011a). Eye contact enhances mimicry of intransitive hand movements. *Biology letters, 7,* 7–10. doi: 10.1098/rsbl.2010.0279

Wang, Y., Ramsey, R., & Hamilton, A. F. de C. (2011b). The control of mimicry by eye contact is mediated by medial prefrontal cortex. *Journal of Neuroscience, 31(33),* 12001e12010. doi:10.1523/Jneurosci.0845-11.2011.

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. Cognition, 191, 103997.

Webb JT (1972) Interview synchrony: An investigation of two speech rate measures in an automated standardized interview In: Pope B, Siegman AW, editors. *Studies in dyadic communication.* New York: Pergamon; pp. 115–133.

Weeks, D. J., & Proctor, R. W. (1990). Salient-features coding in the translation between orthogonal stimulus and response dimensions. *Journal of Experimental Psychology: General*, *119*(4), 355–366. doi: 10.1037//0096-3445.119.4.355

Weeks, D. J., Proctor, R. W., & Beyak, B. (1995). Stimulus–response compatibility for vertically oriented stimuli and horizontally oriented responses: Evidence for spatial coding. *The Quarterly Journal of Experimental Psychology*, *48*(2), 367-383.

White, K., & Argo, J. J. (2011). When imitation doesn't flatter: The role of consumer distinctiveness in responses to mimicry. Journal of Consumer Research, 38(4), 667-680.

Whiten, A., Horner, V., Litchfield, C. A., & Marshall-Pescini, S. (2004). How do apes ape? *Animal Learning & Behaviour*, *32*(1), 36–52. doi: 10.3758/BF03196005

Whiten, A., & Van Schaik, C. P. (2007). The evolution of animal 'cultures' and social intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 603-620.

Wiggett, A. J., Hudson, M., Tipper, S. P., & Downing, P. E. (2011). Learning associations between action and perception: Effects of incompatible training on body part and spatial priming. *Brain and Cognition, 76,* 87–96.

Wilcox, R. R. (2012). Introduction to Robust Estimation and Hypothesis Testing, 3rd Edn. Oxford: Academic Press.

Williams, J. H., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. Journal of Autism and Developmental Disorders, 34(3), 285-299.

Williams, J. H., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. Neuroscience and Biobehavioural Review, 25, 287–295. doi: 10.1016/S0149-7634(01)00014-8

Witmer, L. (1910). Intelligent Imitation and Curiosity in a Monkey. *The Psychological Clinic, 3,* 225-227.

Yarkoni, T., & Braver, T. S. (2010). Cognitive Neuroscience Approaches to Individual Differences in Working Memory and Executive Control: Conceptual and Methodological Issues. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of Individual Differences in Cognition* (pp. 87–108). New York: Springer.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*, 665.

Yee M, Brown R. 1994. The development of gender differentiation in young children. *British Journal of Social Psychology, 33*, 183–196.

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. Nature communications, 9(1), 4288.

Yon, D., & Press, C. (2018). Sensory predictions during action support perception of imitative reactions across suprasecond delays. Cognition, 173, 21-27.

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social cognitive conflict resolution: Contributions of domain general and domain specific neural systems. *Journal of Neuroscience, 30*(25), 8481–8488. doi:10.1523/JNEUROSCI.0382-10.2010

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15,* 675.

Zentall, T. R. (2006). Imitation: Definitions, evidence, and mechanisms. *Animal Cognition*, *9*(4), 335–353. doi: 10.1007/s10071-006-0039-2

Zhang, H. H., Zhang, J., & Kornblum, S. (1999). A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognitive psychology, 38*(3), 386-432.

Zhang, C., Dougherty, C. C., Baum, S. A., White, T., & Michael, A. M. (2018). Functional connectivity predicts gender: evidence for gender differences in resting brain connectivity. *Human brain mapping, 39*(4), 1765-1776.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2017). Making Replication Mainstream. *Behav Brain Sci,* 1-50. doi: 10.1017/S0140525X17001972

APPENDICES

Appendix A - Supplementary Information (Chapter 2)

Development of Stimuli

The imitation inhibition task used in Experiment 2 in the present study is a modified version of a previously existing paradigm (Brass et al., 2000; Catmur & Heyes, 2010). In order to avoid any own-sex bias, we decided to use a hand stimulus which was rated as neutral. 51 participants (19 males, $Mean_{age}$ = 23.49, SD_{age} = 3.12) other than the ones who participated in the current study were asked to rate 18 white Caucasian hand stimuli (9 male hands, 9 female hands) on a scale of 1 to 9, with 1 being "extremely masculine," 9 being "extremely feminine" and 5 being "neutral." An average of the rating score was obtained for each hand stimulus. The hand which had an average rating closest to 5 (Mean = 5.08) was taken as the final stimulus as it was considered to be rated most "neutral" amongst all other stimuli. In the present experiment, we again asked all participants to rate the hand they saw in the imitation inhibition task on the same scale: 1 to 9, with 1 being "extremely masculine," 9 being "extremely feminine" and 5 being "neutral." The mean of the hand ratings was 5.27; thus, on average, participants perceived the hand as 'neutral.'

Table S1.1. *Responses in each MD network fROI for spatial and imitative compatibility when individual contrasts were thresholded at p<.001, uncorrected.*

ROI	ROI size	Inter-subject overlap	Average ROI mask size	Spatial Compatibility			Imitative Compatibility		
				t	p-value	p-FDR	t	p-value	p-FDR
L_SPL	1173	0.92	413	1.71	0.047	**0.063**	1.49	0.072	0.137
L_IPS	287	0.86	85	1.36	0.090	0.097	2.91	0.003	**0.043**
L_IPL	641	0.82	89	1.91	0.032	**0.055**	2.70	0.005	**0.043**
L_MFG	536	0.84	140	2.37	0.012	**0.035**	0.99	0.165	0.240
L_PrecG	338	0.86	103	1.76	0.043	0.062	0.33	0.371	0.396
L_IFG	181	0.74	58	0.11	0.457	0.457	0.77	0.223	0.297
L_Insula	197	0.7	48	2.45	0.010	**0.035**	0.55	0.294	0.336
L_SMA	294	0.88	86	3.00	0.002	**0.019**	0.03	0.489	0.489
R_SPL	1181	0.9	415	1.66	0.052	0.064	2.32	0.012	0.050
R_IPS	227	0.84	71	1.36	0.091	0.097	2.21	0.017	0.053
R_IPL	599	0.76	111	1.89	0.034	**0.055**	2.40	0.011	0.050
R_MFG	535	0.88	144	3.61	<0.001	**0.007**	1.80	0.039	0.093
R_PrecG	269	0.74	56	2.06	0.024	**0.048**	1.57	0.064	0.128
R_IFG	265	0.7	56	2.34	0.013	**0.035**	1.81	0.041	0.093
R_Insula	184	0.78	44	2.22	0.017	**0.037**	0.67	0.255	0.314
R_SMA	328	0.84	101	2.32	0.013	**0.034**	1.05	0.150	0.240

Table S1.1. Responses in each MD network fROI for spatial and imitative compatibility when individual contrasts were thresholded at p<.001, uncorrected. Cells in bold show fROIs which survived correction for multiple comparisons (p<.05, FDR corr.).

Table S1.2.

Responses in each ToM network fROI for spatial and imitative compatibility when individual contrasts were thresholded at p<.001, uncorrected.

ROI	ROI size	Inter-subject overlap	Average ROI mask size (voxels)	Spatial Compatibility			Imitative Compatibility		
				t	p-value	p-FDR	t	p-value	p-FDR
DMPFC	576	1	0.58	-1.23	0.88	0.88	-0.09	0.54	0.82
MMPFC	494	1	0.56	-0.02	0.49	0.88	-0.56	0.71	0.82
VMPFC	382	1	0.44	-0.96	0.82	0.88	-0.93	0.82	0.82
RTPJ	1018	1	0.92	1.20	0.12	0.47	0.78	0.22	0.82

Table S1.2. Responses in each ToM network fROI for spatial and imitative compatibility when individual contrasts were thresholded at p<.001, uncorrected. Cells in bold show fROIs which survived correction for multiple comparisons (p<.05, FDR corr.).

Table S2.1 *Small volume correction (SVC) with MD network mask for the general compatibility effect and the sex*compatibility interaction.*

Region	Cluster Size	P FWE Corr	t-value	MNI coordinates		
				x	y	z
GENERAL COMPATIBILITY (Incompatible>Compatible)						
Left inferior parietal lobule (extending into the left postcentral gyrus)	**382**	**<0.001**	**8.40**	**-39**	**-40**	**43**
			5.57	**-48**	**-40**	**61**
			5.31	**-36**	**-43**	**67**
Left intraparietal sulcus	**79**	**0.005**	**6.50**	**-36**	**-52**	**46**
			4.27	**-30**	**-61**	**49**
Left precentral gyrus	**180**	**<0.001**	**6.38**	**-27**	**-7**	**70**
			5.11	**-33**	**-4**	**52**
			4.78	**-42**	**-1**	**55**
			4.32	**-27**	**-10**	**55**
			4.22	**-39**	**-4**	**52**
Left superior parietal lobule	14	0.249	5.31	-27	-52	70
			3.95	-15	-55	73
			3.80	-12	-58	70
Right middle frontal gyrus	**86**	**0.004**	**5.12**	**27**	**-1**	**70**
			4.04	**42**	**2**	**58**
L insula	21	0.149	4.67	-36	17	-2
Right inferior parietal lobule (extending into the right intraparietal sulcus and postcentral gyrus)	**270**	**<.001**	**4.50**	**48**	**-34**	**37**
			4.47	**39**	**-46**	**67**
			4.46	**48**	**-34**	**43**
			4.41	**45**	**-43**	**64**
			4.41	**42**	**-34**	**40**
			4.39	**33**	**-37**	**37**
			4.38	**54**	**-40**	**40**
			4.35	**45**	**-40**	**52**
			4.29	**36**	**-46**	**46**
			4.28	**39**	**-46**	**52**
			4.27	**45**	**-46**	**55**
Right superior parietal lobule	29	0.086	4.40	36	-52	67
			3.97	42	-52	52
SEX*COMPATIBILITY [Female (Incompatible>Compatible) > Male (Incompatible>Compatible)]						
No suprathreshold clusters						

Table S2.1. Regions surviving a voxel-level threshold of p<.001 and 10 voxels are reported for the general compatibility effect and sex*compatibility interaction, small volume corrected using the MD mask**.** Subclusters at least 8 mm from the main peak are listed. Bold font indicates clusters that survive correction for multiple corrections using a family-wise error (FWE) correction (p < .05). MNI = Montreal Neurological Institute**.**

Table S2.2. *Small volume correction (SVC) with ToM network mask for the general compatibility effect and the sex*compatibility interaction.*

Region	**Cluster Size**	**P FWE Corr**	**t-value**	**MNI coordinates**		
				x	y	z
GENERAL COMPATIBILITY (Incompatible>Compatible)						
Right temporo-parietal junction (supramarginal gyrus)	10	0.124	3.60	57	-43	37
			3.54	51	-40	34
SEX*COMPATIBILITY [Female (Incompatible>Compatible) > Male (Incompatible>Compatible)]						
No suprathreshold clusters						

Table S2.2. Regions surviving a voxel-level threshold of p<.001 and 10 voxels are reported for the general compatibility effect and sex*compatibility interaction, small volume corrected using the ToM mask**.** Subclusters at least 8 mm from the main peak are listed. Bold font indicates clusters that survive correction for multiple corrections using a family-wise error (FWE) correction (p < .05). MNI = Montreal Neurological Institute**.**

Table S3. *Showing Mean RT and SD for each condition of the imitation task for both males and females.*

	Spatially Compatible		**Spatially Incompatible**		**Imitatively Compatible**		**Imitatively Incompatible**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Males	666.20	157.50	699.42	159.76	675.00	158.56	690.62	158.93
Females	736.94	130.34	787.83	134.19	754.88	133.91	769.99	130.65

Table S4.1. *Responses in each MD network fROI for the MD network localiser contrast.*

ROI	ROI size	Inter-subject overlap	Average ROI mask size (voxels)	t	p-value	p-FDR
L_SPL	1173	1	117	11.69	<.0001	**<.0001**
L_IPS	287	1	28	11.15	<.0001	**<.0001**
L_IPL	641	1	64	10.30	<.0001	**<.0001**
L_MFG	536	1	53	11.08	<.0001	**<.0001**
L_PrecG	338	1	33	10.40	<.0001	**<.0001**
L_IFG	181	1	18	9.38	<.0001	**<.0001**
L_Insula	197	1	19	12.26	<.0001	**<.0001**
L_SMA	294	1	29	14.56	<.0001	**<.0001**
R_SPL	1181	1	118	11.36	<.0001	**<.0001**
R_IPS	227	1	22	10.18	<.0001	**<.0001**
R_IPL	599	1	59	9.25	<.0001	**<.0001**
R_MFG	535	1	53	11.61	<.0001	**<.0001**
R_PrecG	269	1	26	9.43	<.0001	**<.0001**
R_IFG	265	1	26	9.13	<.0001	**<.0001**
R_Insula	184	1	18	12.43	<.0001	**<.0001**
R_SMA	328	1	32	11.38	<.0001	**<.0001**

Table S4.1. All MD network fROIs were significantly responsive to the Hard>Easy contrast and survived correction for false discovery rate ($p<.05$).

