Impact of feedback loops on decision-making

Impact of Feedback Loops on Decision-Making

Igor Volzhanin

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

Birkbeck College

Department of Computer Science and Information Systems

University of London

August 11, 2019

I, Igor Volzhanin, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This thesis examines the impact of feedback loops on individual decision-making. This represents a long standing interest of cognitive psychology in how well human beings are able to use external information in individual and group settings to revise their beliefs to control complex systems. This thesis consists of six chapters. Each chapter contains a literature review section, followed by empirical research used to compare theoretical frameworks to actual human performance on a range of tasks. Chapter 1 serves as an introductory chapter by placing the subsequent analysis in the multidisciplinary domain of judgement and decision-making. Chapter 2 represents the first part of the thesis and explores human performance in controlling dynamic physical simulations. It begins by revisiting Berry and Broadbent (1984) research, followed by the exploration of how well humans are able to control dynamic physical systems. The chapter is primarily concerned with exploring the limitations of human control and factors that influence it, ending with the performance comparison between human and generic reinforcement learning algorithms. Chapter 3 extends decision-making into the social domain. It explores the impact of group dynamics on individual belief revision and proposes new models that may better reflect actual belief revision. Chapter 4 looks at the impact of incentivisation on revision and accuracy. It is found that incentivisation has a minor impact on belief revision. Chapter 5 extends group decision-making into the novel domain of rank revision. This chapter seeks to better understand how humans aggregate ranks and revise their beliefs. Finally, Chapter 6 summaries the findings and draws on the research presented in this thesis to provide concluding remarks on human cognitive decision-making processes in dynamic settings.

Contents

List of Figures

List of Tables

Acknowledgements

There is one person without whom this thesis would not have been written. From the day Ulrike Hahn took a chance on me to the day this thesis was submitted, this journey would not have been possible without your tireless dedication and support. The encouragement, patience and time you dedicated surpassed my wildest expectations. I thank you from the bottom of my heart. Onwards and upwards! This thesis would also not have come to be without Dell Zhang and the Department of Computer Science and Information Systems who came in at just the right time both financially and intellectually.

I would also like to extend many thanks to Erik Olsson and Jens Hansen at Lund University for funding some of the research presented in Chapter 3 and graciously providing the dataset and experimental materials from their previous research on which much of the modelling in that chapter is based. Next, I would like to thank Stephan Hartmann and his colleagues at LMU for providing the initial idea and draft paper on rank revision that later become Chapter 5. That was indeed a fortuitous encounter in Munich. I would also like to thank Xiaoling Wang, Guangxuan Song and Junwen Jian for spending many months building the online version of the sugar factory, which was used for Experiments 4 and 5 in Chapter 2.

I would also like to sincerely thank the following individuals for providing their comments on the initial drafts: Nina, Peter, Yana, thank you! Your comments were instrumental in getting this over the finish line.

Finally, I would like to thank my family: Helen, Sergei and Larisa. You understood the importance of this step in my life and your unwavering support made all of this possible.

Chapter 1

Introduction

The world is full of dynamic systems and processes. Something as simple as an interaction between two people can be infinitely complex; full of happiness, misunderstanding, trust and forgiveness. Such complexity starts with a single action, a “hello” perhaps, but soon takes on a life of its own. That *hello* leads to a response, ultimately leading to a reaction, which begets a conversation, leading to any number of possible outcomes. These interactions are inherently *dynamic* – ever changing as a result of individual actions. These successive actions and reactions create loops, more precisely they create feedback loops, which end up creating dynamic systems we find all around us.

Feedback loops are ubiquitous; from a simple (or complicated) friendship, to fundamental principles of economics, to the way a car engine works. Our existence is constantly regulated by positive and negative feedback loops around us (Sterman, 2000). They carry important information about the environment and provide ‘feedback’ on its state, so that the next decision can be reached. That decision, in turn, alters the environment. The next time feedback is sought and received it will be different. Such systems are inherently difficult to study due to the *evolving* nature of the underlying state.

Due to the various interactions, these systems evolve and no two paths through the system are the same. Feedback loops produce a *dynamic* system, which reacts to the user, while forcing the user to adapt to it. Due to the constant flux, dynamic systems are by their very nature non-deterministic. They are constantly changing as a result of the interactions. Since these systems are all around us, human behaviour and decision-making in such systems is of great interest. Are human beings well adjusted to make optimal decisions in complex environments? Is their decision-making optimal? Can they learn from feedback and gain mastery over such systems? What cognitive processes and strategies are used for decision-making? These are the questions that this thesis seeks to answer.

1.1 Introduction

Many disciplines - engineering, biology, political science, economics, cognitive science - incorporate and account for feedback loops and the resulting dynamic systems. For example, Orrell (2011) argues that feedback loops are fundamental to understanding much of the economic decision-making. According to Bernheim and Rangel (2008), behavioural economics is a "research programme that investigates the relationship between psychology and economic behaviour"(Bernheim and Rangel, 2008, p.192). Thus, some of the most recent and relevant research on decision-making in dynamic systems is found in this field.

Behavioural economists study how feedback loops create structures that cause certain behaviour in groups. Akerlof and Shiller (2010) argue that asset prices are susceptible to positive feedback loops where investor confidence is a large determinant of price. Orrell (2011) also attributes many of the modern economic phenomena to feedback loops, from the increasing CEO salaries to the cyclical nature of the stock market. According to Nassim Taleb (2007), the world is becoming increasingly characterised by complex feedback loops that are creating non-linear effects resulting in arbitrary, unpredictable, winner-take-all effects. Indeed, as far back as the 18th century, Adam Smith argued that economics is a series of negative feedback loops that together regulate prices and profits, thereby maintaining equilibrium (Sterman, 2000). The work in this thesis expands on the understanding of feedback loops by conducting research designed to better understand how individuals control and incorporate feedback into their decision-making.

1.1.1 Feedback Loops

According to Sterman (2000), "[a]ll dynamics arise from the interaction of just two types of feedback loops, positive (or self-reinforcing) and negative (or self-correcting) loops" (Sterman, 2000, p. 12). Examples of each are contained in Figures 1.1 (positive), 1.2 (negative). It should be noted that the terms *positive* and *negative* do not carry normative, but substantive connotations, as self-reinforcing and self-correcting.

Figure 1.1 demonstrates a simple positive feedback loop where chickens produce eggs. The more chickens there are, the more eggs they produce. With time this leads to exponential growth that can theoretically be unlimited. On the other hand, Figure 1.2 demonstrates a negative loop, where chickens die as a result of crossing roads and with time, they would disappear, provided no new chickens are born. When these two loops are put together they represent a system that changes over time. A positive loop is constantly producing chickens and a negative loop that is constantly removing them. At a certain point this system would reach an equilibrium, where the rate of chickens being born would be roughly equal the number of chickens dying.

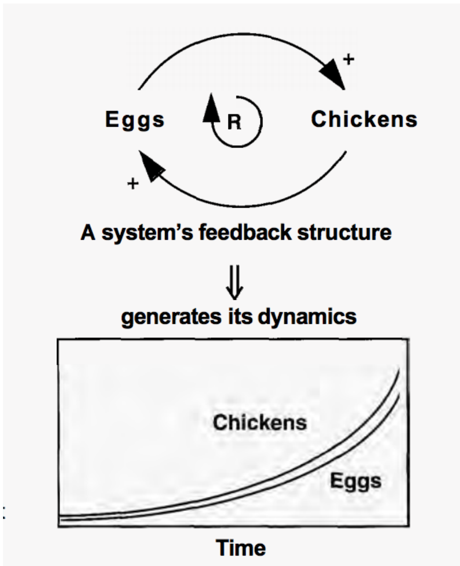


Figure 1.1: Example of a positive feedback loop (Sterman, 2000, p. 13).

This interaction between positive and negative loops is fundamental to much

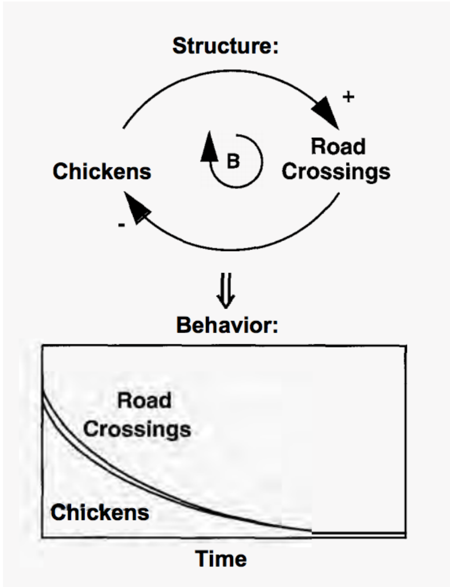


Figure 1.2: Example of a negative feedback loop (Sterman, 2000, p. 13).

of the world and can be found in nature (Neutel et al., 2014), as much as in engineering, economics, or psychology. Neutel et al. (2014) explains how plant growth (positive loop) and the resulting soil erosion (negative loop) create a stable system. Radios rely on the principles of oscillations and bistability, both produced by positive feedback loops (Zeron, 2008). One of the first artificial automatic regulatory devices invented – the water clock – that relied on negative feedback loops dates back to 285-222 BC Hellenistic Greece (Zeron, 2008), and used water level as a trigger for dropping of pebbles onto a drum to make a sound.

In biology, according to Zeron (2008), feedback origins are not as easy to trace because:

> Nature has been using feedback loops for millions of years, and because biologists in general discovered the principles of feedback centuries ago. The concept of negative feedback keeps popping out into our faces day after day. Properly speaking, it literally keeps popping out into our eyes, for the pupils light reflex in our eyes is the perfect example of a negative feedback loop. When light levels are high, the pupil contracts, reducing the light flux onto the retina. The size of the pupil is controlled by circularly arranged constricting muscles, which are activated and

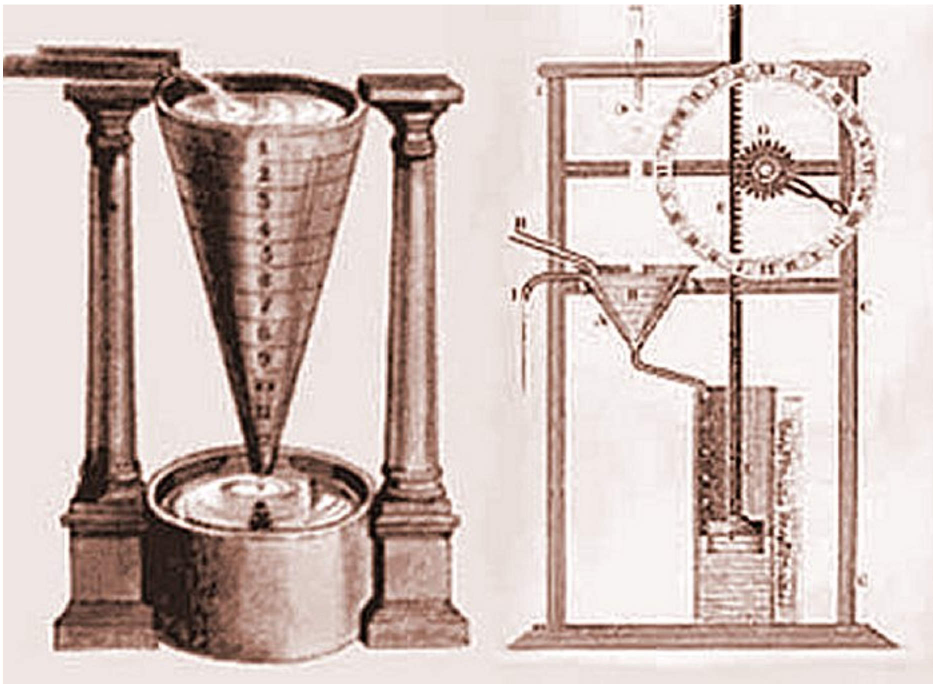


Figure 1.3: An ancient water clock. One of the earliest examples of a human-engineered device using feedback loops

inhibited (left to relax) by control signals from the brain (Zeron, 2008, p.69).

In economics, supply gives rise to a positive feedback loop and demand to a negative one. Given endless resources and time, the product supply would continually increase, leading to an ever greater output, while the demand is a negative loop since given enough time, all production would disappear. The *equilibrium* results at the point where supply and demand meet, creating a relatively stable system, at least for the time being.

1.1.2 System Dynamics

System Dynamics was developed in the 1950s by Jay Forrester to study the behaviour of complex systems over time. The feedback mechanism is at the heart of system dynamics, as feedback and circular causality are the tools for conceptualising systems and predicting outcome. He argued that "social systems belong to the class called multi-loop nonlinear feedback systems" (Forrester, 1971, p.2) and to understand any social system, its feedback loops must be understood.

System dynamics, as popularised by Sterman (2000) in the early 2000's seeks to apply engineering theories to social systems. This field has recast human understanding of social systems as a series of processes and systems that interact with each other creating complex structures requiring *system thinking*. The concept of total, or global understanding, where everything is connected to everything else, is of course important to consider. For example, when discussing policy resistance to change, Sterman writes: "with a holistic worldview, it is argued, we would be able to learn faster and more effectively, identify high leverage points, avoid policy resistance and make decisions consistent with our long-term best interests" (Sterman, 2002, p. 2). System dynamics ultimately calls for scientific (systematic) application of a number of concepts and tools from engineering, biology and technical fields to societal problems. Homer and Hirsch (2006) for example, have used system dynamic principles to model chronic disease prevention. According to Homer and Hirsch (2006): "System dynamics shows promise as a means of modelling multiple interacting diseases and risks, the interaction of delivery systems and diseased populations, and matters of national and state policy" (Homer and Hirsch, 2006, p.452). Currie et al. (2018) have looked at the application of system dynamics modelling in the decision-making processes related to environmental health, while Ibrahim Shire et al. (2018) looked at how system dynamics can help improve industrial safety.

System dynamics is undoubtedly beneficial to understanding complex systems with multiple feedback loops, which explains its growth and popularity. As is clear from Figure 1.4, systematic application of feedback as a way to explain complex systems in a clear manner is indeed powerful. However, as a framework it tends to focus on a macro perspective for the purposes of system-wide optimisation. It rarely focuses on the individual, or smaller groups. This is in contrast to cognitive science, where feedback has been studied in the context of learning, goal-setting and control.

1.1.3 Cognitive Science

The development of understanding of feedback in cognitive science follows the work of John Dewey who recognised the feedback loop character of learning around

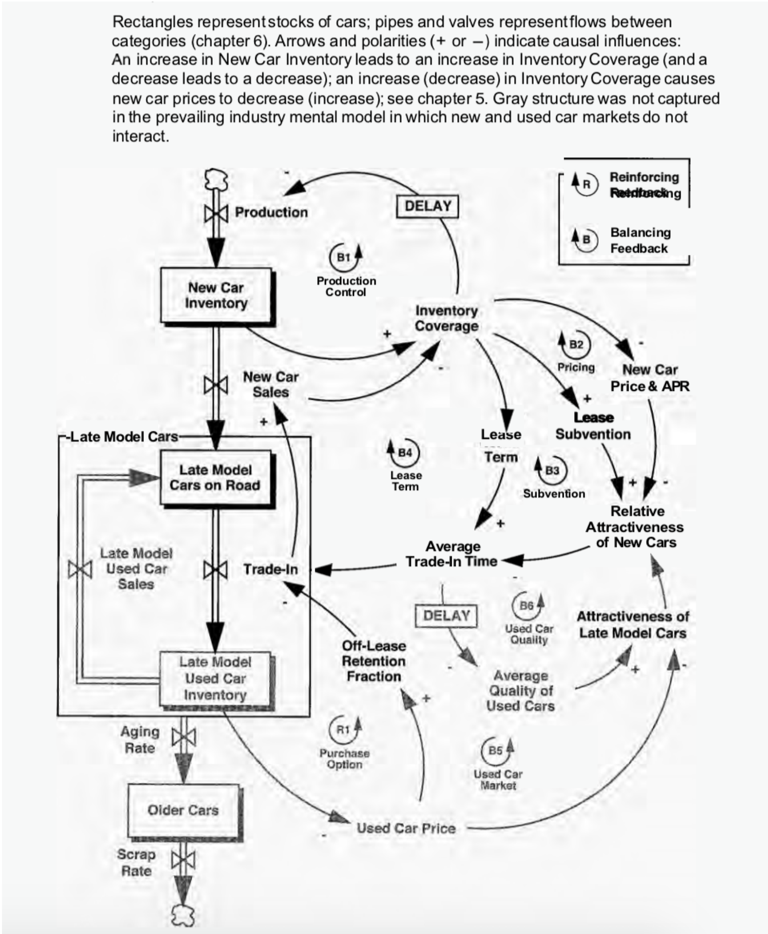


Figure 1.4: Model of automobile market (Sterman, 2000, p. 45).

the beginning of the 20th century when he described learning as an iterative cycle of invention, observation, reflection, and action (Sterman, 2000). In the 1940s, Maurice Merleau-Ponty laid the groundwork for the perceptual control theory, which is a model of behaviour based on the principles of negative feedback. The theory stated that negative feedback control applies to living organisms that do not control their behaviour, but vary their behaviour as the means for controlling perception (Flynn, 2011). Later, Powers (1973) wrote:

> Feedback is such an all-pervasive and fundamental aspect of behaviour that it is as invisible as the air that we breathe. Quite literally it is behaviour we know nothing of our own behaviour but the feedback effects of our own outputs (Powers, 1973, p. 351).

According to Carver and Scheier, behaviour and feedback control are intimately linked. They argue that goals act as reference values for feedback mechanisms, where negative feedback loops seek to diminish or eliminate discrepancy between the target (goal) and input (action) and positive feedback loops as discrepancy-enlarging (Weiner et al., 2012, p. 187).

This literature is extensive and it examines feedback as it pertains to the system providing information to the participant to learn from. However, individual impact on the system is limited. The individual is the receiver of information and adapts their behaviour as necessary, however, the behaviour itself does not modify the environment from which the feedback is received.

1.1.4 Dynamic Control

The dynamic control literature sought to expand the understanding of human learning into the context of more complicated systems that better approximated the real-world, evolving systems. The first simulations in this field were designed by Dörner (1975) and Funke (1992), including "a Beer Distribution Game...[where] subjects seek to minimise costs as they manage the production and distribution of a commodity"(Sterman, 1994, p.304). In their seminal study on dynamic control, Berry and Broadbent (1984) simulated the running of a sugar production facility. These represented a new kind of decision-making tasks, since the output produced by the simulation changed as a result of the actions taken by the participant. This established a feedback loop, whereby learning depended on the actions, which in turned produced the learning signal. Both Dörner (1975) and later Berry and Broadbent (1984) found that participants were actually quite bad at controlling the simulation, with output constantly fluctuating, and performance not improving with practice beyond a few initial trials.

Over time, these early tasks were incorporated into a wider body of research on dynamic decision making (see Cleeremans and Seger, 1994, Gibson et al., 1997, Gonzalez et al., 2003, 2005, Osman, 2010). The field looked into how well individuals are able to control and succeed in dynamic environments that are hypothesised to be more akin to the real world systems Gonzalez et al. (2003), Sterman (1994, 2000). Gonzalez et al. (2003) for example had participants conduct a much more sophisticated simulation with multiple inputs and feedback loops of a water purification plant, whereas Sterman (1994) simulated the workings of a beer factory, again with multiple feedback loops, Brehmer (1992) had participants fight forest fires, and Funke (1988) had perhaps the most sophisticated simulation of them all, with over 200 variables available for manipulation, in the so-called 'Lohhausen' simulation. Each simulation differed in the dynamism of the system (i.e how often it changed), complexity (how difficult the system was to operate) and opaqueness (how apparent the relationships are to the user) (Osman, 2010).

One constant finding is the documented failures of human participants to conduct simulations successfully. In an indicative description of human performance on the beer production task, Sterman (1994) stated that:

> [t]he subjects generated costly oscillations with consistent amplitude and phase relations, even though demand was essentially constant. Econometric analysis of subjects decisions showed that people were insensitive to the time delays in the system. People did not account well, and often not at all, for the supply line of orders that had been placed but not yet received, causing them to overcompensate for inventory shortfalls. Facing an inventory shortfall, many subjects order enough beer to close the gap. Because of the delay in filling orders, inventory remains depressed, and the next period they order more beer. Still deliveries are insufficient, and they order more beer. Finally, the first order arrives, inventory rises to the desired level, and the subjects cut their orders. But the beer in the supply line continues to arrive, swelling their inventory many times above the desired levels and causing emotional reactions from anxiety to anger to chagrin (Sterman, 1994, p.304).

In her review paper on dynamic decision-making Osman (2010) identifies a whole host of undesirable psychological phenomena exhibited by human participants in these tasks, including: "high reliance on biases, high persistence of unsuccessful strategies, poor strategy development/rule-based knowledge, misperception of feedback" (Osman, 2010, p.75). What has not been clearly established in this field is whether human failures occur as the result of the simulations being difficult in of themselves, or whether failures represent a more fundamental problem of human cognition not being able to exert dynamic control.

Outside of the unanswered questions regarding human performance, the main limitation of dynamic control tasks is the fact that they focus on a single participant, controlling these simulations. Given that the world is full of individuals and their interactions give rise to feedback loops, understanding social, or multi-individual, systems is of great importance when one wishes to understand impact of feedback loops on decision-making. Thus, one must go outside of the artificial worlds constructed by researchers to fully comprehend the power of feedback loops.

1.1.5 Network Science

With the advent of social media and increased global interconnectedness, feedback received from social network interactions has only increased. Social interactions have become more complex and more likely to be impacted by others, as well as have a wider impact on the environment thereby creating feedback loops.

Network science is a relatively new field of study, but one that is gaining importance in an increasingly interconnected world. With the advent of the internet, information as well as communications have brought human beings closer together. With Facebook's drive to 'connect the world' this is only likely to increase further. Network science is "the study of the collection, management, analysis, interpretation, and presentation of relational data" (Brandes et al., 2013, p. 2). This field draws on statistical analysis, mathematics, data analysis and management, network and graph theory, among others. Its applications have been numerous, from contagion spread modelling to opinion dynamics. Given that "the roots of network science are particularly strong in social psychology, sociology and anthropology" (Brandes et al., 2013, p. 3), this field has quite a bit to offer in terms of understanding individual behaviour as it relates to the system. Since networks are normally understood as a series of *nodes* and links between them (*edges*), the framework lends itself quite well to the study of social interactions between individuals (for a comprehensive overview on network science see Jackson (2010)).

The earliest social science studies in this field can be attributed to Paul Lazarsfeld, who pioneered opinion research, showing, among other things, how social connections play a role in learning and the formation of opinions and different agents in a society have different influences (Jackson, 2010). David Krackhard extended network science to business environment showing how interconnectedness works within a large organisation (Krackhardt and Hanson, 1993), with informal networks having an outsize impact on performance. Krackhard predicted in 1993 that as traditional institutions flatten, network science and network managers will have an increasingly important job of managing their organisations through the resulting informal networks (Krackhardt and Hanson, 1993, p.111).

More recently, and with increasing frequency, mathematical models have been applied to the study of networks. De Groot's consensus models to this day remain an important cornerstone for analysing consensus (opinion convergence) in networks (see Hegselmann and Krause, 2002, Gonzalez et al., 2003, Das et al., 2014a). De Groot's consensus model, and much of the subsequent research focused on explaining social interactions in mathematical terms, often preferring hypothetical models to empirical research. The examples of such models include: the impact of homophily (the tendency for individuals to seek ties with people who are similar to themselves) to become educated (Jackson, 2009), consensus convergence (Hegselmann and Krause, 2002), and persuasion (DeMarzo et al., 2003). These applications are complex, requiring extensive mathematics to describe fundamental social interactions. The value of such models is to be able to hypothetically explain and ultimately predict outcomes of social interactions.

While original network science has focused on mathematically simple, but empirically untested assumptions about individual strategies for information incorporation and belief revision, much of the current research focuses on observation. Although research has already been done looking at individual and group belief revision (Jönsson et al., 2015, Becker et al., 2017, Moussaïd et al., 2017), the purpose of my research is to examine the *strategies* individuals use to revise their beliefs. The goal is better understand how individuals *actually* revise their beliefs. What information do they consider? How likely are they to change their opinions? What strategies do they use in revising? It is therefore the purpose of this thesis to empirically test existing models on actual human behaviour and if necessary to create a more robust model that would explain actual belief revision of individuals in socially connected networks.

1.2 System Classification

Dynamic systems and games used to understand decision-making can be broadly classified into a two-dimensional table. One axis refers to the nature of the system where the underlying source of knowledge either changes as a result of participant actions (dynamic), or is unchanging (static). In dynamic tasks every action and decision actually changes the underlying structure of the task, creating a feedback loop between the participant and the task. This interaction alters the very structure of the task and feedback that participants receive is a function of the input that they provide. Static tasks on the other hand have predefined logic that cannot be altered by participant behaviour. Examples of each task are presented below.

The other axis represents the collaborative nature of the task at hand. Some tasks are performed by a single individual, working on a task alone, or in isolation from others, not receiving any external information from the other participants (individual). Other tasks involve a group of participants, or indeed entire networks of connected individuals, working and communicating with each other to solve the task either collaboratively, or competitively (group). Importantly, participants are able to communicate and thus receive information from other participants while doing the task.

The traditional cognitive science tasks where a single person is asked to complete a task, or a challenge that does not change as a result of user input falls squarely under the *static* category. An example is the Wason task, where participants are asked to complete a reasoning task using cards (Wason, 1968). Materials are static, in so far as their values do not change as a result of user input. Certain answers may change the path or the level of difficulty of the task, but fundamentally, the task and the materials remain unchanged by users actions. Another example is the artificial language learning tasks for kids and adults (Folia et al., 2010, Uddén et al., 2012). In these tasks participants are often asked to learn an artificial language and answer questions about it. The structure of the language does not change as a result of user input, although the structure of the task itself can give rise to complex systems. Such tasks are quite common in cognitive science, as laboratory experiments tend to be designed to test particular functions, or processes and are required to be quite deterministic in order to remove extraneous effects.

The field of dynamic decision-making was created as a recognition that decisions in the real world are more complex and that the very systems that call for decisions are also altered by them Sterman (1994). For example, the sugar factory simulation required participants to maintain a certain output level over a number of turns Berry and Broadbent (1984). Output at each stage was a factor of two variables, the present user input and previous output (and therefore the previous user input). Thus, users had direct impact on the system and production targets would change over time based on user input. It mattered what user entered two times in a row as both of these inputs would have an impact on production. The environment in the task is non-stationary and changes as a result of user input.

Most of the tasks participants dealt with were simulations of physical processes Funke (1988), Gonzalez et al. (2003), Sterman (1994). The unifying factor always being that participants had some information over the task, did not necessarily explicitly understand the functioning of the system from the onset and had to learn through feedback. Feedback, which was very much a function of user input in the first place. This non-stationary and evolving nature of the system makes it a great paradigm within which to study individual's ability to control feedback systems.

Studying individual behaviour in dynamic context is valuable. However, given the rise of social media and the fact that human beings inherently exist in a social space with others, group dynamics are also important to consider. Understanding how an individual interacts with the group is an important first step. One set of work in this domain comes from the literature on advice (Yaniv, 2004a, Yaniv and

	Static	Dynamic
Physical	Generic tasks Wason (1968), Uddén et al. (2012)	Dynamic decision-making tasks Berry and Broadbent (1984)
Social	Decision-making and belief revision Yaniv (2004a)	Individual belief revision in a group setting Jönsson et al. (2015)

Table 1.1: Task classification of different feedback systems. These tasks are divided into four categories based on their properties.

Milyavsky, 2007) (see Yaniv, 2004b, for a review). The advice literature focuses on mathematical and algorithmic 'rules' of advise incorporation, as well as how context (advisor, confidence, prior knowledge) affects the extent to which the advice takers revise their beliefs in light of new information (Dalal and Bonaccio, 2010, Yaniv et al., 2011, Yaniv and Choshen-Hillel, 2012, Wanzel et al., 2017). An individual is usually asked a question and then is provided with a range of answers that the 'group' has come up with. An individual is then asked to revise their belief and come up with a new answer in light of the information received from the 'group'. However, the 'environment' in this case does not change. Participants' answers do not have an impact on the generated answers. The system remains static. This line of research is relevant given that the range of answers is generated by a 'group' and provides a framework for extending individual studies into a group context, closer to the environment experienced in social media.

The final quadrant of the decision-making tasks are tasks where groups are involved and where the setting is dynamic, constantly changing as a result of user input. This environment reacts to individual user feedback and each user reacts to the feedback received from the environment, thereby creating feedback loops. One line of research where this methodology has been used is in group revision dynamics by Jönsson et al. (2015). They examined the impact of network structures on individual and group accuracy. They had multiple participants participate at the same time, answers the same questions, while also asking them to revise the answers in light of the answers generated by their peers. While the purpose was to understand how network structures impact group and individual accuracy, this paradigm introduces important concepts for studying belief revision. The experiment generated a large data set, tracking individual revisions as well as the answers each individual was exposed to. Multiple rounds of revision allow for the understanding of revision over time. Questions were general knowledge questions, which participants did not necessarily know the answers to, but could guess, making this an accessible task to different people with varied backgrounds. Accuracy as a central measure of success is important as well, allowing for measurement of success, or failure of various strategies for belief revision.

1.3 Hypothesis

So how would individuals behave when presented with feedback? Recent studies take a rather positive view of human ability to incorporate feedback to produce positive outcomes (Granovskiy et al., 2015, Jönsson et al., 2015). According to Jönsson et al. (2015), individual and group accuracy increases as a result of repeated feedback. Furthermore, it is apparent from the everyday life that human beings are able to control a variety of real-world tasks with feedback loops in them. From flying fighter jets to driving cars, human beings have generally displayed a remarkable ability to incorporate feedback and achieve control and mastery. However, it has also been long observed that individuals show limitations in the way they adjust their behaviour based on feedback (Brehmer, 1995, Gibson et al., 1997, Gonzalez et al., 2003, Osman, 2010, Sterman, 1994). This is doubly true in situations where feedback is delayed, or is difficult to interpret.

Given the seemingly contradictory findings, the research in this thesis is focused on dynamic systems and experiments will be used to better understand how individuals make decisions in such environments, focusing on the overall ability to achieve positive results under different conditions. It is hypothesised that factors that impact human performance are: timing of feedback (delayed vs. instant) – where delayed feedback will lead to poorer outcomes; social vs. system interactions – where social interactions will normally lead to increased performance, while interactions directly with the system may lead to more divergent outcomes; randomness – where more randomness would impede mastery; incentivisation – where providing individual incentivisation would lead to greater overall mastery (accuracy) over the task.

1.4 Thesis Structure

There are two parts to this thesis. Part I focuses on the individual's ability to control physical dynamic tasks. Chapter 2 extends the dynamic system control literature to better understand the challenges individuals face when conducting such simulations. It has been well documented that individuals struggle to complete such systems. The chapter starts with a replication of the original sugar factory and moves to explore whether individuals can be taught to better control such systems. It then moves to the discussion around objective performance metrics and ultimately seeks to deconstruct the various parts of the simulation to better understand what exactly causes participants to struggle with control over the simulation.

Part II extends research into the social domain. Specifically, this research looks at feedback in group opinion dynamics. It starts with Chapter 3 and the search for strategies that individuals employ in dynamic revision tasks. The goal is to better model individual belief revision in a group setting to understand what strategies participants use when revising their beliefs and creating feedback mechanisms as a result. Once these strategies are understood, Chapter 4 looks at individual vs group incentivisation performance and what impact it has on the revision strategies, as well as group accuracy. Finally, Chapter 5 looks at a novel application of feedback learning in the context of rank aggregation. Ranked lists are lists where there is an ordered relationship between a set of items such that, for any two items, the first is either 'ranked higher than', 'ranked lower than' or 'ranked equal to' the second. Rank aggregation is a particularly interesting application for belief revision, given that ranked information is often used in computer science and information retrieval. It is the goal of the chapter to explore what strategies are most beneficial when aggregating ranks and then understand what people actually do.

Chapter 6 summarises the findings and provides concluding remarks on the topic of dynamic control and revision.

The research for each Part was done concurrently. While the overall theme of feedback and control remains constant, the concluding remarks and the overall discussion on the hypothesis proposed above are left to the final chapter.

Chapter 2

Part I - Dynamic Control

2.1 Background

To date, the majority of work on learning and decision-making within cognitive science has focused on static environments. For example, seminal work by Herbert Simon (1956) focused on rational behaviour in a typical ‘psychological’ environment, which did not change as a result of the actions taken by the participant. Much later Gigerenzer (2001), continued to build on this work by looking at the environmental and social, as well as psychological factors of human rationality, yet the environment continued to be unchanging and unresponsive to the actions of the individual in these experiments.

Most real world learning and decision-making occurs in a different type of an environment. Typically, individuals face an environment that is not static and changing over time, very often as a response to the interactions within it. Such a *dynamic* environment is by its very nature complex and difficult to predict and requires holistic understanding of the system for the analysis to be useful Sterman (2000). The complexity is further compounded by the fact that the type of feedback one receives – which is essential for learning – is a function of one’s own interactions Sterman (1994).

Given that these systems appear to approximate the real world, concepts introduced by Sterman and his colleagues have been applied to a multitude of real world systems, from public health (Homer and Hirsch, 2006), to environmental sustainability beliefs (Melville, 2010). Homer and Hirsch (2006) for example, have used the concepts of system dynamics to model chronic disease prevention, which allows for incorporation of "all the basic elements of a modern ecological approach, including disease outcomes, health and risk behaviours, environmental factors, and health related resources and delivery systems" (Homer and Hirsch, 2006, p.452). In other words, system dynamics captures nuances and concepts that exist in the real world, thereby allowing for a more accurate representation of the real world.

However useful modelling system-level processes is, learning in dynamic environments is ultimately an individual endeavour. Since a complex system is ultimately a composition of individual choices, there is great value in understanding individual decision-making in a complex system. This area of research has been of cross-disciplinary interest for decades, yet no single theory of learning has emerged (Osman, 2010). An interdisciplinary field of dynamic decision-making has attempted to address this knowledge gap by focusing on learning and control in dynamic, ever evolving systems (Berry and Broadbent, 1984, Brehmer, 1990, 1992, 1995, Gibson et al., 1997, Gonzalez et al., 2003, Gureckis and Love, 2009a, Osman, 2010, Sterman, 1994).

It has long been observed that individuals show limitations in the way they adjust their behaviour based on feedback (Brehmer, 1995, Gibson et al., 1997, Gonzalez et al., 2003, Osman, 2010, Sterman, 1994). This is doubly true in situations where feedback is delayed, or is difficult to interpret. If the underlying system is unknown and feedback is received without any framework to interpret it, it becomes difficult to incorporate or make useful inferences from it. Compounded by the participants' inability to verbalise knowledge attained from the task was the fact that much of the learning was marginal (see Shanks and St John, 1994, on the implicit learning debate). Over the decades of implicit learning experimentation, many of the participants failed to effectively perform and control the various tasks presented to them.

However, *failure* to achieve positive outcomes in these artificial dynamic decision-making environments does not in and of itself constitute an objective cognitive failure on the part of a participant. Without an objective measure against which to measure performance, it is impossible to establish if failure is indeed due to cognitive limitations, or due to poor experimental design, which asks participants to do something that is simply impossible. To our knowledge, no such yardstick has been proposed in the past. Therefore, determination of the cause of performance difficulties fell to the researcher.

The goal is to introduce a more objective measure of performance; namely, a reinforcement learning algorithm tasked with 'solving' decision-making tasks and acting as a yardstick for measuring reasonable performance on a given task. In particular to apply reinforcement learning algorithms to the famous sugar factory designed by Berry and Broadbent (1984) in order to measure human performance against the performance of such an algorithm.

Reinforcement learning has already been used in cognitive science, and decision-making in particular, to better understand human performance. For example, Gureckis and Love (2009a) have used differently programmed reinforcement learning algorithms to conduct a 'Farming on Mars' task in the exploration vs exploitation paradigm, in order to better understand the various cognitive mechanisms supporting learning in such tasks. The task was somewhat similar to the sugar factory in that "the reward structure continually evolved in response to the actions of the individual" (Gureckis and Love, 2009a, p. 2).

Although reinforcement learning has been shown to approximate human learning in various circumstances Dayan and Niv (2008), Gureckis and Love (2009b), Schoenberg et al. (2007), it is used here as a benchmark for understanding the level of difficulty of the task, rather than a model to explain cognitive process of human players doing the task. This is an important distinction, as the aim of this section is to establish and evaluate performance of an artificial agent that can learn to control the simulation. Once such an agent is established it is possible to compare human performance against it, allowing us to make a performance comparison, rather than a cognitive claim.

The goal of this chapter is to focus on developing an artificial agent, powered by a reinforcement learning algorithm, that would play the sugar factory in order to determine which parts of the simulation are objectively more difficult, in an attempt to establish a baseline against which human performance could be measured.