Table S4.2. *Responses in each ToM network fROI for the ToM network localiser contrast.*

ROI	ROI size	Inter-subject overlap	Average ROI mask size	t	p-value	p-FDR
DMPFC	576	1	57	7.097	<.001	**<.001**
MMPFC	494	1	49	7.065	<.001	**<.001**
VMPFC	382	1	38	5.704	<.001	**<.001**
RTPJ	1018	1	101	15.025	<.001	**<.001**

Table S4.2. All ToM network fROIs were significantly responsive to the Belief>Photo contrast and survived correction for false discovery rate ($p<.05$).

Table S5. *Whole-brain analysis (Experiment 2).*

Region	**Cluster Size**	**P FWE Corr**	**t-value**	**MNI coordinates**		
				x	y	z
(A) GENERAL COMPATIBILITY						
Left inferior parietal lobule	**333**	**<0.001**	**5.28**	**-42**	**-28**	**43**
			5.10	**-48**	**-28**	**34**
			3.91	**-30**	**-49**	**43**
Left middle frontal gyrus	**130**	**0.018**	**5.05**	**-27**	**-10**	**49**
Right inferior parietal lobule extending into the right postcentral gyrus	**437**	**<0.001**	**4.85**	**45**	**-34**	**46**
			4.81	**48**	**-25**	**43**
			4.28	**30**	**-46**	**46**
Right middle and superior gyri extending into the right posterior-medial frontal cortex	**259**	**0.001**	**4.85**	**30**	**-4**	**52**
			4.28	**21**	**-7**	**64**
			4.23	**15**	**5**	**52**
Left posterior-medial frontal	29	0.476	4.19	-9	-1	55
Right inferior frontal gyrus and right insula lobe	**96**	**0.050**	**3.87**	**42**	**2**	**19**
				48	**8**	**22**
				42	**17**	**1**
Right precuneus	36	0.374	3.77	12	-64	52
			3.56	18	-70	49
Right supramarginal gyrus	26	0.527	3.75	63	-40	31
			3.68	63	-49	28
Right cerebellum	12	0.805	3.68	36	-49	-35
(B) SPATIAL COMPATIBILITY						
Intraparietal sulcus extending into the right postcentral gyrus	**136**	**0.019**	**4.65**	**33**	**-43**	**46**
			3.87	**48**	**-25**	**43**
Bilateral posterior medial frontal	**117**	**0.033**	**4.55**	**-6**	**-1**	**55**
			4.22	**9**	**5**	**52**
			3.48	**-9**	**5**	**46**
Left precentral gyrus	26	0.538	4.49	-57	5	31
Left precentral gyrus	68	0.139	4.34	-24	-16	58
			4.28	-27	-10	52
Right precentral gyrus	27	0.521	4.30	63	8	28

Right insula lobe	46	0.282	4.23	39	14	1
Right superior frontal gyrus	122	0.028	4.16	27	-7	58
			4.15	27	-7	49
Left inferior parietal lobule	**75**	**0.025**	**3.95**	**-36**	**-37**	**40**
			3.66	**-48**	**-31**	**40**
Left insula lobe	17	0.702	3.92	-33	14	7
Left postcentral gyrus	13	0.780	3.69	-63	-16	34
Right precuneus	22	0.607	3.67	12	-70	46
			3.60	12	-61	46
	10	0.838	3.39	57	-16	22
(C) IMITATIVE COMPATIBILITY						
Right supramarginal gyrus and right postcentral gyrus	40	0.349	3.76	45	-34	43
			3.64	42	-34	55
Left inferior parietal lobule	15	0.666	3.50	-42	-28	40
(D) SEX*COMPATIBILITY (GENERAL)						
Right superior occipital gyrus	10	0.845	3.48	30	-67	28
(E) SEX*COMPATIBILITY (SPATIAL)						
Right superior occipital gyrus	10	0.838	3.65	33	-64	31
(F) SEX*COMPATIBILITY (IMITATIVE)						
No suprathreshold clusters						

Table S5. Regions surviving a voxel-level threshold of p<.001 and 10 voxels are reported for (A) general compatibility (B) spatial compatibility (C) imitative compatibility effects and sex*compatibility interactions separately for (D) general (E) spatial and (F) imitative compatibility effects. Subclusters at least 8 mm from the main peak are listed. Bold font indicates clusters that survive correction for multiple corrections using a family-wise error (FWE) correction (p < .05). MNI = Montreal Neurological Institute.

Table S6. *Detailed table of fMRI studies using the imitation inhibition task and the contributions of ToM and MD networks in imitation inhibition.*

	Sample		No. of trials per condition	Design	Task Instructions	ROI/Whole brain		Regions			Thresholding
	M:F	**Age**		**Block/ Event-related**		**ROI**	**WB**	**mPFC**	**rTPJ**	**MD**	
1. Brass, Zysset, & von Cramon, 2001	10 (4:6)	23.5	80 congruent, 80 incongruent	Mixed	Block1: tap index finger Block2: lift index finger		Y	Y (Frontopolar cortex, BA 10)	N	Y (MFG, Cuneus, Anterior parietal cortex)	p<.001, uncorrected
2. Brass, Derfuss, & von Cramon. 2005	20 (8:12) – 10 for Imi, 10 for Stroop	26	40 congruent, 40 incongruent, 40 baseline, 40 null	Event-related	Index finger for '1' Middle finger for '2'		Y	Y aFMC	Y (BA40)	Y	p<.001, uncorrected
3. Brass, Ruby, & Spengler, 2009	20		35 simultaneous congruent, 35 simultaneous incongruent, 35 delayed congruent, 35 delayed incongruent,	Mixed	Same as above Simultaneous: number cue appeared with irrelevant hand Delayed: response led to appearance of irrelevant hand	Y		Y (for simultaneous incongruent > congruent only)*	Y (for simultaneous incongruent > congruent only)*		P<.001, uncorrected
4. Spengler, von Cramon, & Brass, 2009	18 (9:9)	25	72 incongruent, 72 congruent, 36 null	Event-related	Same as 2.	Y**		Y (overlap with self-referential and ToM tasks; -ve correlation with RT interference)	Y (overlap with agency and ToM tasks)	Y (see supple-mentary material: SII, MFG)	P<.05, corrected for multiple comparisons

5.	Bien, Roebroeck, Goebel, & Sack, 2009	15 (5:10)	23	64 imitative congruent, 64 imi incongruent, 64 spat congruent, 64 spat incongruent	Mixed	Block 1: imitate finger movement (imitative trials) Block2: follow spatial cue for movement (spatial trials)		Y	N	N	Y***(premotor cortex, bilateral posterior parietal and frontal /parietal opercular cortex, right STS)	P<.045, cluster size threshold = 50
6.	Crescentini, Mengotti, Grecucci, & Rumiati, 2011	19 (9:10)	24.6	60 biological congruent, 60 bio incongruent, 60 non-bio congruent, 60 non-bio incongruent	Mixed	Modified version of 1. With biological (human hand) and non-biological (white dot) stimuli; but ppts responded after movement offset instead of onset		Y	N	N	Y (only right Insula)	P<.05, corrected for multiple comparisons
7.	Cross & Iacoboni, 2013	24 (12:12)			Mixed	Block1: imitate finger movements Block2: imitate spatial dot movement		Y	N	N	N	P<.05, FWE corrected
8.	Mengotti, Corradi-Dell'Acqua, & Rumiati, 2012	22 (10:12)	24.4	80 per condition	Mixed	Task1: tap anatomical finger Task2: tap finger on same side of space		Y	N	N/Y (for only AN_NS over all others i.e. for the condition 'imi comp + spat incomp')	Y****	P<.05 FWE
9.	Cross, Torrisi, Losin, & Iacoboni, 2013	25 (5:15); 20 included in analysis	19-39	80 imi congruent, 80 imi incongruent, 80 spat cong, 80 spat incong, 80 nulls	Block	Block1: lift index finger on finger movement Block 2: lift middle finger on finger movement Block ¾: lift index/middle finger when dot moves		Y	Y	N	Y (ACC, bilateral insula extending into frontal pole and orbitofrontal cortex, IFG, PrecG, SPL)	P<.05, FWE corrected

10. Klapper, Rasey, Wigboldus, & Cross, 2014	19 (2:17)	21.95	160 congruent, 160 incongruent	Event related	Index for '1' Middle for '2'	Y	Y	Y (at p<.005, uncorrected: incong>cong, human>non-human)	N/Y (at p<.005, uncorr for human>non-human) and 3-way cong x form x belief at p<.05 FWE corr		
11. Marsh, Bird, & Catmur, 2016	24 (7:17)	23.71	80 imi cong, 80 imi incong, 80 spat cong, 80 spat incong	Event related	Lift index for '1' Lift middle for '2'		Y	N	N	Y (IFG, IPL, ACC); also left TPJ	P<.05, FWE corrected
12. Wang, Ramsey, & Hamilton, 2011	20 (5:15)	23	96 congruent (averted + direct gaze), 96 incongruent, 54 catch trials	Mixed	Block1: Hand open Block2: Hand closed		Y	Y (for averted incong > cong)	Y (for averted incong > cong)	Y (main effect:IPL, Cuneus); (averted: MOG, MTG, STS, temporal pole, IFG, precuneus, MFG, SPL, PMC, IPL, cuneus)	P<.05, FWE corrected

Appendix B – Supplementary Information (Chapter 4)

Questionnaires used in the current paper.

The Mini International Personality Item Pool (mini-IPIP; (Donnellan, Oswald, Baird, & Lucas, 2006), the Short Autism Spectrum Quotient (AQ-10 Adult; (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), the Brief Schizotypal Personality Questionnaire (SPQ-B; (Raine & Benishay, 1995), and the Narcissistic Personality Inventory (NPI-16; (Ames, Rose, & Anderson, 2006) are described in detail in previous work (Butler et al., 2015). The questionnaires new to the current study are described below.

Hypersensitivity Narcissism Scale (HSNS). The HSNS (Hendin & Cheek, 1997) is a 10-item assessment of the vulnerability-sensitivity component of narcissism, as opposed to the grandiosity-exhibitionism component of narcissism which is measured by the NPI-16 (Hendin & Cheek, 1997). Participants answer each item by deciding to what extent it is characteristic of their feelings or behaviour on a 5 point scale from 1 "very uncharacteristic or untrue, strongly disagree" to five "very characteristic or true, strongly agree". An example item is "I can become entirely absorbed in thinking about my personal affairs, me health, my cares or my relations to others". Scores for each participant on all 10 items were summed so that scores ranged from 10 being very low narcissism to 50 being very high narcissism.