The chapter is structured as a series of experiments designed to reproduce and expand on the original sugar factory and explore the nuances of human performance on the task. First experiment aims to replicate the original study, while the second experiment is programmed with a different underlying equation to allow for more refined control, while also examining the efficacy of training materials for the participants. A third study extends the domain and the cover story from a fictional sugar factory and towards a more topical area of climate change to see what impact the cover story has on participants' ability to control the simulation. It then moves away from lab experimentation altogether and into a web-based game, allowing for a larger participant pool and a different way for individuals to interact with the study. Finally, reinforcement learning algorithms are introduced to conduct the original sugar factory and simulation results are compared to those of human participants.

2.2 Sugar Factory

The sugar factory was originally designed by Berry and Broadbent in 1984 and was initially used to test the limitations of learning and feedback control. It also provided evidence for implicit learning."The task has been widely used to investigate hypotheses about how decision makers learn on-line in dynamic decision environments" (Gibson et al., 1997, p. 1), and is widely cited as one of the premier simulations of its time (Berry, 1991, Berry and Broadbent, 1984, 1987, 1988, Buchner et al., 1995, Dienes, 1990, Dienes and Fahey, 1995, Gibson, 1996, Hayes and Broadbent, 1988, McGeorge and Burton, 1989, Sanderson, 1990, Stanley et al., 1989) Over time, the sugar factory simulation evolved into a broader set of experiments and simulations on dynamic decision making (Cleeremans and Seger, 1994, Gibson et al., 1997, Gonzalez et al., 2003, 2005, Osman, 2010). Although not the first simulation of its kind, it became the foundation upon which a multitude of other, more complex simulations were built upon. Its popularity and persistence in academic literature can be explained by a combination of factors. It was one of the first simulations. Its authors were influential voices in the implicit learning debate. The results it produced were surprising, and therefore illuminating to the limitations of human cognition. Structurally, the simulation is very simple. There is only one variable that can be influenced and feedback given by the system is timely and correct. There is no additional noise, parameters, or hidden state shifts (see Gureckis and Love, 2009a, for a more complex simulation example). Despite its simplicity, researchers have consistently found that participants are unable to successfully complete the task. This has given rise to a number of conclusions on human cognitive limitation, such as ‘misrepresentation of feedback’, ‘poor resource allocation’, ‘poor attention to feedback’ and others (see Osman, 2010, for overview of the limitations and biases).

The basic simulation is straightforward. Participants are tasked with controlling a computer simulation of a sugar production facility by adjusting the number of workers assigned to production. Each turn a new number of workers is assigned and output is displayed on the screen. The task is to achieve and maintain a certain level of output. After each assignment, participants are provided with immediate feedback on the actual factory output for the turn and are asked to enter a new number. This is repeated for a set number of turns. However, as noted by Gibson et al. (1997) and later Gonzalez et al. (2005) this simulation represents a problem of significant difficulty due to the way time figures in the underlying equation: the system requires two independent inputs – at different times – to stabilise production. Otherwise, production tends to oscillate between two extremes, and the resulting feedback only serves to further confuse the participants. Immediate feedback produced by the task is only marginally useful and most participants fail to learn from it (at least initially) (Berry and Broadbent, 1984, 1988, Gibson et al., 1997). Individuals do get better with practice, but remain unable to correctly describe the underlying system (Berry and Broadbent, 1984, 1988, Gibson et al., 1997), with most of the participants appealing to factors and mechanisms that are not present in the actual simulation to describe its inner workings. This is consistent with the instance-based learning theory (IBLT) suggested by Langley and Simon (1981) and adopted by Gonzalez et al. (2005). This theory suggests that: "people learn with the accumulation and refinement of instances, containing the decision-making situation, action, and utility of decisions. As decision makers interact with a dynamic task, they recognize a situation according to its similarity to past instances, adapt their judgement strategies from heuristic-based to instance-based, and refine the accumulated knowledge according to feedback on the result of their actions" (Gonzalez et al., 2003, p.591). Mastery over dynamic decision-making tasks revolves around learning the heuristics and pattern recognition, rather than holistic understanding of the system and its underlying processes.

2.2.1 Experiment 1: Replication

Much time has passed since the original sugar factory was introduced. With the advent and proliferation of personal computers, it is quite possible that the current generation may be more adept at controlling the simulation. Because of the amount of time that has elapsed since Berry and Broadbent's original study, and the change in human activities and experiences, such as proliferation of gaming, it may be wise to start with a replication to provide a baseline of modern performance. All attempts have been made to stay as close to the parameters of the original study as possible, using the same underlying equation.

2.2.2 Method

Participants were assigned to one of two conditions based on the instructions they received. In the first condition, participants were told that they should reach and maintain a target output of 9,000 tons of sugar per month, which corresponds to the instructions in the original task. Given that Berry and Broadbent's evaluation of performance involved a range, whereby any output between 8,000 and 10,000 counted as 'on-target', we also examined a version where participants were told of this evaluation criterion. For both sets of instructions, output in this range was considered to be 'on-target'. All participants completed the simulation twice.

2.2.2.1 Participants

All participants (n=20) were students at the University of London who volunteered to do the study. They were all between 18 and 28 years old. No further demographic information was collected on the participants. Participants were randomly assigned to one of the two conditions and asked to read a slightly different script regarding the task. One group of participants were told that only output at 9,000 would count on-target, whereas participants in the second group were told that any output between 8,000 and 10,000 would be counted as being on target.

2.2.2.2 Materials

The ‘sugar factory’ was coded in MATLAB, and participants saw a simple graphical interface, with a graph in the middle of the screen, with one horizontal line displaying target output and another showing factory output. The underlying equation Berry and Broadbent governing factory output each month was:

$$s_{t+1} = (2a_t - \frac{s_t}{1000} + r) * 1000 \tag{2.1}$$

where s was the output, a was the number of workers, t was the turn number, and r was a random element (noise), that could assume a discrete value of -1,0 or 1 and was of uniform distribution. Each participant was given the following instructions, which were based on the original (Berry and Broadbent, 1984, p.24) instructions. The full set of instructions can be found in Appendix A:

> You are in charge of running a sugar production factory in an underdeveloped country. You control the rate of production by simply changing the size of the work force, ignoring all other factors. You start with 600 workers that produced 6000 tonnes of sugar in the previous month. Your task is to reach and maintain a target output of 9,000 tons per month. To help with the task, the maximum output of the factory has been set at 12,000 and the minimum to 1,000. You will have to run the factory for 30 months. Each month you will assign a number between 1-12 representing the number of workers that would work in the factory

that month. The computer will multiply your number by 100 to get the actual number of workers. Example: 6 is 600 workers.

2.2.2.3 Procedure

Participants were asked to control the factory for 30 turns (months). Each turn they entered a number between 1-12 and receive instant feedback on the factory output. A sample complete simulation of the factory (run) can be seen in Figure 2.1. After completing the factory once, all settings were reset and participants were asked to run the simulation again, with the same instructions, for the second time.

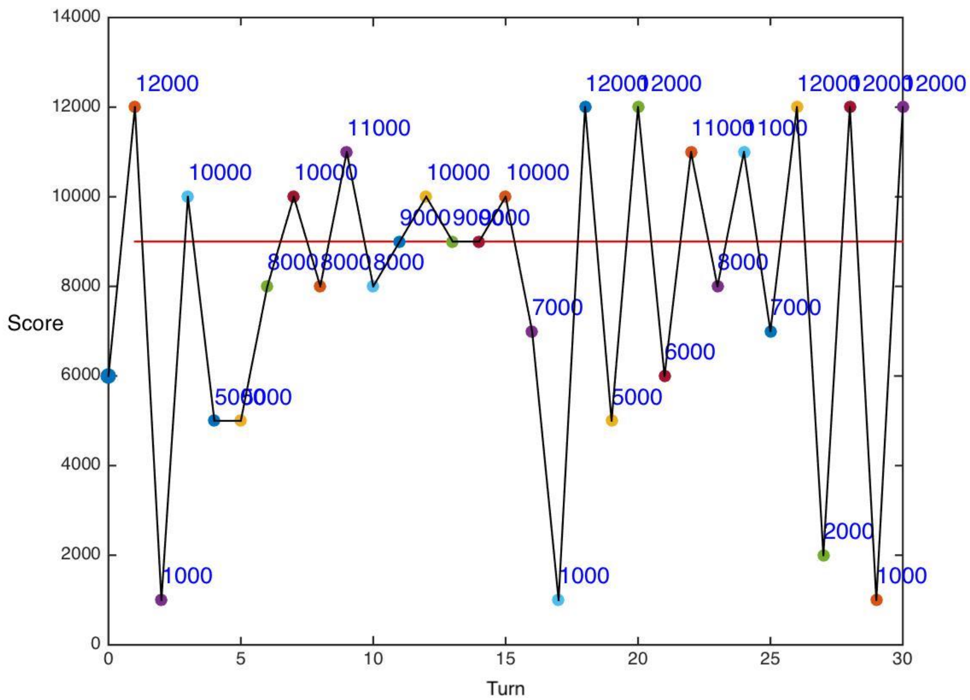


Figure 2.1: A sample output of a player completed basic sugar factory. Each dot represents factory output.

2.2.3 Results

Output that fell between 8,000 and 10,000 was counted as ’on target’. The number of times participants managed to achieve this output over 30 turns represented their performance on the task. As there was no significant difference in performance depending on the instructions received, the data was combined from both groups.

Table 2.1 shows the mean performance on our replication study compared with the original Berry and Broadbent study. The replication results are in line with the original findings, which indicates that there has not been a notable shift in skills, or ability to control this type of a game over the past 30 years.

Table 2.1: Replication study average performance of all participants over 30 trials. It is split by the first and second run of the simulation.

Run	Original Result	Replication Result
1	5.58	6.1 (SD 4.0)
2	6.83	7.6 (SD 3.1)

Also, as in the original study, a significant practice effect was present, $t(19) = 2.5$, $p = .021$, $d = 0.57$, confirming that participants did significantly better on the second run.

2.2.4 Experiment 2: 'Advanced' Sugar Factory

Based on the results of the replication study, it was clear that control of the original sugar factory remains as difficult as ever. As suggested by Gibson et al. (1997) and Gonzalez et al. (2005), difficulty appears to lie in the underlying equation. But, is it the delayed nature of the feedback that creates this difficulty?

It is difficult to delve deeper into nuanced participant behaviour due to the noisy output of the equation. For example, entering '6' could produce a factory output of between 4000 and 9000, depending on previous output and random noise. This impact of the random variable (r) produces final output that can be up to forty percent different; coupled with the fact that participants could only control the factory by assigning workers in hundreds, making gradual increase and decrease virtually impossible. These characteristics also prevent meaningful interpretation of individual strategies used by the participants.

In order to expand and elaborate on where exactly the difficulty lies, the governing equation was modified for the second experiment. The new equation governing the factory gave participants a much greater range of control over the output. In the original set up, a range of possible inputs was limited to 12 and each worker was deemed to represent 100. The new equation allowed participants a much greater range of inputs of between 0-1200. Furthermore, the new underlying equation was changed to make the function with a fixed exponent in order to control for the new ability to set a more exact number of workers, while maintaining the difficulty level, as greater inputs would still produce a disproportionately large output (but the variance was still less than the original equation). The delayed nature of feedback was preserved in the new set up. However, borrowing from the literature on dynamic decision-making an attempt was made to create explicit ‘training’ materials directing participants, in general terms, on how they may be able to control the factory with greater precision. This was a 2x2 between participant factorial design experiment with manipulations of the governing equation (original/power) and training (present/absent).

2.2.5 Method

Participants started with one type of the factory, and then completed two sets of the other, with participants in the training group receiving a two-page handout with explicit instructions after completing the first factory.

2.2.5.1 Training

The strategies outlined for the participants were based on a combination of work on intuition training (Hogarth, 2001), heuristic competence (Brehmer, 1992, Sterman, 1989, Dörner, 1997), and decision competence (Gonzalez et al., 2005). Hogarth (2001) argued that applying the scientific method to decision-making improves outcomes. The four stages of this method are: observation, speculation, testing and generalisation (Hogarth, 2001, p.24). This framework was combined with Brehmer’s observation that performance on such tasks is greater for “subjects who collect more information, who collect it more systematically, who construct adequate goals, who evaluate the effects of their decisions, and who generally behave in a systematic fashion tend to perform better than those who do not” (Brehmer, 1992, p.225). Lastly, Gonzalez et al. (2003) observed that: “decision makers improve their performance by following heuristics less closely and more inconsistently. Experienced decision makers show a lower fit to heuristics and higher standard deviation compared to their own behaviour at the beginning of their practice" Gonzalez et al. (2003). Combining these insights, it was important to get participants to experiment, and to use a wider range of strategies to tackle the simulation. The strategies presented to participants in the two-page document were based on the idea that inducing a negative feedback loop-like behaviour would allow for greater control over the system. The two-page handout provided to the participants included information about feedback loops and hints such as: attempt to understand the underlying processes; test various theories; make changes gradually when testing your theories; account for noise; and use your answers to create convergence on the goal. Participants were asked to read the two pages, paying particular attention to the strategies outlined on the second page. See Appendix A for a copy of the training materials.

2.2.5.2 Participants

Participants (n= 40; 10 per condition) were all students at the University of London who either volunteered to do the study, or did so for course credit. This was a between participant (2x2) study where each participants was provided with a different set of instructions. Hence, there were 4 groups with 10 participants in each group. No demographic information was collected on the participants.

2.2.5.3 Materials and Procedure

The *power* version of the sugar factory was governed by the following equation:

$$b_t = \frac{a_t^3}{10+r} - a_{t-1} \tag{2.2}$$

with each variable representing the same assignments as above. The lower and upper bounds were set at 1,000 and 250,000 respectively.

As in the first experiment, participants were asked to control the factory for 30 turns (months). For the *original* Berry and Broadbent version (labelled as 'basic' in the analysis), participants entered a number between 1-12 (representing multiples of 100 workers), whereas in the *non-linear* version (labelled as 'advanced') participants could enter any number between 0-1200, giving them much greater control over production. In addition, to help with the task, the advanced version of the sugar factory contained two further pieces of information. First, the factory output was displayed above the black line representing the output for the turn. Secondly, the number of workers assigned by the participant was displayed above a dot on the chart. This was intended to aid participants in keeping track of the number of workers they assigned each month. Figure 2.2 demonstrates a successful run in the advanced factory.

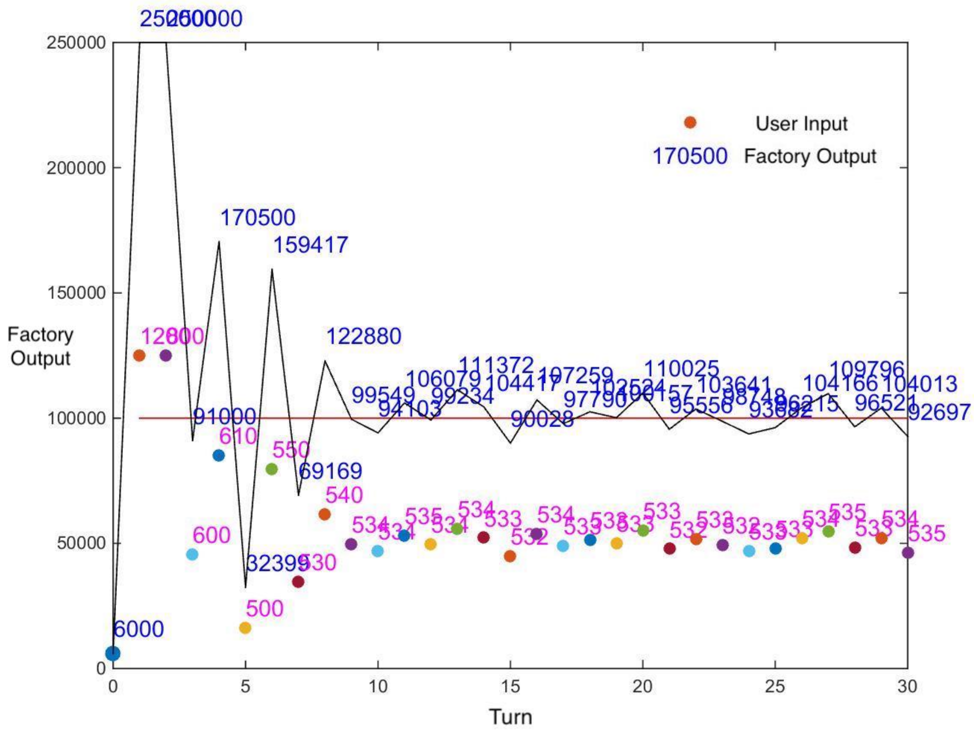


Figure 2.2: Example of a successfully player-controlled advanced sugar factory. Each dot represents player output and the continuous line represents factory output.

Participants were explicitly told that they would be controlling a different type of a factory when switching from one type to the next. They were told that one factory was located in a developed country (advanced factory), or an underdeveloped country (basic factory). Target output and number of possible workers to assignment differed significantly between the two factories, and served as a cue that the two factories operated differently. Finally, the output was presented differently in the two factories further highlighting the differences. However, participants were not told that each factory was governed by a different equation.

It was expected that participants would incorporate the training materials and perform better in those conditions that included them. Furthermore, it was expected that participants would find the advanced version of the factory more difficult initially, but would also perform better with time, as the governing equation allowed for more varied strategies, and more explicit feedback, which should be more conducive to learning.

2.2.6 Results

The target output of the basic factory was the same as in Experiment 1: 9,000 tons, with any response between 8,000 and 10,000 scored 'on-target'. The advanced factory had a target output of 100,000, with output between 90,000 and 110,000 deemed 'on-target'.

The number of outputs that each participant scored as on-target was summed to determine the overall performance. Table 2.2 provides a summary of the mean number of hits per condition for all participants.

Table 2.2: Average performance on the basic and advanced sugar factory split by simulation run.

	No Training Mean (SD)	Training Mean (SD)
	Basic	
Run 1	6.7 (2.8)	8.6 (3.3)
Run 2	5.9 (1.4)	11.4 (5.3)
	Advanced	
Run 1	8.8 (6.5)	4.6 (3.3)
Run 2	12.8 (9.1)	9.2 (6.6)

2.2.6.1 Practice

A two way ANOVA showed a significant effect of practice on performance in the advanced sugar factory, $F(1,18)$= 10.427, p= .005, η_p^2= 0.578, as well as the basic version, $F(1,18)$= 7.097, p=.016, η_p^2=0.374.

However, provision of training materials did not produce a statistically significant effect, nor was there significant interaction between training materials and practice for either of the versions of the factory.

2.2.7 Discussion

Given the somewhat artificial nature of the basic factory, one might suspect that limitations displayed by participants, including the failure to benefit from explicit training, may be brought about by its specific characteristics. Unlike the basic factory, the advanced factory allowed for greater control over the number of workers being assigned, lower noise levels and a range of solutions that could stabilise the system and reach the desired output. Nevertheless, most participants still did not exhibit a high degree of control over the advanced factory. Nor did they benefit from training.

In order to successfully reach and maintain a target output of 100,000, participants would have had to *gradually* assign up to between 520 and 540 workers each month. Provided they did so for a few turns, the volatility in the output would have decreased significantly and it would have settled in the target range. The exponential growth of the output, however, added to the complexity of finding the optimal solution. Entering numbers above 650 would yield a significant growth in output and entering numbers under 400 would have yielded significantly lower output (with neither output particularly helpful to solving the problem). Consequently, participants who only entered numbers in large ranges did not do particularly well on this task.

In the advanced factory, those participants who were able to stay on-target the most tended to exhibit behaviour that was outlined in the training material. In particular, the most successful individuals tended to test extreme values early on and then slowly increase, or decrease the amount of workers until they oscillated around the target. They learned it themselves, not from the training materials, as there was no significant effect of the training materials on performance. This suggests that either the method of delivery, or the information itself in the training materials is not an efficient method of imparting knowledge on how to best control dynamically changing, nonlinear systems.

2.3 Experiment 3: Climate Change Simulation

Practice had a statistically significant impact on performance, but participants were on-target less than 50% of the time, with the average performance on the second run being 12 out of 30 on-target 'hits'. Changing the original equation by giving participants more control over assignment, did not produce any difference in performance, neither did the provision of training materials.

In the reasoning domain, it has been observed that participants are better at tasks when they are moved from abstract manipulations to concrete real-world tasks. For example, experiments using the Wason Selection task materials "add credence to the conclusion that framing the task in a thematically meaningful way can facilitate performance"(Nickerson et al., 2017, p.134). There is general agreement that it is an oversimplification to treat conditionals as if they were conjunctions, and it has been suggested that the conflation of the two occurs less often when using real world materials (Over et al., 2007, Singmann et al., 2014)

It is then worth considering whether participants would interact differently with a system that more closely resembles a real world system, rather than a more abstract 'sugar factory'. In this case participants may be able to draw on their previous knowledge and experience about a topic to improve their performance, better incorporating the strategies outlined in the training materials. This would also have an added benefit of measuring the extent to which behaviour might be affected by general attitudes, opinions and beliefs.

One of the most high-profile systems in current public debate is Earth's climate and the issue of anthropogenic climate change (Revkin and Seelye, 2003, Rosenzweig et al., 2008). Climate change is a complex issue with multiple feedback loops, which seems well suited for the study of systems control that seeks to understand the impact of prior beliefs, while being relevant to the participants. Basing a simulation on this topic would also have an added benefit of measuring the impact of interactions with a climate simulation on beliefs about climate change. A new climate change simulation was developed, complete with a new cover story, interface and a governing equation. The simulation was based on a simplified version of the Kaya identity (see Kaya, 1990), which seeks to express the relationship between various social and economic factors and CO_2 emissions.

The Kaya identity incorporates several high level concepts, such as greenhouse gas emissions, wealth, energy intensity and population in a single equation. The equation brings together diverse fields of study - physics, economics, engineering, demographics into a simple equation.

$F = P * \frac{G}{P} * \frac{E}{G} * \frac{F}{E}$ where:

F is global CO2 emissions from human sources

P is global population

G is world GDP

E is global energy consumption

As Mann and Gaudet (2018) explain, the interpretation of the equation is that:

> by projecting the future changes in population (P), economic production (G/P) , energy intensity (E/G) , and carbon efficiency (F/E) , it is possible to make an informed projection of future carbon emissions (F). Obviously, population is important as, in the absence of anything else, more people means more energy use. Moreover, economic production measured by GDP per capita plays an important role, as a bigger economy means greater use of energy. The energy intensity term is where technology comes in. As we develop new energy technologies or improve the efficiency of existing energy technology, we expect that it will take less energy to increase our GDP by an additional dollar, i.e., we should see a decline in energy intensity. Last, but certainly not least, is the carbon efficiency. As we develop and increasingly switch over to renewable energy sources and non-fossil fuel based energy alternatives and improve the carbon efficiency of existing fossil fuel sources (e.g., by finding a way to extract and sequester CO2), we can expect a decline in this quantity as well, i.e., less carbon emitted per unit of energy production (Mann and Gaudet, 2018, para.2).

The simulation based on the Kaya identity involved maintaining the Gross Domestic Product (GDP) levels of an economy in such a way as to keep voters happy, while simultaneously maintaining acceptable CO_2 targets. The concept of voter satisfaction was added as a mechanism of tying the various variables together and providing a way to end the simulation. If voter satisfaction dropped below a certain level, the player was 'voted out of office' and the simulation ceased. In addition to completing the new simulation, participants were asked to complete two rounds of the advanced sugar factory to test task transfer between the two simulations. It was expected that the new domain and its proximity to the real world issues would induce greater interest and ultimately lead to better performance on the task.

2.3.1 Method

In the first stage of the experiment, both groups were asked to complete a short questionnaire (see sample questionnaire in Appendix C) on their beliefs and opinions regarding climate change. This was followed by one run of the advanced sugar factory as described in Experiment 2. In the second stage, depending on the condition, participants either received no instructions, or were provided with training materials similar to those used in Experiment 2, along with a description of the climate simulation task and asked to control the new system, with the goal of reaching 40 turns. After completing the simulation four times, participants again completed the climate change questionnaire to gauge the post-intervention beliefs and attitudes towards climate change. Lastly, participants completed another run of the advanced sugar factory.

The climate questions were designed to gauge participants' overall attitudes towards climate change, their behaviours and intent, and their opinions of collective and government action. The questionnaire was based on Whitmarsh (2011) climate scepticism scale. Additional questions probing specific attitudes toward climate action were drawn from the Yale and George Mason University Climate Change study (Leiserowitz et al., 2012) and the Cardiff University report: 'Public Perceptions of Climate Change and Energy Futures in Britain' (Spence et al., 2010). In particular, questions dealing with behaviour and intent were taken from the Yale study, while questions related to government were taken from the Cardiff report. Additional questions regarding participants' own future behavioural intentions were simple modifications of these questions.

These questions were selected based on their broad coverage of the topic as well as their previous use in academic studies on climate change beliefs and are considered to be robust for understanding individual beliefs on the topic (Spence et al., 2010). The final questionnaire contained questions on climate beliefs, opinions on climate action, views on the role of government in dealing with climate change and future action intentions. In line with the research from Clark et al. (2013), it was believed that participant behaviour would be impacted by the simulation and would become more attuned to the issue of climate change, specifically by its link to human activities. It was also hypothesised that participants who have stronger belief in human-caused climate change would perform better on the task, as the task was designed in such a way to promote CO2 reduction.

The advanced sugar factory was added to this experiment in order to introduce participants to the notion of dynamic decision-making tasks, to see if skill transference may occur as a result of completing the climate simulation, and to corroborate the findings in the previous experiment on the lack of impact of training on the ability to complete the simulation.

2.3.1.1 Participants

Participants were a diverse group of individuals, and either students at the University of London, or professionals who either volunteered to do the study, did so for course credit, or were paid £5. 30 individuals took part in this study, distributed evenly between the two conditions.

2.3.1.2 Materials

The Climate Change simulation was programmed in MATLAB, with the interface similar to the previous simulations. It was governed by the following equation:

$$y_t = (x1_{t-1} * 1.05) * (x2_{t-1} * 0.98) * (x4_{t-1} * (a_t/100) + 1) * 100$$

where y is the final output of the round, t is turn number and a is user input. $x1$ was a positive feedback loop that represented carbon content of the economy and increased over time, while $x2$ was a negative feedback loop that represented energy intensity and diminishes over the course of the simulation; $x4$ represented GDP growth.

Where possible numbers corresponding to the actual state on Earth were used for the simulation, although they were modified to make the simulation run for more turns. 1.05 was used as the world GDP growth, and approximates the true global GDP growth number (IMF, 2019), and 0.98 was used as the constant for energy consumption, which is expected to decline at about an annual rate of 2% (Ritchie and Roser, 2019). Both of these values were picked to allow for the simulation to run for more turns. Additionally, while participants were able to indirectly control $x4$ by inputting the value for a, $x1$ and $x2$ could not be directly influenced. These values grew and declined automatically, reflecting that the carbon content of the energy is expected to grow, while energy intensity of the economy is expected to decline over time and would magnify the choices made by participants.

While participants were asked to control the simulation for 40 turns, it was not expected that anyone would actually reach that target. Rather, it was important to see how long participants would last and what choices they would make along the way. 40 turns were chosen as a benchmark because that would give participants ample time to try different strategies before they simulation ended. It should be noted, that it was possible to finish the simulation successfully if a particular strategy focused on early CO2 reduction was chosen.

2.3.1.3 Procedure

In the climate change simulation, participants could enter any number between 10 and -10, indicating GDP growth or decline. The system provided immediate feedback on a graph showing the corresponding CO_2 growth or decline. The simulation also included a voter 'approval' bar, which reflected popularity. The bar would change colour and get progressively redder if negative GDP growth was entered by the user, or if emission levels became too high. Reaching severe negative approval levels removed the player from office, ending the simulation prematurely. Participants had to maintain a balance between the ever rising emissions, which could be stopped by GDP growth cuts, and the political unpopularity of such moves, which would ultimately end the simulation if the economy grew too slowly.

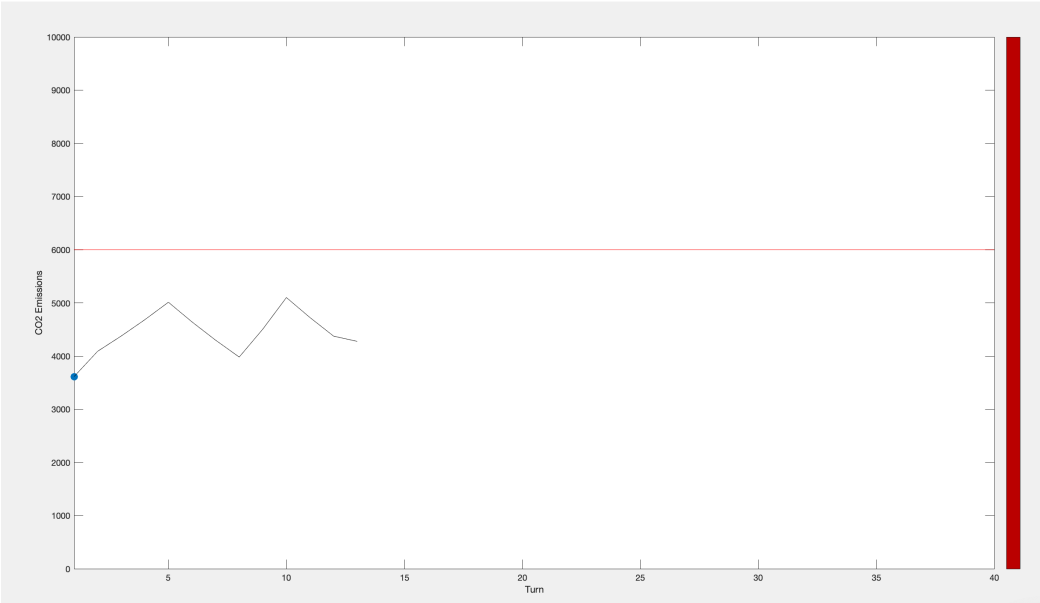


Figure 2.3: Completed Carbon Emissions Simulation. The red line represents emissions limit, the line in the middle represents the current CO_2 output and the red bar on the right represents popularity with the voters.

After completing the simulation four times participants completed the same version of the initial climate change questionnaire, but the order of the questions was randomised. The second administration of the questionnaire was designed to gauge how conduct of the simulation may have impacted climate change beliefs.

It was postulated that participants would indeed be impacted by the simulation and would become more attuned to the issue of climate change and specifically its link to human activities. This is in line with the findings of Clark et al. (2013), who found that a quick introduction to climate change through simulations or instructions increases climate change acceptance. It was also hypothesised that participants who have stronger belief in human-caused climate change would perform better on the task. Finally, it was predicted that the individuals who received training will do better on the task. Since the training sheet contained information regarding the importance of initial choices and determinism thereof it was expected that the participants would incorporate this into their behaviour.

2.3.2 Results

2.3.2.1 Climate Change

Unlike the sugar factory, the carbon simulation could end before turn 40 (only one participant reached the stated goal). The performance metric on this task was the number of rounds a user was able to keep the simulation going. Table 2.3 summarises the results, broken down by run, and whether or not training was provided.

Table 2.3: Carbon Emissions Simulation Results. It shows the average number of turns participants were able to control the simulation for, split by the run of the simulation

Category	No Training	Training
Run 1	13.2 (SD 7.5)	13.3 (SD 7.8)
Run 2	14.0 (SD 8.3)	16.0 (SD 9.3)
Run 3	14.9 (SD 8.8)	15.0 (SD 9.1)
Run 4	15.3 (SD 8.9)	15.3 (SD 9.2)

A 4x2 repeated mixed measures ANOVA with run (1-4) and group (training, no training) as within and between subject factors showed the main effects of run which refers to the number of times a simulation was done by a participant, $F(3,84)$= 3.223, p= .027, η_p^2= 0.103. This was a post-hoc anova test comparing different runs 1 and 2 vs 3 and 4. There was no effect of training, nor was there a significant interaction between the different variables.

2.3.2.2 Beliefs, Opinions and Actions

Questions administered to participants covered four broad categories of climate change: climate beliefs, opinions on climate action, views on the role of government in dealing with climate change, and future action intentions. None of these categories showed significant effects of engaging with the simulation in the before-after comparisons. This was unexpected in light of findings by Clark et al. (2013) and Ranney et al. (2012) who suggested that interaction with climate change knowledge leads to greater belief in climate change. Clark et al. (2013) in particular has conducted a series of experiments designed to influence climate change beliefs in a laboratory setting. They found that carefully presented scientific information in a form of a 400-word letter can, "clearly change the public's understandings and opinions [on climate change]" (Clark et al., 2013, p.2070). Furthermore, participants in past studies have been shown to revise their beliefs and attitudes when presented with conflicting information (Horne et al., 2013). While participants in our study did not receive a detailed explanation of scientific evidence for climate change, they did receive basic explanation of the Kaya identity and their instructions stated the following:

> In front of you is a simulation of the Kaya identity, which economists use to express the relationship between several social and economic factors and CO2 emissions. This simulation starts in 2000. You have just been appointed as the Prime Minister of a developed country and your job is to stay in power by carefully balancing economic growth with the rise in CO2 emissions. Climate change is about to become a major political issue as it is becoming clear that if carbon emissions are not curbed, the global temperatures will rise, with potentially unpredictable consequences. As such, you are to navigate a path between economic growth and carbon emissions. Your task is to maintain healthy GDP growth, while ensuring that your country's emissions stay below the critical threshold of 6 billion tons of CO2 a year.

Regression analysis showed no significant correlation between questionnaire scores and task performance. Performance did not appear to be influenced by participants' prior beliefs and opinions. Resistance to the use of background knowledge and beliefs is surprising as increased belief in climate change and need for personal and governmental action was expected. After all, the system was based on what has been advanced as a genuine characterisation of the relationship between GDP and CO_2.

2.3.2.3 Sugar Factory

Finally, performance on the advanced version of the sugar factory was analysed, which participants completed at the beginning and the end of the experiment. Table 2.4 below provides the summary of the results, broken down by run and whether or not participants received training between first and second run.

Table 2.4: Advanced Sugar Factory Results showing the average number of times participants were able to stay on target in the thirty turns. It is split by simulation run, as well as showing the no training and training conditions, compared to the second run of the second study.

Category	No Training	Training
Run 1	9.3 (SD 5.7)	13.1 (SD 5.0)
Run 2	13.4 (SD 7.0)	13.1 (SD 7.1)
Run 2 (Table 2.2)	12.8 (SD 9.1)	9.2 (SD 6.6)

The results for the second run are very similar to the results presented in Table 2.2 and are reprinted at the bottom of Table 2.4 for reference. This suggests that the climate change simulation did not have an effect on participant performance in the sugar factory task. No skill transfer between the two systems was observed. As in the previous section, there was no statistically significant effect of training on performance.

2.4 Experiment 4: Beyond the Lab

Given the inherent shortcomings of conducting lab-based experiments, namely, participants having limited time to understand instructions, limit to the number of times a simulation could be run, and maintaining high motivation, it was important to take it beyond the lab. Overall, the structured nature of the laboratory environment may well contribute to the lack of concrete results observed so far. Allowing participants to play the simulation beyond the lab was the only way to eliminate these constraints.

To take the study outside the lab is to bring it closer to the way we interact with feedback loops in the real world. Indeed, as anyone who has played Candy Crush, or a similar crop of games may have noted, mastery over a task usually comes naturally, but requires time. To explore whether such mastery may occur with the sugar factory task it was necessary to program a web-based interface for the simulation. The goal was to create a simulation that would be ‘addictive’ enough for participants to do it over a period that is longer than possible in the lab and to complete more simulations overall.

The original version used in Experiment 1 of Berry and Broadbent (1984) sugar factory was selected for the experiment as it provided the most direct link to the original studies on the subject and performance could be easily compared, both to the replication studies conducted in Experiment 1 and original research. Furthermore, the original sugar factory is a simple simulation, which could be easily explained in a remote setting, increasing the chances of an individual playing it for longer.

2.4.1 Method

In the first instance, the sugar factory was re-coded in Java and deployed on the web.[1] All attempts were made to retain the original look and feel of the application, but make it usable 'as-is', without needing to assist, or further explain the task to participants. The simulation was hosted on a Digital Ocean server under the following link: http://178.62.60.219:8080/SimpleServlet/

The link to the simulation was distributed over Facebook among friends of the researchers to seed distribution, but it could also be publicly accessed with no restrictions.

Each participant was provided with a link to the website where they had to register and provide consent to participate in the study. Afterwards, they were able to login, view a tutorial on the basic controls of the simulation and begin to play the sugar factory (See Figure 2.5)

2.4.1.1 Participants

In this case participants were simply those who registered and played the web application. Demographic information, such as age and level of education was collected during the registration phase. However, since registration was anonymous it was impossible to independently verify it and all data was self-reported. Also, due to a software bug, data was not recorded correctly in the database.

There were 23 unique registrants, but there were actually 20 users who played

[1] A special thank you to Xiaoling Wang, Guangxuan Song and Junwen Jian for coding the web-based version of the simulation.

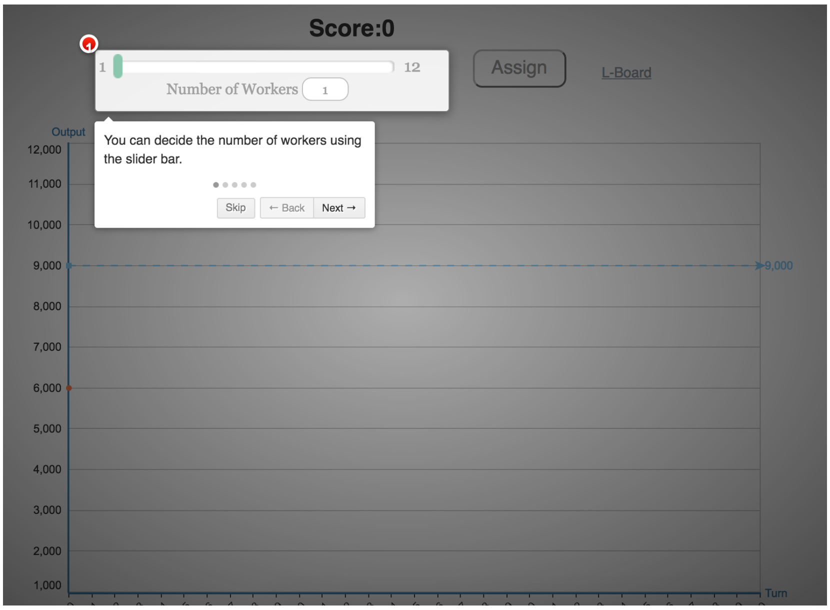


Figure 2.4: Web Sugar Factory Tutorial. Players would this screen after logging into the simulation for the first time.

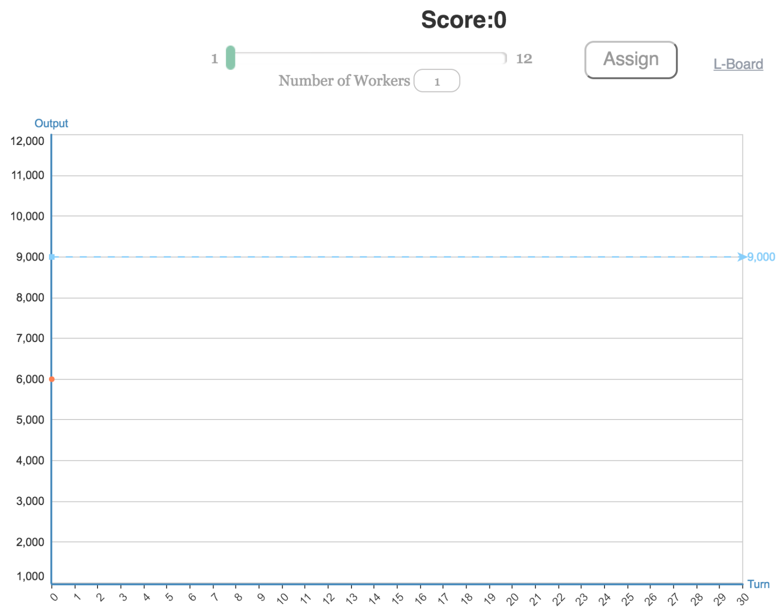


Figure 2.5: Sample Web Sugar Factory Interface. Players would see this screen at the beginning of the simulation, with x-axis representing the turn and y-axis representing factory output

the simulation at least once. Participants were not paid for participation and no incentives were provided, beyond a leader board that displayed the highest scores of the day and the week.

2.4.1.2 Materials

The web version of the simulation was based on the same set of instructions and the underlying equation as the original sugar factory featured in Experiment 1. All instructions were posted online and participants were given an opportunity to play a tutorial before doing the actual simulation.

There were no restrictions on time or the number of runs that a participant could do. The target was the same at 9000 tons of sugar per month. The only addition was the score that participants would see at the top of the screen. This score reflected their success in maintaining the target and was governed by the following set of rules:

- 1000*x points for hitting the target of 9,000 tons of sugar, where x is the number of successive on-target hits
- 500 points for being within 1000 tons from the target (8,000 or 10,000)
- 100 points for being within 2000 tons from the target (7,000 or 11,000)

Although the scoring logic was not readily displayed to participants for review, it closely matched the instructions and rewarded the type of play where participants maintained output at the target level for several turns in a row. The more consecutive turns on target the bigger the multiplier grew, generating a higher overall score. The score was displayed above the main area and was updated after each turn. At the end of 30 turns, the final score was displayed, along with the leader board containing top 5 players of the day and their highest scores, as well as the top 5 weekly players and their respective scores. This added an element of competition and showed participants the range of scores that was possible to obtain in the simulation, encouraging additional attempts at beating their own score, as well as those of others.

2.4.1.3 Procedure

The link to the registration page was distributed widely through Facebook and email, starting with the network of the lead researcher. Beyond that, initial recipients were encouraged to share the link with their friends as well. The intention was to distribute and encourage participation as widely as possible.

Upon registration, participants were able to access the simulation screen. After completing a short tutorial they were asked to play the simulation for 30 turns. After each set of 30 turns they were shown the final score and asked to play again. There was no limit to the number of times one could play. Indeed, while most participants only played 5-10 times, others played as many as 30 to 40 times. There were no time limits and participants could log in at any time to play.

2.4.2 Results

2.4.2.1 Performance

The basic measure of performance on this version of the sugar factory was the same as for the previous two: the number of times participants were able to be 'on-target', defined as having achieved the output of between 8 and 10 thousand tons in a given month. Summary statistics, including minimum, maximum and mean for the number of tries, number of hits on target and highest scores.

Table 2.5: Web Sugar Factory Results showing the mean, minimum and maximum statistics for all players on the following characteristics: number of runs, number of times players were on target in each run, as well as the higher score achieved.

Category	Runs	On Target	Highest Score
Mean	9.9	9.23	17,875
Min	1	1	1,000
Max	40	18.02	47,600

Table 2.5 shows that an average user played about ten times and was on target 9.2 times out of 30 turns. This is comparable to the findings in the first experiment. There is, however, high variability in these measures, with some participants quitting after a single run, while others playing up to 40 times, which is significantly higher than the number of runs in our lab-based studies, where participants had a maximum of four chances to complete the simulation.

2.4.2.2 Practice Effects

Overall, there was a statistically significant and strong correlation between the highest score achieved by each individual and the number of trials they conducted. This relationship continues to hold for the individual highest scores as a function of the

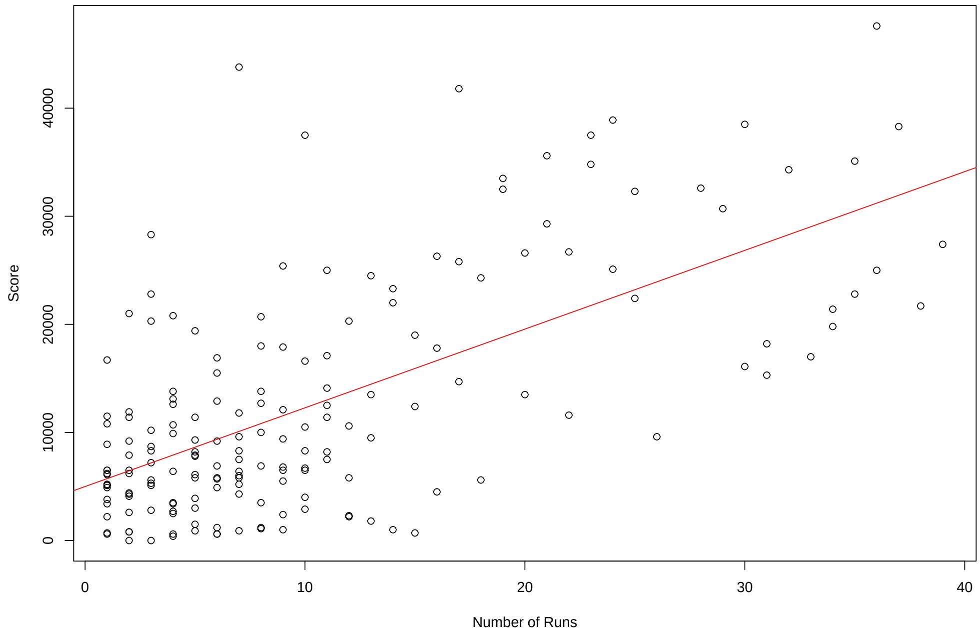


Figure 2.6: Number of runs plotted against the highest score achieved on the run.

number of runs. The regression correlation of this relationship is: $R^2 = .63$, $F(1, 18) = 33.56$, $p < .001$ (see Figure 2.7 for visualisation).

Participants who played the simulation the most, achieved significantly greater mastery over the task. Their top scores were significantly higher (44,400 to 13,194) than those who played only a few times and achieved an average number of hits per trial that was substantially higher (6.77 to 2.57) than the beginners. See Table 2.6 for comparison between the top three most active players and the rest.

There is a clear practice effect, where those who played the most also performed better over time. This is in line with the original findings that demonstrated presence of the practice effect. On the extreme side, three participants played the simulation up to 40 times and became competent at it, far surpassing the other 17 who played only a few times, or those who conducted the simulation in the lab. While there is some divergence in the higher number of trials, this is largely due to the sparse number of participants who actually completed more than 15 trials.

Although some mastery was achieved through practice, there was also a limit to this effect. Even top performers could on average be on target only 6.7 times out

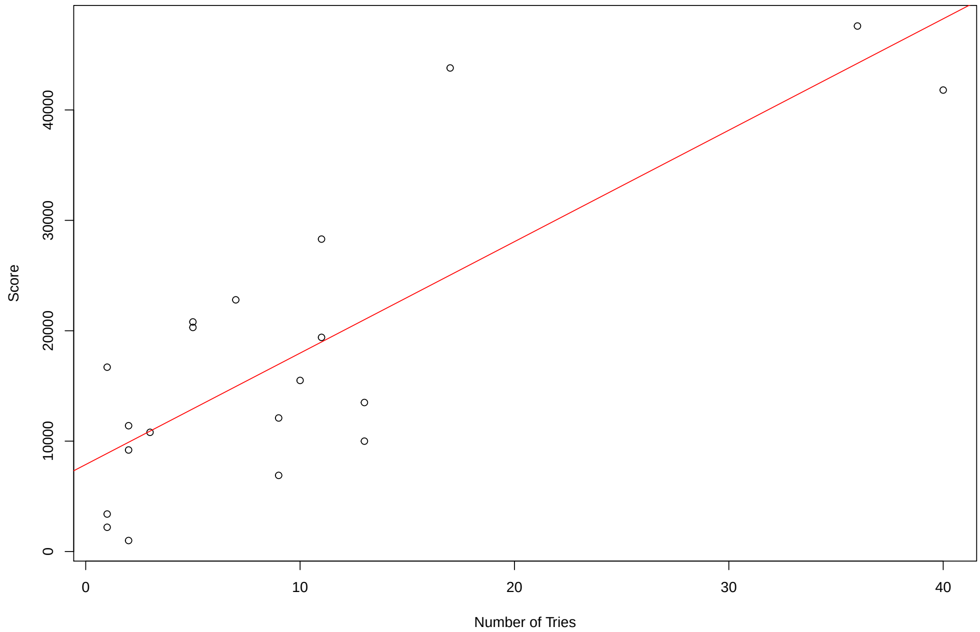


Figure 2.7: Highest score per participant compared against the number of tries attempted by the participant.

of 30, or just over 20%. And the top score of 44,400 represents just over 10% of theoretically maximum score of 420,000. And although overall mastery increased with practice, participants continued to struggle to maintain a stable system, constantly oscillating, and perhaps most interestingly, diverging widely between runs, achieving some stability on one run, while failing to achieve the same score in the next.

Table 2.6: Revised Online Sugar Factory Results

Summary Statistics	Top 3 Players	The Rest
Average Number of Trials	31	6.18
Average Top Score	44,400	13,194
Average Hits per Trial	6.77	2.57

2.4.3 Discussion

To summarise, there was a successful replication of the original sugar factory, with results suggesting that participants continue to struggle to master the task, although they do get better with practice. Providing training materials does nothing to increase performance and neither does giving participants an ability to more finely control the number of workers in the factory. There did not appear to be any difference in how they conducted the simulation.

Furthermore, there did not appear to be any transfer learning, or greater mastery displayed when the domain was switched to better approximate the real world system of climate change. Lastly, providing participants with the ability to play outside of the lab setting, does lead to some performance gains, but does not ultimately result in the participants being able to be on target more than 20-30% of the time.

So far all findings suggest deep limitations in the ability to control the sugar factory. These findings match closely the original results (Berry and Broadbent, 1988), which has led many researchers to conclude that human mastery in feedback tasks are difficult to achieve, representing a cognitive limitation (Brehmer, 1992, Sterman, 1994, Gibson et al., 1997). However, how can one speak about cognitive limitations control in the absence of an independent measure of success?

2.5 Making a Robot

Are human beings simply bad at doing the sugar factory, or is there a more fundamental problem with the sugar factory? How good of an example is the sugar factory of real world dynamic control tasks that humans perform every day? Clearly human beings successfully navigate a complex environment that is their daily life, which is full of feedback loops on the daily basis.

Unfortunately, without a model to compare human behaviour in dynamic tasks to, this is impossible to answer. Indeed, without a standard against which to compare performance, it is impossible to say if the ability to do the sugar factory is something that one could possess at all.

2.5.1 Reinforcement learning

There are a number of approaches that can be taken to model human behaviour. According to Kieras and Meyer (1998), "there are two communities that have been interested in modelling and predicting human performance. [One looks at] analyzing the task that the system operator performs, using systematic task analysis methods that have developed over many years of practical experience in system analysis and design. [While analysis of the second group] are based on computational modelling packages that implement an overall structure for human cognition, a cognitive architecture, analogous to the hardware architecture of a computer" (Kieras and Meyer, 1998, p.2) (see Howes et al., 2009, for a more overview of systemic approach to evaluating human behaviour).

As computational power has grown, so has the popularity of computational modelling of human behaviour. As Thomas and Van Heuven (2005) argues, such models: "force clarity on theories because they require previously vague descriptive notions to be specified sufficiently for implementation to be possible. The implemented model can then serve as a test of the viability of the original theory, via quantitative comparisons of the model's output against empirical data. This is a particular advantage where the implications of a theory's assumptions are difficult to anticipate, for instance, if behaviour relies on complex interactions within the model. Models also allow the generation of new testable hypotheses and permit manipulations that are not possible in normal experimentation" (Thomas and Van Heuven, 2005, p.203). In the case of the sugar factory and indeed the wider dynamic decision-making field, few computational models of cognition have been proposed. The field has largely focused on human performance on the task, without necessarily focusing on modelling such behaviour in computational form to establish the baseline to uncover cognitive processes behind solving of such tasks.

Reinforcement learning has shown significant promise in cognitive science for explaining human behaviour and decision-making in particular. In most general terms, reinforcement learning is "learning what to do so as to maximise a numerical reward signal. The learner is not told which actions to take ... but instead must discover which actions yield the most reward by trying them"(Sutton and Barto, 1998, p.3). Reinforcement learning traces its roots to two distinct fields, which did not necessarily communicate until the advent of modern reinforcement learning appearing in the 1990s: 1) psychology of learning, focused on animal and human learning from the environment, and 2) machine control theory and its attempts to use value functions and dynamic programming to achieve maximum performance, which did not necessarily involve learning (Sutton and Barto, 1998, p.16).

According to Dayan and Niv (2008) the modern field of reinforcement learning "studies the way that natural and artificial systems can learn to predict the consequences of and optimise their behaviour in environments in which actions lead them from one state or situation to the next, and can also lead to rewards and punishments"(Dayan and Niv, 2008, p.185). Such environments can vary greatly and can be found in ethology, economics, psychology, and control theory. Modern applications of reinforcement learning can be found throughout, from the seminal papers published by Google's DeepMind showing their ability to play complex video games (Mnih et al., 2013), to contributing to solving complex real-world tasks, such as autonomous car driving (Bojarski et al., 2017).

Reinforcement learning provides a comprehensive framework that at the very least approximates human behaviour in decision-making in a variety of ways. It is also growing body of research, with established methodologies that provide an interesting, if not an intuitive way to judge human performance against. Therefore, a properly trained reinforcement learning algorithm could conceivably learn to play the sugar factory and would approximate human cognitive processes while doing so. If such algorithms can be trained to perform well on a dynamic decision-making task, it would provide some measure against which human performance can be judged. After all, one could at least say that humans perform worse than a machine approximating human cognitive processes. It is also important to keep such algorithms as generic as possible. If a very specific algorithm is designed to solve a specific task, it could only be said that an algorithm exists that could solve the task, not that human performance is worse than that of an algorithm with a similar cognitive set up.

2.5.2 Q-Learning

Q-learning is a type of reinforcement learning algorithms, belonging to a class of unsupervised machine learning algorithms, where an agent tries to learn the optimal way to behave through the positive and negative feedback it receives from the environment (Watkins, 1989, Watkins and Dayan, 1992, Keerthi and Ravindran, 1994). "It provides agents with the capability of learning to act optimally in Markovian domains by experiencing the consequences of actions, without requiring them to build maps of the domains... By trying all actions in all states repeatedly, it learns which are best overall, judged by long-term discounted reward"(Watkins and Dayan, 1992, p.279). In the simplest case, when success is achieved, an agent is rewarded with positive reinforcement ($+1$), and when negative outcome is achieved negative reinforcement is received (-1). All feedback is stored in an array, which is updated each time new feedback is received. Each time an action is required the array is consulted and each time feedback is received the array and values are updated. Based on this feedback, an agent learns over time what constitutes positive and negative actions and learns to successfully navigate a variety of environments. Q-learning agents have been trained to play complex video games (Mnih et al., 2013), as well as simpler tasks. The complexity of q-learning algorithms can range from a simple two dimensional array to complex neural networks with multiple hidden layers, and beyond (for more information see Sutton and Barto, 1998, Gaskett et al., 1999, Keerthi and Ravindran, 1994, Barto and Singh, 1991).

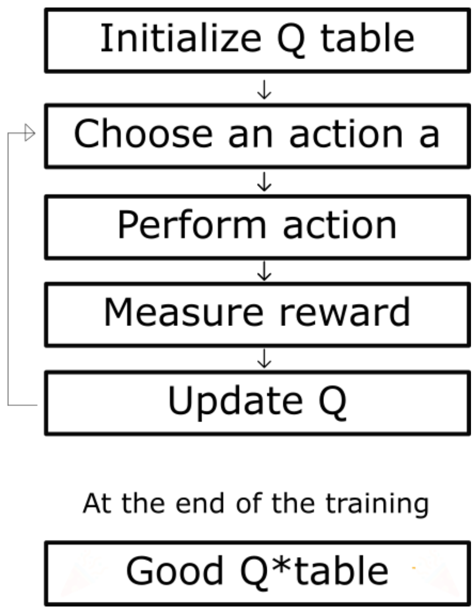


Figure 2.8: Q-learning Process ((Simonini, 2018))

When decision-making space is limited, as is the case with the sugar factory,

$$NewQ(s,a) = Q(s,a) + \alpha[R(s,a) + \gamma\, max\, Q'(s',a') - Q(s,a)]$$

Figure 2.9: Q-learning Update Equation (Simonini, 2018)

all states can be mapped into a single two-dimensional look up table (Simonini, 2018)). This table represents all possible states along with the corresponding rewards learned from previous trials. Figure 2.8 outlines the entire process. Initially, there is an initialised q-table, which is used to store the information. An action is chosen and feedback is received. The table is updated according to the feedback. This is repeated as many times as necessary for the agent to begin to successfully navigate the environment. The Bellman Equation is used to update the q-table (Watkins, 1989) (see 2.9).

The sugar factory simulation, in its original form could be formally specified as a finite Markov Decision Process (MDP), which means that a Q-learning algorithm could be applied to study it (Sutton and Barto, 1998) $\langle \mathcal{S}, \mathcal{A}, P, R \rangle$.

- The state space $\mathcal{S} = \{1, 2, \ldots, 12\}$ where each state $s \in \mathcal{S}$ is a natural number representing how many thousands of tons of sugar output are currently produced by the factory, e.g., $s = 3$ means that the current size of sugar ouput is $3 \times 1,000 = 3,000$ tons.

- The action space $\mathcal{A} = \{1, 2, \ldots, 12\}$ where each action $a \in \mathcal{A}$ is a natural number representing how many hundreds of workers have been assigned to this job, e.g., $a = 4$ means that the current size of work force is $4 \times 100 = 400$ workers.

- The transition kernel $P : \mathcal{S} \times \mathcal{S} \times \mathcal{A} \to [0, 1]$ gives the conditional probability that action a_t in state s_t at time t will lead to state s_{t+1} at time $t + 1$:

$$P_{a_t}(s_t, s_{t+1}) = \Pr[s_{t+1} | s_t, a_t] = \begin{cases} \frac{1}{3} & \text{if } s_{t+1} = 2 \times a_t - s_t, \\ \frac{1}{3} & \text{if } s_{t+1} = 2 \times a_t - s_t + 1, \\ \frac{1}{3} & \text{if } s_{t+1} = 2 \times a_t - s_t - 1, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the next state is completely determined by the current state and the action taken: $s_{t+1} = 2 \times a_t - s_t + r$ where r is a random noise being 0, 1, or -1 with equal probabilities[2].

- The reward function $R : \mathcal{S} \times \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ gives the immediate reward received after transitioning from state s_t to state s_{t+1}, due to action a_t:

$$R_{a_t}(s_t, s_{t+1}) = \begin{cases} 1 & \text{if } s_{t+1} = 9, \\ 1 & \text{if } s_{t+1} = 9 + 1 = 10, \\ 1 & \text{if } s_{t+1} = 9 - 1 = 8, \\ 0 & \text{otherwise.} \end{cases}$$

  In other words, the immediate reward r_{t+1} is completely determined by the next state s_{t+1} only: a sugar output within the range between $8,000$ and $10,000$ tons would be regarded as "on target" and be rewarded 1 point.

This MDP has a finite horizon of $T = 30$, i.e., each "episode" τ of the game consists of 30 time steps. Moreover the initial state is set to $s_1 = 6$, i.e., $6,000$ tons of sugar production. The goal of our MDP problem is to find a deterministic "policy" — a function $\pi : \mathcal{S} \to \mathcal{A}$ that specifies the action $\pi(s)$ which the decision maker should choose when in state s. Under the given policy π, the probability of a particular episode $\tau = (s_1, a_1, \ldots, s_T, a_T)$ can be calculated as

$$\Pr[\tau|\pi] = \prod_{t=1}^{T} P_{\pi(s_t)}(s_t, s_{t+1}) \ , \tag{2.3}$$

while its cumulative reward is given by

$$G_\pi(\tau) = \sum_{t=1}^{T} R_{\pi(s_t)}(s_t, s_{t+1}) \ . \tag{2.4}$$

[2]The real transition kernel P is a bit more complicated, as the sugar output would be clipped at both ends by the fixed lower bound $1,000$ and the fixed upper bound $12,000$.

The optimal policy π^* is the one that maximizes the expected return:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\tau \sim \Pr[\tau|\pi]} G_{\pi}(\tau) \ . \tag{2.5}$$

2.5.2.1 Teaching the agent

As a first step, the sugar factory simulation was recreated in Python. Python was chosen for a number of reasons, namely for its fast prototyping qualities, availability of machine learning libraries and a sizeable number of tutorials on reinforcement learning available on the web.

Once the full simulation was recreated in Python, a Q-learning reinforcement-learning agent was built that could play the sugar factory and most importantly, learn and improve over time. The initial agent was powered by an epsilon-greedy Q-learning algorithm. All code was built from scratch and no outside libraries were used for the sugar factory and the reinforcement learning agent.

Since the sugar factory has a rather simple decision-making space of 12 ((1-12) possible inputs (actions) and 12 (1000-12000) possible outputs (states), a 2-dimensional vector could easily hold all of the information the agent needs to learn to navigate this space. Simple Q-learning algorithms have an added bonus of storing all of its information in a table with the corresponding weights, making it human-readable and consequently interpretable.

For the purposes of the simulation a number of decisions and assumptions was made in creating the agent:

- Type of learning - a simple q-learning table was used to revise and store the weights (feedback). One axis represented the range of potential outputs, while the second axis represented the sample space of possible inputs.

- Feedback - although a number of reward mechanisms was considered, it was found that passing -1 when output was not 9000 (i.e. not on target) and 1 for when the agent hits the target of 9000 tons of sugar in the previous turn provided for the best results.

- Exploration vs exploitation - a q-learning algorithm relies on some degree of randomness to make a decision between 'exploring' the environment, i.e. picking a random value and getting more information about the environment and 'exploiting' it, or picking the value from the look-up table based on previous feedback. There are a number of different 'exploration vs exploitation' mechanisms available. Ultimately, an epsilon-greedy way of determining this trade-off was picked, which makes the agent explore more initially, and decrease exploration with time. Significantly, this is similar to the training instructions participants received in the lab experiments. The fundamental tension between exploration and exploitation is inherent in decision-making (March, 2008, Laureiro-Martinez et al., 2015, Hills et al., 2015, Gureckis and Love, 2009a). Without some degree of exploration, an optimum strategy cannot be arrived at. On the other hand, at some point, in order to achieve high scores, existing knowledge should be used. Q-learning, and indeed, many of the machine learning algorithms rely on initial exploration to build the action-state table, but also need to be prevented from excessive exploration since exploration in this context is random and continuous random actions cannot lead to optimal outcomes.

- Governing equation - the algorithm was governed by the following equation:
  $Q_{s,a} = Q_{s,a} + lr * (r + y * np.max(Q_{s1,:}) - Q_{s,a})$

  where **s** is number of workers, **a** is output, **lr** is learning rate, **r** is the reward and **y** is the second learning rate.

2.5.2.2 Running the simulation

With the agent in place, it started to repeatedly play the sugar factory. The starting conditions of the simulation were the same as those for human players in the previous studies.

In order to make learning and testing easier the element of randomness was removed from the sugar factory equation. It was important to make the initial simulation as easy as possible for the agent to learn, and to increase the difficulty with time once the initial training parameters were working. The first results are displayed in Figure 2.10. The y-axis represents the total score, while x-axis represents the number of simulated runs. It is apparent from the figure that the agent learns for the first 500 or so trials, and subsequently is able to reach the top score. However, due to the initial exploration parameters it continues to oscillate between high and low scores for each trial.

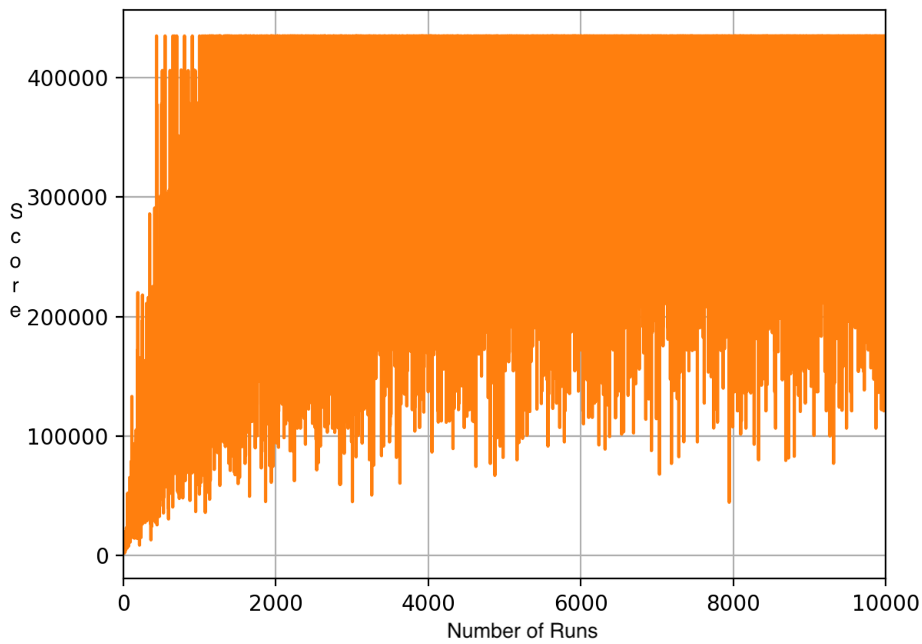


Figure 2.10: Score achieved on each run over 10,000 trials by the high greedy exploration algorithm.

In the next step, the exploration parameter was turned down, which made the agent explore less and less in subsequent simulations. The result is displayed in Figure 2.11, where after about 5,000 trials the agent is able to consistently reach the highest scores of the simulation. It is worth noting that this score represents the maximum limit of the simulation (420,000). Therefore, it can be concluded that the agent learns to reliably and successfully conduct the simulation using the q-table approach, provided randomness is turned off after approximately 5,000 trials.

With the agent able to consistently and successfully play the simulation, the next step was to better understand the resulting q-table. The table represents a lookup table with all of the possible current states (output) and all of the possible

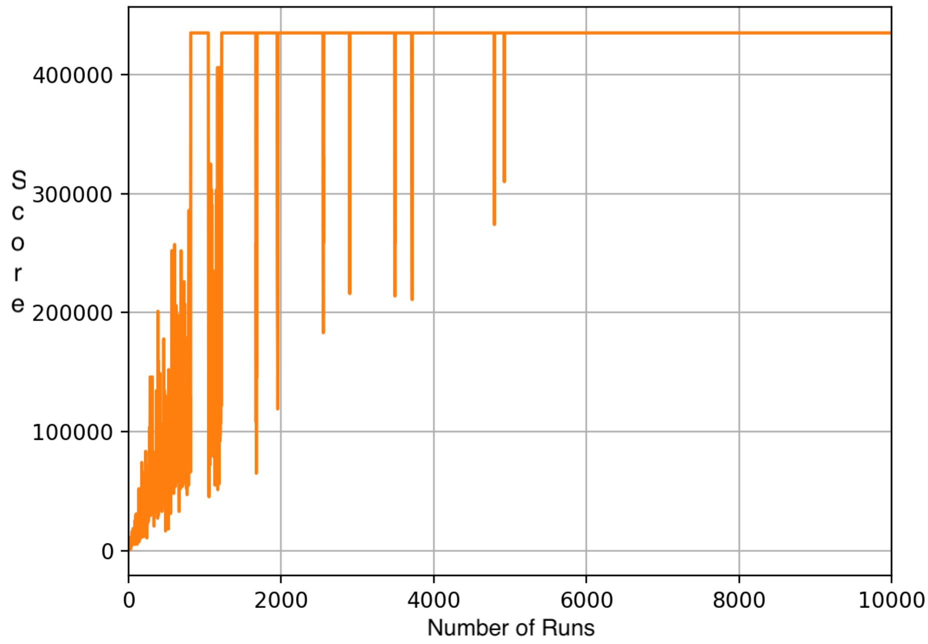


Figure 2.11: Score achieved on each run over 10,000 trials by the low greedy exploration algorithm.

decisions (workers to be assigned). Each pair has a corresponding expected benefit associated with the action. Figure 2.12 shows the full version of the q-table after 10,000 simulations.

The agent learns that the optimal combination for achieving the highest score on the simulation is: 1,5,9,9,9...(repeated). Given that the starting point of every simulation is 6, the most optimal choice is to select the lowest possible number of workers to 'reset' the factory. After output is reduced, the simulation becomes solvable by picking 5 and then 9 workers. This patterns allows the agent to achieve the maximum score.

2.5.2.3 Adding noise

It is clear that an agent can play the sugar factory successfully without noise. Next the agent played the same version of the sugar factory that human participants had to do (i.e. with noise). Figure 2.13 shows the results of the agent after playing the original sugar factory for over 10,000 simulations with randomness turned on.

There are a few notable findings here. Firstly, the agent was not able to achieve

	Number of Workers											
Output	1	2	3	4	5	6	7	8	9	10	11	12
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0000	-0.8885	-0.8889	-0.8889	-0.8889	1.1111	-0.8889	-1.0887	-1.0882	-1.0876	-1.0887	-1.0888
2	0.0000	-0.5091	-0.5350	-0.5349	-0.5468	-0.5372	-0.5336	-0.5599	-0.6144	-0.5076	-0.5657	-0.5541
3	0.0000	-0.6953	-0.7281	-0.5852	-0.7786	-0.3749	1.1111	-0.4855	-0.8837	-0.4973	-0.6990	-0.5561
4	-0.5453	-0.5451	-0.5100	-0.5377	-0.5304	-0.5350	-0.5382	-0.5266	-0.5604	-0.5641	-0.5525	-0.5659
5	0.0000	-0.6382	-0.6109	-0.5522	-0.5260	-0.6504	-0.4728	1.1111	-0.6578	-0.8107	-0.8841	-1.0058
6	-0.8889	-0.8889	-0.8889	-0.8889	-0.9644	-0.9646	-0.9241	-0.9356	-0.9309	-0.9957	-0.9985	-0.9655
7	0.0000	-0.7723	-0.5790	-0.4644	-0.4165	-0.6258	-0.6921	-0.6527	1.1111	-0.6987	-0.7738	-0.8552
8	-0.5448	-0.5529	-0.5174	-0.5447	-0.5508	-0.5356	-0.5812	-0.5500	-0.5345	-0.5331	-0.5574	-0.5632
9	0.0000	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	1.1111	-0.8889	-1.0889
10	-0.4663	-0.4746	-0.4337	-0.4657	-0.4700	-0.4691	-0.5381	-0.4759	-0.4941	-0.5390	-0.4801	-0.4377
11	0.0000	-0.7070	-0.6126	-0.7063	-0.7266	-0.7408	-0.6855	-0.6987	-0.7640	-0.6699	1.1111	-0.6183
12	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-0.8889	-1.0120	-1.0189	-1.0353	-1.0297	-0.9950

Figure 2.12: Averaged Q-Table for simulation with no randomness

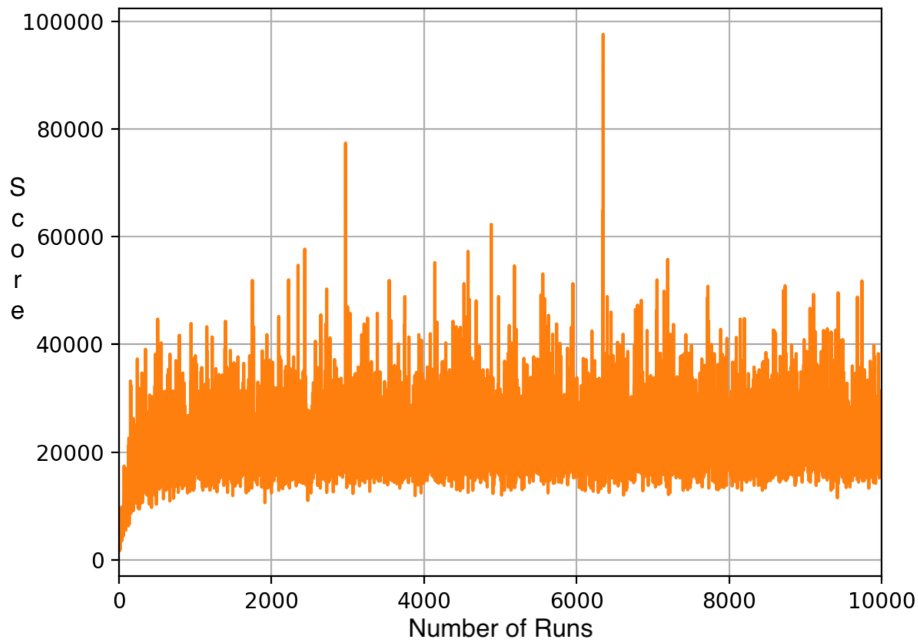


Figure 2.13: Results of the algorithm conducting the simulation with randomness turned on.

the same high score once noise was turned on. Secondly, agent's performance appears to *closely approximate that of human players.* There is an initial learning curve, followed by oscillations between the floor of about 10,000 points and the ceiling of approximately 30,000, with certain simulations achieving scores higher than that; and with the maximum being about 100,000. There were a few participants who reached the score above 40,000, although the average was closer to 14,000. While no human player achieved any scores above 50,000, it is entirely conceivable that given enough simulations it may be possible to see this level of one-off performance. Importantly, the lower and upper bounds appear to mirror human performance, suggesting that the agent approximates quite well the actual human performance, allowing us to explore human performance and boundaries through the simulation. Finally and perhaps most importantly, this result demonstrates that human performance is not necessarily 'poor'. Indeed, the agent's performance suggests that there is some learning that can be done initially, however, the impact of randomness is such that it renders *the simulation impossible to complete at the same levels as without randomness*, to any degree of repeatability in the long run.