Interpersonal Reactivity Index (IRI). The IRI (Davis, 1980) is a 28-item measure of dispositional empathy whereby an assumption of the measure is that empathy comprises a set of separate but related constructs. As such, responses on items from each of the four subscales are summed separately. Participants indicate how well each item describes them, using the letters A-E where A indicates that the item does not describe them well, and E indicates that the item does describe them well. An "A" response scores 0, and an "E" response scores 4. There are seven items per subscale, thus scores on each subscale range from zero to 28, where 28 indicates high levels of that particular subscale of empathy. Subscales are perspective taking (example item: "I sometimes try to understand my friends better by imagining how things look from their perspective"), empathic concern (example item: "I often have tender, concerned feelings for people less fortunate than me"), personal distress (example item: "Being in a tense emotional situation scares me"), and fantasy (example item: "When I am reading an interesting story

or novel, I imagine how I would feel if the events in the story were happening to me"). The subscales of empathic concern and alexithymia were used in the current experiments.

Toronto Alexithymia Scale (TAS-20). The TAS-20 (Bagby, Parker, & Taylor, 1994) is a 20-item measure of alexithymia which is characterized by difficulty identifying and describing emotions as well as minimization of emotional experience with an externally focused attention. On each item, participants indicate to what extent they agree with each statement where 1 indicates strong disagreement and 5 indicates strong agreement. Scores are summed so that final scores range from 20 being very low alexithymia to 100 being very high alexithymia. The TAS-20 comprises three subscales; difficulty describing feelings (5 items; e.g., "It is difficult for me to find the right words for my feelings"), difficulty identifying feelings (7 items; e.g., "I am often confused about what emotion I am feeling"), and externally-oriented thinking (8 items; "I prefer to talk to people about their daily activities rather than their feelings"). The total score of alexithymia was used for the current study.

Development of Stimuli for Experiment 3 (Chapter 4).

Forty-eight participants (18 males) outside of those recruited for the tasks in Experiment 3 filled out an online stimuli rating form. Participants were presented with 18 hand images (9 female hands and 9 male hands). They were asked to rate them on a scale of 1 to 9 (with 1 = extremely masculine and 9 = extremely feminine). Out of the nine male stimuli, four stimuli rated as most masculine were chosen as male stimuli in the current experiment. Average rating for the male stimuli was 3.57. Out of the nine female stimuli, four stimuli rated as most feminine were chosen as female stimuli in the current experiment. Average rating for the female stimuli was 6.63. The eighteen stimuli and participant ratings can be found here (https://osf.io/fsh9b/; folder named "hands_stimuli_exp3").

Table S1.
Descriptive Statistics for Experiment 1.

A. Imitation Task - Accuracy				
Sex	**Compatibility**	**N**	**Mean**	**SD**
female	compatible	116	98.21	2.68
female	incompatible	116	90.28	6.90
male	compatible	59	98.18	2.45
male	incompatible	59	90.97	7.09
B. Imitation Task – Reaction Time				
female	compatible	116	488.02	63.76
female	incompatible	116	584.45	91.34
male	compatible	59	487.59	86.20
male	incompatible	59	571.62	101.58
C. Flanker Task - Accuracy				
female	compatible	123	94.73	7.75
female	incompatible	123	95.59	7.88
male	compatible	59	95.07	4.83
male	incompatible	59	95.64	6.085
D. Flanker Task – Reaction Time				
female	compatible	123	434.52	54.26
female	incompatible	123	491.31	63.97
male	compatible	59	421.58	60.27
male	incompatible	59	472.74	59.67

N.B. Abbreviations: N=sample size, SD=standard deviation

Table S2.
Complete ANOVA information for Experiment 1.

A. Imitation Task – Accuracy					
Effect	**DFn**	**DFd**	**F**	**p**	**pes**
Sex	1	173	0.2226	0.6377	0.0013
Compatibility	1	173	258.0956	<.001	0.5987
Sex*Compatibility	1	173	0.6016	0.439	0.0035
B. Imitation Task – Reaction Time					
Sex	1	173	0.2598	0.6109	0.0015
Compatibility	1	173	669.7751	<.001	0.7947
Sex*Compatibility	1	173	3.1637	0.0771	0.018
C. Flanker Task – Accuracy					
Sex	1	180	0.0362	0.8493	0.0002
Compatibility	1	180	2.2379	0.1364	0.0123
Sex*Compatibility	1	180	0.0944	0.7591	0.0005
D. Flanker Task – Reaction Time					
Sex	1	180	3.0965	0.0802	0.0169
Compatibility	1	180	334.1472	<.001	0.6499
Sex*Compatibility	1	180	0.9059	0.3425	0.005

Table S3. *Complete information of the multiple regression models with personality traits (Experiment 1).*

		Model 1 (Base Model); f^2 = 0.51					Model 2 (Alexithymia); f^2 = 0.001					Model 3 (Empathy); f^2 = 0.01					Model 4 (Grandiose Narcissism); f^2=0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	90.18	2.88	31.25	<.001	84.49, 95.88	90.19	2.89	31.19	<.001	84.49, 95.90	90.13	2.91	30.95	<.001	84.38, 95.88	89.98	2.89	31.17	<.001	84.28, 95.68
	Mean RT	0.27	0.03	8.02	<.001	0.20,0.34	0.27	0.03	8.00	<.001	0.20, 0.34	0.27	0.03	7.93	<.001	0.20, 0.34	0.28	0.034	8.10	<.001	0.21, 0.34
	Sex	5.40	2.88	1.87	.06	-0.29, 11.10	5.39	2.89	1.86	.06	-0.31, 11.10	5.72	3.03	1.89	.06	-0.26, 11.70	6.08	2.93	2.07	.04	0.28, 11.87
	Mean RT * Sex	0.10	0.03	3.01	.003	0.03, 0.17	0.10	0.03	2.98	.003	0.03, 0.17	0.10	0.03	3.02	.003	0.03, 0.17	0.10	0.03	2.98	.003	0.034, 0.17
Personality variables	**Alexithymia**						-0.13	0.25	-0.51	.61	-0.63, 0.37										
	Empathic Concern											-0.05	0.7	-0.07	.95	-1.45, 1.36					
	Perspective Taking											-0.80	0.73	-1.09	.27	-2.24, 0.64					
	Grandiose Narcissism																17.72	14.75	1.20	.23	-11.40, 46.84
	Vulnerability Narcissism																				
	Autistic-like Traits																				
	Schizotypal Traits																				
	Agreeableness																				
	Extraversion																				

APPENDICES		Model 5 (Vulnerability Narcissism); f^2 = 0.001					Model 6 (Autism); f^2 = 0.01					Model 7 (Schizotypy); f^2 = 0.001					Model 8 (mini-IPIP); f^2=0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	90.19	2.89	31.17	<.001	84.48, 95.91	90.66	2.89	31.39	<.001	84.96, 96.36	90.22	2.89	31.17	<.001	84.5, 95.94	90.05	2.91	30.92	<.001	84.30, 96.80
	Mean RT	0.27	0.03	7.87	<.001	0.20, 0.34	0.28	0.03	8.17	<.001	0.21, 0.34	0.27	0.03	8.00	<.001	0.20, 0.34	0.28	0.03	8.06	<.001	0.21, 0.34
	Sex	5.42	2.89	1.87	.06	-0.29, 11.14	4.98	2.88	1.73	.09	-0.71, 10.68	5.37	2.89	1.85	.06	-0.35, 11.08	5.65	2.98	1.89	.06	-0.24, 11.54
	Mean RT * Sex	0.10	0.03	3.01	.003	0.03, 0.17	0.10	0.03	3.08	.002	0.04, 0.17	0.10	0.03	3.01	.003	0.03, 0.17	0.10	0.03	3.01	.003	0.03, 0.17
Personality variables	**Alexithymia**																				
	Empathic Concern																				
	Perspective Taking																				
	Grandiose Narcissism																				
	Vulnerability Narcissism	-0.14	0.47	-0.30	.76	-1.07, 0.79															
	Autistic-like Traits						-2.60	1.64	-1.58	.11	-5.83, 0.64										
	Schizotypal Traits											-0.22	0.59	-0.37	.71	-1.39, 0.95					
	Extraversion																2.88	3.01	0.96	.34	-3.07, 8.83
	Agreeableness																-3.49	4.27	-0.82	.41	-11.91, 4.93

Table S4.

Correlation matrix showing zero-order correlations (Pearson coefficients and p-values) of the predictor variables with the compatibility effect and mean reaction time (Experiment 1).

	Compatibility Effect	Mean RT	Perspective Taking	Empathic Concern	Alexithymia	Autistic-like traits	Schizotypal traits	Grandiose Narcissism	Vulnerability Narcissism	Extraversion	Agreeableness
Compatibility Effect	1	0.54	-0.07	0.07	-0.05	-0.05	-0.02	0.01	-0.08	0.07	0.05
Mean RT	0.54	1	0.02	0.16	-0.02	0.09	-0.01	-0.08	-0.14	-0.01	0.12
Compatibility Effect (p-value)	NA	<.001	0.38	0.32	0.52	0.50	0.75	0.84	0.28	0.37	0.48
Mean RT (p-value)	<.001	NA	0.79	0.03	0.83	0.23	0.89	0.31	0.07	0.91	0.10

Table S5. *Complete information of the multiple regression models with personality traits and their sex*trait interaction terms (Experiment 1).*

		Model 1 (Base Model); f^2 = 0.51					Model 2 (Alexithymia); f^2 = 0.02					Model 3 (Empathy); f^2 = 0.05					Model 4 (Grandiose Narcissism); f^2=0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	90.19	2.89	31.17	<.001	84.48, 95.91	90.26	2.87	31.39	<.001	84.58, 95.94	88.06	3.08	28.57	<.001	81.97, 94.14	89.76	2.96	30.31	<.001	83.92, 95.61
	Mean RT	0.27	0.03	7.87	<.001	0.20, 0.34	0.27	0.03	8.10	<.001	0.21, 0.34	0.27	0.03	8.09	<.001	0.21, 0.34	0.28	0.03	8.09	<.001	0.21, 0.34
	Sex	5.42	2.89	1.87	.06	-0.29, 11.14	5.33	2.87	1.85	.06	-0.35, 11.00	7.14	3.08	2.32	.02	1.06, 13.23	6.19	2.96	2.09	.04	0.34, 12.03
	Mean RT * Sex	0.10	0.03	3.01	.003	0.03, 0.17	0.10	0.03	2.99	.003	0.034, 0.17	0.09	0.03	2.66	.01	0.02, 0.16	0.10	0.03	2.93	.003	0.03, 0.17
Personality variables	**Alexithymia**						-0.33	0.28	-1.18	.24	-0.88, 0.22										
	Alexi * Sex						0.48	0.28	1.71	.09	-0.07, 1.03										
	Empathic Concern											-0.61	0.75	-0.82	.41	-2.09, 0.86					
	EC * Sex											1.63	0.75	2.17	.03	0.15, 3.11					
	Perspective Taking											0.014	0.81	0.02	.99	-1.60, 1.62					
	PT * Sex											-1.91	0.81	-2.34	.02	-3.52, -0.30					
	Grandiose Narcissism																19.09	15.31	1.25	.21	-11.14, 49.32
	GN * Sex																-5.29	15.31	-0.34	.73	-35.52, 24.94

APPENDICES		Model 5 (Vulnerability Narcissism); f^2 = 0.01					Model 6 (Autism); f^2 = 0.02					Model 7 (Schizotypy); f^2 = 0.01					Model 8 (mini-IPIP); f^2=0.02				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	90.12	2.89	31.23	<.001	84.43, 95.82	91.00	2.92	31.20	<.001	85.24, 96.76	90.35	2.90	31.14	<.001	84.63, 96.08	90.42	3.13	28.85386	<.001	84.23, 96.60
	Mean RT	0.27	0.034	7.75	<.001	0.20, 0.33	0.28	0.03	8.15	<.001	0.21, 0.34	0.27	0.03	7.93	<.001	0.20, 0.34	0.27	0.03	8.06264	<.001	0.20, 0.34
	Sex	5.38	2.89	1.86	.06	-0.31, 11.08	4.62	2.92	1.58	.11	-1.14, 10.37	5.23	2.90	1.80	.07	-0.50, 10.96	5.32	3.13	1.69889	.09	-0.86, 11.51
	Mean RT * Sex	0.11	0.03	3.20	.002	0.04, 0.18	0.10	0.03	3.00	.003	0.03, 0.17	0.10	0.03	3.04	.003	0.04, 0.17	0.10	0.03	2.92802	.004	0.03, 0.17
	Vulnerability Narcissism	-0.30	0.48	-0.61	.54	-1.25, 0.66															
	VN * Sex	0.67	0.48	1.39	.17	-0.28, 1.63															
	Autistic-like Traits						-3.10	1.73	-1.78	.07	-6.53, 0.33										
	AT * Sex						1.53	1.73	0.88	.38	-1.89, 4.96										
	Schizotypal Traits											-0.42	0.64	-0.66	.51	-1.68, 0.84					
	ST * Sex											0.53	0.64	0.84	.40	-0.72, 1.80					
	Agreeableness																-3.06	4.91	-0.62	.53	-12.75, 6.63
	Agree * Sex																0.83	4.91	0.17	.21	-8.86, 10.52
	Extraversion																4.14	3.32	1.25	.86	-2.41, 10.68
	Extra * Sex																-3.63	3.32	-1.09	.27	-10.18, 2.92

Table S6.

Descriptive Statistics for Experiment 2.

A. Imitation Task - Accuracy				
Sex	**Compatibility**	**N**	**Mean**	**SD**
female	compatible	116	95.7474337	3.39595366
female	incompatible	116	92.3880697	5.3978881
male	compatible	107	96.3342589	2.55139733
male	incompatible	107	93.2219602	6.10469398
B. Imitation Task – Reaction Time				
female	compatible	116	448.423802	46.1914623
female	incompatible	116	481.195323	57.5410987
male	compatible	107	438.679551	44.0087827
male	incompatible	107	464.465991	49.3650946
C. Flanker Task - Accuracy				
female	compatible	116	97.6054849	2.61173661
female	incompatible	116	88.721221	12.4278259
male	compatible	101	98.346959	1.79034068
male	incompatible	101	91.7085145	6.58600053
D. Flanker Task – Reaction Time				
female	compatible	116	404.986331	56.1742997
female	incompatible	116	499.853571	54.545204
male	compatible	101	404.193441	53.7338331
male	incompatible	101	498.077058	50.0164191

Table S7.

Complete ANOVA information for Experiment 2.

A. Imitation Task - Accuracy					
Effect	**DFn**	**DFd**	**F**	**p**	**pes**
Sex	1	221	1.8695	0.1729	0.0084
Compatibility	1	221	96.2253	<.001	0.3033
Sex*Compatibility	1	221	0.1402	0.7084	0.0006
B. Imitation Task – Reaction Time					
Sex	1	221	4.235	0.0408	0.0188
Compatibility	1	221	293.1808	<.001	0.5702
Sex*Compatibility	1	221	4.1716	0.0423	0.0185
C. Flanker Task - Accuracy					
Sex	1	215	5.7793	0.0171	0.0262
Compatibility	1	215	151.3335	<.001	0.4131
Sex*Compatibility	1	215	3.1677	0.0765	0.0145
D. Flanker Task – Reaction Time					
Sex	1	215	0.0336	0.8547	0.0002
Compatibility	1	215	1986.8939	<.001	0.9024
Sex*Compatibility	1	215	0.054	0.8165	0.0003

Table S8. *Complete information of the multiple regression models with personality traits (Experiment 2).*

		Model 1 (Base Model); f^2 = 0.19					Model 2 (Alexithymia); f^2 = <0.001					Model 3 (Empathy); f^2 = 0.005				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	29.24	1.69	17.32	<.001	25.91, 32.56	29.24	1.69270	17.28	<.001	25.91, 32.58	29.11	1.70	17.16	<.001	25.77, 32.46
	Mean RT	0.19	0.03	5.37	<.001	-1.19, 5.45	0.19	0.03509	5.36	<.001	-1.19, 5.48	0.19	0.03	5.42	<.001	0.12, 0.26
	Sex	2.14	1.69	1.27	.21	0.12, 0.26	2.15	1.69243	1.27	.21	0.12, 0.26	2.70	1.86	1.45	.15	-0.96, 6.37
	Mean RT * Sex	0.07	0.03	1.90	.06	-0.002, 0.13	0.07	0.03509	1.89	.06	-0.32, 0.27	0.07	0.03	1.91	.06	-0.002, 0.14
Personality variables	**Alexithymia**						-0.02	0.14997	-0.16	.87	-0.002, 0.13					
	Empathic Concern											-0.31	0.42	-0.74	.46	-1.13, 0.51
	Perspective Taking											0.43	0.43	0.99	.32	-0.42, 1.28

Table S9.

Correlation matrix showing zero-order correlations (Pearson coefficients and p-values) of the predictor variables with the compatibility effect and mean reaction time (Experiment 2).

	Compatibility Effect	Mean RT	Alexithymia	Perspective Taking	Empathic Concern
Compatibility Effect	1	0.37	0.004	0.03	0.05
Mean RT	0.37	1	0.05	-0.06	0.06
Compatibility Effect (p-value)	1	<.001	0.95	0.68	0.48
Mean RT (p-value)	<.001	1	0.46	0.38	0.40

Table S10.

*Complete information of the multiple regression models with personality traits and their sex*trait interaction terms (Experiment 2).*

		Model 1 (Base Model); f^2 = 0.19					Model 2 (Alexithymia); f^2 = 0.01					Model 3 (Empathy); f^2 = 0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	29.24	1.69	17.32	<.001	25.91, 32.56	29.19	1.69	17.26	<.001	25.85, 32.52	29.12	1.87	15.56	<.001	25.43, 32.81
	Mean RT	0.19	0.03	5.37	<.001	-1.19, 5.45	0.19	0.03	5.42	<.001	0.12, 0.26	0.19	0.039	5.33	<.001	0.12, 0.26
	Sex	2.14	1.69	1.27	.21	0.12, 0.26	2.08	1.69	1.23	.22	-1.25, 5.41	2.67	1.87	1.43	.15	-1.01, 6.36
	Mean RT * Sex	0.07	0.03	1.90	.06	-0.002, 0.13	0.06	0.03	1.82	.07	-0.005, 0.13	0.07	0.03	1.92	.06	-0.002, 0.14
Personality variables	**Alexithymia**						-0.03	0.15	-0.20	.84	-0.32, 0.27					
	Alexi * Sex						0.18	0.15	1.21	.23	-0.11, 0.48					
	Empathic Concern											-0.31	0.43	-0.73	.46	-1.16, 0.53
	EC * Sex											-0.004	0.43	-0.01	.99	-0.85, 0.84
	Perspective Taking											0.41	0.44	0.94	.35	-0.45, 1.28
	PT * Sex											0.15	0.44	0.33	.74	-0.72, 1.01

Table S11.

Descriptive Statistics for Experiment 3.

A. Spatial Compatibility - Accuracy					
Sex	**Spatial Compatibility**	**Stimulus Sex**	**N**	**Mean**	**SD**
female	Spatial Compatible	female	92	96.35413	3.90492256
female	Spatial Compatible	male	92	96.1696395	3.95753095
female	Spatial Incompatible	female	92	86.1331652	6.65471092
female	Spatial Incompatible	male	92	92.122774	7.09003734
male	Spatial Compatible	female	97	96.0115023	5.22027902
male	Spatial Compatible	male	97	96.1057142	4.97390812
male	Spatial Incompatible	female	97	85.7965344	7.58244008
male	Spatial Incompatible	male	97	92.0855767	7.70425654
B. Spatial Compatibility – Reaction Time					
Sex	**Spatial Compatibility**	**Stimulus Sex**	**N**	**Mean**	**SD**
female	Spatial Compatible	female	92	431.198365	52.678426
female	Spatial Compatible	male	92	430.740814	50.2367702
female	Spatial Incompatible	female	92	469.846832	54.5336491
female	Spatial Incompatible	male	92	466.64636	58.2847811
male	Spatial Compatible	female	97	424.35871	63.624757
male	Spatial Compatible	male	97	419.902758	57.3842241
male	Spatial Incompatible	female	97	453.505451	64.8873388
male	Spatial Incompatible	male	97	452.242503	65.736253
C. Imitative Compatibility - Accuracy					
Sex	**Spatial Compatibility**	**Stimulus Sex**	**N**	**Mean**	**SD**
female	Spatial Compatible	female	92	96.35413	3.90492256
female	Spatial Compatible	male	92	96.1696395	3.95753095
female	Spatial Incompatible	female	92	86.1331652	6.65471092
female	Spatial Incompatible	male	92	92.122774	7.09003734
male	Spatial Compatible	female	97	96.0115023	5.22027902
male	Spatial Compatible	male	97	96.1057142	4.97390812
male	Spatial Incompatible	female	97	85.7965344	7.58244008
male	Spatial Incompatible	male	97	92.0855767	7.70425654
D. Imitative Compatibility – Reaction Time					
Sex	**Spatial Compatibility**	**Stimulus Sex**	**N**	**Mean**	**SD**
female	Spatial Compatible	female	92	447.044795	51.1427468
female	Spatial Compatible	male	92	443.816099	52.9400457
female	Spatial Incompatible	female	92	454.000402	54.208876
female	Spatial Incompatible	male	92	453.571075	54.4368132
male	Spatial Compatible	female	97	434.023539	60.997859
male	Spatial Compatible	male	97	433.959945	60.0391406
male	Spatial Incompatible	female	97	443.840622	66.7307088
male	Spatial Incompatible	male	97	438.185316	62.0975039

Table S12. *Complete ANOVA information for Experiment 3.*

A. Spatial Compatibility - Accuracy					
Effect	**DFn**	**DFd**	**F**	**p**	**pes**
Sex	1	187	0.0617	0.8041	0.0003
Spatial Compatibility	1	187	563.3472	<.001	0.7508
Stimulus Sex	1	187	335.474	<.001	0.6421
Sex * Spatial Compatibility	1	187	0.0007	0.9783	0
Sex * Stimulus Sex	1	187	0.7548	0.3861	0.004
Spatial Compatibility * Stimulus Sex	1	187	202.3158	<.001	0.5197
Sex * Spatial Compatibility * Stimulus Sex	1	187	0.0006	0.981	0
B. Spatial Compatibility – Reaction Time					
Sex	1	187	2.1401	0.1452	0.0113
Spatial Compatibility	1	187	459.7064	<.001	0.7108
Stimulus Sex	1	187	5.6324	0.0186	0.0292
Sex * Spatial Compatibility	1	187	4.2416	0.0408	0.0222
Sex * Stimulus Sex	1	187	0.2721	0.6026	0.0015
Spatial Compatibility * Stimulus Sex	1	187	0.0102	0.9197	0.0001
Sex * Spatial Compatibility * Stimulus Sex	1	187	1.7706	0.1849	0.0094
C. Imitative Compatibility – Accuracy					
Effect	**DFn**	**DFd**	**F**	**P**	**pes**
Sex	1	187	0.0617	0.8041	0.0003
Imitative Compatibility	1	187	205.6503	<.001	0.5237
Stimulus Sex	1	187	335.474	<.001	0.6421
Sex * Imitative Compatibility	1	187	1.8084	0.1803	0.0096
Sex * Stimulus Sex	1	187	0.7548	0.3861	0.004
Imitative Compatibility * Stimulus Sex	1	187	162.977	<.001	0.4657
Sex * Imitative Compatibility * Stimulus Sex	1	187	1.3429	0.248	0.0071
D. Imitative Compatibility – Reaction Time					
Sex	1	187	2.1401	0.1452	0.0113
Imitative Compatibility	1	187	54.9629	<.001	0.2272
Stimulus Sex	1	187	5.7021	0.0179	0.0296
Sex * Imitative Compatibility	1	187	0.4153	0.5201	0.0022
Sex * Stimulus Sex	1	187	0.2721	0.6026	0.0015
Imitative Compatibility * Stimulus Sex	1	187	0.4982	0.4812	0.0027
Sex * Imitative Compatibility * Stimulus Sex	1	187	3.8582	0.051	0.0202