2.5.3 Discussion

It is possible to conclude two things from the agent's performance on the simulation. Firstly, the optimal strategy is to reset the simulation by assigning the least amount of workers possible. Secondly, the randomness factor in the simulation makes control extremely difficult, not just for human subjects, but also for an agent. This appears to be the main source of difficulty in the simulation. It has long been established that the difficulty of the simulation is "due to the lag term [where] two separate, interdependent inputs are required at times t and t +1 to reach steady-state production (Gibson et al., 1997, p.4). However, the agent's inability to control the factory with randomness suggests that it is actually the factor originally placed to "to ensure that subjects would exercise continuous control"(Berry and Broadbent, 1984, p.212), which makes the simulation impossible to complete with any degree of competence.

The agent appears to mimic the initial learning that is observed in human players, but after the minimum score is reached, the maximum score tends to be the function of randomness much more than skill, showing that consistent control is all but impossible. A simple reinforcement algorithm performs as well as human participants in the lab, suggesting that human performance is actually quite good, considering that it takes an agent almost 500 tries to get to the same level as a human player after 10 or so tries. Indeed, when the two are compared, there is no real difference. Human players are quite adept at learning the initial strategies required to control the factory and then spend most of the time combating randomness and its effects on control.

Berry and Broadbent's original findings, and subsequent analysis by the broader decision-making community suggesting that human performance is sub-par on these tasks can be challenged. It is not human performance that should be questioned, but rather the task itself and in particular the addition of randomness as an element of the decision-making paradigm. It is therefore not surprising that participants only learned *implicitly* since randomness cannot be learned. Neither practice, nor preparation, or strategies can have a meaningful impact in an environment where randomness plays such an important role.

This has important implications for the field of dynamic decision-making. It is imperative that dynamic simulations are fundamentally solvable, or *learnable* by humans. However, these simulations, especially the ones with multiple feedback loops are designed to be non-linear and potentially impossible to solve from the onset.

2.6 Experiment 5: Reconsidering the Sugar Factory

In light of the new findings, there appears to be a need to test the new hypothesis, i.e. that if randomness is removed, human participants can also learn to control the sugar factory. To test the hypothesis, randomness was removed from the online version of the sugar factory. It was believed that without randomness the simulation would become 'solvable' and human performance would dramatically improve and become consistent over time, in line with our findings from the agent-based simulations.

2.6.1 Method

A sugar factory identical to the one described in 'Experiment 4' was set up, but with the minor tweak of randomness factor being removed from the equation.

2.6.1.1 Participants

Participants were University of London students (n=24) who participated in the study for credit. They had to register online to receive a link to the simulation. They were asked to complete each simulation at least 10 times – the base number for improving performance derived from previous studies – after which they were granted credit. Three days after registering for the study each participant received a reminder email to do at least 10 trials and suggesting that there were many participants who achieved quite high scores on the simulation to encourage participation.

2.6.1.2 Materials

The web version of the simulation was based on the same set of instructions and the underlying equation as in Experiment 4. All instructions were posted online and participants were given an opportunity to play a tutorial before doing the actual simulation.

2.6.1.3 Procedure

Upon registering and creating an account, participants were able to access the simulation screen. After completing a short tutorial they were asked to play the simulation for 30 turns. After each set of 30 turns they were shown the final score and asked to play again. There was no limit to the number of times each participant could play. If participants did not complete at least 10 trials in a 3 day period, they were sent a reminder email asking them to play at least 10 times. They were also advised that the simulation was not as difficult as it looks and that many participants had achieved very high scores of above 300,000.

2.6.2 Results

In total, 24 participants completed the simulation. On average, participants played 14 times before quitting. Two participants were removed because they played only twice and neither time was done to completion. There were 22 participants whose results are included in the analysis below. Table 2.7 summarizes the overall performance results.

In general, participants could be divided into two groups based on their performance and persistence. Those who played less than 13 times and achieved an average highest score of just under 10,000, and those there who played more than 13 times and achieved an average highest score of over 200,000. Overall, it took participants about 14 tries to achieve a score above 100,000.

Table 2.7: Table showing the results of performance for participants doing the online version of the sugar factory with randomness turned off. It is split between participants who did more, or less than 13 trials, as this represents the mean point at which mastery substantially increased

Number of Tries	Average Highest Score	Percentage of turns on target
<13	9,909	8.6 %
>13	256,000	47 %

2.6.2.1 Online Performance Comparison

Figures 2.14 and 2.15 compare performance between the group of participants who did the online simulation with and without randomness. From Figure 2.14 it is possible to see that there is very little difference in performance during the first 10 tries between the two groups. That changes immensely when we extend this analysis to 40 tries in Figure 2.15. It is quite clear that participants that do not have to deal with randomness learn to control the simulation in a much more effective way than their peers in the 'randomness' group.

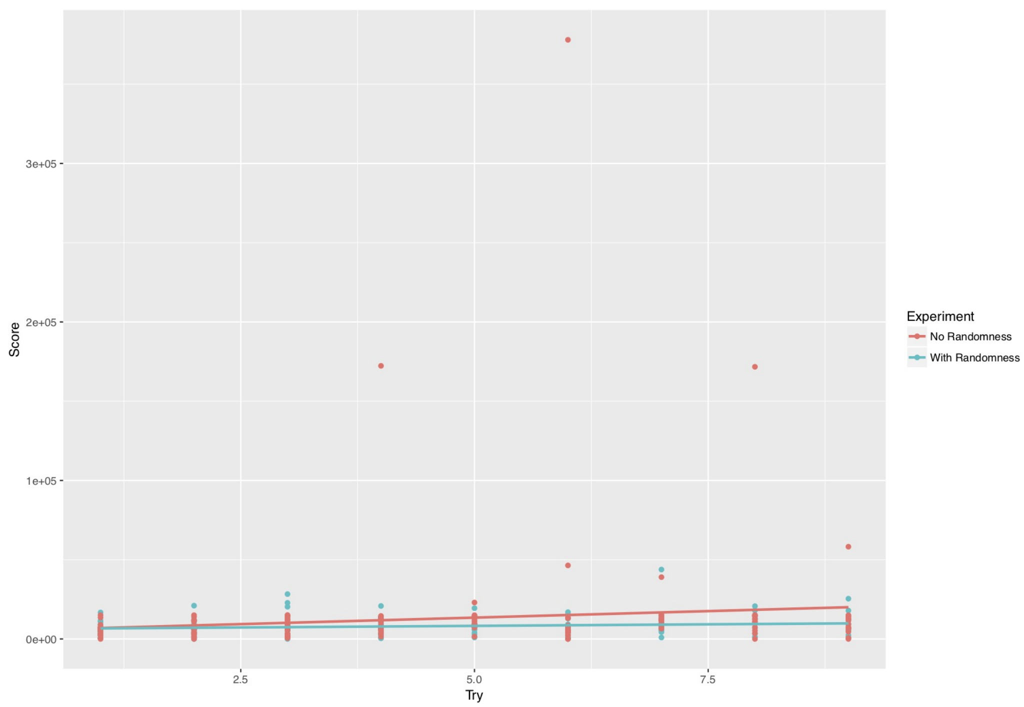


Figure 2.14: Online Individual Performance With and Without Randomness.

Based on these findings, randomness does play the greatest impact on performance, in-line with the expectations derived from the simulations. Participants who did not have to deal with randomness in their simulation learned to effectively control it after approximately 10 to 14 trials. It took the simulated agent on average of approximately 500 trials to reach the highest score. This, of course, is *significantly faster* than the reinforcement algorithm used in the simulations. However, it is worth highlighting that the reinforcement algorithm used in the simulation was quite generic. It was not specifically designed with the sugar factory problem in mind. The algorithm did not include complex neural networks, or machine learning algorithms that would have been *biased* towards a type of problem that sugar factory was. It is entirely conceivable that a more specialised algorithm may have been able to learn to do the simulation much faster, or account for the randomness in the simulation better. However, a specialised algorithm would not have represented participants' experience on the task. This would have simply shown that the sugar factory is solvable, but not necessarily through generic reinforcement learning.

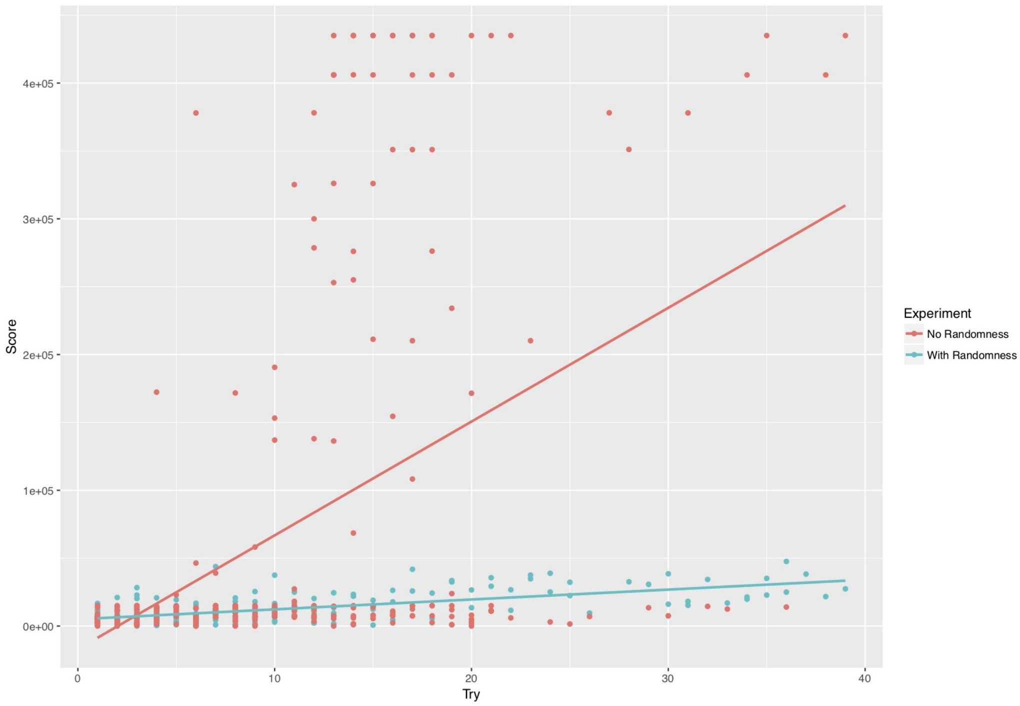


Figure 2.15: Performance on Initial Trials With and Without Randomness.

The algorithm was chosen specifically due to its generality. Much like a human participant who encountered the task for the first time and had to apply his previous knowledge and current skills to solve it, the algorithm had to be generic enough to be able to learn the task in a 'naive' way. It is, therefore, less important that humans outperformed the algorithm by learning to control the simulation about 50 times faster, than the fact that using reinforcement algorithm led to the formulation of a hypothesis regarding the difficult point of the simulation and allowed to prove it in an experimental fashion.

The argument here is that in the future machine learning algorithms should be used more often in an attempt to evaluate the solvability and the difficulty of dynamic decision-making games before human participants are subjected to them. This would allow for a more objective evaluation of human performance.

2.7 Overall Discussion

A new way is required to evaluate human performance in this domain. It is not enough to simply put a seemingly simple dynamic decision making simulation in front of a human participant and then declare the limit of human cognition when failure is observed. Without an objective yardstick against which to measure performance, simple failure cannot be treated as proof of cognitive limitation, especially where ability to achieve success is not obvious, or indeed possible.

Application of artificial agents, such as the one presented here could be such a measure. There are two possible ways to apply reinforcement learning agents to evaluate dynamic decision-making tasks. In one instance, such agents could be used to show that simulations are indeed 'solvable'. On the other hand, they can be used to evaluate the cognitive requirements of the simulation (Thomas and Van Heuven, 2005, Gureckis and Love, 2009a). Application of artificial agents can further provide hypothetical limits to what is possible, as per Howes et al. (2009, p.721).

Given a rather large number of machine learning approaches and algorithms available, we suggest a few points of guidance to select the most appropriate one to evaluate a simulation.

- Simple - starting with the simplest algorithm first. More generic algorithms are more likely to reflect human interaction with the simulation and have an added benefit of being interpretable;
- Interpretable - Ideally, algorithm's output should be interpretable. Much like an autopsy, being able to 'dissect' the final weights in a model may hold cues to how the algorithm learned to deal with a particular task and allow the researcher to alter it accordingly;
- Working - perhaps the most important criteria is that the algorithm should be capable of solving the simulation in the first instance.

2.7.1 Impact of Feedback Loops

The general nature of the task allows us to draw some conclusions regarding a general impact of feedback on learning and performance. This Chapter shows that artificial randomness can prevent learning. It is important to also note that delayed nature by itself does not necessarily prevent the achievement of mastery over a task. Experiments contained in this Chapter show that individuals can and do learn from delayed feedback, but when such feedback is combined with randomness it becomes much more difficult to control. Other factors such as cover stories and prior beliefs did not appear to have a significant impact on learning. This suggests that dynamical tasks can take many forms and still be free from previous participant-specific bias.

2.8 Conclusion

The chapter started with a series of replication studies designed to understand whether the original sugar factory results remain valid today. Given that Berry and Broadbent (1984) have introduced the original sugar factory over 30 years ago, it was important to understand if the proliferation of computer games, many of which are quite complex, could have had a positive impact on the human sugar factory performance. Unfortunately, despite societal changes, there was no evidence to support this hypothesis. Participants continue to struggle with controlling the sugar factory simulation. There was some practice effect observed, whereby performance did improve in the first few trials, but participants continued to under-perform, reaching the target only about a third of the time. This effect was consistent with the existing studies utilising the sugar factory and the field of dynamic decision-making in general (Berry, 1991, Gibson et al., 1997, Gonzalez et al., 2003, 2005, Osman, 2010).

The next step was to try and improve performance through training. Various training materials were distributed to the participants in an attempt to get them to use the 'scientific method' or another way of *exploring* the various strategies in the hope of finding a single strategy that would lead to better control and higher performance on the simulation. These materials were consistent with the various hypothesises on task control and designed to overcome the heuristics demonstrated in the scholarship Brehmer (1992), Gonzalez et al. (2003), Osman (2010). Unfortunately, these methods did not result in any noticeable improvement in performance. Participants continued to struggle to control the sugar factory beyond the initial practice effect.

The next hypothesis that was tested concerned more explicit control over the factory. In the original sugar factory setup, participants could only enter numbers in the range of one to 12, which were multiplied to get the total number of workers. It was believed that providing participants with the greater ability to control the number of workers would lead to better control. In this case again, participants improved their performance as a result of practice, but continued to struggle to master the simulation. Furthermore, there did not appear to be any impact of the training materials, or indeed, skill transfer between the two version of the simulation.

Next, the cover story was changed, moving away from a hypothetical sugar factory and towards a more realistic climate simulation. It was believed that a more explicit and realistic cover story would provide participants with an ability to apply prior knowledge and experience to perform better on the simulation. This also allowed for measurement of impact of prior beliefs on the conduct of the simulation and performance. Here again, there was a negative result. There did not appear to be a measurable impact of prior beliefs on the performance in the simulation and there did not appear to be a statistically significant impact of skill transfer onto the sugar factory.

At this point, there did not appear to be a significant impact of prior beliefs, explicit training, more realistic cover story and simulation, or the ability to control the exact assignment of workers on performance. Continued poor performance, coupled with the inability to articulate learning, has usually led to the conclusion that such tasks are simply too difficult for human beings to do (Gibson et al., 1997, Gonzalez et al., 2003).

However, research presented in this chapter challenges this notion. Previous authors relied on data obtained from the simulations to judge human behaviour and did not necessarily use the modelling methodology suggested by Howes et al. (2009) of using computational models to build a wider framework for understanding human behaviour. Following research by Thomas and Van Heuven (2005), Gureckis and Love (2009a) of using reinforcement learning algorithms to model human behaviour, an artificial agent was created to do the sugar factory simulation.

The simulated agent became quite skilled at conducting the simulation provided randomness was turned off. However, once randomness was turned back on, the agent struggled to get the same high scores. Its performance became quite comparable to those of human players. In fact, once human participants had a chance to play the simulation without randomness, their performance markedly improved and they were able to control the sugar factory with extreme precision. Most human players learned to conduct the simulation after approximately 20 tries. This compares quite favourably to the reinforcement learning agent, which took on average 500 tries to reach the same scores. It can be concluded that human performance on dynamic decision-making tasks can become quite good, and randomness appears to be the variable that is most difficult to overcome. Human participants were able to figure out the delayed nature of feedback, which once was hypothesised as the difficult part of the simulation, once randomness was turned off. This suggests that it is in fact the randomness inherent in many of the dynamic decision-making games that makes them so difficult to solve.

Implementation and use of a simulated agent represents an important step in understanding the limitations of human dynamic system control. The general discussion section contains a set of guidelines that could be used to construct such agents and measure human performance. Hopefully, the use of simulated agents will become more widespread and hypothesis on human limitations will gain an important objective standard against which such performance could be measured.

These findings also challenge the common notion that human performance is poor (Gibson et al., 1997, Gonzalez et al., 2003, Osman, 2010). When compared against the artificial agent, human performance is on par. Furthermore, research in this chapter suggests that randomness embedded in the original simulation is the source of difficultly, rather than the delayed nature of feedback as previously argued by Berry and Broadbent (1984, 1988), Gibson et al. (1997).

In order to further generalise these findings, it would be a worth-while endeavour to take other dynamic decision-making tasks and train a reinforcement learning agent to play them to determine their difficulty to compare human performance against. Additionally, given the difficulty presented by randomness, it would also be advisable to better understand the impact of computer-generated randomness in this domain. This suggests two separate strands of research activity: 1) application of AI-powered algorithms to establish performance benchmark, and 2) removal of randomness from the simulations to determine the true difficulty of the given system.

Chapter 3

Part II - Individual Belief Revision in Groups

Chapter 2 focused on individual performance and the impact of feedback loops that arise as a result of the interaction between an individual and a physical system. However, there is another area where feedback loops play an important role: social interactions. Here the interaction is between an individual and a wider group, which creates the very feedback system. Unlike systems in the previous chapter, the feedback in these systems is fully a function of the individuals, with no pre-programmed 'system' with its own governing logic and equations. The final quadrant of the decision-making tasks discussed in the introductory chapter are tasks social where the setting is dynamic, constantly changing as a result of user interactions within it. This environment reacts to individual user feedback and each user reacts to the feedback received from the environment, thereby creating feedback loops.

Part II of this thesis is interested in social systems and the impact of social interactions on individual belief revision and group. Chapter 3 focuses on the strategies that individuals use in group revision tasks. The goal is to better understand and model individual belief revision behaviour. The second goal is to better understand what strategies individuals use when revising their beliefs. These strategies give rise to the feedback system itself and ultimately have an impact on the feedback each participant receives from the system. This interaction at its simplest form is expressed in Figure 3.1. An individual has an impact on the group answer (and accuracy), which then has an impact on the individual's answer in the future. The more individuals are involved the more the group answer becomes a function of multiple inputs and the more complex the system becomes (see Figure 3.2). Therefore, there are two things that need to be understood: 1) what do individuals do (Chapter 3); and 2) what impact it has on accuracy, as accuracy is one of the main expressions of group interaction (Chapter 4).

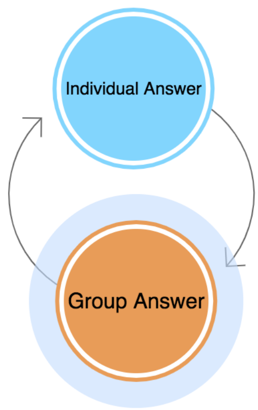


Figure 3.1: Simple Group Interaction System.

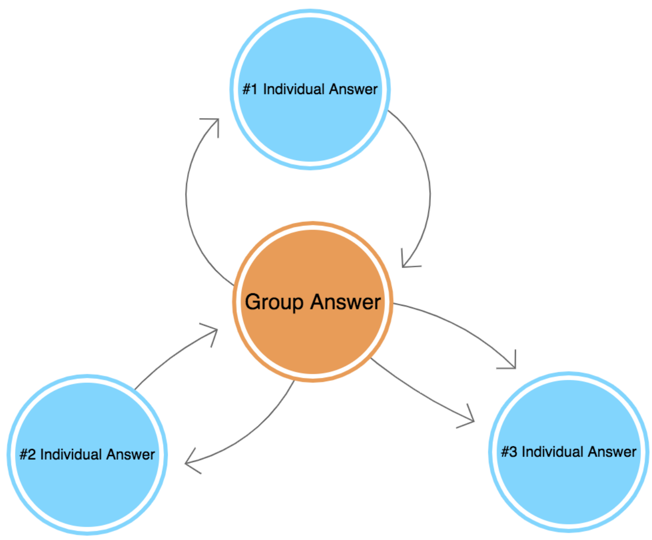


Figure 3.2: Complex Group Interaction System.

3.1 Introduction

Social psychology has a long standing interest in group performance as it relates to individual competence. This line of research traces its roots to the early 20th century, with the famed Galton (1907) 'wisdom of the crowd' experiments. Indeed, understanding how 'wisdom' arises in groups has been an important research area for social psychologists for a long time (for review see Lorge and Brenner, 1958, Hill, 1982, Gigone and Hastie, 1997).

This line of research has received renewed relevance and prominence in light of recent developments in network science, which have shown, both through real world data analysis and through simulation, how individual behaviour is shaped by the structure of social networks (see Jackson, 2010, Becker et al., 2017, Centola, 2018). Understanding whom one knows is in many cases the single best predictor of what they are likely to do (Pentland, 2014). In keeping with this, philosophers concerned with the nature of knowledge have become increasingly interested in social epistemology (Goldman, 1999). There have even been studies revisiting the famed Galton experiments, which are now over a century old, in order to understand networks dynamics and their impact on the overall wisdom of the crowds (Becker et al., 2017).

As the world becomes increasingly interconnected – through the advent of the internet and the relatively recent rise of social networks – individual cognition cannot be fully understood without considering the impact of the networks one belongs to. As Bakshy et al. (2015) show:

> "the composition of our social networks is the most important factor limiting the mix of content encountered in social media. The way that sharing occurs within these networks is not symmetric – liberals tend to be connected to fewer friends who share conservative content than conservatives (who tend to be linked to more friends who share liberal content)"(Bakshy et al., 2015, p.2).

At the same time, full understanding of social network influence on individuals is impossible without considering how individuals respond to information and cues provided by others. Is simply being exposed to the information enough to change opinion? Is it the amount of information that matters? Is it the strength of conviction? Could it be that the network structure itself plays a role? For example, recent simulations suggest that network structure plays an important role in contagion and diffusion (Kretzschmar and Morris, 1996, Watts, 1999, Lazer and Friedman, 2007, Jönsson et al., 2015, Hahn et al., 2018) (for an overview see Jackson, 2010, Centola, 2018, Doer et al., 2012).

However, much of the modelling research relies upon mathematical formulas borrowed from network science rather than empirical studies to model individual behaviour (Hegselmann and Krause, 2002, Hu, 2017). These simulations rest upon empirically untested assumptions about individual strategies for information incorporation and belief revision. Unless these assumptions match, at least crudely, behaviours of the actual people, the generalisability of these models remains limited. At the same time, the cognitive science literature on judgement and decision-making contains remarkably little empirical work on how individuals *actually* revise their beliefs over time in light of information they receive from those they are connected to. The belief revision literature tends to focus on group dynamics (Jönsson et al., 2015, Becker et al., 2017), while individual belief revision and strategies in this context remain largely unexplored. Given that group revision is a combination of individual revisions, this underrepresented area may provide important clues for understanding group dynamics as well.

When one considers belief revision literature, and human cognition modelling in particular, Bayesian inference and updating comes to mind (Dubois and Prade, 1997). As far back as 1980, basic modes of individual belief revision have been studied and modelled using the Bayes rule (Levi, 1983). More recently Bayes' theorem has been used to understand probabilistic reasoning (Pearl and Russell, 2011), and belief revision in light of new evidence (Ma et al., 2010). Much of the work to understand belief revision using Bayesian modelling involves the creation of mathematical models to approximate cognitive models (Ma et al., 2010, Pearl and Russell, 2011). The question this approach seeks to answer revolves around the understanding of the cognitive structures of decision-making and their approximation of the actual cognitive processes.

However, the advice and opinion dynamics literature tends to take a different approach to understanding human belief revision and aggregation. Especially in the context of multiple opinions being shared in a group setting, the research tends to be more experimental, relying on observed behaviour and modelling thereof. More complex computational methods based on Bayes' theorem are less common in psychological research on opinion aggregation and are not part of this thesis (Yaniv, 2004b). Furthermore, this thesis seeks to understand and model *actual* human behaviour and the impact of others on it.

The one set of work in this area comes from the studies by Yaniv and colleagues (Yaniv, 2004a, Yaniv and Milyavsky, 2007) (see also Yaniv, 2004b, for a review in the context of the literature on advice). In these studies, participants were asked general knowledge questions such as "in what year was the Suez Canal first opened for use?" (Yaniv and Milyavsky, 2007). Participants provided an initial "best estimate" and then received 'advice' from several advisors (e.g., "the best estimate of advisor #33 was 1905"). Participants were then asked to provide the final "best estimate". The main finding of these studies was that participants overweighed their own opinion over to that of (unknown) others: when participants revised their estimates they weigh their own answer more strongly than they did the answers of others. This is consistent with the results from other advice paradigms such as cue-learning (Harvey and Fischer, 1997) or forecasting (Lim and O'Connor, 1995). In contrast to these studies, however, Yaniv and Milyavsky also examined the effects of receiving multiple pieces of evidence, each from a different agent. In their 2007 study, each participant received an estimate from 2, 4 or 8 advisors (in actual fact these "advisors" estimates were drawn randomly from a pool of initial estimates provided by participants in an earlier study). Participants' accuracy improved in all conditions as a result of incorporating outside 'advice', but the benefit of receiving additional estimates seemed to decrease with their number. In the same study Yaniv and Milyavsky (2007) also examined a range of possible models of participant strategy, finding evidence for models that discounted opinions that were too distant from participants' own initial guesses. In general, participants seemed sensitive both to their own degree of knowledge and to how far other opinions were from their own.

While these studies make an important start in seeking to pin down, on a process level, individual belief revision in light of information from others, much work remains to be done. Given the limited scope of the earlier research, it would also be desirable to extend the paradigm in a number of ways. For instance, source reliability and individual confidence seem two obvious lines of inquiry, as is understanding the impact of prior knowledge on revision.

This seems particularly important because one plausible reason for the greater weight placed on participants' own judgements could lie in considerations of source reliability. Participants know a fair bit about themselves and nothing about other sources in these studies, including whether they even exist other than as an experimental manipulation. This is especially important given that information received from less reliable sources normatively *should* (and empirically does) have less impact on beliefs (see also Bovens & Hartmann, 2002; Bovens & Hartmann, 2003; Hahn, Harris & Corner, 2009). It would therefore be interesting to conduct such a study in a context where it is clear that advisors are human beings who are genuinely engaged in the task at hand, ideally in the same room, and incentivised to perform well. Moreover, Yaniv and Milyavsky examined only one cycle of advice and revision, while many social contexts involve repeated exchanges. These exchanges create dynamic and evolving interactions, where initial opinions change beliefs and opinions of others and these influence us in return. Consequently, an experiment focused on repeated revisions over time, with real people, would yield a better insight into belief revision dynamics.

3.2 Belief Revision

An experimental investigation along these lines needs a design which is as experimentally controlled as Yaniv and Milyavsky's study, yet enables participants to see that advice is coming from others. In particular, the experimental context should not introduce a plethora of other factors that may impact the perceived reliability of various information sources. An experimental context in which a group of participants is present in the same room at the same time, but interact with each other only through a computer terminal in front of them, seems most appropriate. After providing their initial answers, each participant would see the answers of some of their peers on screen.

One such dataset comes from a study that was conducted by researchers at Lund University, Sweden. Its primary aim was to examine the impact of network topology (structure) on the accuracy of participants' beliefs, both individually and collectively. The results with respect to individual and collective accuracy has been published elsewhere (Jönsson et al., 2015). The interest here, however, is in trying to understand individual belief revision strategies. The original study involved gathering individual trial-by-trial changes in participants' answers, which allows for a detailed analysis of individual participant behaviour. The analysis performed on the same dataset presented here has a different focus. The focus is on determining what strategies participants used in the Lund study to revise their answers over a number of rounds. Various models are tested against the actual performance to determine the goodness of fit of such models.

3.2.1 The Lund Study

In the Jönsson et al. (2015) study participants signed up for one of four sessions, forming in groups of 9, 9, 7 and 13 participants respectively. The testing conditions, procedure, materials and instructions were the same across the four groups. There were 38 undergraduate, University of Lund students (15 male and 23 female) who participated in the study. In addition to a flat payment for participation (100 SEK), in each group participant with the most accurate answers received a reward of 300 SEK (Jönsson et al., 2015). In total, there were 7 answers collected for each participant for 20 questions. Therefore, the dataset contained 5,320 individual answers.

During the experiment, each member of a group was seated at a computer and given two sheets of paper with instructions. When everyone in a group stated that they had understood the instructions, a NetLogo-based program was used to send out questions to all of the participants. There was an initial warm-up question, followed by ten questions. Each question was repeated eight times over the course of eight consecutive rounds. During the first round, each participant answered independently. In the subsequent seven rounds, each participant received information about what (some) other participants had answered on the previous round and asked to revise their answer.

For each question, either a random small world (where each participant would be connected to a few other participants) or a complete network (where participants saw everyone's answers) was generated to connect participants, such that it contained one node corresponding to each participant. Participants could then, on the subsequent rounds, see the answers of the participants corresponding to the nodes they were immediately connected to. Participants did not, however, know which answers belong to which person present in the room. Each participant answered the same question 7 times.

Questions were drawn from a set of 21 questions derived from reports by Statistics Sweden ('Statistiska Centralbyrån') and included questions on Swedish demographics, agriculture and geography. Except for the warm-up question, questions were presented in a random order. All questions asked participants to provide a percentage estimate. Example questions include: "What percent of the Swedes are between 15-24 years?" and "What percent of Sweden is covered by agricultural land?" All questions share a common scale and all groups received the same warm up question followed by 20 questions.

The results can be summarised as follows:

> clear evidence of effects of network structure on both collective and individual accuracy, whereby less densely connected groups outperform groups where every members judgements are accessible to all. However, in all groups we find clear evidence against the claim ... that access to others judgements is detrimental to performance because of the reduction in diversity that it brings, as we found that individuals average accuracy rose in response to information about others esti-

> mates. Moreover, in the less densely connected networks even collective competence rose as a result of information exchange. Though collective competence (wisdom of the crowds) is necessarily a function of both individual competence and group diversity, less fully connected groups may increase individual accuracy sufficiently to offset the decrease in diversity information exchange brings about (Jönsson et al., 2015, p.25).

A large data set was collected from the study which reflected individual answers for each round of revision. This dataset provides insights into individual accuracy and behaviour. The data is used to analyse and understand *individual* behaviour as it relates to the revision and accuracy of belief revision over time.

3.2.2 Analysis

Of the thousands of collected responses four scores were eliminated as likely errors. Three were zero-answers that probably resulted from the participant accidentally clicking submit before choosing an estimate (which was done on a sliding bar next to the submit-button); the fourth was a very large number in a sequence of identical low numbers which was also probably due to a mis-click. The rest of the results were left intact and analysed.

3.2.2.1 Rounds

Rounds represent discrete time periods that formed the basis of our analysis. Participants entered their initial estimate independently, and were shown the answers of others in the subsequent rounds. Therefore, belief revision as a result of increased information could be observed in rounds two to eight. During the seven rounds of change, participants had an opportunity to enter revised answers, observe others revise their answers and so on.

3.2.2.2 Change

Two measures were used to determine the magnitude of change exhibited by each participant: absolute change and percent change. Absolute change was defined as the absolute difference between the initial and revised answers, while percent change refers to the percentage of the change in the subsequent round compared to the previous answer. Given that different questions had different true answers in order to be able to compare revisions across questions and round, percentage change was calculated for each answer. For example, if the initial answer was 5 and the revised answer was 10, this would be an absolute answer change of 5, but a percentage change of 100%.

The histogram in Figure 3.3 shows that the most prevalent behaviour was to not change at all. The histogram is the total count of all revisions by round. In the rounds where changes were made, it was mostly by 1 or 2 points. This was true across all groups. As shown in Table 3.1, the mean absolute change was between 1.1 and 2.3, depending on the group.

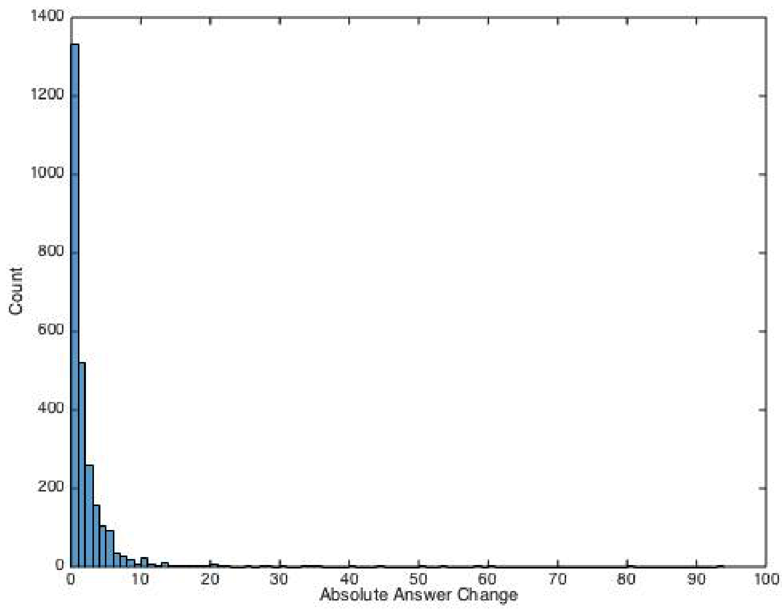


Figure 3.3: Tally of answer changes counted across all rounds and participants.

Table 3.1: Mean absolute answer change for all groups split by group.

Group	Mean Change
1	1.1 (SD 2.8)
2	2.3 (SD 3.5)
3	2.2 (SD 2.6)
4	2.1 (SD 2.8)

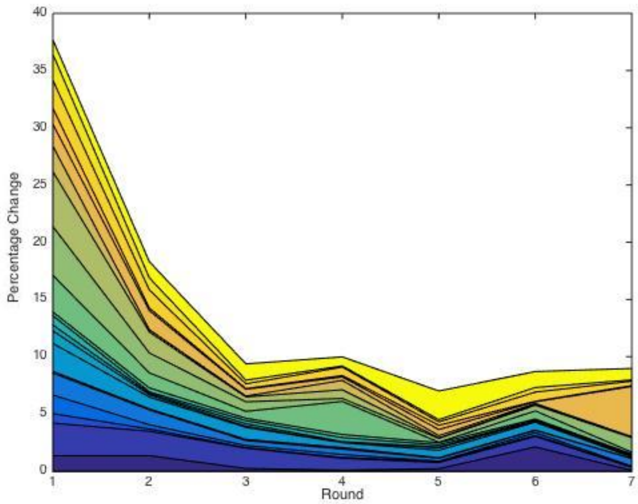


Figure 3.4: Percentage of Answer Changes by Revision Round where round one represents the first revision round (Groups 1-3). Each line represents a different question.

Most changes tended to occur in the first round, dropping off sharply and stabilising in later rounds. As Figure 3.4 demonstrates, some 35 percent of all changes occurred in the first round. This drops off to just under 20 percent in the second round and remains at 10 percent for the later rounds. This held true for three of the four groups.

3.2.2.3 Percentage Change

The percentage change allowed for cross question and group comparison, by normalising these effects to the particular question and group. Two outliers aside, generally changes were quite mild, which made percentage change a representative and informative measure for reporting participant behaviour. With respect to the magnitude of change, there were two notable outliers. In Group 4, participant 3 revised their answer by 4700 percent in round five, from 3 to 96 (with the correct answer being 96). In Group 3, participant 1 changed their answer by 2010 percent, going from 3 to 63 in round two (with correct answer being 55). These two instances are the only changes of this magnitude across all rounds and players. Moreover, these players did not exhibit similar behaviour on other questions. While these changes were not excluded from the analysis, they were excluded from the graphs as including them in the graphs would significantly distort the overall picture.

There was great variability in the magnitude of change in the answers both for individuals and for questions. As an example, Figure 3.5 breaks down the overall percentage change in answers across rounds for Group 1. Some participants in this group changed their answers by almost 160 percent, however, many also did not change their answers at all. The mean value of change was around 20 percent for this group. Figure 3.6 shows that these same participants also responded very differently to different questions.

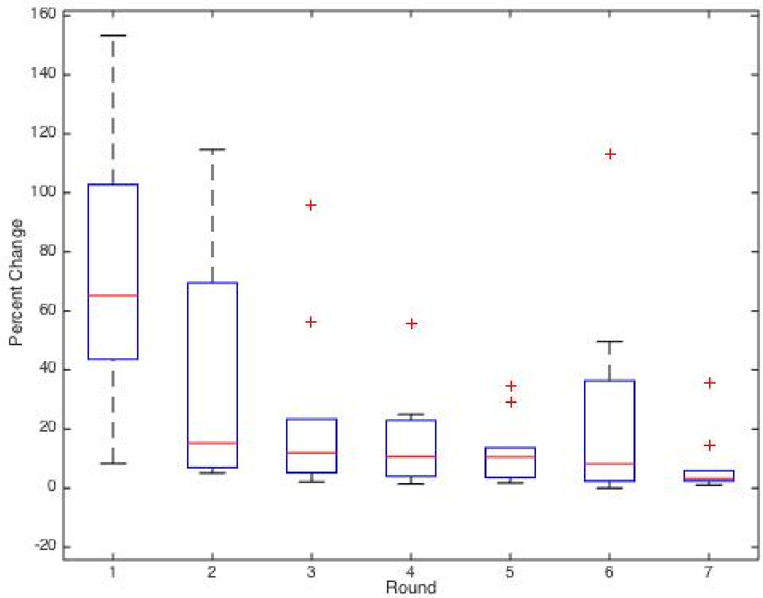


Figure 3.5: Percent Change of Answers by Round (Group 1).

3.2.3 Revision Analysis

A so called ‘absolute answer change’ measure was used to understand which strategies participants were using to revise their beliefs. The measure was calculated as the absolute difference between the expected answer given by the model and the actual answer supplied by the participant. The absolute difference was taken as the measure of fit. Models that had the lowest absolute difference were considered to be more reflective of the actual participant behaviour. This measure provides a direct indication as to how far away the participant was from answering the question correctly. The absolute difference, rather than the more typical squared mean er-

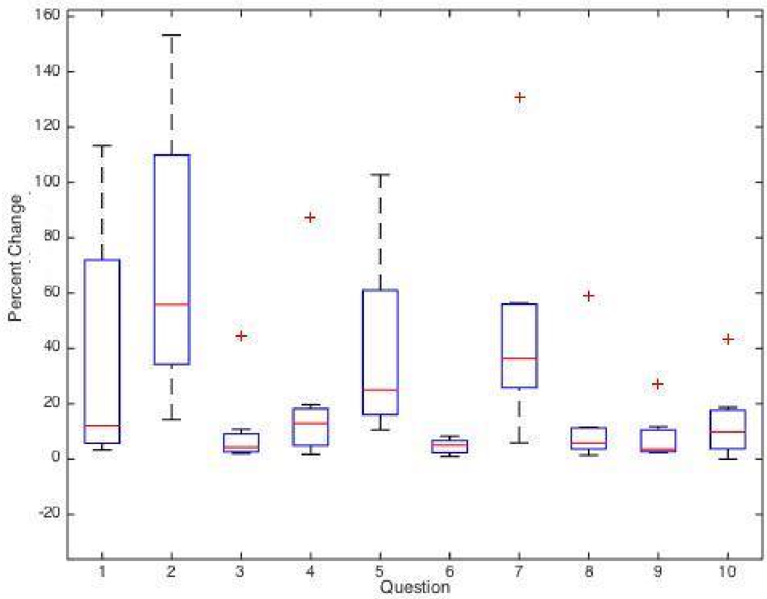


Figure 3.6: Percent Change of Answers by Question (Group 1).

ror used in statistics, was used as it has been found to be a more statistically robust measure where the error is not necessarily Gaussian in its distribution (Willmott and Matsuura, 2005, Chai and Draxler, 2014). Willmott and Matsuura (2005), argue that mean squared error is "inappropriate because it is a function of 3 characteristics of a set of errors, rather than of one (the average error). RMSE varies with the variability within the distribution of error magnitudes and with the square root of the number of errors ($n1/2$), as well as with the average-error magnitude (MAE). Our findings indicate that MAE is a more natural measure of average error, and (unlike RMSE) is unambiguous" (Willmott and Matsuura, 2005, p.79). Given that there is no Gaussian distribution assumption of the error term in this study, the mean absolute error provides a good measure of individual and group performance, is easy and transparent to communicate and analyse. Although absolute difference is used in most of the reporting below, in cases where statistical significance is observed, squared error was also used to verify significance or lack thereof.

3.2.3.1 Predictive Models

The starting point for explaining observed behaviour are the existing models of opinion revision. These models typically focus on the individual adopting some combination of the mean or median values derived from the group and can be found in diverse literature on networks and belief revision. These models are widely used to calculate opinion diffusion, belief revision, and other modelling exercises where units meaningfully interact with each other. For example, the early studies by Yaniv (1997) looked at the importance of participants' weighing (anchoring on own answer) and trimming (discarding other answers). Lorenz et al. (2011) looked at the importance of confidence in belief revision, while Jackson (2010) has described at length the models where participants, at some pace, move towards the group mean.

Drawing on these sources, it is possible to come up with several models to test against the dataset:

- The *weighted average* model predicts that an individual will adopt the group mean, but in adopting the mean, will weight their previous answer as an anchor, and will give the initial answer a higher weight than the group mean thus adopting an answer between the two values. In our model we set the weight at two: a participant would 'count' their own answer twice, before averaging it with the others.

- The *"split the difference"* model assumes that an individual will take the mean of the answers and then average that with their previous answer. This is simply a more aggressive version of the weighted average model.

- The *median* model predicts that participants will simply adopt the median value of the available answers (including their own).

- The *"previous answer"* model simply predicts that the answer in the next round will be exactly the same as in the previous round, thus no revision will have occurred.

Finally, the two remaining models are taken from the Yaniv and Milyavsky (2007) studies. Yaniv and Milyavsky (2007) found that participants were 'highly egocentric', and gave less weight to advice the further it is from their initial opinion. They were also sensitive to the variability within the groups' judgements. This *egocentric trim* model seeks to capture that participants will,"weigh the opinions that are close to their own, while ignoring those that are distant from their own prior opinion" (Yaniv and Milyavsky, 2007, p.105). In this particular model, an individual will dismiss value(s) most distant from her own, and adopt the mean value of all remaining answers (including her own). The *consensus trim* model is similar, but here an individual first takes the group mean and then discount the answer most distant from that mean. They will then take another group mean and adopt that value as their own (Yaniv and Milyavsky, 2007).

Yaniv and Milyavsky found that the egocentric trim and the median models most closely accounted for the actual belief revision strategies adopted by their participants. Indeed, egocentricity was found to be the overarching principle employed by the participants. This factor was found to be mediated by prior knowledge (confidence in own answer) (Yaniv and Milyavsky, 2007, p.117). Similarly to the Lund study there was also a substantial number of trials where participants" adhered to [the initial answer] in roughly 35-40% of the trials, changing their opinions only in the remaining 60-65%" (Yaniv and Milyavsky, 2007, p.116).

3.2.4 Results

Each of the models above was compared to the actual revision done by the participants in the Lund study, with the focus on the mean absolute error per turn – the mean deviation between the predicted value and the actual responses for an individual participant in a given round for a single question. Table 3.2 summarises the results for every model. Higher numbers indicate greater deviation between the predicted and the observed behaviour and therefore less of a fit between the predicted and the observed behaviour.

3.2.4.1 First round revision

Firstly, model performance was examined on the initial round of change, where comparison is most direct with Yaniv and Milyavsky's study which gave partici-

Table 3.2: Model Performance

Model	First Round of Revision Average Across Groups	Across All Rounds Group 1	Group 2	Group 3	Group 4	Average Across Groups
Weighted Average	7.25	6.39	7.79	8.74	7.61	7.57
Split the Difference	6.33	5.85	7.11	7.93	6.41	6.72
Median	8.13	5.21	7.77	8.66	6.04	6.73
No Change	5.05	4.03	6.41	6.92	6.06	5.82
Egocentric Trim	6.20	4.45	6.72	8.16	6.21	6.27
Consensus Trim	9.24	6.49	8.42	10.12	7.43	7.94

pants only one set of advice and gave an opportunity for only one revision. In contrast to their results, the 'no change' model had the lowest predictive error for our data. It was 20% more accurate than the second best performing model of egocentric trim. This is a notable difference from the findings of (Yaniv and Milyavsky, 2007, p.111). In their study the so-called 'self-initial' or no revision-model was outperformed by the median and egocentric trim models.

The notable difference in how participants responded in the Lund study may be due to the different experimental paradigm. After all, advice was experimenter-provided in Yaniv and Milyavsky's (2007) studies, while in Lund other students were supplying opinions. This difference emphasises the need to examine social belief dynamics in a broader range of experimental paradigms. Secondly, the fact that the 'no change' model is the best predictor suggests that none of the other models are particularly good at explaining participant behaviour. From Figure 3.3, it is clear that participants changed their answers about 50% of the time, and from Figure 3.5 it is also clear that revision occurs most often in the first round of revision, yet the no-change model outperformed all of the other models. It appears that the models seeking to capture this change do not perform well.

Where Lund results do fit with Yaniv and Milyavsky's findings is in the order of the other three models they test (median, egocentric trim, consensus trim). The consensus trim model performs worst, with the median model coming second, and Yaniv and Milyavsky's (2007) egocentric trim model being the best in both studies. Finally, the two additional averaging models place last.

Interestingly, Yaniv and Milyavsky noted a relationship between revision and a number of pieces of advice presented to participants; more pieces of advice induced greater change. This was particularly true with the median and egocentric trim model performance. Jönsson et al. (2015) add to this finding by introducing the concept of network topology and the impact it has on revision. Participants in the Lund study not only saw different number of answers, as a function of whom they were connected to, but the network itself was wired differently. However, network topology does not come into play in the first round of revision, given that the first revision is simply a function of the initial estimate of each participants. The small-world networks are simply an example of a smaller number of advisers in the Yaniv and Milyavsky studies. The network dynamics, however, do come into play in the subsequent revision rounds.

3.2.4.2 Multiple Rounds of Revision

How then do these models fare in predicting repeated rounds of revision in a dynamically changing environment? Again, the 'no change' model is the best predictor, followed somewhat more closely by the egocentric trim model. In fact, moving from first round to all rounds, the rank order of models changes only between two of the poorer performing models (median and weighted average). Consensus trim, again, comes last suggesting that this model fails to capture participants' approach to opinion variability in a meaningful way. Again, the no-change model outperforms all of the other models.

Where do the failures of the models lie? Firstly, all models, with the obvious exception of the no change model, over-predict change for the first round of revision, that is, the transition from participants' initial answer to their second answer. Despite the fact that most change in responses occurs in the first round of revision, participants still change less than the various models suggest they should. This can be seen by comparing Figure 3.7, which displays round-on-round change for Group 1 participants with Figures 3.8, 3.9, and 3.10. Even the best performing models, the Yaniv and Milyavsky's egocentric trim model (Fig. 3.8), predicts noticeably more change in this round than actually occurs. However, the same models then *under-predict* change on the second round of revision. It appears that repeated feedback encourages participants to take a comparatively greater note of others' opinions on the second revision round. This is particularly clear in comparison with the weighted average model (Figure 3.10) which predicts a sharp, monotonic decrease in round-on-round change.

In other words, there is some evidence from these comparisons that 'weights' placed on the opinions of others are *dynamic*, as opposed to static, across the subsequent rounds. Seeking to probe the nature of such dynamic changes further seems imperative given that the most common models of belief and opinion dynamics assume constant weights and relationships between agents.

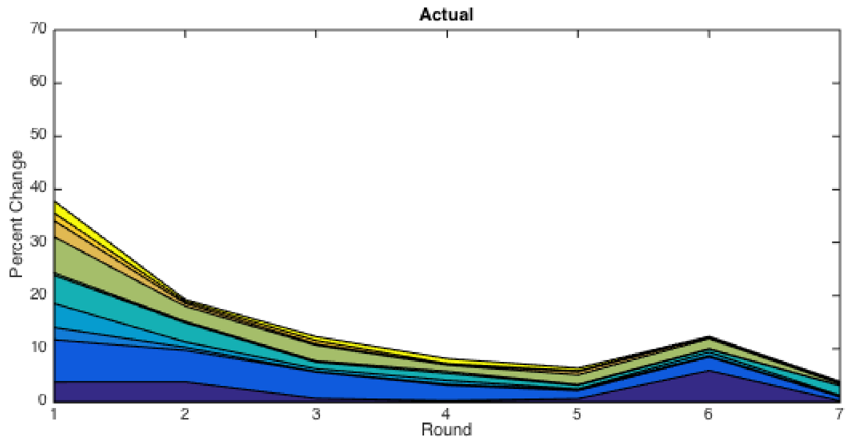


Figure 3.7: Actual Percentage Change by Round (Group 1).

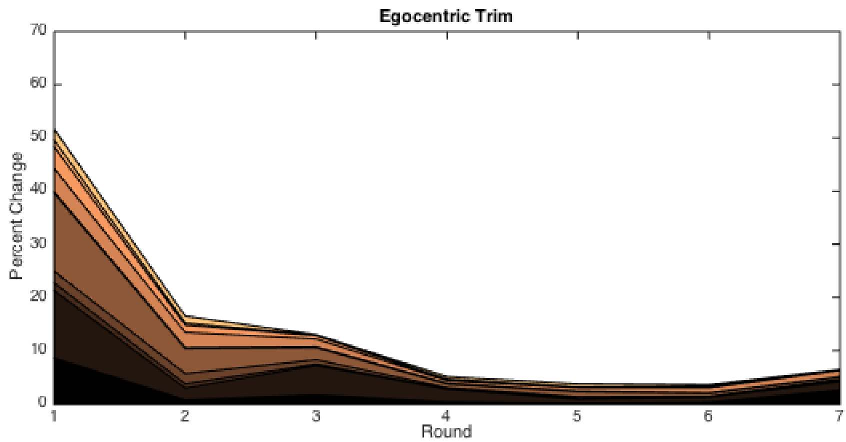


Figure 3.8: Predicted Percentage Change by Round for Egocentric Trim Model (Group 1).

In order to better account for actual human revision behaviour and overcome the shortcomings of other models, a new model called, *neighbour inclusion* was

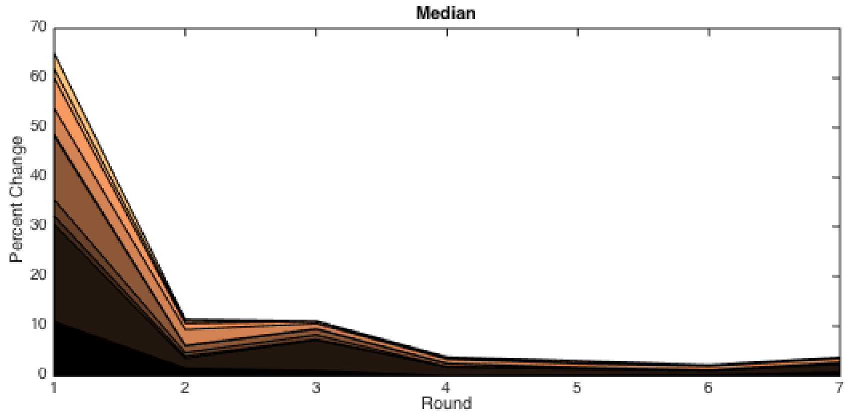


Figure 3.9: Predicted Percentage Change by Round for Median Model (Group 1).

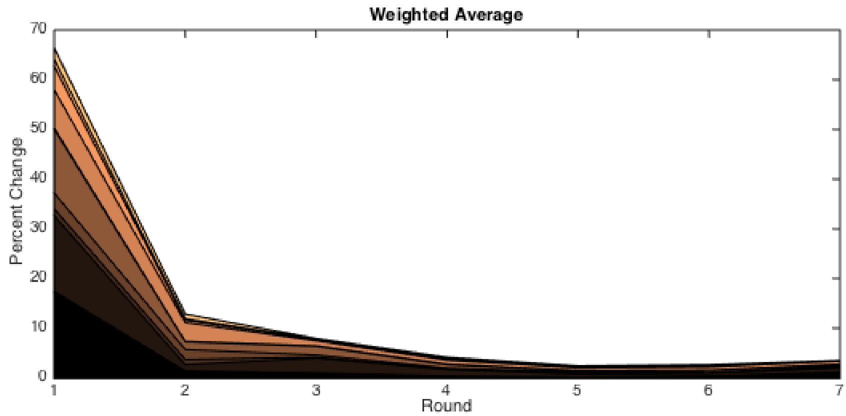


Figure 3.10: Predicted Percentage Change by Round for Weighted Average Model (Group 1).

created. The model is a mixture of the *egocentric mean* and *split the difference* models, but with a more egocentric bias (i.e. it more aggressively discounts other answers). This model predicts that individuals would be more sensitive to answers closest to their own and exclude those further away. It essentially predicts that a participant will adopt the mean between their own answer and that of the closest neighbour.

When applied, it can be seen from Figure 3.11 that the neighbour inclusion model performs slightly worse than the no change model. However, it significantly outperforms all of the other models suggested in the literature. Furthermore, by adjusting the weight factor further towards a respondent's own answer, the model essentially performs on par with the no change model. As Figure 3.11 shows, the best performing model is the no change model, and no model is able to explain any unique variance beyond that individuals simply do not change their answers.

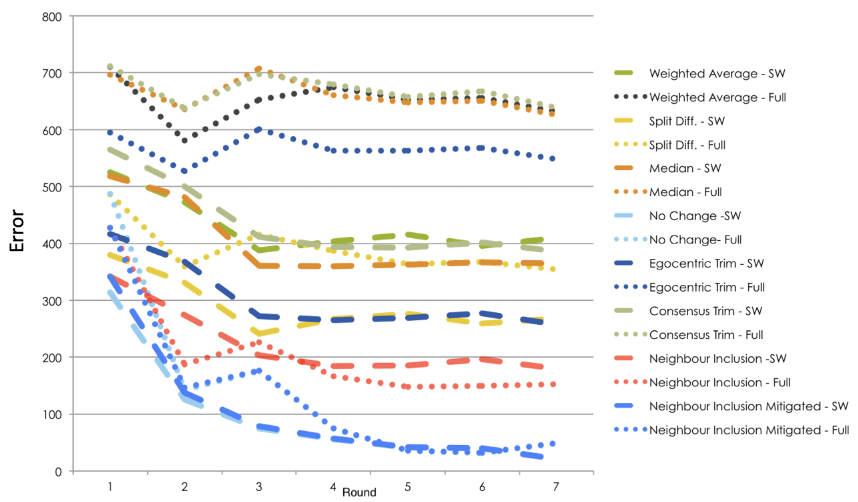


Figure 3.11: Cumulative error by round for different revision models, averaged across all groups for both small and full World Networks.

Notably, some models perform significantly better in the small world context, while others perform similarly in both contexts. Models that rely on the adoption of the group's answer – average, median and egocentric trip models – tend to perform significantly worse in the full world context, whereas models that use the initial answer as the starting point have less of a divergence in performance, suggesting that using the participant's initial answer better reflects the actual opinion revision dynamics. This also suggests that overall, less change occurs in the small world context, as is evident from the performance of the no change model in round one, where the total error is significantly higher in the full world context.

3.2.4.3 Accuracy

As Yaniv (2004a), Yaniv and Milyavsky (2007), Jönsson et al. (2015) note, participants almost always get more *accurate* with each revision. This is true of both individual account accuracy, as feedback loops in the system act to facilitate more optimal behaviour. Yaniv and Milyavsky (2007) found that participants' initial answer was correct in only 5% of the cases. Moreover, in their study, median, mean, consensus and egocentric trim produced more accurate final estimates than the no change model. This suggests that individual accuracy always improved when information was incorporated and beliefs were revised. Jönsson et al. (2015) note that groups overall also get more accurate as a result of revision. Given that existing models fail to explain this revision, *there is a tension between existing models' ability to explain revision strategies and the fact that revisions do occur, and are actually beneficial.*

3.3 Experiment 6: Individual Belief Revision Part I

The paradox of groups getting more accurate through revision despite every model failing to account for it, as evidenced by the performance of the no-change model, was unsatisfactory to say the least. With this, began the quest to better understand and model individual belief revision in the hope of finding some unique features of belief revision beyond the no-change model.

The empirical interest is in unpacking the unexplained variance to better understand change dynamics in order to create a model that can better reflect participants behaviour. In order to better understand the underlying dynamics a series of experiments was created to artificially control the generated answers which participants would see in order to understand the factors that go into the decision-making process. This was an attempt to take opinion back to *basics* and vary what may reasonably affect opinion revision. A basic tenet of all consistency theories of attitude change is that individuals seek to resolve discrepancies among their beliefs. Such theories predict that attitude change should decline with distance (Aronson et al., 1963, Sherif and Hovland, 1961, Yaniv, 2004a, Yaniv and Milyavsky, 2007). Distance is a function of multiple opinions then, which are governed by three parameters: how far they are from each other (standard deviation), how far they are from the participant (mean), and how far the closest answer is to the participant (closeness). The follow up study seeks to vary the mean, standard deviation and closeness to the actual participant response in order to induce and measure the degree of revision.

3.3.1 Method

3.3.1.1 Participants

31 University of London students were paid £5 each to participate in the study. Of the 31 participants, 13 were female and 18 male (n=31). There was no age, or education information collected from the participants. The study manipulated the distance and standard deviation of the generated answers in a 2x2x2 factor, within subject factorial design.

3.3.1.2 Materials

Participants were asked general knowledge questions based on the 2011 UK census data. All questions were general knowledge questions chosen to allow participants to develop an intuitive guess (estimate), but not necessarily know the correct answer. All questions had an answer range between 1 and 100. A sample question was: 'What is the percentage of adults in the UK who report not drinking alcohol at all?' A full list of questions can be found in Appendix D. Questions were chosen to be engaging, but not necessarily easily answerable. They were designed to be 'general knowledge' in a sense of an average individual having an intuition about the answer, without necessarily knowing the exact answer. Most importantly, each question had a numeric true answer (i.e. a fact, with evidence in the natural world), so that the accuracy of answers could be assessed.

After participants provided an initial estimate, they were provided with responses by 'other participants' (computer generated) and given a chance to enter a new, revised answer. The 'other responses' were generated based on the participant's own initial response. Three factors were manipulated depending on the condition. The first factor was the mean of the generated responses, which was either near or far from the participant's own response. Mean refers to the distance between the generated answers and the actual answer. High meant that the answer distance was 10 percentage points away and low meant 5%. The second factor was the standard deviation of the generated responses, which was either low or high. In high conditions variance of the generated answers was 20%, in low variance it was 10%. The third factor was the generation of a special answer. This answer was generated based on the distance from the participants answer. In the close condition this answer was within three points from participant's answer. In the far condition the answer was more than three points away.

25 questions were presented to participants. This included an initial warm up question and three questions per condition. Not all questions could appear in each condition with equal plausibility (e.g. a question related to average age could not have generated answers that are implausibly high or low without making it seem incredible that these were genuine responses from other participants). To minimise the impact of question-based variability, however, there were two different assignments of question to condition with the first 15 participants receiving one assignment, and the remainder the other.

The experiment was conducted using a MATLAB-based user interface that displayed the pre-programmed questions. See Figure 3.12 for an example of a user interface. Upon entering their estimate, participants were presented with their answers as a red dot and computer generated answers as blue dots on a graph. Participants were told that the blue dots are in fact the answers of participants who had done this study prior to them. In actual fact, blue answers were computer-generated to match the experimental condition for each question. Figure 3.13

3.3.1.3 Procedure

For each question, participants provided a numeric answer between 1 and 100. In a separate box they also indicated their confidence on a scale from 1 (not at all) to 10 (completely) in the accuracy of their estimate. Once participants entered their estimates, they were shown five other answers (see Figure 3.13 for a sample interface with the generated answers).

Once participants saw other responses, they were prompted to answer the same question again as well as enter a new confidence estimate. Each question was answered twice: once before participants saw other answers and once after. Once a question was answered twice, a new question was displayed and the procedure repeated. In total 25 questions were asked.

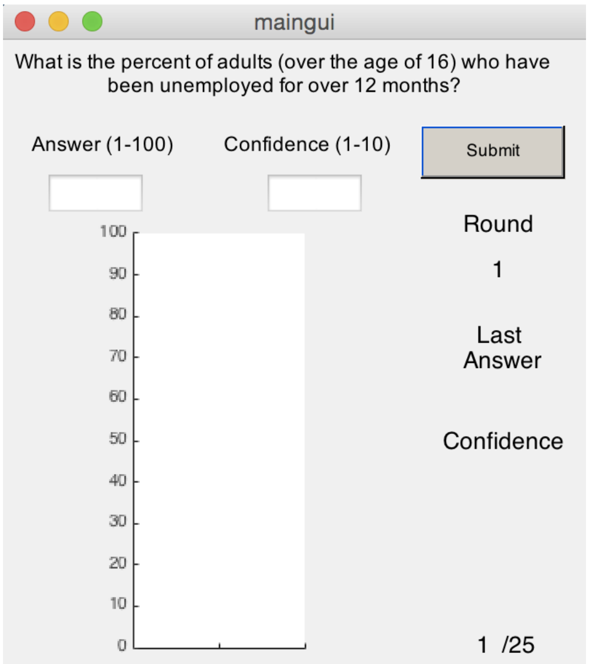


Figure 3.12: Sample Participant Interface.

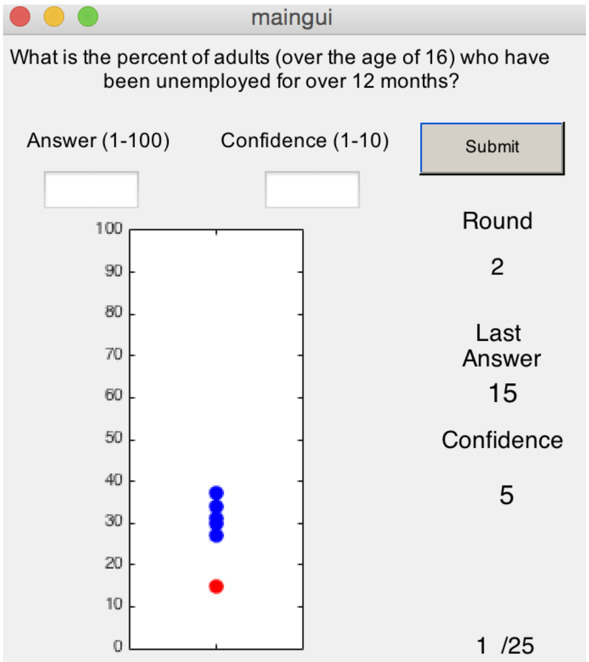


Figure 3.13: Sample Interface with Generated Answers.

Table 3.3: Breakdown of the mean, standard deviation and distance to the closest answer of the generated answers by condition.

Measure	Mean: Near Std: Low Closest: Close	Mean: Near Std: High Closest: Close	Mean: Near Std: Low Closest: Far	Mean: Near Std: High Closest: Far	Mean: Far Std: High Closest: Close	Mean: Far Std: Low Closest: Close	Mean: Far Std: High Closest: Far	Mean: Far Std: Low Closest: Far
Mean (SD)	1.52 (3.18)	0.74 (6.93)	4.93 (5.81)	2.72 (7.68)	7.92 (11.46)	13.26 (9.1)	11.28 (17.03)	16.56 (6.28)
Distance to Closest Ans.	1.00	2.01	6.80	5.20	1.60	1.64	12.30	12.02

The *other* responses were generated based on the participant's initial answer. For example, in the first condition where the mean was 'near' the initial answer, the standard deviation was 'low' and the closest answer was 'close', the distance of the generated answers to participant's initial answer was on average only 1 point away, the mean of the generated answers was only 1.52 points away and the standard deviation was 3.18. This contrasts with the condition where the mean was far (11.28 points away), standard deviation was high (17.03) and the closest answer was far (12.30 point away). Table 3.3 summarises generated answer statistics, where 'Mean' refers to the mean distance of the generated answers to that of the user. Near refers to the condition where the distance to the generated answer was within 5 points, whereas far meant that the distance was beyond the five points. 'Std' refers to the standard deviation of the generated answers. Low means the standard deviation was around 6 points and high meant that the standard deviation of the generated answers was above 6. 'Closest' variable refers to how close the closest answer was to that of the participant. In close condition the answer was one or two points away, whereas in the far condition it was above 5 points away. Each column in the table refers to an experimental condition where three variables were manipulated. As is clear from Table 3.3, it was not always possible for each condition to be exactly balanced as the three variables are intertwined, however, all efforts were made to make each condition as balanced as possible.

3.3.2 Results

The primary measure of participant behaviour was the degree of absolute change between their initial judgement and their second, revised judgement, that is, the absolute difference between the first and second estimates for each question. This measure was chosen to make analysis similar to the Lund study analysis described above. The purpose was to determine which of the experimental factors caused participants to change their judgements, and to see whether there were any interactions between these. The secondary aim was to determine the relative importance of each of these factors, and gather a data set that could be used to mathematically compute the impact of each of the factors allowing to better understand human revision.

3.3.2.1 Outliers

There was great variability in the amount of revision displaced by the participants. There were three notable outliers. There were two participants (13 and 24) who did not change their answers at all and participant 5 only changed their answer once (by 5 points). Although participant 13 had the lowest confidence, their average confidence was within two standard deviations of the group mean; the other two participants did not report unusual confidence levels. Since their behaviour could not be explained by any available analysis, the three participants were excluded from further analysis due to likely non-engagement with the task.

3.3.2.2 Descriptive Statistics

Table 3.4: Participant mean and standard deviation of answer change by condition.

Condition	Mean: Near Std: Low Closest: Close	Mean: Near Std: High Closest: Close	Mean: Near Std: Low Closest: Far	Mean: Near Std: High Closest: Far	Mean: Far Std: High Closest: Close	Mean: Far Std: Low Closest: Close	Mean: Far Std: High Closest: Far	Mean: Far Std: Low Closest: Far
Mean (SD)	0.516 (1.241)	0.973 (2.959)	2.237 (3.160)	1.194 (2.896)	2.118 (3.448)	3.882 (5.042)	3.946 (7.594)	5.183 (6.587)

The primary point of interest was the impact of experimental condition on answer change (belief revision). Table 3.4 summarises answer change by condition. The structure of the table is similar to Table 3.3. Each column refers to a different experimental condition and the three variables that were manipulated for the generated answers. From Table 3.3 it is apparent that certain conditions had more of an impact on participant behaviour change than others. Conditions where most of the generated answers were close to participants' initial answers induced the mean answer change of less than 0.5, while conditions 7 and 8, where many of the generated answers were far away from the initial answer, induced the average answer change of more than 5 points. This suggests that distance and dispersion of the generated answers had an impact on the way participants revised their answers.

Given that certain conditions were hard coded to certain questions to make generated answers more plausible, it was important to evaluate the impact of accuracy in the initial answer. If initial accuracy was greater for certain conditions, lower revision independently of manipulations in the condition would likely occur. Table 3.5 shows mean distance between the correct answer and the initial guess, by condition. Conditions 1 and 2 had lowest initial accuracy, while the rest of the conditions had similar initial accuracy.

It was expected that lower initial accuracy would lead to higher answer change – since an initial answer that is further from the correct answer should prompt participants to change more. However, results did not show this. As can be seen from Table 3.4, participants changed the least in conditions one and two, where initial accuracy was the *lowest*. There was no statistically significant relationship between initial accuracy and revision between rounds, nor was there a statistically significant relationship between initial accuracy and condition, as well as initial confidence and condition. It is therefore possible to conclude that experimental approach in itself did not bias participants towards greater change and it was the manipulation of the experimental variables that led to change.

Table 3.5: Participant initial accuracy for each condition

Condition	Mean: Near Std: Low Closest: Close	Mean: Near Std: High Closest: Close	Mean: Near Std: Low Closest: Far	Mean: Near Std: Low Closest: Far	Mean: Far Std: High Closest: Close	Mean: Far Std: Low Closest: Close	Mean: Far Std: High Closest: Far	Mean: Far Std: Low Closest: Far
Mean	22.486	20.57	17.48	17.658	16.541	18.823	19.551	17.054

3.3.2.3 Multivariate Modelling

Given the apparent relationship between between experimental manipulations and degree of change, it was important to understand a more exact measure of this relationship for future modelling. To better understand this relationship multivariate regression modelling was used.

The initial model was based on the significant factors identified by the ANOVA model – absolute distance from the mean, standard deviation of the generated answers and distance between participant's answer and the closest generated answer.

In the first instance the actual distance between the mean of the generated answers and that of the participant, the standard deviation of the generated answers, as well as the distance between participant's answer and the closest generated answer was calculated.

Table 3.6 shows that only distance to the mean significantly predicted estimate revision b = 0.299, t(672) = 9.383, p <.001. The other two factors (SD of the answers and distance to the closest answer) were not statistically significant. The overall model with adjusted R2 = 0.16, F (672) = 43.502, p<.001 is quite predictive of revision.

However, the fact that SD of the answers and distance to the closest answer does not appear to impact revision is important. This result diverges from the ANOVA results, suggesting that while some factors may have correlated with belief revision, their impact is limited, as demonstrated by further modelling. Given that experimental cells were not entirely balanced, regression results are more robust and more likely to be indicative of real behaviour.

3.3.2.4 Hierarchical Model

While the base model provides a starting point there is a deeper level of analysis that can be done. Based on the plots below (see Figures 3.14, 3.15, 3.16, 3.17), it is clear that different participants behaved differently on the task. From Table 3.17 it is possible to see that some participants revised less when their confidence was high, while others did the opposite. While Table 3.16 shows that most participants revised more when the closest answer was further away, however the slopes are clearly different, reflecting individual dynamics not captured by ANOVA and multi-level models above. In order to account for this variability, a technique called hierarchical modelling was used. Modelling was done in R, with the help of a package called 'lme4'.

Table 3.6: Multivariate regression model for the absolute answer change looking at mean distance, standard deviation and distance to the closest answer.

	Dependent variable:
	Absolute Answer Change
Distance to the Mean	0.299***
	(0.032)
SD of generated answers	−0.016
	(0.038)
Distance to Closest	0.069
	(0.042)
Constant	0.341
	(0.403)
Observations	672
R^2	0.163
Adjusted R^2	0.160
Residual Std. Error	4.812 (df = 668)
F Statistic	43.502*** (df = 3; 668)
Note:	*p<0.1; **p<0.05; ***p<0.01

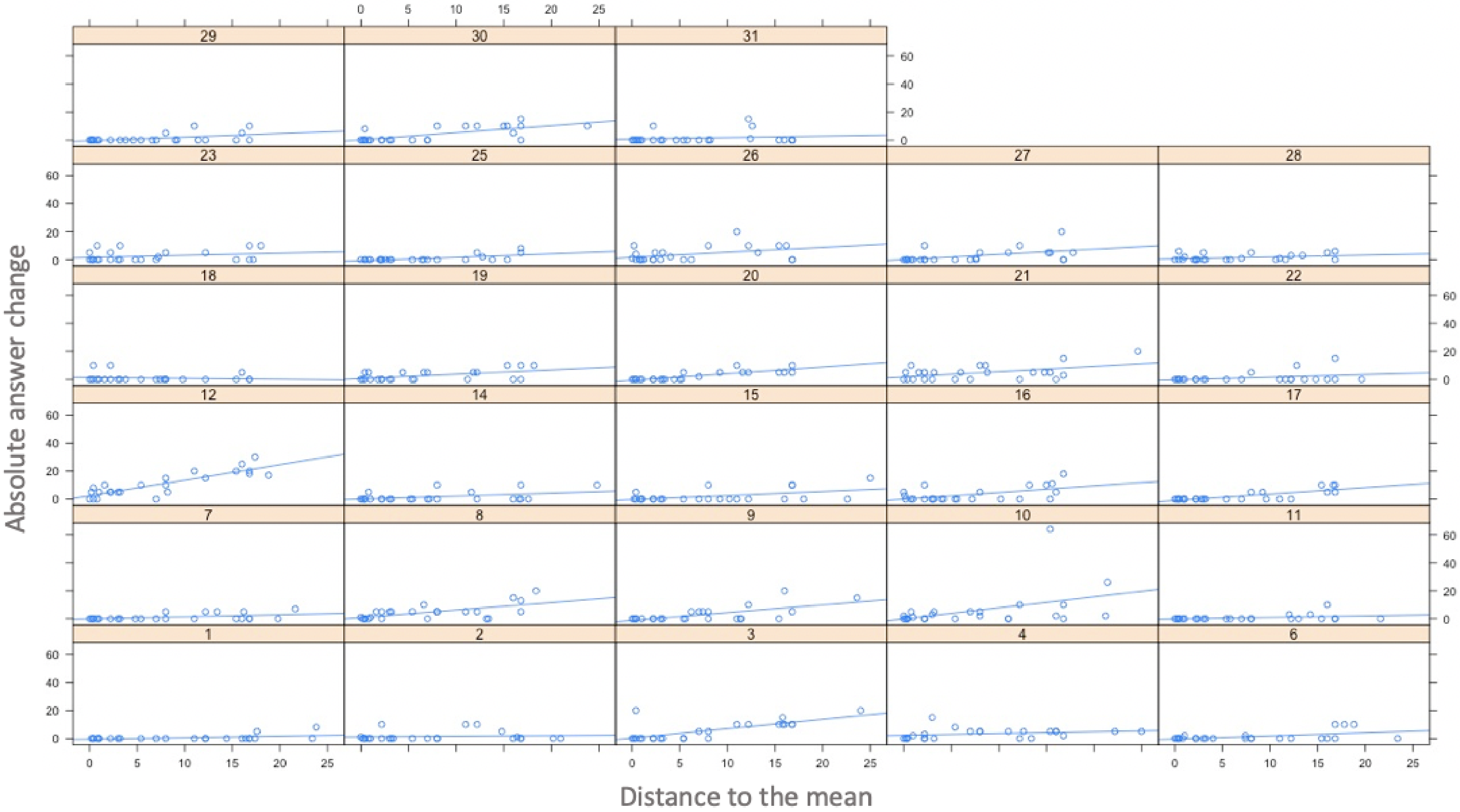


Figure 3.14: Distance from the group mean and belief revision per player.