Table S13. *Complete information of the multiple regression models with personality traits. (Experiment 3 – Imitative Compatibility)*

		Model 1 (Base Model); f^2 = 0.05					Model 2 (Alexithymia); f^2 = 0.001					Model 3 (Empathy); f^2 = 0.03					Model 4 (Grandiose Narcissism); f^2=0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	7.59	1.04	7.27	<.001	5.53, 9.65	7.58	1.05	7.24	<.001	5.51, 9.64	7.58612	1.03682	7.31672	<.001	5.54, 9.63	7.63	1.0	7.31	<.001	5.57, 9.69
	Mean RT	0.05	0.02	2.79	.006	-1.83, 2.29	0.05	0.02	2.74	.007	0.01, 0.09	0.05292	0.01823	2.90220	.004	0.02, 0.09	0.05	0.02	2.79	.006	0.01, 0.09
	Sex	0.23	1.04	0.22	.82	0.01, 0.089	0.26	1.05	0.25	.80	-1.80, 2.33	0.73463	1.07862	0.68108	.50	-1.39, 2.86	0.64	1.11	0.58	.56	-1.55, 2.84
	Mean RT * Sex	-0.00	0.02	-0.21	.83	-0.04, 0.03	-0.003	0.02	-0.19	.84	-0.04, 0.03	-0.00381	0.01824	-0.20880	.83	-0.04, 0.03	-0.004	0.02	-0.24	.81	-0.04, 0.03
Personality variables	**Alexithymia**						0.05	0.09	0.51	.61	-0.14, 0.23										
	Empathic Concern											-0.50825	0.24371	-2.08551	.04	-0.99, -0.03					
	Perspective Taking											0.31568	0.23599	1.33765	.18	-0.15, 0.78					
	Grandiose Narcissism																6.04	5.70	1.06	.29	-5.21, 17.29
	Vulnerability Narcissism																				
	Autistic-like Traits																				
	Schizotypal Traits																				
	Agreeableness																				
	Extraversion																				

Predictors		Model 5 (Vulnerability Narcissism); f^2 = <0.001					Model 6 (Autism); f^2 = <0.001					Model 7 (Schizotypy); f^2 = 0.002					Model 8 (mini-IPIP); f^2=0.03				
		B	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	7.58	1.05	7.23	<.001	5.51, 9.65	7.59	1.05	7.25	<.001	5.52, 9.65	7.62	1.04	7.28	<.001	5.56, 9.69	7.61	1.04	7.33	<.001	5.56, 9.66
	Mean RT	0.05	0.02	2.79	.006	0.01, 0.09	0.05	0.02	2.78	.006	0.01, 0.09	0.05	0.02	2.69	.008	0.01, 0.09	0.05	0.02	2.58	.01	0.01, 0.08
	Sex	0.24	1.05	0.23	.82	-1.83, 2.30	0.23	1.05	0.23	.83	-1.84, 2.29	0.11	1.06	0.10	.92	-1.99, 2.20	0.51	1.04	0.49	.62	-1.55, 2.57
	Mean RT * Sex	-0.004	0.02	-0.21	.83	-0.04, 0.03	-0.004	0.02	-0.22	.83	-0.04, 0.03	-0.004	0.02	-0.25	.80	-0.04, 0.03	-0.003	0.02	-0.17	.86	-0.04, 0.03
Personality variables	**Alexithymia**																				
	Empathic Concern																				
	Perspective Taking																				
	Grandiose Narcissism																				
	Vulnerability Narcissism	0.05	0.18	0.28	.78	-0.31, 0.41															
	Autistic-like Traits						0.09	0.56	0.16	.87	-1.01, 1.19										
	Schizotypal Traits											-0.14	0.21	-0.68	.50	-0.56, 0.27					
	Extraversion																1.82	1.09	1.67	.10	-0.33, 3.96
	Agreeableness																-2.68	1.53	-1.75	.08	-5.70, 0.34

Table S14.

Correlation matrix showing zero-order correlations (Pearson coefficients and p-values) of the predictor variables with the imitative compatibility effect and mean reaction time (Experiment 3).

	Imitative Compatibility Effect	Mean RT	Perspective Taking	Empathic Concern	Alexithymia	Autistic-like traits	Schizotypal traits	Grandiose Narcissism	Vulnerability Narcissism	Extraversion	Agreeableness
Imitative Compatibility Effect	1	0.21	0.04	-0.10	0.05	0.01	-0.08	0.05	0.01	0.11	-0.10
Mean RT	0.21	1	0.005	0.07	0.07	0.02	-0.12	-0.05	-0.02	0.09	-0.02
Compatibility Effect (p-value)	1	0.003	0.54	0.18	0.49	0.83	0.30	0.46	0.83	0.13	0.15
Mean RT (p-value)	0.003	1	0.95	0.35	0.37	0.80	0.09	0.49	0.79	0.23	0.77

Table S15. *Complete information of the multiple regression models with personality traits and their sex*trait interaction terms (Experiment 3 – Imitative Compatibility).*

		Model 5 (Vulnerability Narcissism); f^2 = 0.01					Model 6 (Autism); f^2 = 0.005					Model 7 (Schizotypy); f^2 = 0.003					Model 8 (mini-IPIP); f^2=0.04				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	7.62	1.04	7.29	<.001	5.56, 9.68	7.60	1.05	7.26	<.001		7.57	1.06	7.11	<.001	5.46, 9.67	7.73	1.04	7.39	<.001	5.66, 9.79
	Mean RT	0.05	0.02	2.79	.006	0.01, 0.09	0.05	0.02	2.82	.005		0.05	0.02	2.66	.008	0.01, 0.09	0.04	0.02	2.42	.02	0.01, 0.08
	Sex	0.17	1.04	0.16	.87	-1.89, 2.23	0.25	1.05	0.24	.81		0.12	1.06	0.11	.91	-1.98, 2.22	0.64	1.04	0.62	.54	-1.42, 2.71
	Mean RT * Sex	-0.003	0.02	-0.19	.85	-0.04, 0.03	-0.004	0.02	-0.20	.84		-0.005	0.02	-0.28	.78	-0.04, 0.03	-0.01	0.02	-0.35	.73	-0.04, 0.03
Personality variables	**Alexithymia**																				
	Alexi * Sex																				
	Empathic Concern																				
	EC * Sex																				
	Perspective Taking																				
	PT * Sex																				
	Grandiose Narcissism																				
	GN * Sex																				
	Vulnerability Narcissism	0.04	0.18	0.23	.82	-0.32, 0.40															
	VN * Sex	0.26	0.18	1.43	.15	-0.10, 0.62															
	Autistic-like Traits						0.15	0.56	0.26	.79											
	AT * Sex						-0.54	0.56	-0.97	.33											
	Schizotypal Traits											-0.14	0.21	-0.68	.49	-0.56, 0.27					
	ST * Sex											-0.06	0.21	-0.31	.76	-0.48, 0.35					
	Extraversion															2.06	1.10	1.87273		.06	-0.11, 4.24
	Extra * Sex															1.71	1.10	1.55202		.12	-0.46, 3.88
	Agreeableness															-2.91	1.53	-1.89570		.06	-5.93, 0.12
	Agree * Sex															0.19	1.53	0.12756		.89	-2.83, 3.22

		Model 1 (Base Model); f^2 = 0.05					Model 2 (Alexithymia); f^2 = 0.02					Model 3 (Empathy); f^2 = 0.03					Model 4 (Grandiose Narcissism); f^2=0.01				
Predictors		**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**	**B**	**SEB**	**t**	**p**	**CI**
Base model predictors	**Constant**	7.59	1.04	7.27	<.001	5.53, 9.65	7.60	1.05	7.24	<.001	5.53, 9.68	7.78	1.09	7.14	<.001	5.63, 9.93	7.93	1.14	6.97	<.001	5.68, 10.17
	Mean RT	0.05	0.02	2.79	.006	-1.83, 2.29	0.05	0.02	2.75	.007	-1.82, 2.33	0.05	0.02	2.93	.004	0.02, 0.09	0.05	0.02	2.76	.006	-1.45, 3.04
	Sex	0.23	1.04	0.22	.82	0.01, 0.089	0.25	1.05	0.24	.81	0.01, 0.09	0.71	1.09	0.65	.51	-1.44, 2.86	0.79	1.14	0.69	.49	0.01, 0.09
	Mean RT * Sex	-0.00	0.02	-0.21	.83	-0.04, 0.03	-0.004	0.02	-0.22	.82	-0.04, 0.03	-0.003	0.02	-0.19	.85	-0.04, 0.03	-0.004	0.02	-0.25	.80	-0.04, 0.03
Personality variables	**Alexithymia**						0.05	0.09	0.54	.59	-0.13, 0.24										
	Alexi * Sex						0.04	0.09	0.40	.69	-0.15, 0.22										
	Empathic Concern											-0.51	0.28	-1.84	.07	-1.06, 0.04					
	EC * Sex											-0.09	0.28	-0.32	.74	-0.64, 0.46					
	Perspective Taking											0.31	0.26	1.20	.23	-0.20, 0.82					
	PT * Sex											-0.14	0.26	-0.53	.60	-0.65, 0.37					
	Grandiose Narcissism																7.75	6.27	1.24	.21	-4.62, 20.13
	GN * Sex																4.14	6.27	0.66	.51	-8.23, 16.52
	Vulnerability Narcissism																				
	VN * Sex																				
	Autistic-like Traits																				
	AT * Sex																				
	Schizotypal Traits																				
	ST * Sex																				
	Agreeableness																				
	Agree * Sex																				
	Extraversion																				
	Extra * Sex																				

Table S16. *Complete information of the multiple regression models with personality traits. (Experiment 3 – Spatial Compatibility)*

	Base Model			**Alexithymia**			**Empathy**			**Autistic-like traits**			**Schizotypal traits**			**Grandiose Narcissism**			**Vulnerability Narcissism**			**Mini IPIP (Agreeableness & Extraversion)**		
	B	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***
Intercept	33.76	30.66 – 36.87	**<0.001**	33.79	30.68 – 36.90	**<0.001**	33.78	30.72 – 36.85	**<0.001**	33.81	30.70 – 36.91	**<0.001**	33.82	30.71 – 36.93	**<0.001**	33.69	30.59 – 36.79	**<0.001**	33.82	30.71 – 36.93	**<0.001**	33.76	30.64 – 36.89	**<0.001**
Sex	2.74	-0.37 – 5.84	0.086	2.66	-0.45 – 5.78	0.096	3.48	0.29 – 6.67	**0.034**	2.77	-0.33 – 5.87	0.082	2.51	-0.64 – 5.67	0.120	1.99	-1.32 – 5.30	0.241	2.70	-0.41 – 5.81	0.090	2.80	-0.34 – 5.95	0.082
Mean RT	0.08	0.02 – 0.13	**0.008**	0.08	0.02 – 0.13	**0.007**	0.08	0.02 – 0.13	**0.005**	0.08	0.02 – 0.13	**0.007**	0.07	0.02 – 0.13	**0.011**	0.08	0.02 – 0.13	**0.008**	0.07	0.02 – 0.13	**0.008**	0.07	0.02 – 0.13	**0.009**
Sex*Mean RT	-0.01	-0.06 – 0.05	0.813	-0.01	-0.06 – 0.05	0.798	-0.01	-0.06 – 0.05	0.792	-0.01	-0.06 – 0.05	0.855	-0.01	-0.06 – 0.05	0.775	-0.01	-0.06 – 0.05	0.844	-0.01	-0.06 – 0.05	0.815	-0.01	-0.06 – 0.05	0.823
Alexithymia				-0.11	-0.38 – 0.17	0.458																		
Perspective Taking							0.73	0.03 – 1.43	**0.041**															
Empathic Concern							-0.86	-1.58 – -0.14	**0.021**															
Autistic-like traits										-0.98	-2.64 – 0.67	0.244												
Schizotypal traits													-0.26	-0.89 – 0.37	0.424									
Grandiose																-10	-27.84 – 6.04	0.209						

Narcissism															.90								
Vulnerability Narcissism																		-0.22	-0.77 – 0.32	0.417			
Extraversion																					0.37	-2.91 – 3.65	0.825
Agreeableness																					-0.69	-5.30 – 3.91	0.768
Observations	185			185			185			185			185			185			185				
R^2 / R^2adjusted	0.063 / 0.048			0.066 / 0.045			0.099 / 0.074			0.070 / 0.050			0.066 / 0.046			0.071 / 0.051			0.067 / 0.046			0.064 / 0.038	

Supplementary Table 17. *Multiple regression models with personality traits and their sex*trait interaction terms (Experiment 3 – Spatial Compatibility).*

	Base Model			**Alexithymia**			**Empathy**			**Autistic-like traits**			**Schizotypal traits**			**Grandiose Narcissism**			**Vulnerability Narcissism**			**Mini IPIP (Agreeableness & Extraversion)**		
	B	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***	***B***	***CI***	***p***
Intercept	33.76	30.66 – 36.87	**<0.001**	33.75	30.63 – 36.87	**<0.001**	33.93	30.70 – 37.15	**<0.001**	33.84	30.74 – 36.93	**<0.001**	33.57	30.41 – 36.72	**<0.001**	32.88	29.51 – 36.25	**<0.001**	33.82	30.70 – 36.94	**<0.001**	33.91	30.74 – 37.08	**<0.001**
Sex	2.74	-0.37 – 5.84	0.086	2.68	-0.45 – 5.80	0.095	3.52	0.30 – 6.75	**0.034**	2.81	-0.28 – 5.91	0.076	2.57	-0.59 – 5.73	0.112	1.59	-1.78 – 4.96	0.356	2.71	-0.41 – 5.83	0.090	2.88	-0.28 – 6.05	0.076
Mean RT	0.08	0.02 – 0.13	**0.008**	0.08	0.02 – 0.13	**0.007**	0.08	0.02 – 0.13	**0.005**	0.08	0.02 – 0.13	**0.006**	0.07	0.02 – 0.13	**0.013**	0.08	0.02 – 0.13	**0.007**	0.07	0.02 – 0.13	**0.008**	0.07	0.02 – 0.13	**0.011**
Sex*Mean RT	-0.01	-0.06 – 0.05	0.813	-0.01	-0.06 – 0.05	0.820	-0.01	-0.06 – 0.05	0.804	-0.00	-0.06 – 0.05	0.875	-0.01	-0.07 – 0.04	0.696	-0.01	-0.06 – 0.05	0.855	-0.01	-0.06 – 0.05	0.815	-0.01	-0.06 – 0.05	0.751
Alexithymia				-0.11	-0.39 – 0.17	0.442																		
Alexithymua*Sex				-0.06	-0.34 – 0.22	0.673																		
Perspective Taking							0.77	0.00 – 1.54	0.050															
Empathic Concern							-0.91	-1.74 – -0.09	**0.031**															
Perspective Taking*Sex							0.05	-0.72 – 0.81	0.907															
Empathic							-0.12	-0.95 – 0.70	0.767															

Concern *Sex																							
Autistic-like traits										-0.87	-2.53 – 0.79	0.306											
Autistic-like traits*Sex										-1.14	-2.80 – 0.52	0.179											
Schizotypal traits													-0.26	-0.89 – 0.36	0.410								
Schizotypal traits*Sex													-0.30	-0.92 – 0.33	0.357								
Grandiose Narcissism																-15.58	-34.15 – 2.99	0.102					
Grandiose Narcissism*Sex																-11.34	-29.91 – 7.23	0.233					
Vulnerability Narcissism																			-0.22	-0.77 – 0.32	0.420		
Vulnerability Narcissism*Sex																			-0.03	-0.57 – 0.51	0.915		
Extraversion																						0.60	-2.73 – 3.94	0.723

Agreeabl ness																						- 0.8 7	- 5.52 – 3.77	0.7 12
Extraver sion*Sex																						1.3 6	- 1.98 – 4.70	0.4 26
Agreeabl eness*Se x																						- 0.7 6	- 5.41 – 3.89	0.7 49
Observat ions	185			185			185			185			185			185			185			185		
R^2 / R^2 adjust ed	0.063 / 0.048			0.067 / 0.041			0.099 / 0.064			0.080 / 0.054			0.071 / 0.045			0.079 / 0.053			0.067 / 0.041			0.067 / 0.030		

Figure S1. Compatibility Effect by Emotion Type. Illustrating general invariance of the general compatibility effect (measured in milliseconds on the Y axis) to the emotional expression signalled by the interacting partner.

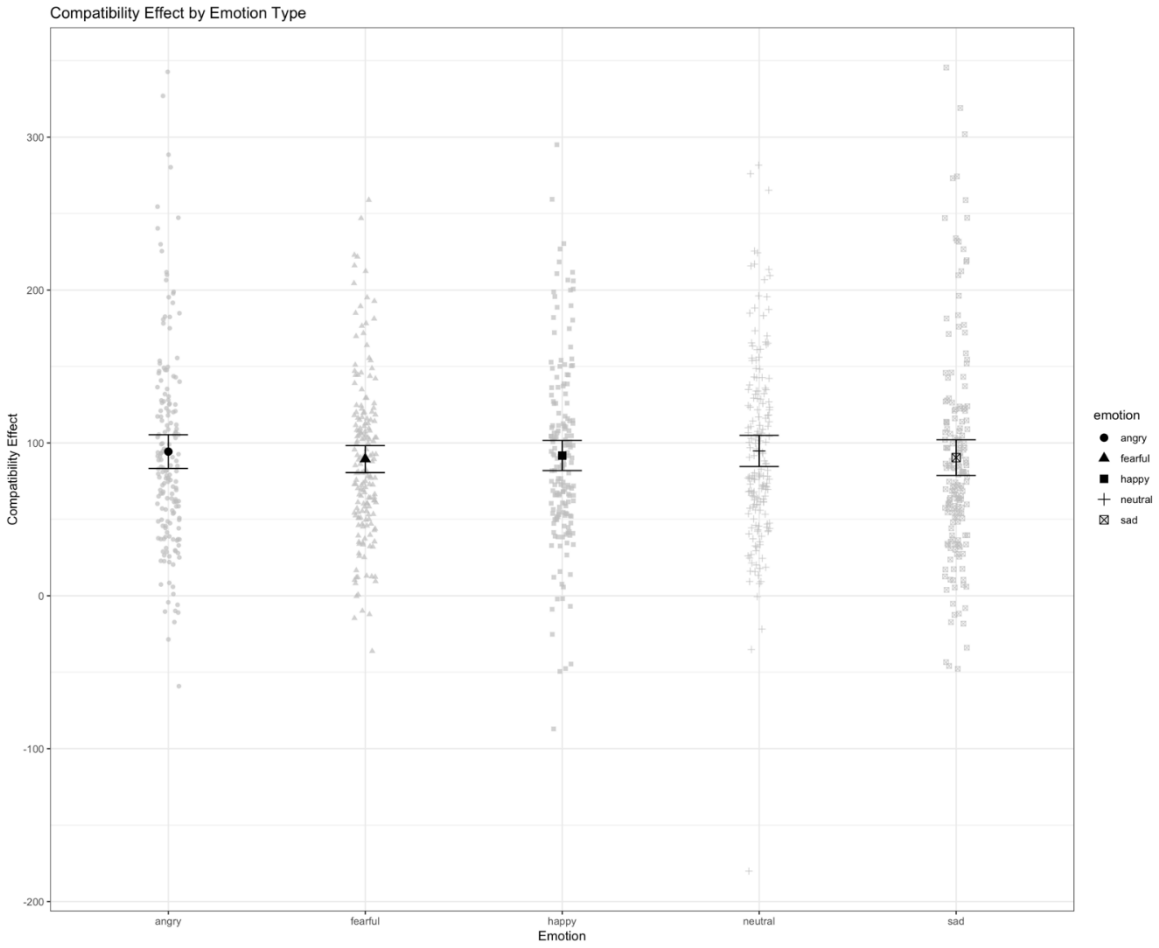


Figure S2. Zero-order correlations illustrating relationship between predictor variables (personality traits) and the general compatibility effect. The X axis denotes the general compatibility effect in milliseconds. The Y axis denotes mean-centred scores on personality traits. Abbreviations: RTms = reaction time in milliseconds.

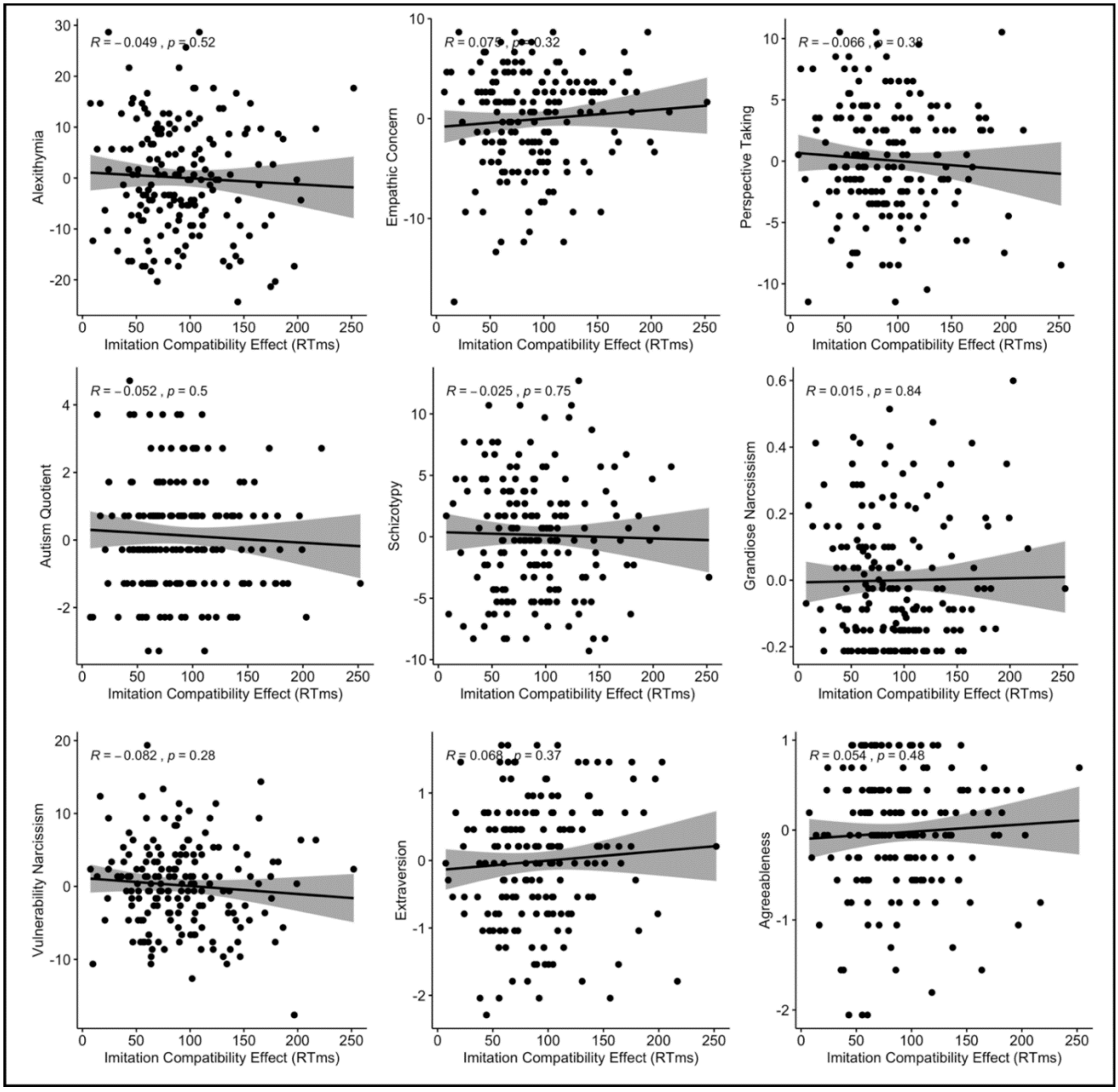


*Figure S3. Experiment 1 – Multiple Regression Analyses, Sex*Trait Interactions.* Values of standardised coefficients are plotted for each predictor variable (personality trait), and the sex*trait interaction, along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