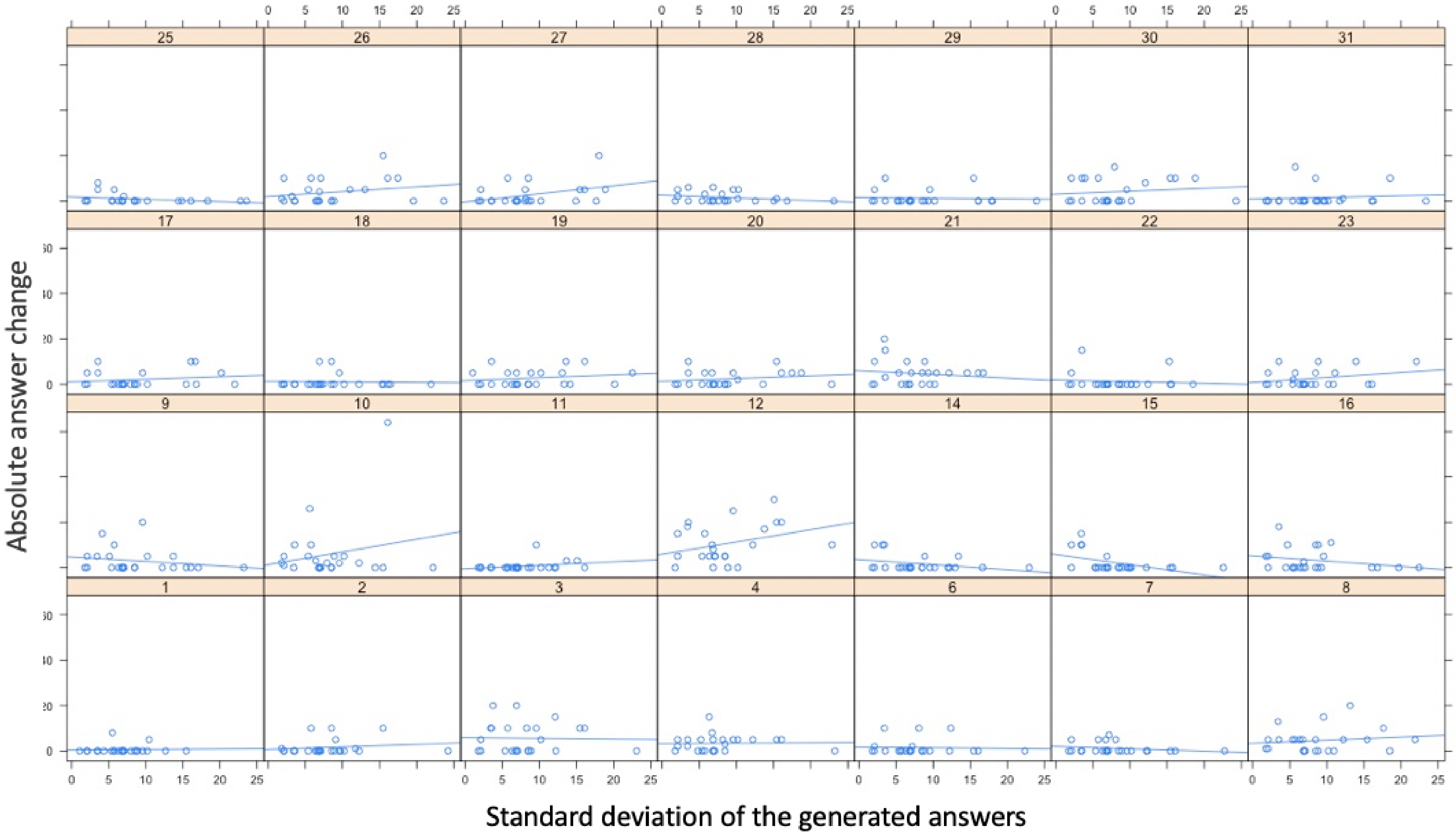


Figure 3.15: Standard deviation of the generated answers and belief revision per player.

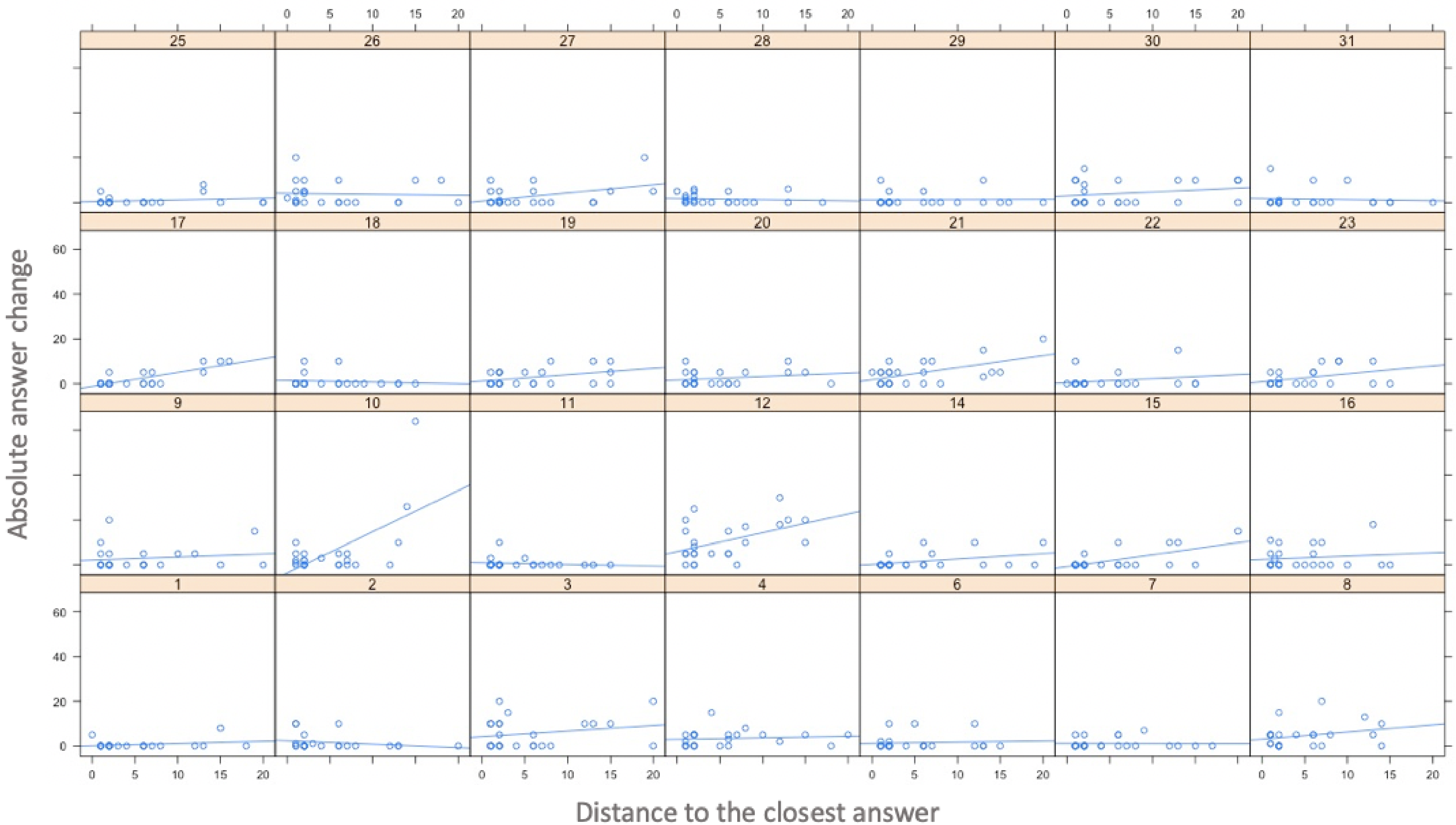


Figure 3.16: Distance from the closest generated answer and belief revision per player.

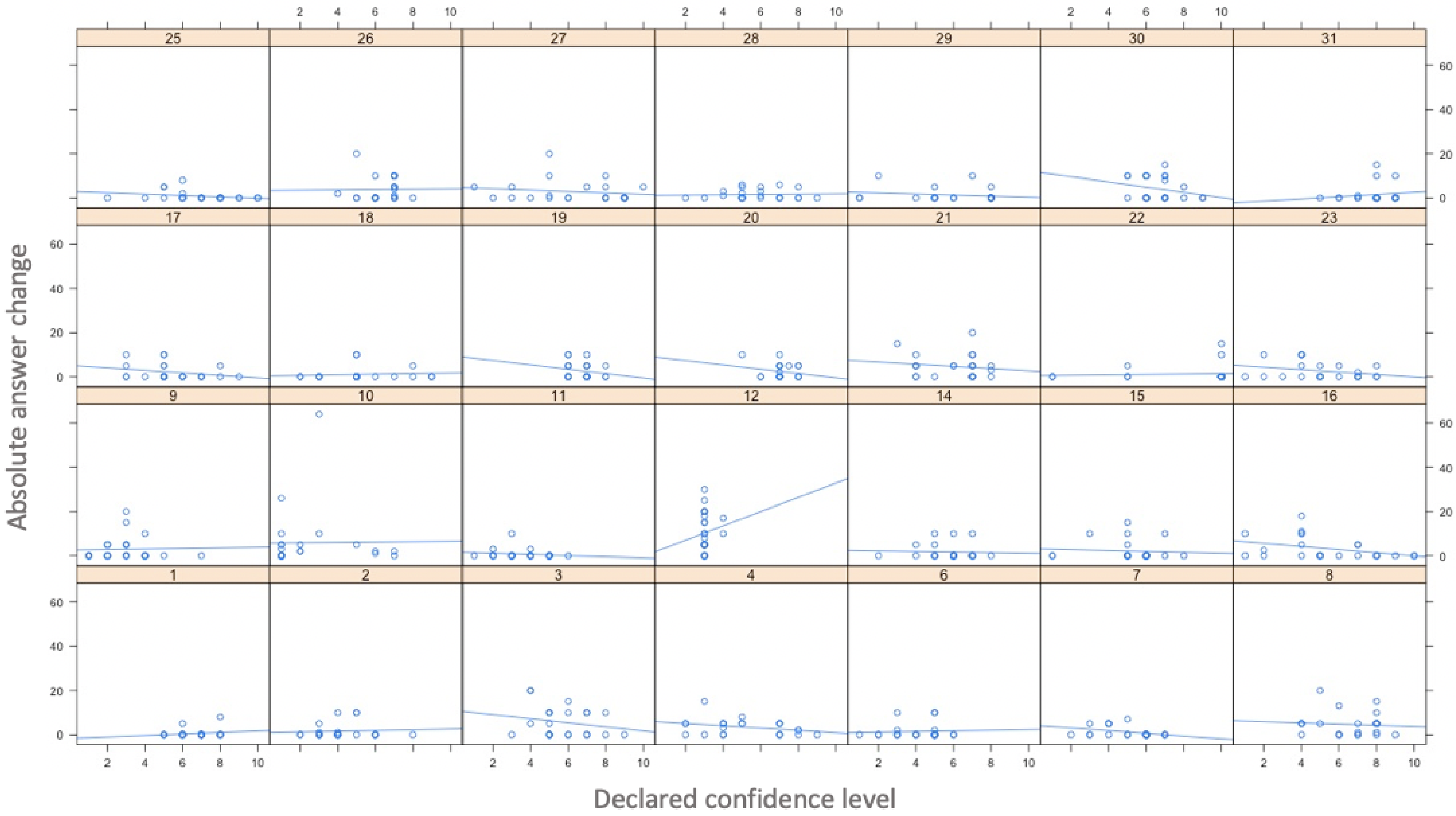


Figure 3.17: Initial confidence and belief revision per player.

In hierarchical regression models individual regression coefficients – here participants – are given their own probability models, which can often better capture the underlying dynamics of behaviour, not captured by standard ANOVA models (Gelman and Hill, 2007). In this case multilevel modelling allows probability models for the coefficients at the lower levels, thereby increasing its explanatory and predictive power, estimating hyper-parameters from the data itself.

As is evident from the graphs, there is considerable difference between individual participants in terms of their revision. In order to understand individual differences, hierarchical modelling was used to create a more robust model for the data. Since we obtained multiple measures from the same participants (nested within subject) the data lent itself well to multi-level modelling.

As in the previous regression model, the goal was to determine which factors influence the magnitude of answer change. The dependent variable was the amount by which participants changed their answer. Level-1 unit was the actual answer changes, while level-2 unit was the answer change by participant. This allowed the model to accommodate individual differences among the participants. The standard measure of estimates is the inter-rater reliability, "or the amount of variance in individual level responses that can be explained by group level properties" (Castro, 2002, p.70), which is also called the Interclass Correlation Coefficient. It is not influenced by group size or by the number of groups. The ICC(1) measure was about 15%. While not extremely high, it suggests that hierarchical model may add additional explanatory power by allowing for individual slopes for each participant.

Hierarchical modelling works by building successive models, with additional factors and interactions, which evaluating their AIC and BIC scores, which are goodness of fit measures. The lower the scores, the better the model fits the data. Table 3.7 outlines each model, along with the AIC and BIC scores, along with the Chi Squared statistic and its corresponding P value.

The basic model includes Mean Difference to the generated answers and participant Confidence as the two factors, in addition to the intercepts for each participant. The second model adds an interaction term between dean Difference and

Table 3.7: Multilevel Model showing summary statistics for the variables impacting belief revision.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
MeanDiff + Confidence + (1 — Player)	5	3949.9	3972.4	-1970.0	3939.9			
MeanDiff + MeanDiff * Confidence + Closest + (1 — Player)	7	3937.7	3969.2	-1961.8	3923.7	16.241	2	0.0002974 ***
MeanDiff + MeanDiff * Confidence + (MeanDiff — Player)	8	3878.7	3914.8	-1931.3	3862.7	60.976	1	<0.001***
MeanDiff + MeanDiff * Confidence + (MeanDiff + Closest — Player)	11	3862.3	3911.9	-1920.1	3840.3	22.421	3	<0.001***

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Confidence, as well as the distance to the closest answer. The next model also allowed the slope for random effects to vary with Mean Difference. The last model added the distance to the closest answer as another random effect. While several other models were tried, such as a model with standard deviation of the generated answers as a factor, they did not produce better fit models (ie models with lower AIC scores).

The four models presented in Table 3.7, show the progression of the factors that in successive sequence produced models that fit the data, with increased closeness. An ANOVA test was used to compare the four models and determine the relative goodness of fit of each model. As can be seen from Table 3.7, each model performed statistically better, with AIC being reduced in each model.

The model with the lower AIC was the fourth model. The factors in the model were: distance between participant's answer and that of the mean of the group, initial confidence and the interaction term between Mean Difference and Confidence, as well as individual slopes for each participant. All other variables and terms were found to be to extraneous and were excluded.

The model contained three fixed effects: the mean difference, confidence and the interaction term of mean difference and confidence. Table 3.8 summarises beta statistics for the fixed effects. The random effects in the model were the MeanDifference and the Closest Answer, as these varied by participant. Table 3.9 summarises random effects.

To conclude, participant revision appears to be influenced by a combination of factors. It is strongly influenced by the distance to the mean (coefficient of 0.52516). The further away the group mean, the more revision there is. Confidence appears to have a moderating influence on this relationship, suggesting that less confidence

Table 3.8: Fixed Effects Statistics of the model showing factors that impact belief revision.

Fixed Effects	Estimate	Std. Error	t-value
(Intercept)	0.61438	0.65719	0.935
Mean Difference	0.52516	0.08449	6.215
Confidence	-0.04143	0.10970	-0.378
MeanDiff:Confidence	-0.03535	0.01271	-2.781

Table 3.9: Random Effects Statistics of the model for belief revision

Random Effects	Variance	Std.Dev	Corr
(Intercept)	0.70207	0.8379	0.935
Mean Difference	0.03980	0.1995	0.43
Closest Answer	0.07853	0.2802	-0.93

produces more revision, in combination with the distance to the group mean. The coefficient for the interaction between the distance to the mean and initial confidence is statistically significant, albeit small (-0.03535). Finally, revision is impacted by a number of effects which depend on the individual participants, namely: the distance to the mean and distance to the closest answer. The results suggest that some participants reacted very strongly to these variations, while others less so.

3.3.3 Discussion

Modelling revealed a number of factors that impact revision. In all models distance to the mean was the most important predictor of revision. Hierarchical modelling also revealed a mediating effect of confidence and was further improved by the inclusion of distance to the closest answer as a random effect. Perhaps the most important finding is that variability in participant behaviour is also an important factor when considering belief revision. The discussion about the various factors and individual differences follows.

3.3.3.1 Distance to the Group Mean

Individual distance to the group mean has the most significant impact on whether an individual will change the initial answer. This has proved to be true for most participants. The mean difference is a significant factor in both ANOVA and hierarchical regression models. This suggests that individuals are indeed sensitive to the answers of others.

The significance of the 'group mean' partly supports the earlier models. Although, individuals do not adopt the group mean as their answer in the second round, they do usually move towards it. For every one point in distance, on average, there is a 0.26511 move towards the mean. Therefore, individuals tend to move about a quarter of the way towards the mean. This suggests that the weighted averaging models would work better, although the factor that weights individuals' 'own' answers would have to be quite high.

3.3.3.2 Individual Differences

Although some individuals are clearly sensitive to the distance to the group, others appear to be much less so. This is apparent both from the individual revision (Table 3.14) and from the hierarchical modelling. Out of 28 participants, some display positive correlation between the distance to the mean and revision, while other do not follow this relationship at all, by either rarely revising, or revising in the opposite direction. The same finding applies to the other experimental manipulations, SD and distance to the closest answer, although they appear to be weaker.

More research exploring individual differences in needed to produce a more complete account of divergence in participants' behaviour. This would also be instrumental in building a more holistic model of belief revision in the future. This research could be extended by providing participants with an individual differences inventory such as Big5 personality questionnaire, or examining more closely prior beliefs on a range of topics, or a level of initial knowledge on the topic to better understand other factors that may influence individual behaviour.

3.3.3.3 Confidence

There is a statistically significant interaction between confidence and distance to the mean, as a fixed factor in the hierarchical model. Although the interaction term is small (-0.03535), it suggests that confidence moderates the adoption of the group mean.

Previous studies such as Moussaid et al. (2013), Lorenz et al. (2011), have looked at confidence as an important explanatory variable in decision-making. They suggest that confidence has important influence on belief revision in situations where participants are exposed to both the opinion and the confidence level of another participant. The impact of confidence on belief revision was much more muted in our study. While self-reported initial confidence does seem to play a role in belief revision, as suggested by the modelling, its effects appear to be difficult to detect. Furthermore, Moussaid et al. (2013) observed that "judgements of high confidence are good indicators of accuracy before social influence occurs" (Moussaid et al., 2013, p.7). Yet, no such relationship between confidence and accuracy was observed in our study.

3.4 Experiment 7: Individual Belief Revision Part II

Although the effects on revision discussed above appear to be in line with the findings from the Lund study, there seems to be a combination of factors at play, that are not well captured by the existing models. None of the findings from the previous experiment are helpful in building a model that can outperform the no-change model. Given that individuals are sensitive to the distance to the closest answer (the neighbourhood inclusion model accounts for this) and are sensitive to the distance to the mean and move towards the mean (the weighted average model accounts for this), these behaviours are far from being consistent. Looking at the data from the experiment, there appeared to be a factor that has not been previously discussed in the literature however. Participants tended to change their answers more when their answers were an outlier of the group. In other words, whenever they were on the outside of the group, not surrounded by other answers, they changed their answers more; moving towards the group mean and the closest answer.

In order to confirm that this relationship was indeed present, a followup study was conducted focusing on a single manipulation. The generated answers either *surrounded* a participant's initial estimate, or put it on the *outside*. If the participants did indeed revise more in the outside condition, this would help tie together the various insights gained from the previous experiments, such as the observed move towards the mean, and the significance of the closest neighbour in the original study.

3.4.1 Method

Given that the hypothesis was based on the previous study, the original methodology was preserved and replicated with minor variations. The same computer interface was used and participants were given the same set of instructions: to answer general knowledge questions and to provide a confidence score for each estimate.

3.4.1.1 Participants

18 (12 female and 6 male) University of London students participated in the study. They did so as part of another larger study and were paid £5 for their participation.

3.4.1.2 Materials and Procedure

The same MATLAB-based interface as described in Experiment 6 was used. As this was a follow up study, intended to test a single hypothesis, the number of questions asked of each participant was reduced from 25 to 11. Each participant did a practice question, before doing 10 actual questions. It was noted that there was some variability among the 25 original questions in terms of difficulty, as measured by the distance between respondents initial estimates and the correct answer. With this in mind, an effort was made to ensure that the 11 questions selected for the follow-up study were of comparable difficulty, so as to eliminate, to the extent possible, the impact of the difficulty factor and focus on the manipulation itself. This was based on participants initial distance to the true answer in Experiment 6.

There was a single manipulation of the experimental condition. This was a within-subject study where some questions were programmed to generate answers that placed participant's initial answer in the middle ('inside') of the generated answers, while other answers placed participant's answer outside of the generated answers ('outside'). In the *inside* condition, participants would see their own answers enveloped by the generated answers, while in the *outside* condition, all generated answers would be either above or below participants estimates (see Figures 3.18 and 3.19 for examples of what participants saw in either condition). The hypothesis was that the 'outside' condition would induce a greater answer change, as participants would feel more pressure to conform to the group mean. In the inside condition, the

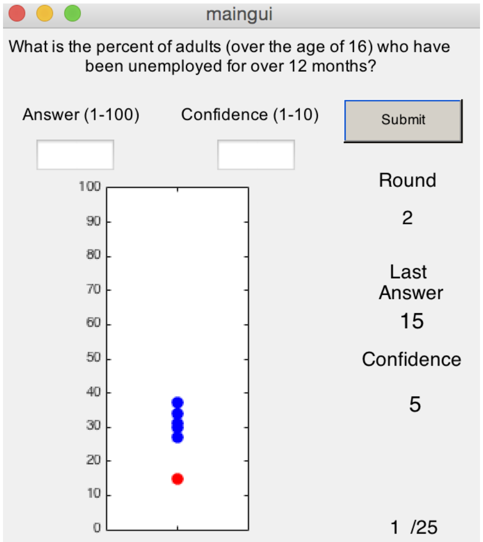


Figure 3.18: Outside condition.

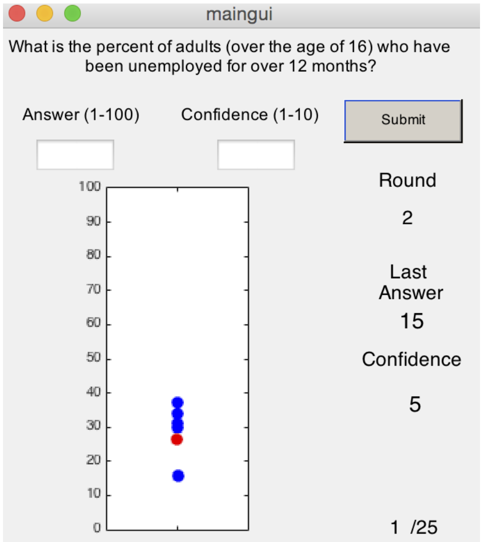


Figure 3.19: Inside condition.

opposite was expected. It was predicted that fewer answers would be revised.

3.4.2 Results

In order to determine whether participants acted differently in the two conditions, a new measure of answer change was created. It took into account whether the group mean was above or below a participant's initial answer and multiplied the answer change by either -1, if the group mean was below, or by 1 if it was above. This created a measure of change that took into account the expected direction of change, allowing the comparison of answer change between the two conditions.

The mean change of the two groups was 7.03 (SD 6.3) for the 'outside' condition and 2.1 (SD 6.1) for the 'inside' condition. Boxplot 3.20 below demonstrates the difference in behaviour between the two conditions. The outside condition has a clear positive impact on revision towards the group mean. In other words this condition had an effect of encouraging participants to move closer to the mean. At the same time, in the inside condition, the median is equal to 0, with some participants moving closer to and some moving further away from the group mean, with no coherent direction. An ANOVA showed significant difference between the two conditions, $F(1, 177) = 31.26$, $p < 0.001$, $\eta^2 = 0.15$. There were no significant differences between the two groups in terms of initial confidence levels.

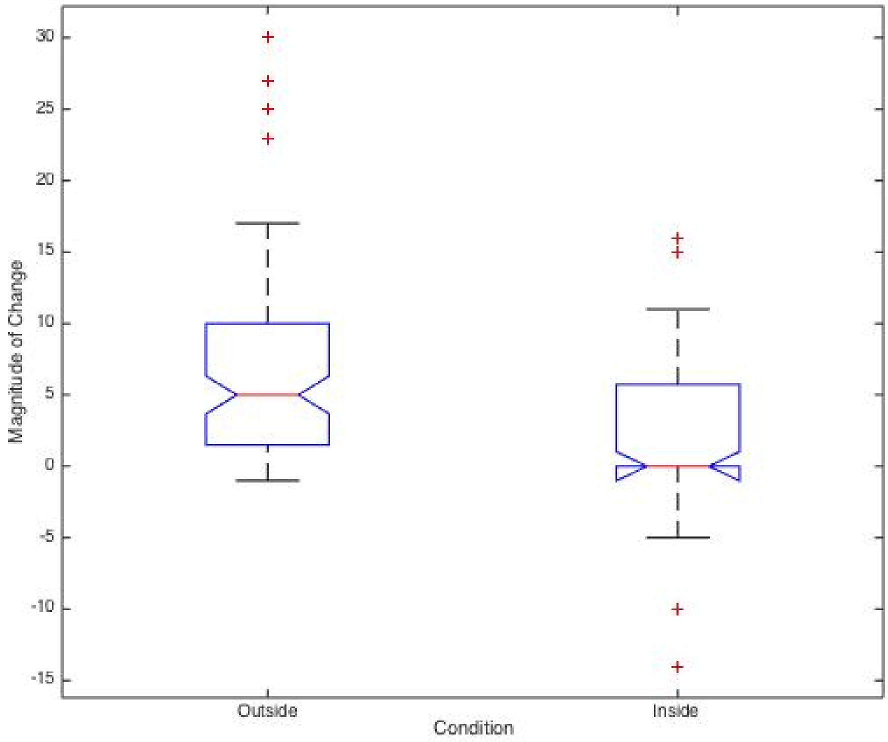


Figure 3.20: Distance from the group mean and belief revision summed across all players.

Further regression analysis indicates that a model which includes distance to the mean and condition as the interaction term is able to account better for the answer change, with both variables contributing to the explanatory power of the model. Although distance to the mean of the generated answers is the main explanatory variable (R^2=.30, F(1,177)=79.06, p<.01), the model that in addition includes condition as an interaction term is more robust (R^2=.34, F(3,175)=30.21, p<.01) (see Table 3.10 for the full regression model). This analysis confirms earlier findings that distance to the mean plays an important role in inducing revision, however, this experiment also confirms that being outside the group influences revision behaviour.

When looking at absolute revision and the absolute distance between participants' answers and the mean, the impact of revision becomes even more clear. As can be seen from Figure 3.21, in the inside condition participants who are further away from the mean are actually *less* likely to revise their answers, whereas dis-

Table 3.10: Multivariate regression model for the participant answer change based on the distance to the mean and condition as an interaction term.

	Dependent variable:
	Answer Change
Distance to the Mean	-0.674^{***}
	(0.073)
Condition	2.358^{**}
	(1.130)
Distance to the Mean:Condition	0.601^{***}
	(0.217)
Constant	-0.709
	(0.814)
Observations	179
R^2	0.341
Adjusted R^2	0.330
Residual Std. Error	6.265 (df = 175)
F Statistic	30.214^{***} (df = 3; 175)
Note:	*p<0.1; **p<0.05; ***p<0.01

tance from the mean induces more change in the outside condition. Regression analysis further confirms the statistical significance of the relationship: R^2=.04, F(2,176)=5.572, p<.05.

This finding suggests that where participants initial answers lie in relation to other answers has a substantial influence on individual revision behaviour. In cases where participants own answers were ‘outside’ the range of answers offered by others, a a significant and deliberate move towards the group mean was observed, whereas the ‘inside’ condition produced a much more haphazard and undirected change. To summarise, being outside the group caused participants to move towards the group, while being inside the group produced undirected, and comparatively small change, either away or towards the group, centering around previous answer.

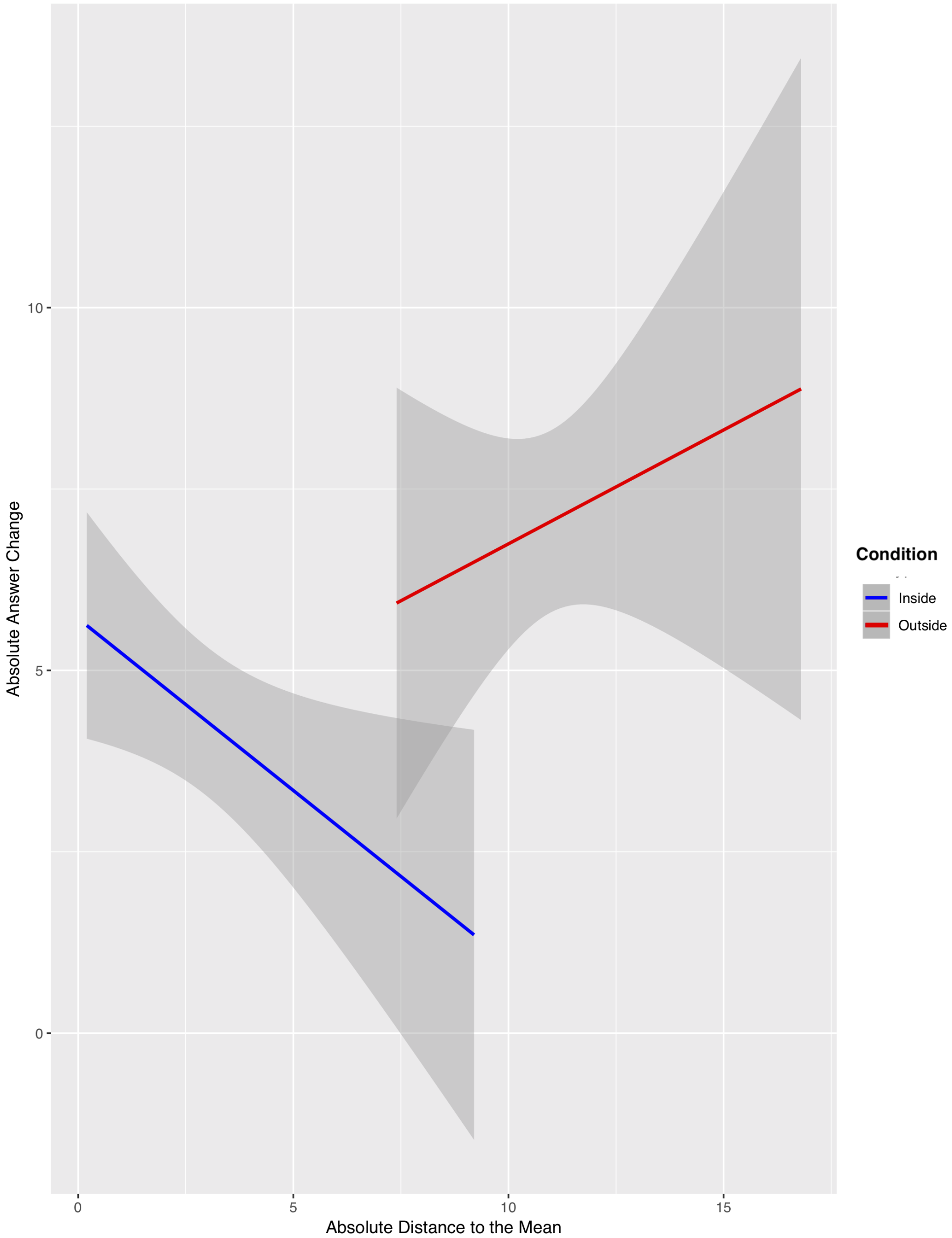


Figure 3.21: Absolute answer change plotted against absolute distance to the mean compared by condition.

3.5 Towards a Better Model

Taking the new insight about participant revision into account, it was important to revisit the original models and see if it is possible to improve existing models to make them more predictive of the Lund data. If this could be the case it would confirm the finding on a different data set and will hopefully lead to the creation of a model that is more predictive of participant behaviour and more predictive than the no-change model.

The new model is a combination of several insights gained from the two lab studies. It includes the check on whether participant's answer is inside, or outside of the group. It also predicts that participants will move towards the group mean, but will weigh their initial answer more heavily.

The updated model can be summarised as follows:

1. Revision is estimated for the first revision round only. While change does occur in subsequent rounds, these rules only apply for the first round of revision. Answers in subsequent rounds equal to the previously stated answer.

2. The decision to move towards the group mean is predicated on whether the initial guess is inside, or outside of the minimum and maximum values of the group.

3. The formula for revision is as follows: predicted change = absolute difference between the initial answer and the group mean $*\alpha$ 0.26, iff user's answer was outside of the group.

The α coefficient refers to the elasticity of revision displayed by the group. 0.26 is taken from Experiment 6 where participants moved about a quarter of the way (0.26511) towards the group mean. This factor could be changed to reflect different revision elasticity of the different groups.

Figure 3.22 and 3.23 apply the new model and compare it to the performance of the 'no-change' and 'split the difference' models on the Lund dataset. Two different revision coefficients are applied. 0.26 is taken from Experiment 6 and a the more extreme 0.1 where participants change their opinions even less, which appears to work better where fewer opinions are available.

When applied to the Lund data, the new model performed marginally better than the no change model across all groups and in small world and full world networks. The differences were small, but consistent. The overall improvement was approximately 3%. This was not a statistically significant difference however. Although, the gains are small, this is the first model to consistently perform better, in both small and full world contexts, than the no change model.

There were two α coefficients that were applied to the Lund data. The original 0.26 coefficient appears to perform better in predicting revision in the full world network group. This explanation also fits with the earlier finding from Figure 3.11, where it is hypothesised that there is less overall change occurring in the small world context. This supports the theory that much of the revision occurs in the context of a social pull where more opinions are 'forcing' participants on the extreme ends to move closer to the group mean.

The smaller 0.1 α coefficient model performs better on the small world network group. It should be noted when this coefficient is set to 0, it would converge on the no change model (participant would be predicted to not revise at all). However, when the coefficient was reduced to below 0.1 the overall error increased, indicating that the model was indeed accounting for unique behaviour. In the small world context, where participants saw few answers revision appears to be rarer, which is why the reduced coefficient performed better. This finding does stand in contrast to Yaniv and Milyavsky (2007), where just a few opinions induced great change. Although, it is possible that there was enhanced source reliability introduced by their study, by labelling other opinions as 'expert'.

Notably, the new model also outperformed the split the difference model with the same α coefficient, especially in the full world context, suggesting that the inside/outside rule adds additional explanatory power, beyond the general rule of splitting the difference, which the new model is based on. Furthermore, when the rule was applied for multiple rounds of revision, rather than only the first one, the overall error increased, suggesting that revision could only be accounted for in the first round. There was simply too little revision occurring in subsequent rounds, so any revision predicted by these models only increased the overall error.