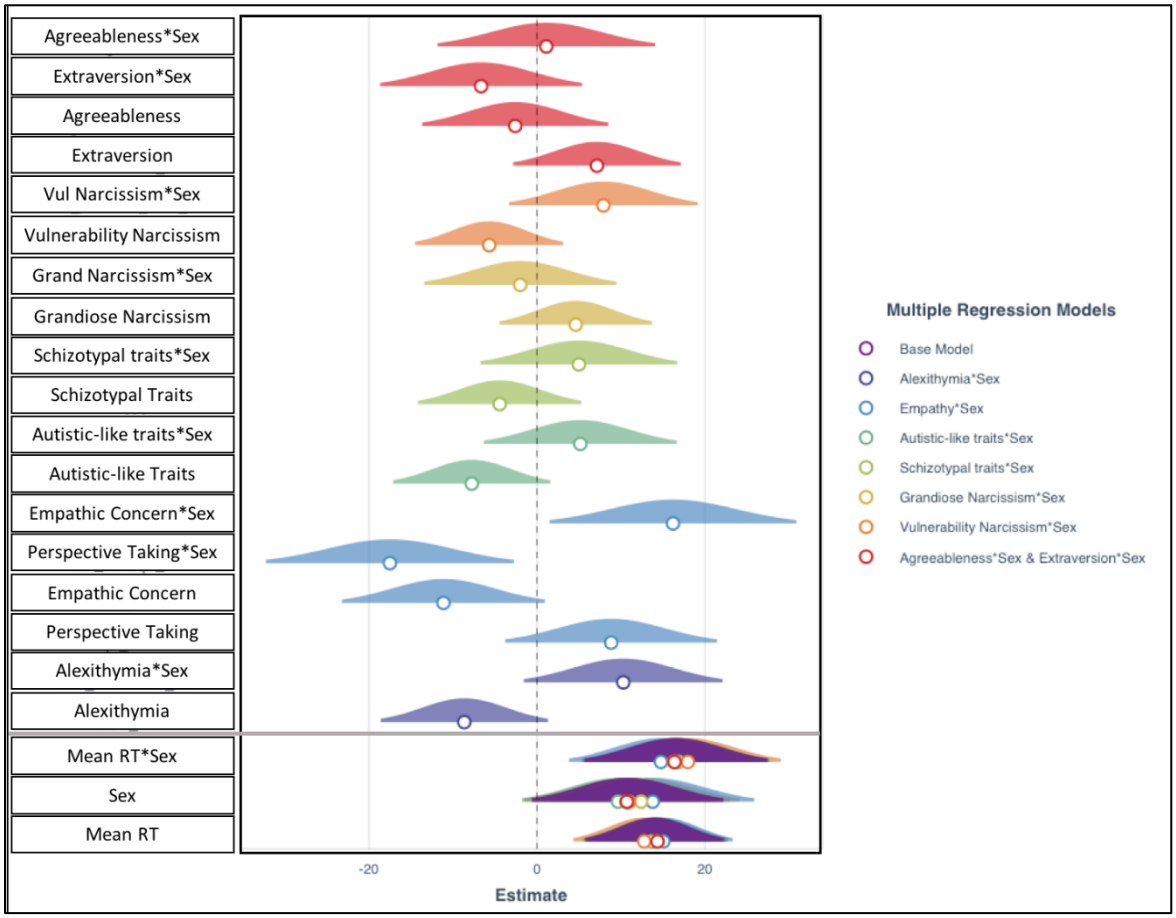


Figure S4. Zero order correlations illustrating the relationships between each predictor variable (personality trait) and the orthogonal compatibility effect. The X axis denotes the orthogonal compatibility effect in milliseconds. The Y axis denotes mean-centred scores on personality traits. Abbreviations: RTms = reaction time in milliseconds.

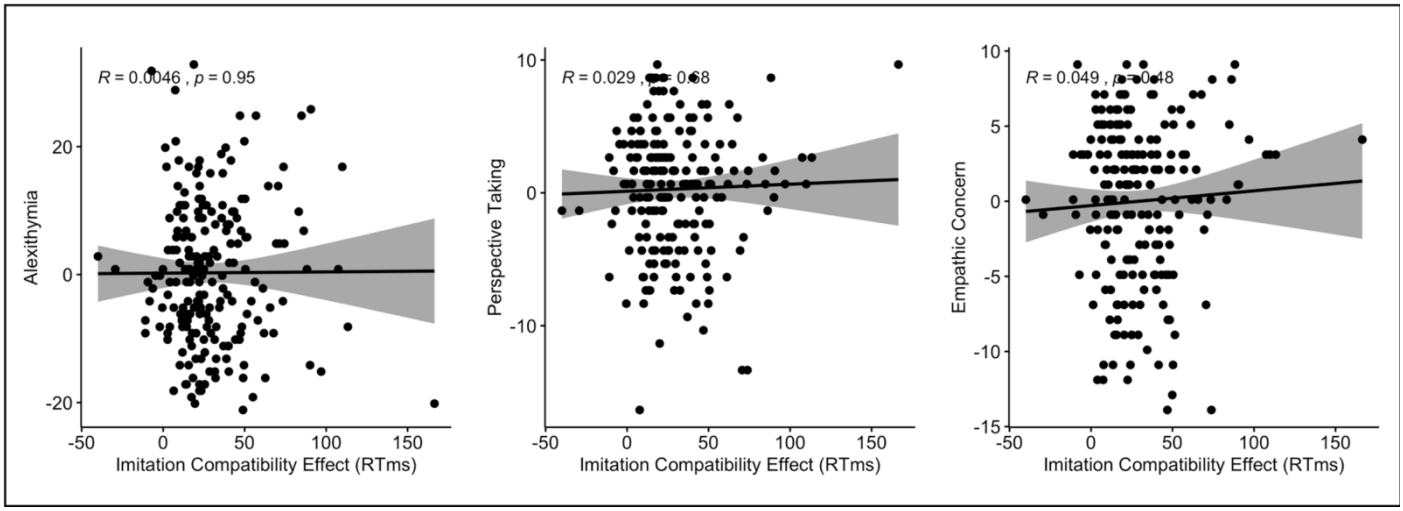


Figure S5. Experiment 2 – Multiple Regression Analyses, Sex*Trait Interactions. Values of standardised coefficients are plotted for each predictor variable (personality trait), and the sex*trait interaction, along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

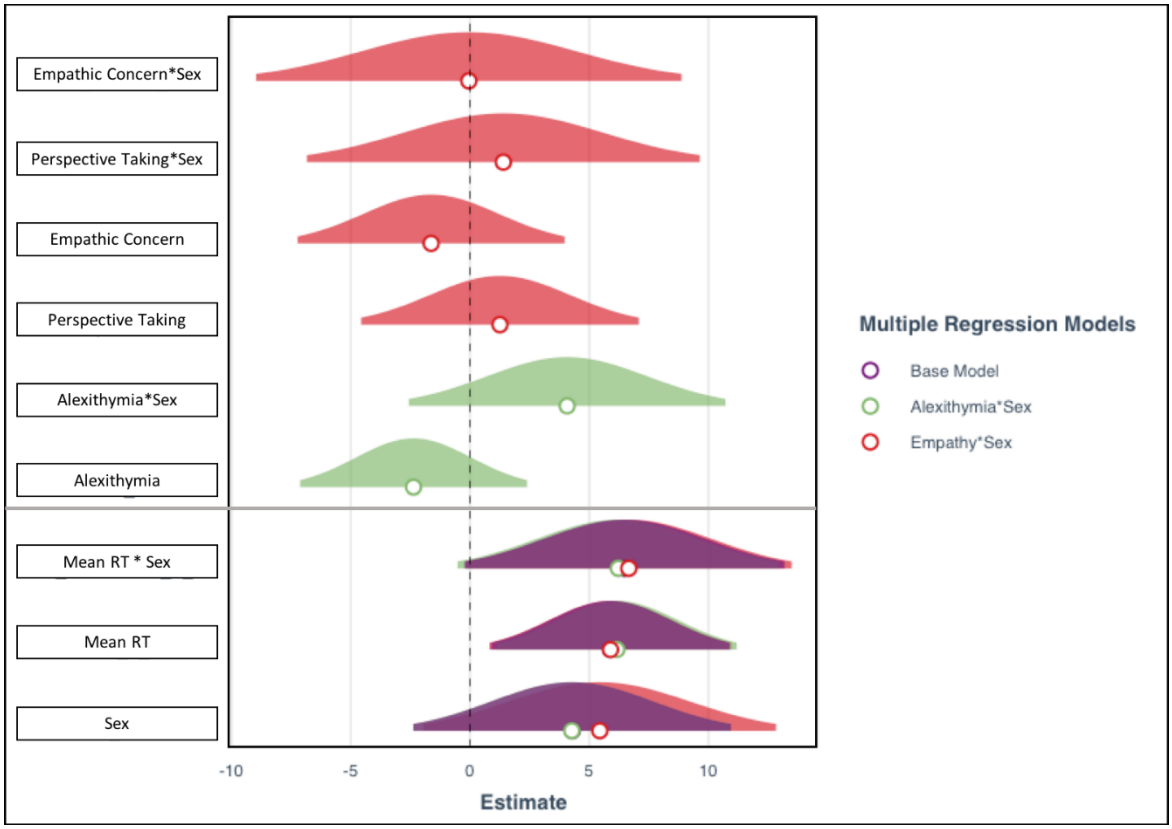


Figure S6. Zero order correlations illustrating the relationships between each predictor variable (personality trait) and the imitative compatibility effect. The X axis denotes the imitative compatibility effect in milliseconds. The Y axis denotes mean-centred scores on personality traits. Abbreviations: RTms = reaction time in milliseconds.