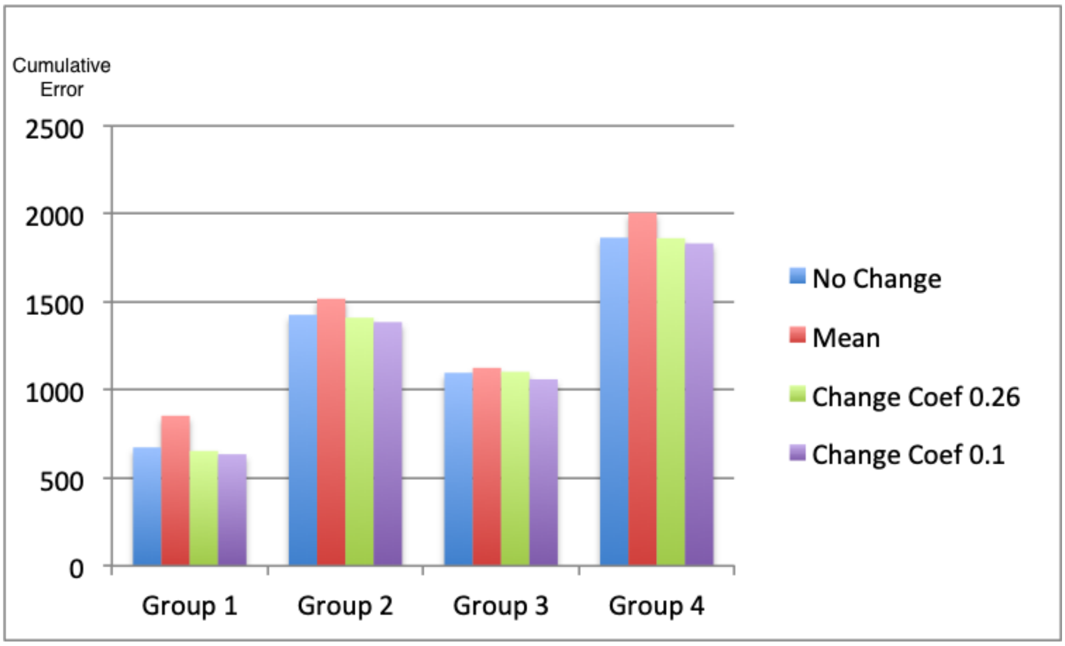


Figure 3.22: Top Four Model Comparison (Small World Network).

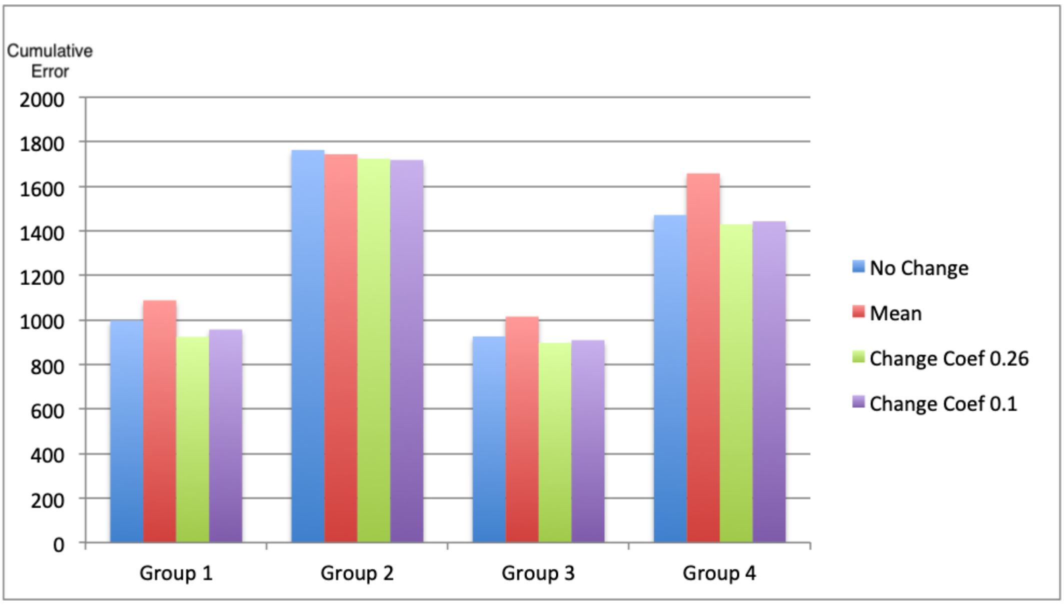


Figure 3.23: Top Four Model Comparison (Full World Network).

The α coefficient is an interesting variable. It appears to be responsive to the group size, and signifies the groups' overall susceptibility to belief change. Based on the analysis of Lund data, it appears that at least the number of opinions influence this term, but there may be other variables that can impact it. Indeed, it would be quite interested to look at α and group accuracy, as well as α and group size, source reliability, incentisivisation, etc.

3.6 General Discussion

The Lund data set provided an important empirical backdrop against which different models of human belief revision in a group environment could be compared. It also provided an interesting puzzle, as none of the existing models could accurately predict belief revision.

A number of such models, from a variety of fields were applied to the data set. These models are widely used to calculate opinion diffusion and belief revision (for examples, see Yaniv, 1997, Jackson, 2010, Lorenz et al., 2011). However, it was found that none of the current models predicted revision with great accuracy. The best performing model was the ‘no change’ model that suggests that participants do not change their answers from one round to another. However, this was not satisfactory, as it was clear that participants did chance their answers from round to round and they became more accurate as a result. The group as a whole was getting closer to the true answer, yet no model could explain where this ‘wisdom of the crowd’ was coming from.

A deeper analysis of the Lund dataset revealed that most change tended to occur in the first round, dropping off sharply and stabilising in the later rounds. Some 35 percent of all change occurred in the first round. This drops off to just under 20 percent in the second round and remains at 10 percent for the later rounds. Furthermore, participants tended to revise their answers slowly, mostly sticking closely to their original answers, except in a few rare cases.

Two experiments were conducted in order to better understand individual factors that may influence belief revision. In the first experiment the generated responses by ‘other participants’ were manipulated. Of all the factors, distance to the mean was the most significant predictor of belief revision. Including the distance to the closest answer, as well as initial confidence improved the as well model. Although, relative impact of either variable was small.

The second experiment went a step further by simplifying the design and only manipulating whether the generated answers enveloped the initial participant answer, or placed it outside. It was found that when participants' answer was outside they moved towards the group mean, further confirming our initial findings and discovering the new factor to belief revision.

A number of discoveries were made as a result of these experiments. It was found that participants tended to move towards the group mean, but were strongly anchored to their initial answer, which supports findings by Das et al. (2014b), Granovskiy et al. (2015). It was also found that individuals were strongly affected by the fact they were either inside the group, and not changing their answers, or outside the group, where change occurred much more readily.

Based on these findings a new model of revision was built and applied to the original Lund data set. The final model predicts that most changes occurs in the first round, opinions moves towards the group mean group, albeit remaining heavily weighted by the initial response, and is influenced by whether is initial estimate is inside, or outside of the group of opinions presented.

One of the most important findings to come from the work is the so-called $alpha$ factor of group revision. This factor has been shown to reflect group's opinion elastisity, or the propensity to revise answers. When applied to Lund's small world groups, it appeared to be much smaller (i.e. participants changed their answers less), while a larger $alpha$ factor appeared to better describe revision where more opinions were available to each participant in the full world groups.

Future research of individual revision will need to focus on expanding on this model and evaluating the conditions that influence the $alpha$ factor. The features of our model include: answer 'stickiness' ($alpha$), revision when outside of the group, revision towards the group mean, some mediation of the initial confidence – although this could not be tested in the Lund data, as no confidence was collected, and importance of the number of opinions (network structure). The findings suggest that individuals are particularly resistant to change and only incorporate information when they are sufficiently outside of a *sufficiently* large group to be compelled to move closer to it.

The work in this chapter builds upon a number of studies conducted in recent times, looking at modelling similar dynamics. A recent set of studies on belief revision comes from Das et al. (2014b) who modelled update opinions based on neighbour opinions and introduced important concepts of ‘stubbornness’, ‘compromise’, and ‘biased conformity’. Importantly, Das et al. (2014b) argues that individuals are quite reluctant to change their opinions and their models rely on adjusting the ratio of ‘stubborn’ nodes to recreate experimental results. They propose a biased-voter model, which has ”only two parameters, one capturing the stubbornness in behaviour, and the other capturing the conformity bias of the user” (Das et al., 2014b, p.2). More recently, Goldstone and his colleagues looked at the integration of social information and its impact on belief revision. Their best performing model had two parameters, the probability of changing an answer and the magnitude of change (Granovskiy et al., 2015).

In contrast to the Granovskiy et al. (2015) model, which provides a probabilistic term for changing the answer, and creates a stochastic model that is difficult to apply, our model is procedural in nature, and it goes further by unpacking the parameters into more defined constituent parts. The stochastic nature of their model, is elaborated by our successive studies to better determine what makes an individual change their answer at all. Our free parameter is similar to their’s and could be adjusted for a particular group. It can also be used as an indicator of group’s ’stickiness’, or unwillingness to chance their answers. Building upon the Granovskiy et al. (2015) findings, research in this Chapter outlines the reason for why people change their opinions in the first place.

There are two parts to every revision. Our model unpacks stickiness in two ways. Whether you change at all and how much you change. The Goldstone has a probability of change, while our model digs deeper and suggests a reasons for whether people will change at all. The model presented in this chapter unpacks where revision comes from in the first place, going further than the Granovskiy et al. (2015) model.

Undoubtedly, further research is required in this area. Although our model performs better than the no-change model, much remains to be understood. In particular, the relationship between source reliability, incentivisation, group size and willingness to change needs to be explored, along with the individual differences, such as personality, intelligence and prior knowledge. The dynamics of constant updating also need to be better understood. The final model assumes no revision beyond the first round, and empirical data suggests a big drop, however, some revision does occur. The nature and dynamics of this revision is beyond the scope of this chapter.

3.6.1 Impact of Feedback Loops

In terms of the feedback loop elements that impact belief revision, research in this chapter shows that being further away from the group mean and in particular being outside of the group leads to greater revision. This behaviour also leads to more accurate estimates in successive rounds. Interestingly, in a social system randomness, or indeed, the distribution of participants' initial estimates does not in itself prevent the group from becoming more accurate. Given that this randomness is natural in a sense that it is produced by the participants themselves, rather than by the system, this suggests that not all noise is detrimental to performance. In contrast to Chapter 2, where noise impeded performance, here it can be seen as neutral, or even enhancing in a sense of providing revision direction to other participants.

Lastly, it should be noted that much of the revision relies on the diversity of opinions. If all opinions were exactly the same, or quickly converged towards the mean, there would not have been the accuracy enhancing element that can be seen with subsequent revision. The temporal nature of the task actually allows participants to get better, suggesting that time is an important element of feedback learning.

3.7 Conclusion

This chapter makes important contribution to our understanding of belief revision. It builds upon the research of Yaniv (2004b), Yaniv and Milyavsky (2007) and uses the research paradigm first introduced by Jönsson et al. (2015) to better understand what strategies individuals use to revise their beliefs in a group setting. So where does Galton's 'wisdom of the crowds' come from? Why do individuals get more accurate while revising their beliefs?

Conformity appears to be an important factor in this interaction. Individuals whose answers are outside of the group tend to become more conformist, by moving closer to the mean of the group. This behaviour appears to conforms with the Hegselmann and Krause (2002) models. Individuals tend to be sensitive to how far away they are from the group and how close they are to the closest answer. However, they are most sensitive to whether they are inside or outside of the group. This particular behaviour tends to disappear once the participant's answer is corralled by other answers and revision ceases. Furthermore, full convergence towards the group mean does not occur largely because participants barely revise their answers after the initial revision round.

The entire feedback system in group revision is created by the individuals within the system. Interestingly, each individual acts as both a positive and a negative feedback loop. In the first round of revision, participants tend to readily move towards the group mean, especially those with the outlying views. This produces a positive loop in terms of the amount of revision, whereby new information is introduced into the system, further encouraging participants to revise their answers on subsequent rounds. At the same time, participants act as negative feedback loops in subsequent rounds, where the probability of each person revising drops precipitously, which leads to lack of new information, which leads to further reduction in revision. The overall revision largely disappears in the later rounds. This leads to virtual extinction of revision. Without outside influence, the negative feedback loop essentially takes over and brings the dynamic system to an end.

Importantly, individual confidence had only a minor impact on belief revision.

This stands in contradiction to Lorenz et al. (2011) who found that confidence increased over time and had a significant impact on revision. Research presented in this chapter suggests that confidence had a moderating impact on how much individuals moved towards the group mean. However, it was the group size, and in particular the effective number of opinions that mattered most, with larger effective number of opinions inducing more change among the participants. It should be noted that this finding on the impact of confidence should be considered in the particular paradigm in which is has been derived from. It is entirely possible that under different circumstances, confidence may play a more significant role in belief revision, or lack thereof.

Taking these insights, this chapter outlined a more comprehensive model of belief revision that closer reflects true human performance on belief revision tasks. This follows closely the work and findings of Granovskiy et al. (2015), but also goes further by understanding *why* participants revise their answers, rather than just *how*.

The hope is that this research leads towards more accurate models of belief revisions than are currently used across multiple disciplines to model, or incorporate human decision-making in a group setting. This research seems particularly important due to the proliferation of computational models that incorporate human decision-making in order to explain a range of real-world phenomena, from opinion dynamics to network science. More empirical research will hopefully lead to more accurate models of human performance, which could then be used as more accurate models of human agents in the relevant tasks.

Chapter 4

Belief Revision and Incentivisation

4.1 Introduction

Chapter 3 dealt with feedback loops in a system that is constructed by the interactions of participants within the system. Participants in Experiments 6 and 7, were only paid for participation and there was limited influence on participants to conform to certain goals, or beliefs. In fact, participant motivation in these experiments was all *intrinsic* rather than extrinsic (n.b. for a meta review on rewards and motivation see Deci et al., 2005). Beyond each participants' innate desire to perform well on the task, there was little outside influence on this desire. In other words, the system created by the feedback loops is in some way unrepresentative of the real world system. Often, in the real world, how one behaves is not only influenced by the system one is in, but also by the incentives: economic and social, (Bonner and Sprinkle, 2002, Guala, 2005), which also influence individual behaviour and, ultimately, are part of any social system.

This chapter builds upon the research presented in Chapter 3 by comparing participant performance under different incentivisation frameworks. Incentivisation should alter how participants behave and lead to substantively different group dynamics since it alters the very structure of the feedback system by introducing an additional source of influence.

There are two types of incentives introduced in this chapter. Individual incentivisation – where each participant has a monetary incentive to perform better, and group incentivisation – where incentive is aligned in such a way that every participant only benefits if the group does well. As can be seen from the feedback system created by providing performance incentives to individuals in Figure 4.1, each individual gains personal incentive to perform better, which may not necessarily help the group. The individual modifier acts to mitigate the actions of every participant, introducing another source of influence on the answer to be provided. It is not clear what this impact may be. Given the wealth of research that suggests that incentivisation has an important effect on performance in laboratory research (Brase, 2009, Cadsby et al., 2016, Bowen and Kensinger, 2017), and individual and group incentivisation can be found in every day life, it becomes an important empirical question of how exactly incentivisation influences individual and group dynamics.

In the previous studies by Jönsson et al. (2015), a simple incentivisation mechanism was used, whereby participants with the most accurate answers in each group received a performance bonus. This was done largely so that participants would take the task seriously. It is the purpose of this Chapter to better understand how this and other incentivisation mechanics impact participant behaviour and accuracy.

As Jönsson et al. (2015) noted:

> “Participants in the complete network condition could have fairly easily (roughly) calculated the mean answer, and, on average, if they had adopted this answer, their individual accuracies would have been much higher... However, had they done so, they would not have improved as a group at all, and missed out on the collective improvement that they did in fact obtain. In other words, over-weighting of their own opinions led participants to less accurate individual responses than they could have otherwise obtained, but it is only due to that selective weighting that collective competence improved (Jönsson et al., 2015, p.202).

If all individuals adopted the group mean after the first round, convergence would be achieved and learning would cease. Although each individual might improve their immediate accuracy by adopting the group mean, they would sacrifice all possible future learning by becoming essentially a homogeneous group. By not

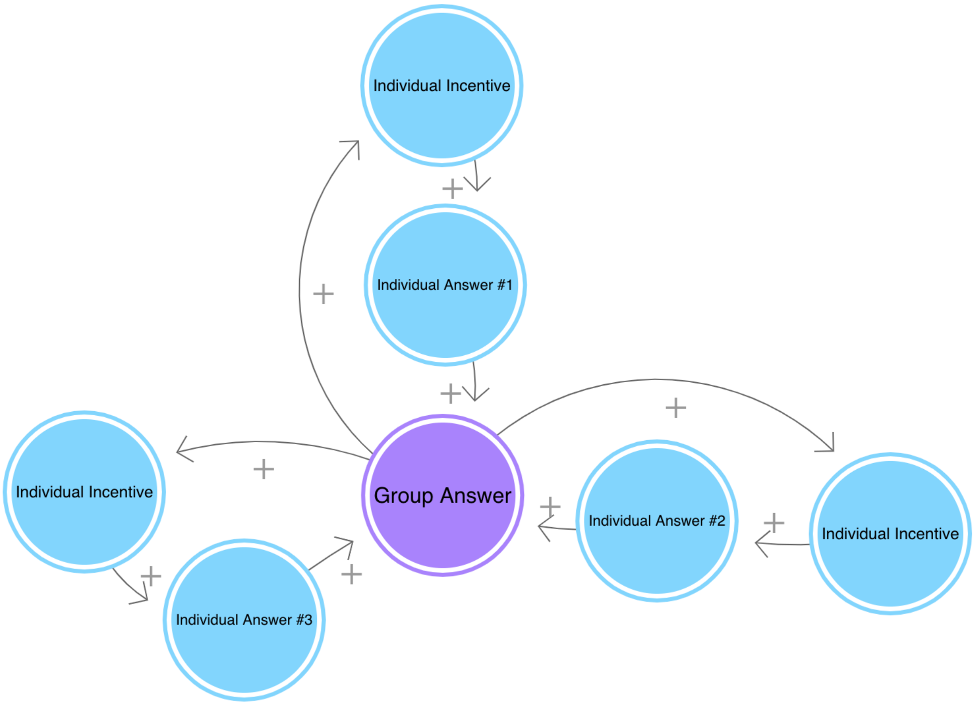


Figure 4.1: Individual Incentivisation System.

adopting the group mean, individuals sacrifice individual accuracy, but the group retains the ability to learn new information from the subsequent rounds and improve overall as a group. However, it remains unclear whether this accuracy enhancing mechanism was due to the performance incentivisation, whereby participants did not want to adopt the group mean, fearing it would lead them to a less accurate answer, or due to some other factors inherent to the group dynamics.

In addition to exploring the new incentivisation domain, replicating the larger group experiment in the UK would serve to further broaden the applicability of the models derived in Chapter 3.

4.2 Literature Review

4.2.0.1 Incentivisation

Inventivisation has been shown to directly effect participant performance in psychological experiments (Deci et al., 2005). In tasks where payment is tied to performance, participants have often displayed risk aversion (Bowen and Kensinger, 2017). Other researchers found that participants who were paid for their time outperformed those who merely received credit (Brase et al., 2006). The debate around incentivisation of performance crosses disciplinary boundaries as well.

For example, it is well established in behavioural economics that economic incentivisation is imperative in laboratory studies to bring performance to the level of the real-world systems, with direct impact on effort and performance. As Guala (2005) notes: ”monetary incentives has become de-facto a prerequisite for publication in economics journals – and, conversely, the lack of incentives is considered a sufficient condition for the rejection of an experimental study”(Guala, 2005, p. 231). In psychology, monetary incentivisation, especially for performance, is not as widespread. It is nonetheless an important feature of many experiments.

Naturally, incentivisation can also be applied outside of the laboratory grounds to job performance. As Bonner and Sprinkle (2002) suggest, a combination of monetary incentives and goal setting can have significant positive impact on performance.

It is perhaps telling that much of the literature on belief revision does not take into account incentivisation. Most of the Yaniv and Miyavsky studies do not incentivise participants beyond compensating for participation. Indeed, incentives may well alter the very behaviour one is seeking to measure; an argument that behavioural economists echo (Guala, 2005). Furthermore, in the real world, much of the behaviour is insentivised and ‘putting your money where your mouth is’ is a common belief in investment and elsewhere for identifying how serious one is about the advice one is giving (Guala, 2005). Borrowing from *mechanism design* theory, this chapter seeks to explore how altering the insentivisation structure of the group belief revision studies may alter participant behaviour.

4.3 Experiment 8: Group vs Individual Incentivisation

The purpose of this experiment is to determine how individual and group accuracy may change under different incentive schemes. Incentivising either individual, or group performance, is expected to produce the employment of different revision strategies. As discussed in the introduction to the chapter, incentivisation potentially creates an additional factor for participants to consider when revising their answers. Given that real-world systems often contain different incentivisation mechanisms the potential difference in revision strategies is important to understand. Furthermore, it is important to understand if different incentivisation schemes may have an impact on individual and group accuracy. It is hypothesised that group rewarding should lead to faster convergence to consensus and group accuracy would decline as a result, since early group cohesion constricts the available information in the system, leading to less information and less accuracy (for discussion on group accuracy see Jönsson et al., 2015).

4.3.1 Method

4.3.1.1 Participants

Participants were all students from Birkbeck College, University of London (n=37). Eight groups of five participants each were assembled in the study. Given that this was a group experiment and all participants had to be in the room at the same time, in some cases the experiment was conducted even if not all participants came on time. As a result five groups had five participants and three groups had four participants each, with three participants not showing up at the designated time for the experiment. Each participant received £5 for participation.

4.3.1.2 Materials & Procedure

The experiment was conducted using the same software and methodology as in Experiment 6 and 7 described in Chapter 3. Participants were presented with a NetLogo interface connected in a network to all other computers in the room. Once participants read the instructions, they were presented with a series of questions, which they used the interface to answer. A sample interface can be found in the Figure 4.2 below.

To answer each question participants used a slider bar to enter an initial, independent, guess. After entering their answer they would see estimates of the other participants. Following the first round, participants had a chance to revise their answers in light of what their peers entered. This procedure was repeated for three rounds for each question. In total each person answered the same question four times.

In total, each participant answered 20 different questions and provided a total of 80 answers (20x4rounds). The questions were the same as in Experiments 6 and 7 presented in Chapter 3. All questions were general knowledge questions based on the 2011 UK census data, chosen to allow participants to have an intuitive guess, but not necessarily know the correct answer. All questions had an answer range between 1 and 100. A full list of questions along iwth the correct answers can be found in Appendix D.

All participants were answering the same set of questions at the same time in the same order and had to wait for everyone in their group to finish providing their estimate before moving on to the next round and question. All participants answered the same set of 20 questions.

It was a between group manipulation study. In the individual reward condition, participants were told that they would be rewarded for individual performance. They were told that the top five individuals across all groups would receive an additional £10 payment if, individually, they had the lowest overall error, defined as the distance between the true and stated answers. In the second condition participants were told that all of participants of *the group* that achieved the highest accuracy (lowest cumulative error) would receive the £10 bonus payment. The payment amount was double the participation amount, and was considered to be a sufficient incentive to have an impact on participant behaviour.

4.3.2 Results

Given the creation of additional incentivisation components, it was important to understand whether treatment influenced individual belief revision. The first section focuses on understanding the impact of incentivisation on revision, while the second part focuses on understanding the impact of incentivisation on accuracy.

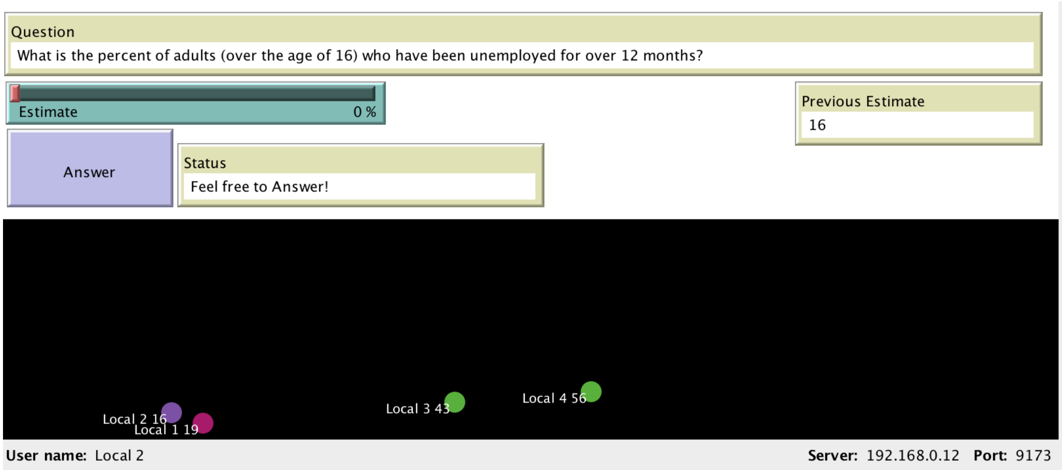


Figure 4.2: An example of a NetLogo interface. Participants communicated with each other through such an interface.

4.3.2.1 Revision

In this study there was an already familiar distribution of revision among the participants. As Figure 4.3 shows, in about half of the cases, participants did not revise their answers. This is a similar pattern that can be observed in studies in the previous chapter.

Participants revised their opinions each round. Most of the revision occurred in the first round, with the magnitude of revisions in the last round being about half of that of the first revision round. When revision is broken down by condition, there was no statistically significant relationship between the group and the magnitude of revision (t = -1.6858, df = 2211.8, p-value = 0.09198).

Table 4.1: Sum of all participants answer changes split by individual and group reward conditions.

Condition	Answer Change
Individual Reward	251.167
Group Reward	282.947

4.3.3 Revision Analysis

Regression analysis was then conducted to better understand the underlying dynamics. The aim was to confirm previous findings of revision dynamics and determine if incentivisation had an impact on the revision strategies.

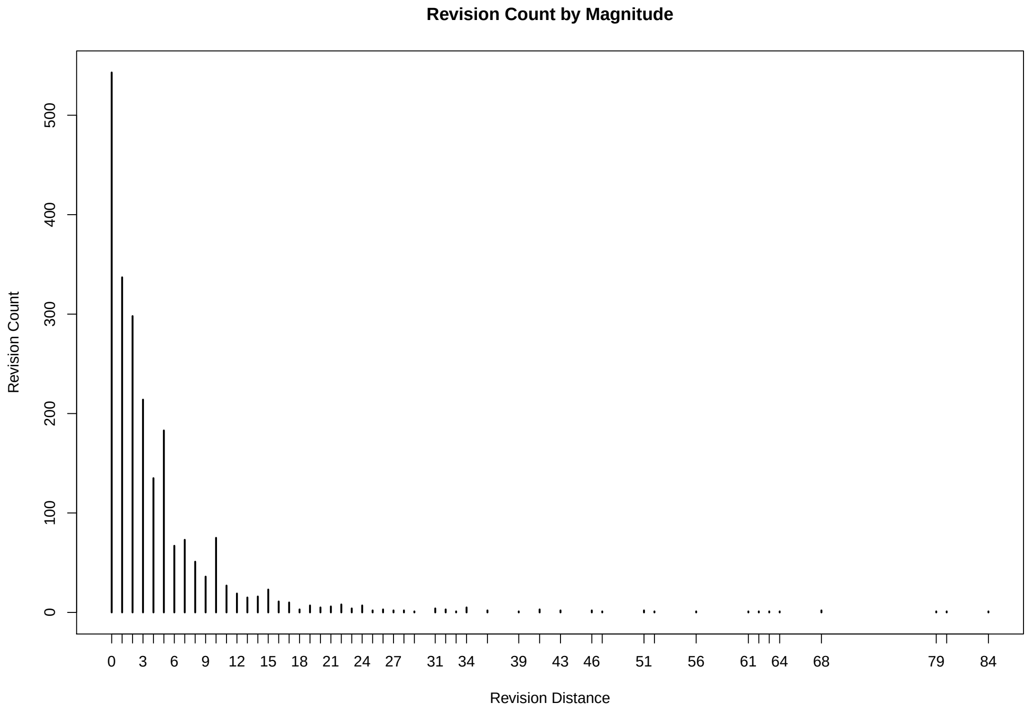


Figure 4.3: Total cumulative distribution of individual answer revision by the magnitude of revision.

4.3.3.1 Linear Regression

Linear regression was applied to the revision data, with absolute revision (absolute difference between the previous and current answers) as the dependent variable. Simple regression models taking into account, individually, condition, participant group, and group size did not show a statistically significant relationship with revision.

However, a combined linear regression model with group and condition as independent factors was statistically significant in predicting belief revision (see Table 4.2). A significant regression equation was: $R^2 = .17, F(2,34) = 3.483, p < .01$. The statistical significance of the results was confirmed when re-tested using the mean squared error instead of the absolute mean error, suggesting robustness of the finding. This suggests that condition and the group to which participants belong to did indeed have some impact on the way they revised their answers.

This model is incomplete, however. As is evident from Figure 4.4, there is also considerable difference between individual participants in terms of how much and when they revised their answers. These individual dynamics were very much the factor in the previous chapter and required hierarchical regression modelling in order to better understand the underlying dynamics. Similarly, in order to understand individual differences, hierarchical modelling was used to analyse and create a more robust model for the data. The data lent itself well to be modelled with multi-level modelling given that multiple measures were obtained from the same participants (nested within subject).

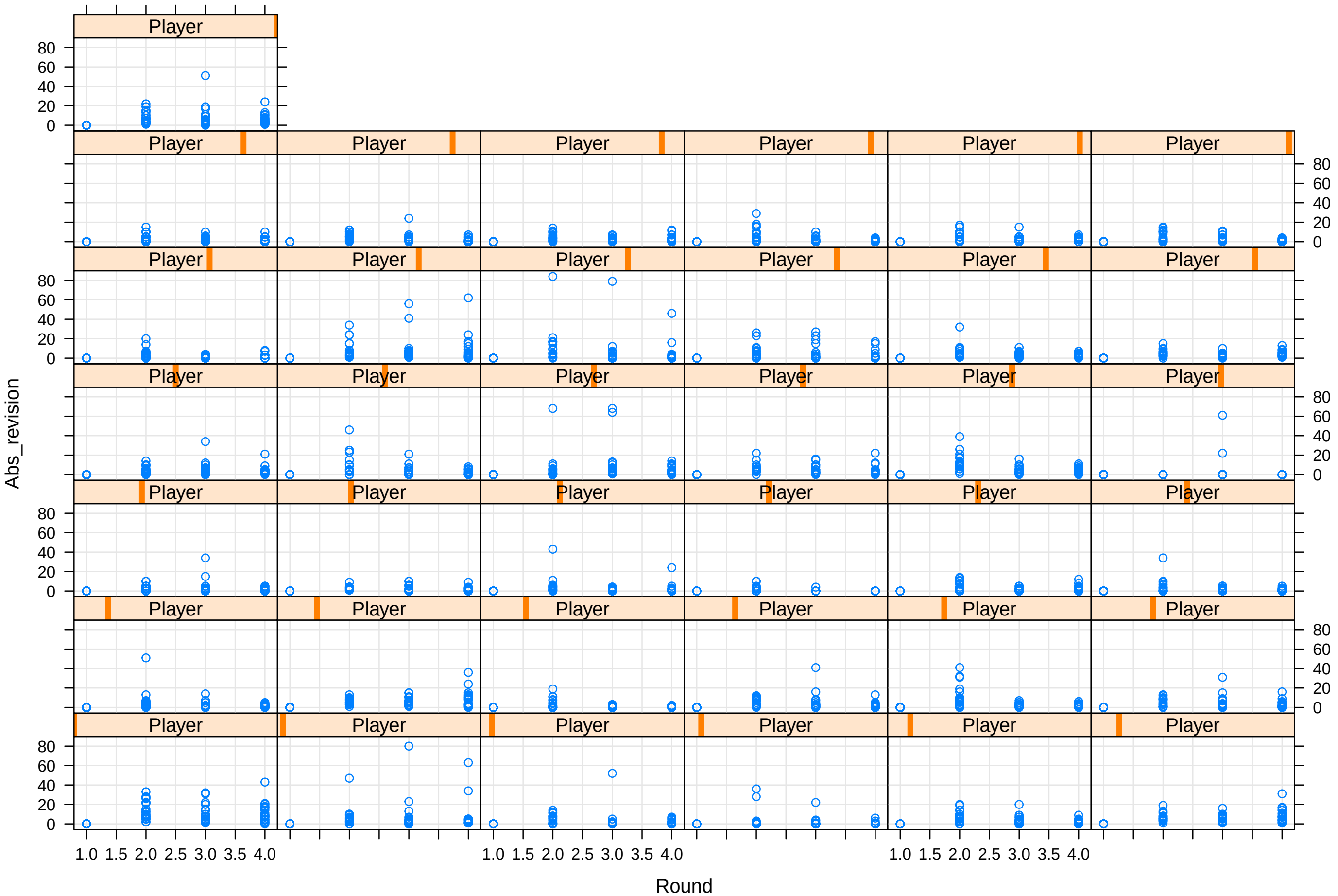


Figure 4.4: Magnitude of individual participant revision plotted again each round of revision.

Table 4.2: Regression model for revision by group and condition.

	Dependent variable:
	Revision
Group	-48.624^{**}
	(19.213)
Condition	232.815^{**}
	(90.059)
Constant	137.209^{*}
	(75.020)
Observations	37
R^2	0.170
Adjusted R^2	0.121
Residual Std. Error	128.996 (df = 34)
F Statistic	3.483^{**} (df = 2; 34)
Note:	*p<0.1; **p<0.05; ***p<0.01

4.3.3.2 Hierarchical Modelling

Hierarchical regression modelling was used in order to quantify the relationship between experimental variables and revision. The level-1 unit was the overall answer change, while the level-2 unit was the answer change by each participant. This allowed the model to accommodate individual differences among the participants. Models were built in a sequence starting with a Group variable that was found to be statistically significant in the regression analysis. Additional variables were added and an ANOVA goodness of fit test was used to compare the relative performance of each variable. Some of the variables that did not add explanatory power, such as participant confidence, were removed from the final model. Each step in building the model is detailed below.

The basic model included the Group variable, which was found to be statistically significant in the regression analysis above, as well as allowed for random intercepts for each participant. 'Condition' was added as an interaction term, which was an important factor from the linear model as well. Next, the difference to the true answer was added (represented as ‘Abs diff’) as an interaction term to the model. The last version of the model included ‘Round number’ as another interaction variable. All other variables were applied, but did not add to the model. An ANOVA goodness of fit test was used to compare the four models and determine the relative performance of each. As can be seen from Table 4.3, each model performed statistically better, with AIC being reduced in each model, meaning that these variables all had some impact in explaining revision.

Table 4.3: Hierarchical Multilevel Modelling of Factors Influencing Belief Revision.

	Dependent variable:			
	Absolute Revision			
	(1)	(2)	(3)	(4)
Abs_diff			0.076^{***}	0.074^{***}
			(0.010)	(0.010)
Round				-1.390^{***}
				(0.181)
Condition		3.880^{***}	3.539^{**}	3.550^{**}
		(1.439)	(1.440)	(1.440)
Group	-0.080	-0.810^{***}	-0.721^{**}	-0.724^{**}
	(0.158)	(0.307)	(0.307)	(0.307)
Constant	4.824^{***}	2.287^{*}	1.101	5.309^{***}
	(0.812)	(1.199)	(1.210)	(1.328)
Observations	2,220	2,220	2,220	2,220
Log Likelihood	-7,543.573	-7,540.254	-7,514.332	-7,485.145
Akaike Inf. Crit.	15,095.150	15,090.510	15,040.660	14,984.290
Bayesian Inf. Crit.	15,117.970	15,119.030	15,074.900	15,024.230
Note:				*p<0.1; **p<0.05; ***p<0.01

What this suggests is that participants were sensitive to a number of factors when revising their beliefs. In the first instance both condition and the group which participants belong to had an impact. Participants revised more in the group incentivisation condition and also were impacted by the group they belong to, but not necessarily the group size. The round of revision was also important in whether participants revised or not. Finally, distance to the true answer also had a minor

(0.076), but statistically significant coefficient. It can be concluded that treatment had at least some part in altering participant behaviour, although it was mostly in combination with the other factors outlined above.