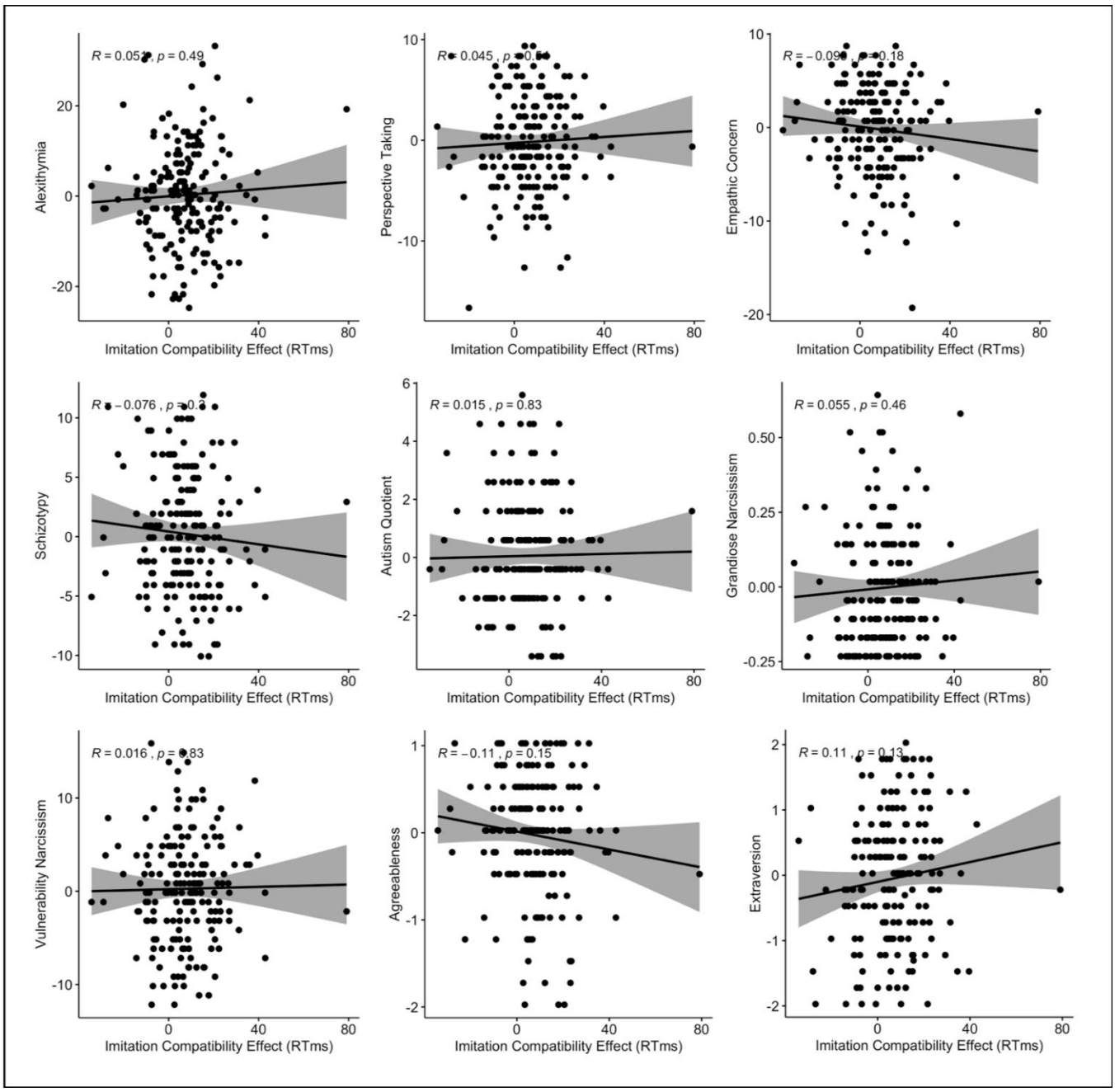


Figure S7. Experiment 3 (Imitative Compatibility) – Multiple Regression Analyses, Sex*Trait Interactions. Values of standardised coefficients are plotted for each predictor variable (personality trait), and the sex*trait interaction, along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

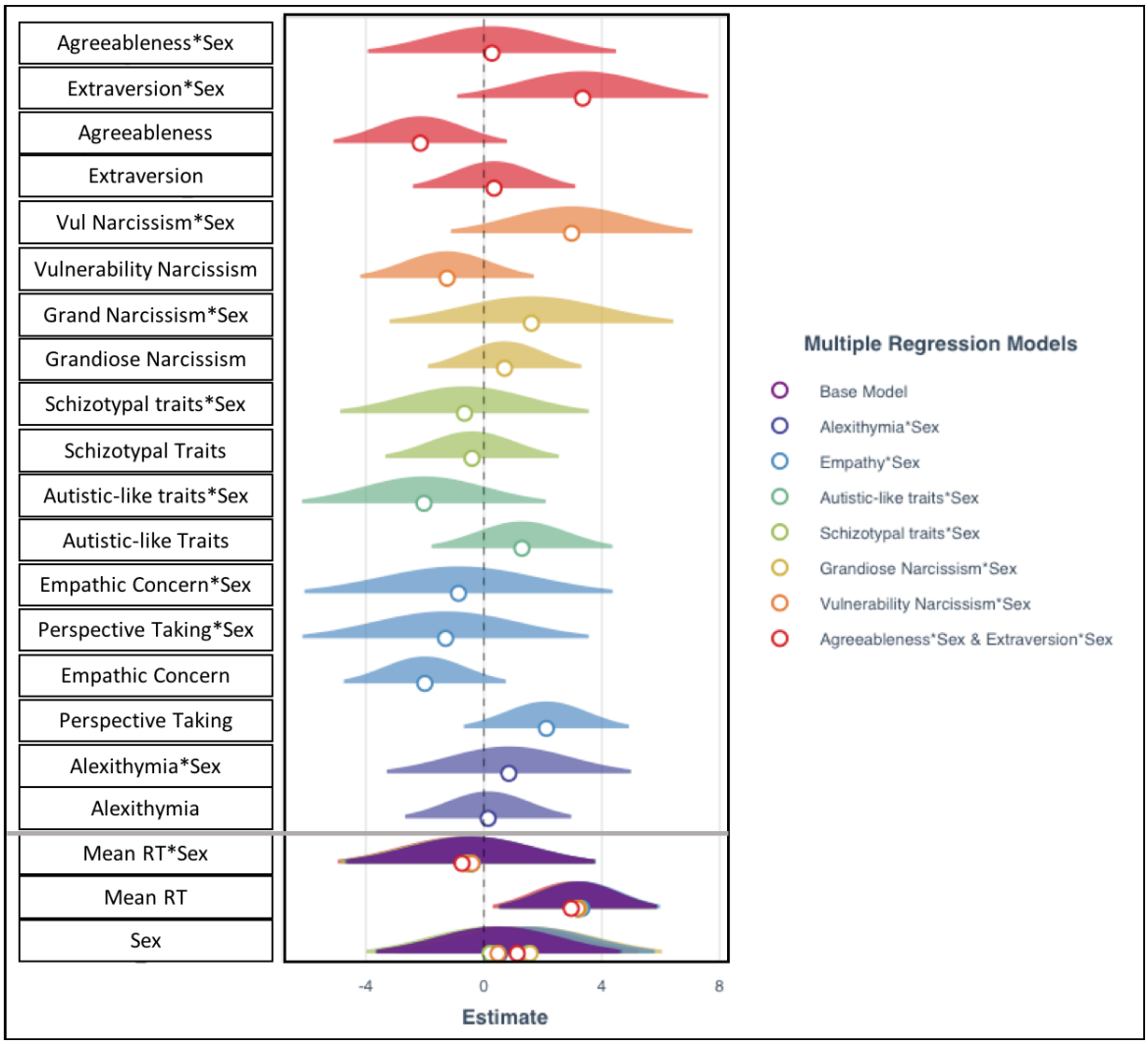


Figure S8. Experiment 3 (Spatial Compatibility) – Multiple Regression Analyses. Values of standardised coefficients are plotted for each predictor variable (personality trait), along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.

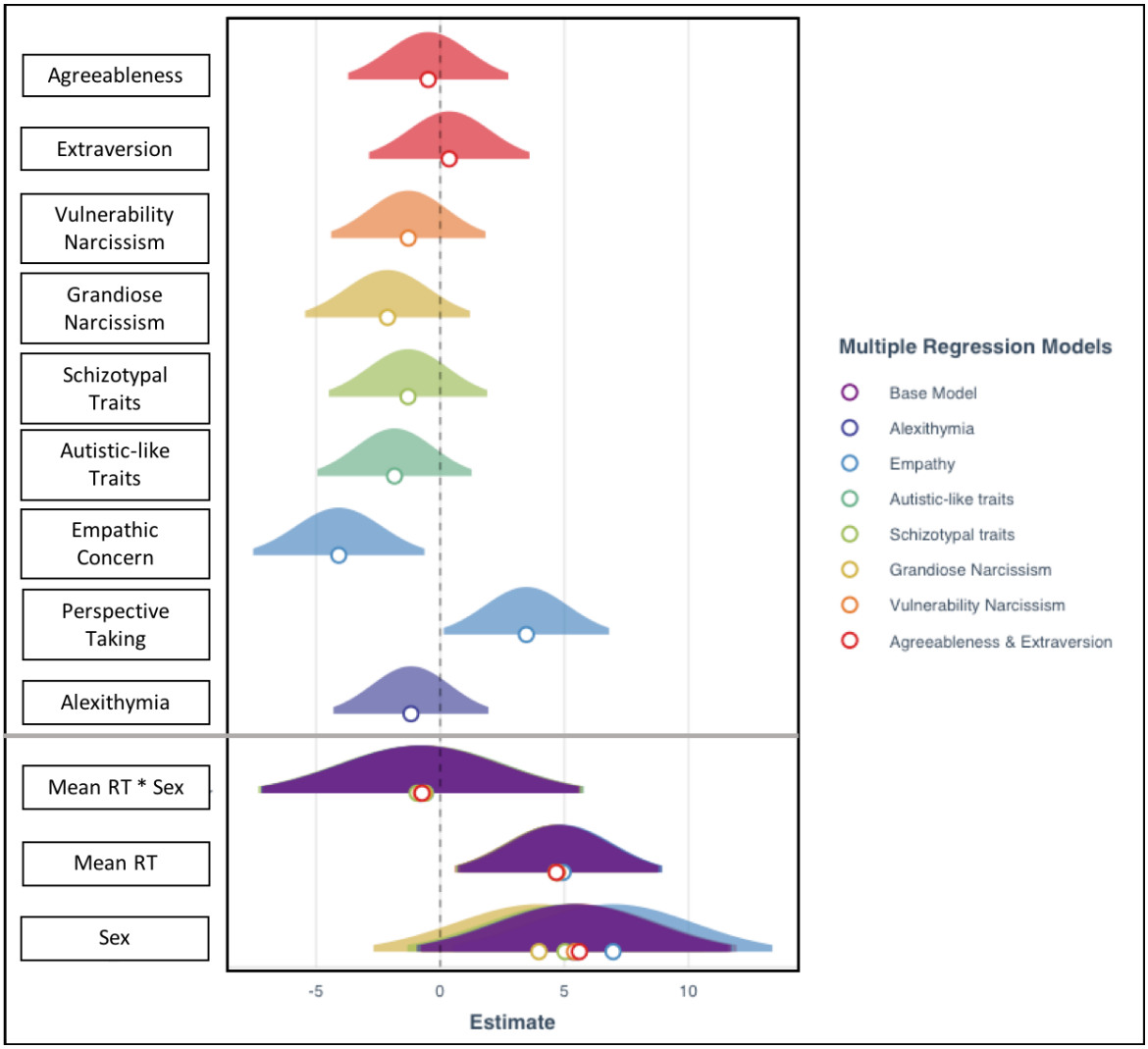


Figure S9. Experiment 3 (Spatial Compatibility) – Multiple Regression Analyses, Sex*Trait Interactions. Values of standardised coefficients are plotted for each predictor variable (personality trait), and the sex*trait interaction, along with their corresponding uncertainties (95% confidence interval width for a normal distribution for each estimate). Coefficients are standardised by dividing by two standard deviation units according to Gelman (2008). The base model consists in the bottom three predictor variables (depicted in violet) – mean RT, Sex, and meanRT*Sex. Abbreviations: RT = Reaction Time.