4.3.3.3 Accuracy

But, if condition had an impact on revision, did it also have an impact on accuracy? Belief revision in a social context relies on feedback from the system created by the individuals to improve individual and group accuracy. Individual and group accuracy improvements have been a subject of some debate however. Lorenz et al. (2011) have found that the so-called wisdom of the crowds effect can be undermined by even mild social influence effects, "diminishing the diversity of the crowd without improvements of its collective error, moving the position of the truth to peripheral regions of the range of estimates so that the crowd becomes less reliable in providing expertise for external observers... boost[ing] individuals confidence after convergence of their estimates despite lack of improved accuracy" (Lorenz et al., 2011, p.9020). On the other hand, Farrell (2011) found that social influence is in fact accuracy enhancing, finding methodological errors in Lorenz et al. (2011) analysis. More recently, work by Jönsson et al. (2015) also showed increase in group accuracy as a result of social interactions.

The measure of accuracy for this study is the same as in the earlier opinion dynamics tasks presented in Chapter 3; the absolute difference between the stated and actual answers. This measure provides a direct indication as to how far away the participant was from answering the question correctly. It has also been found to be a more statistically robust measure where the error is not necessarily Gaussian in its distribution (Willmott and Matsuura, 2005, Chai and Draxler, 2014). Given that there is no Gaussian distribution assumption of the error term in this study, the mean absolute error provides a good measure of individual and group performance, and is easy and transparent to communicate and analyse. Although, absolute difference is used in most of the reporting below, in cases where statistical significance is observed, squared error was also used to verify significance, or lack thereof.

As can be seen from Figure 4.5, in aggregate, participants in the individual

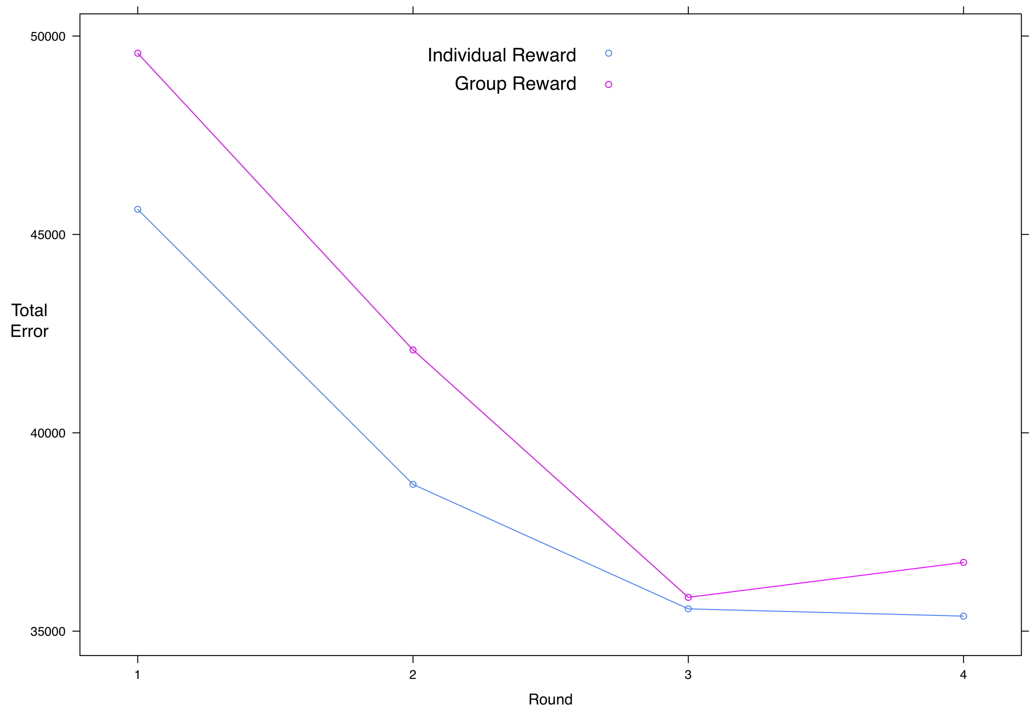


Figure 4.5: Cumulative error for each round of revision broken down by the individual and group reward conditions.

reward condition were more accurate in their estimates. Similarly to the observed behaviour in the Lund study described in Chapter 3, the overall accuracy continually increased with each round of revision. Both conditions performed more or less equally on the task and a t-test on the accuracy by condition confirmed that there was no statistically significant difference between the two conditions($t = 0.20158$, $df = 34.98$, $p\text{-value} = 0.8414$). Therefore, condition did not have statistically significant impact on participant behaviour or accuracy.

4.3.4 Discussion

Altering incentivisation did not appear to cause differences in group accuracy, or in the way participants behaved. Although *individually*, participants in the group reward condition were more accurate and four of the top five participants were from this condition, the relationship was not statistically significant when extended to all participants.

Hierarchical regression was used to analyse the factors that may have induced belief revision. This was done because of the high variance of change exhibited by individual participants. It showed that experimental condition had some impact on participant behaviour, although other factors had a stronger impact. Overall, four statistically significant factors emerged: distance to the true answer, round number, condition and group.

The fact that distance to the true answer positively impacts revision, suggests that as participants get closer to the true answer, they are less likely to revise their beliefs. Overall, participants became more accurate with each revision, therefore it is not surprising that they revised less. This drift towards the correct answer is well documented (see Galton, 1907, Jönsson et al., 2015, Hill, 1982), and results here continue to support the finding that groups as a whole become more accurate. This interplay also suggests that individual participants get more accurate, which causes the rate of change to slow down. This means that as participants get closer to the true answer there is less of a chance for them to become less accurate since there is less information in the system overall.

Round number negatively impacts revision, which is also corroborated by the results. In all of the previous studies discussed in Chapter 3, it was found that participants change their answers less frequently and by a smaller magnitude with each revision round. This holds true for this study as well. This further contributes to the observed dynamic of slowing change over time. The first round of revision is the most significant, as participants are much more likely to change during this round.

The incentive condition did not appear to have a particularly strong impact on revision. Condition did not have a statistically significant impact on accuracy, or revision on its own. However, when combined with other factors, it was a statistically significant factor in the multi-level the regression model, suggesting that condition had at least some effect on belief revision, but only in combination with the other variables. Its inclusion improved the predictive power of the model, suggesting that at least some participants did in fact react differently to incentivisation.

Finally, the group that participants belonged to also had a statistically significant impact on behaviour. Interestingly, it was not the group size, but the actual group membership that was statistically significant. Since participants self-selected the time slots, it was impossible to control for this variable.

These findings broadly conform with the revision factors discussed in Chapter 3. Indeed, there was significant overlap between this study and the data observed in Lund. As far as replication goes, this study confirmed many of the findings; namely that revision occurs most often in the first round and groups get more accurate and revise less with time. This study also found that incentivisation had minor impact on revision and virtually no impact on group accuracy.

4.4 General Discussion

As a final step in understanding belief revision dynamics in this study, it was important to apply previous models of belief revision onto the new data obtained in this study. The goal was to see if previous findings would hold and if the model created in Chapter 3 would perform better than the no-change model in explaining participant behaviour.

Two models were applied to the data to test which strategies participants used to revise their answers. Both models came from the analysis done in Chapter 3. The first model applied was the no change model, which was originally the best performing model. This model predicts that participants will not change their answers at all. The second model used was the custom revision model developed in Chapter 3, which was shown to explain revision better than the no change model. As a reminder, the custom model developed in Chapter 3 assumed that only the outliers – two of the most extreme answers – would move some distance towards the group mean, taking their initial answer as the anchor point. In the model the magnitude (distance) of the move is determined by the change coefficient (0 representing no change and 1 representing the adoption of the group mean). The experimental findings show that a chance coefficient of 0.26 was representative of the behaviour in the Lund study. Its application on a different data set is an important step in validating its applicability beyond the very specific data set in which it was developed.

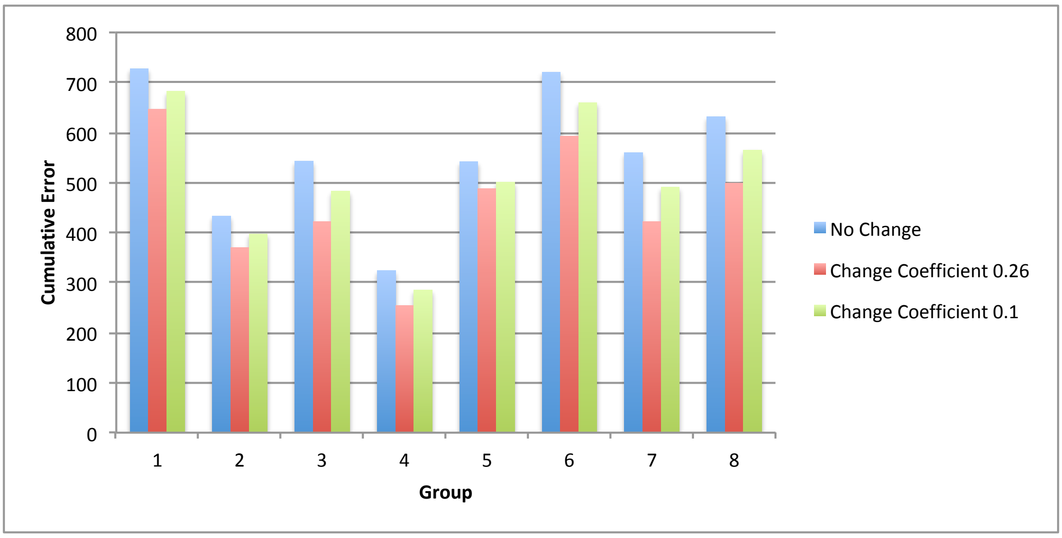


Figure 4.6: Comparative Model Accuracy for the first round of revision.

The methodology of applying the models is the same as in Chapter 3. Each model predicted what the next answer would be, had participants used the same rule as the model. For example, the no change model assumes participants do not change their answers from one round to another and that the next answer will be the same as the previous answer. The fit of the model is determined by the the absolute difference between the model predicted answer and the actual participant answer. A model with the lower difference has a better *fit* to the actual data and better represents the observed participant behaviour.

As can be seen in Figure 4.6, in the first round of revision, the custom model outperforms the no change model in all groups. There were two coefficients 0.26 and 0.1 that were tested in Chapter 3. When applied to the data gathered in this experiment, the model with the stronger coefficient of 0.26 outperforms both the no change model and the model which predicts less change with the 0.1 change coefficient. However, a t-test between groups shows that this difference is not statistically significant (t = 1.5059, df = 14, p-value = 0.1543). Perhaps a larger sample size would make these findings more robust.

Given that this study included an incentivisation condition, it was important to look at the aggregate performance of the models by condition. Figure 4.7 shows that the change coefficient models outperformed the no change models for both condi-

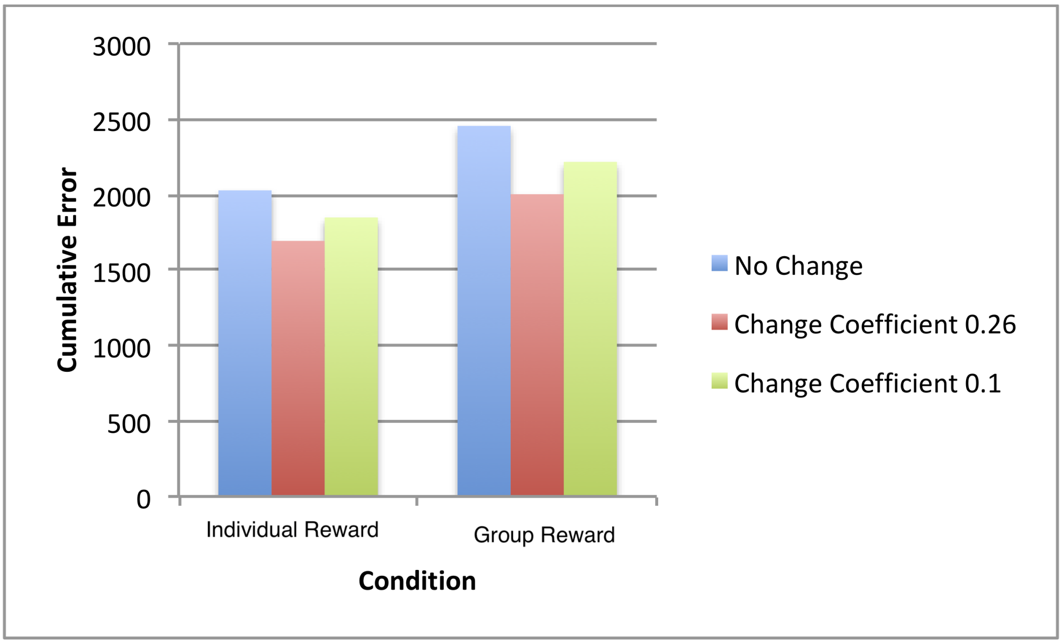


Figure 4.7: Comparative Model Accuracy by condition.

tions. However, a chi-squared test did not show statistical significance, suggesting that a larger sample size may be necessary to make these results more robust.

4.5 Conclusion

The clear relevance and applicability of the custom change model in the UK suggests its applicability beyond the original data set. The fact that it is able to outperform the no-change model for an entirely new data set is promising. This further suggests its usefulness as a model to understand belief revision that is more closely matches actual human behaviour on such tasks. As a first of its kind, cross-country application of the models to understand belief revision, this analysis and findings provide an important contribution towards the understanding of belief revision dynamics.

The fact that the custom model better explains behaviour in the group incentivisation condition is also important. Monetary incentivisation, and its relatively modest impact on revision, suggests that monetary rewards for performance do not lead to more accuracy or greater revision. It does, however, lead to greater conformity, particularly among the outliers. While there was no statistically significant relationship between individual and group incentivisation and performance on the belief revision task and participants did not answer questions more accurately, or, indeed revise their beliefs more as a result of the different incentivisation schemes employed in this study, they did revise differently. Participants in the group incentivisation condition were much more likely to revise towards the group mean, suggesting that incentivisation lead to greater conformity. It is notable then, that greater conformity did not lead to increase in accuracy.

When talking about feedback systems, participants in this study sought much greater conformity than in the Lund study. It appears to be be at least partly due to the incentivisation introduced in the study. Notably, the difference between the two types of incentivisation (individual vs group) was not particularly big. But, the fact that overall participants revised more towards the group mean, than in the Lund study is important. It is possible that incentivisation in general alters the system in such a way that greater conformity is induced. It could be due to fear of blame, or a desire to be more helpful to the group that leads to this behaviour. Further research will be required to better understand what impact incentivisation has on behaviour, however, the research here represents the first, but very important step in understanding its impact.

Chapter 5

Rank Aggregation and Belief Revision

5.1 Introduction

The research in the second part of this thesis thus far has dealt with absolute values and questions which have a singular answer. This chapter extends this analysis by looking at ranks and ranking as a sub-domain of belief revision and group accuracy. This chapter looks at a special application of feedback learning in the context of rank aggregation. Rank aggregation is a particularly interesting application for belief revision, given that ranked information is often used in computer science and information retrieval. It is the goal of this chapter to explore what strategies are most beneficial when aggregating ranks and then understand what people actually do.

Chapters 3 and 4 have demonstrated repeatedly the effect of group accuracy increase under different circumstances. Individuals create dynamic systems through repeated interaction with others which leads to increased accuracy. In the previous chapter incentivisation was used to alter the underlying structure to introduce greater conformity among the participants. This chapter extends this research into the new domain of rank aggregation.

The problem of *rank aggregation* where ranked lists from a diverse set of "judges" are combined into a single "consensus" ranked list, is an active research area in computer science. Rank aggregation has found successful applications in meta-search (Dwork et al., 2001, Renda and Straccia, 2003, Fernández et al., 2006), crowd-sourcing (Niu et al., 2015), and recommender systems (Baltrunas et al., 2010). Aggregation can take many forms. In its simplest form aggregation can mean taking the most popular list (i.e. list that appears most often). The most complicated forms of aggregation can include calculating the distances between the lists in order to come up with an aggregate list that most closely matches all other lists. This requires computing new measures for rank distance (i.e. how far away each list is from each other) and then calculating the one list that is the closest to all of the others. Computational complexity is increased as a result.

Although extensive studies have already been conducted on this topic by computer scientists, these largely concern only the algorithmic issues, i.e., how to produce the "optimal" ranked list, without questioning the very concept of *optimality*. Computer scientists working on the topic normally use the more computationally intensive methods to create the aggregate list, in order to optimise a particular project. Dwork et al. (2001) for example were looking for aggregation measures that would produce, "rank aggregation techniques that can effectively combat 'spam,' a serious problem in Web searches"(Dwork et al., 2001, p.613).

5.2 Literature Review

Rank aggregation is a study in condensing complex information into its most common and appropriate form. It is of interest to a diverse group of research communities dating back to the beginning of the 20th century. For example, rank aggregation was an important concept in information retrieval dating back to the middle of the 20th century. Rank aggregation is key in information retrieval as information is ranked and displayed based on its relevance to the user (Kumar and Vassilvitskii, 2010, Brancotte et al., 2015). Since all information needs to be ranked, various rank aggregation and rank distance measures provide for a way to do so. Some are more optimised than others to deal with the various complexities that arise from dealing with large volumes and diversity of information.

In bioinformatics rank aggregation is used for a variety of purposes. For example to aggregate gene lists, which are essentially lists of DNA strands that need to be compiled to determine the 'common' strand (Kolde et al., 2012). In the wider field of biology, which often uses clustering techniques for micro-array data analysis to discover relevant biological groupings, rank aggregation is used to reconcile the different results of the clustering methods (Pihur et al., 2007). Thus, biology relies heavily on rank aggregation to solve some of its most complex problems.

Arguably, the most prevalent use of rank aggregation today can be found in website indexing – a field similar to information retrieval that deals with displaying relevant search results in the most optimal way; defined as the order most relevant for a given query. Given that there is a proliferation of website indexing services, rank aggregation is most commonly used in "building meta-search engines, combining ranking functions, selecting documents based on multiple criteria, and improving search precision through word associations" (Dwork et al., 2001, p.613).

Interestingly, aggregation has always been firmly in the domain of mathematics and computational sciences, with little regard for how humans aggregate ranks. Since ranks are everywhere, from football table leagues, to voting, it is quite an interesting empirical question indeed. However, before we launch into testing, it is similarly important to understand the performance of some of the more common aggregation techniques under various conditions and for that the use of modelling is required.

5.3 Modelling

The purpose of this section is to simulate various aggregation techniques in order to better understand how well the different methods perform under different conditions and to determine the most salient variables that have the greatest impact on the relative performance of the various methods.[1]

Modelling provides a way of analysing the task, and providing better understanding of the relative performance of the different aggregation methods. Given

[1] The work in this section is based on the original concept paper by Stephan Hartmann and colleagues at LMU.

that these methods are not normally compared against each other in a standardised way, it is the aim of this section to establish their relative performance to each other. It is not designed to model human cognition, or to make any claims about human cognitive processes in rank aggregation.

In the first instance, a theoretical simulation was developed to test the accuracy of the various rank aggregation methods. This simulation can be thought of as a hypothetical meta-search engine, which aggregates the results from various other search engines that generate search results. Given that rank aggregation is heavily used in information retrieval and website indexing this seems to be a particularly fitting thought experiment.

Imagine there is a new super web indexing website, which queries Google, Bing and other search engines on the same keyword and then aggregates and returns a single list. Importantly, lets imagine that there is indeed a single list that is ‘the best’. This list would contain the items in just the right way. The first item would be most relevant to the search term, followed by the second, third, and so on. The objective of the meta-engine is to query various search engines, combine their results and return a single list that would rank items in the best order possible. The assumption that there is a single list superior, or *true* list, is quite important here, as it allows us to test the various aggregation methods against this *truth*. The goal is to see which of the aggregation methods produce the optimal list most frequently and under what conditions.

Given a set of m items (e.g., web pages), we consider n ranked list of them, $\{r_1, \ldots, r_n\}$, each of which is given by a judge (e.g., search engine). One, and only one, of the possible ranking orders (permutations) r_* is deemed to be true (the ‘best list’).

Each judge is characterised by his “competence” which is defined as the probability of providing the true list. Imagine that each search engine would use different algorithms to generate the list and that the algorithm would have material impact on the ability of the engine to produce optimal lists.

Once each judge generates a list, the simulation takes these generated lists and aggregates them into a single list using one of the rules outlined further in this section.

Unlike previous work on this subject, such as Fernández et al. (2006), for each item in a list only its rank position vis-a-vis other items is known. No other properties are available. Indeed, only the the relative order of the items is known, no information on how or why each judge chose that particular order is revealed.

It is often impossible or unrealistic to obtain the scores of individual items and only their relative positioning to each other is available (Dwork et al., 2001, Renda and Straccia, 2003). This is true of many of the real-life problems such as gene lists (Kolde et al., 2012). More importantly, a wealth of psychological research suggests that in many domains, humans represent faithfully only ranking order information and more detailed information is unhelpful (Stewart et al., 2006). We simulate the relative rank order and not the underlying information about each item on the list.

For the sake of simplicity, modelling considers that each judge will produce a complete list and no ties are possible. So when ranking items it will rank all of the available choices and will rank them relative to each other in such a way that each item will occupy a unique position. The choices available will be uniform for all of the judges. Furthermore, it is assumed that every judge in the model has the same level of competence $c \in [0,1]$. Finally, when a certain rule produces multiple lists that are equally optimal, one of them is selected at random to break the tie. This work could be generalised straightforwardly in the future by relaxing these assumptions and constraints.

The following rank aggregation methods and rules have been proposed in previous literature and are widely used in practice, and will be modelled here:

- **majority**: the consensus list is just the ranked list that appears most frequently (Dwork et al., 2001).
- **average**: the consensus list is generated by ranking the items according to their average rank positions, which is essentially same as the Borda's count (Dwork et al., 2001).

- **Spearman**: the consensus list is the one with the minimum sum of Spearman's footrule to entire ranked list. Spearman's footrule is defined as the total number of displacements needed to achieve parity between two lists (Diaconis and Graham, 1977)

- **Kendall**: the consensus list is the one with the minimum sum of Kendall's tau to the ranked list. Kendall's tau is defined as the total number of inversions required to achieve parity between two lists (Kendall, 1938).

- **Kemeny-Snell**: the consensus list is the one with the minimum sum of Kemeny-Snell distance to all the given ranked list. The Kemeny-Snell (KS) distance is similar to Kendall's tau, but more robust when dealing with ties (Heiser and D'Ambrosio, 2013).

While the first two methods are simple and easy to compute, the other three are based on distance measures and have high computational complexity. It has been shown that finding the optimally ranked list based on Kendall's tau (known as the Kemeney optimal aggregation) is an NP hard problem with just four full lists (Dwork et al., 2001). This naturally places a limit on the size of the lists, which we will explore in more detail further on.

The research question is then: "which rank aggregation method would produce the most accurate results?" Accuracy is defined as the frequency with which the consensus list is returned by a rank aggregation method and is indeed the true list. Each time rank aggregation returns a list, which is then compared to the true list and if the two match it is considered to be correct and if not, the opposite result is recorded. At the end, it is possible to judge each aggregation method on its ability to produce the true list under different conditions.

5.4 Computer Simulations

The simulation was coded in *R*. It is programmed to generate a number of lists from different judges and uses different aggregation methods to determine the list reflective of the group, or the 'average list.' The generated consensus lists are then compared with the true list to calculate accuracy, which is used as the "performance" measure for the aggregation method. This procedure is repeated while sampling and increasing number of judges with each iteration. In order to smooth out effects of randomness, bootstrapping is performed at each number of judges and the average value is taken. Therefore, each set of judges was simulated multiple times, before adding additional judges.

The simulation had a number of free parameters that could be altered:

- **list size**: number of unique items in a list
- **competence level**: individual probability of picking the correct list
- **aggregation method**: methods of aggregation described above
- **number of simulations**: a number of repeats of the same simulation with the same conditions to smooth out any noise due to randomness

We began with a list size of 4, meaning there were 4 unique items in the list that could be arranged in order. With no ties there were 24 possible permutations. In the simulation k groups consisting of n number of judges would draw a single list from the full list of permutations. Using one of the aggregation methods, a single list would be selected for each group as the aggregate product, and then compared to the true list. Each group of judges would be re-sampled a number of times to boostrap the results to get a smoother result. Scores reported below are the average results sampled over multiple trials for the same group.

5.4.0.1 Error Models

One important consideration in the study was the underlying error model that governed a judge's probability of picking the wrong list among all possible permutations. The error distribution was important since it governed the probability of picking any of the permutations, outside of the true list. We modelled each judge to have a competence measure which reflected the probability of picking the true list. The rest of the probability was distributed among the remaining possible choices. We modelled three different error distribution models to reflect different assumptions about each judge's ability to pick the various lists. These include: the linear-decay model where the probability of picking the correct list diminishes with the distance from the true list; fastest-decay model where the selection probability drops rapidly as a function of distance so only the closest lists to the true list are selected; and, none-decay, all lists have the same probability of being selected. Each model is described in more detail below.

The first modelled assumption was that judges know something about the domain in question and as such, the probability of picking a wrong list is likely to be an inverse function of the distance from that list to the true list. Without the loss of generality, we used the Kemeny-Snell distance measure $d(\cdot,\cdot)$ to determine the probability of a given list being selected. It is defined as follows:

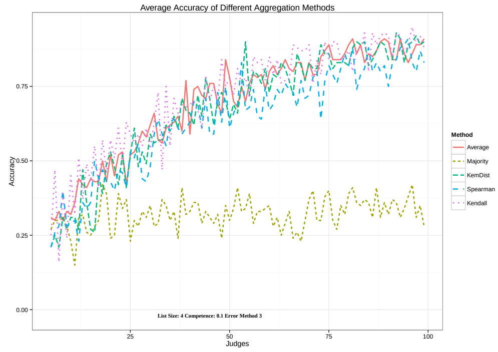
$$\Pr[r_i] = \begin{cases} c & \text{if } r_i = r_*, \\ (1-c)\frac{1/d(r_i,r_*)}{\sum_{j\neq *}\left(1/d(r_j,r_*)\right)} & \text{otherwise.} \end{cases} \tag{5.1}$$

In effect, lists that are closer to the true list would be more likely to be drawn than the lists further away. This was called a *linear decay* error distribution since the probability of picking lists further away from the true list decays with distance.

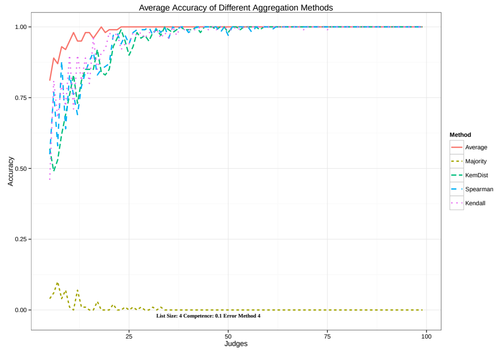
We wanted to see relative performance of the various aggregation methods as the number of judges increased. For all results, a constant competence level of $c = 0.1$ was maintained, which meant a 10% chance for a judge to pick the true list r_*. We selected the simulation range from 5 to 100 judges. After running several different simulations we produced a number of interesting and insightful results, summarised in a series of graphs below (see Figure 5.1).

5.4.0.2 Results

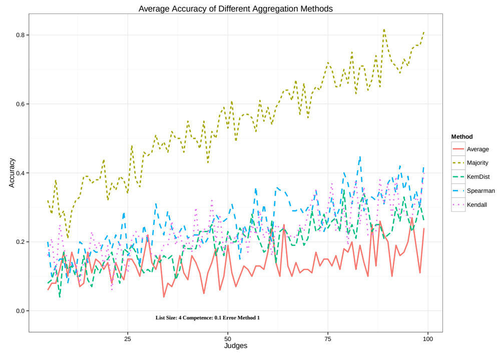
Each graph in Figure 5.1 summarises a simulation, which models sampling by 5 to 100 judges (x-axis) and demonstrates the relative performance of the various aggregation methods, as defined by their accuracy (y-axis). Figure 5.1a summarises the linear-decay error model. As can be observed from the graph the majority rule performs significantly worse than the alternatives and does not increase in accuracy as the number of judges increases. This aggregation method produces correct result



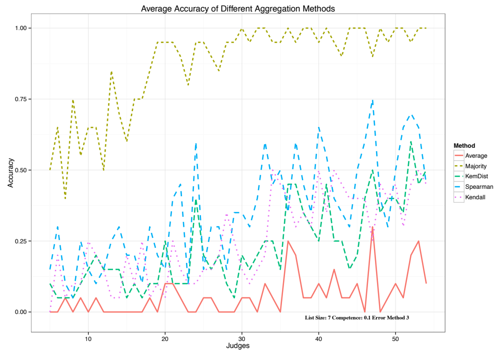
(a) linear-decay error model (list size 4)



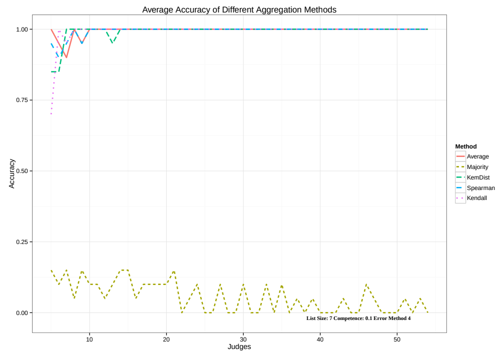
(b) fastest-decay error model (list size 4)



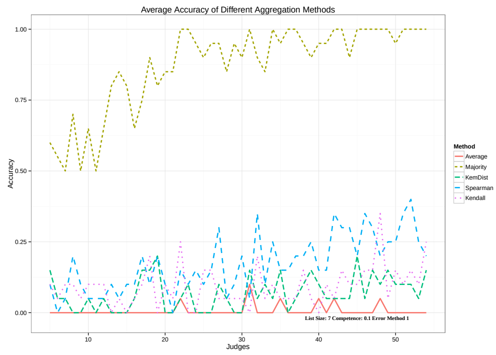
(c) none-decay error model (list size 4)



(d) linear-decay error model (list size 7)



(e) fastest-decay error model (list size 7)



(f) none-decay error model (list size 7)

Figure 5.1: Comparison of aggregation methods plotted over multiple conditions.

just about a quarter of the time. On the other hand, the other four methods perform similarly to each other and their accuracy increases as the number of judges goes up.

It is important to note that the Kemeny-Snell aggregation method does not perform significantly better than the other distance-based methods, despite it being used as the distance measure in the error distribution! Furthermore, averaging, which is a very simple method (both computationally and cognitively), performs at least on par with the more computationally complex distance-based methods.

A minor comment regarding the competence measure is due at this point. It was found that once the probability of picking the true list was higher than 20%, there was a quick rise towards perfect accuracy of *all* methods. This is true because beyond this threshold the probability of picking the true list was higher than the probability of picking any other list, especially for the larger list sizes. Given that there is no differentiation between the performance of the various methods at or above this level of competence, there is little analysis that can be performed. Therefore, the competence level was kept well below this threshold in order to understand how robust different rank aggregation methods would be under the more challenging condition of lower individual competence.

In the *linear-decay* error mode, with a 10% chance of picking the correct list, the remaining 90 points are divided among the remaining 23 lists, based on their distance to the true list. That means alternative lists, on average, have a 4% (90/23) probability of being picked. However, due to the non-linear nature of the error distribution, the probability is more than double for lists closer to the true list, exceeding the probability of the true list being picked. This effect is demonstrated by the low performance of the majority list. The true list simply does not come up more often than any of the other lists. As the number of judges increases, the overall accuracy of the majority model does not increase. However, the performance of the other models does increase. This suggests that models that look for averages, or take into account distances between the lists and look for a common list, are more adaptive in the circumstances where the probability of picking the true list is not highest.

Linear decay, as represented by the equation above (5.1), is just one way of converting the underlying KS-distance to the targeted true list into a probability of erroneous list selection. Actually any monotonic decaying transformations – such as an exponential decay – of those distances could be utilised to pick the non-true lists. To generalise the results, two extreme cases of monotonic decay of distance were considered: at one end (*fastest-decay*), the selection probability drops so rapidly as a function of distance that only the closest lists to the true list stand a chance of being selected; at the other end (*none-decay*), the selection function is flat and the lists of all distances are equally likely to be selected. We have examined both of these extreme cases.

First, let us consider the case where only the lists closest to the true list had a non-zero selection probability (with each list at that distance equally likely to be picked). The results of this error distribution are represented by Figure 5.1b. From the results of the simulation it is apparent that the majority method plummets almost immediately towards zero accuracy. This is due to the fact that the competence level, i.e., the probability of picking the true list (10%), is significantly lower than the probability of picking any of those closest lists (which is 30% in this example as there are three lists that are similarly close).

The other methods appear to quickly move towards perfect accuracy. All 4 lists that could be picked are very similar and the average list of the 4 is indeed the true list by definition. The averaging aggregation method, in particular moves quickly to perfect accuracy and outperforms the other methods. There also appears to be little difference among the other distance-based methods and they all behave similarly to the average method.

Secondly, we consider the case where a judge is equally likely to pick any of the wrong lists, regardless of their distance to the true list. The *none-decay* model equally distributes the error among all possible lists.

The results of this simulation stand in stark contrast to the other two simulations. The majority method performs significantly better and improves with the number of judges, which is exactly reverse of what was observed in the earlier simulations.

Although the observation was initially quite surprising, the explanation is fairly intuitive. Since the probability of picking the true list is 10%, the remaining probability would be distributed evenly over 23 other possible permutations, which leads to only 3.9% per permutation. This means that the true list is more than twice as likely to be picked, as any other list. Therefore, the ranked list occurring most frequently is almost guaranteed to be the true list, and the majority method would always perform the best.

Just as importantly, the other aggregation methods appear to falter at this stage. Although there is some improvement along with the increase in the number of judges, the accuracy stays well below 0.5, even for groups with 100 judges. Notably, the average method performed the worst, while the Spearman method performed the best among the three distance-based methods.

5.4.0.3 Larger Lists

How general are these findings? To explore this, we applied the same simulation to larger list sizes. Beyond a certain list size, the number of possible permutations becomes unwieldy and computationally intensive. However, we were able to extend the simulation to list size of 7 and still compute the results. At this list size, there are 5,040 possible permutations.

We tested the performance of the various aggregation methods on the larger list size, using the same methodology as above; applying the three different error distribution models on the larger list size. The results are summarised in Figures 5.1d, 5.1e, and 5.1f. In order to be comparable, the competence level remained the same. Given the larger sample space, this would have obvious implications on the performance described below.

The most important differences can be found in Figures 5.1d and 5.1f. It is not surprising that the majority model does better, given the relatively high competence level. However, what is surprising is the terrible performance of the averaging model. Unlike its performance in the list size 4 simulations, with larger simulations this method begins to falter, especially as more judges are added. As Figure 5.1d shows that the other methods begin to outperform the averaging method. This trend was observed for list sizes 5 and 6 and is most pronounced in list size 7, which suggests that if the list size is further increased, the relative performance distance would only continue to increase, making averaging impractical for all, but the smallest list sizes.

5.4.1 Discussion

A few key insights emerge from our modelling efforts. The first and most important one is that there appears to be little benefit of using computationally-expensive distance-based methods to conduct rank aggregation for smaller list sizes. However, as list size increases, along with the possible number of permutations, averaging falters, while the more computationally intensive methods continue to be robust, with their effectiveness increasing with the number of judges. The number of judges also appears to play an important role. As the number of judges increases, the accuracy goes up across most conditions. This suggests that extra information is picked up in larger lists and robust aggregation methods 'collect' this information as a result of a larger number of overall guesses in the system. The only method that appears to not improve with the number of judges is the averaging method. It can be concluded that averaging is a good method when the list size is small, but as the space increases, as a result of larger list size, this method does not work.

The last notable finding is around the majority method. There is a rather binary nature to this aggregation method. It appears to work well when competence is high, but fails miserably when the probability of picking the true list is less than those of the other lists. It is a rather simple heuristic method of aggregating and its success and failure is evident in the modelling. If one believes that judges are competent, there is no better way to aggregate their opinions (see 5.1c, 5.1d, 5.1f) but to go with the majority, but in virtually every other situation this method fails (see Figures 5.1a, 5.1b, 5.1e). It also does not increase in any meaningful way with a larger number of judges suggesting that groups, or larger 'crowds', do not produce more accurate lists with this aggregation method. We therefore, must look elsewhere for what gives rise to the 'wisdom' of the crowds, as majority is simply not it.

This supports Condorcet's jury theorem (Condorcet, 1785) where p (the probability of picking the correct answer) is more than 1/2 makes juries more accurate and where p is less than 1/2 adding more jury members makes the probability of the correct decision less probable.

So why do the computationally intensive methods perform better under the same conditions, with larger list size? Distance between individual lists appears to hold the key. Distance is clearly an important measure of relevance. Lists which are closer to each other all signal that the *correct* answer must be in that vicinity. Methods that take into account distance between lists are able to take this additional information into account. What is perhaps most interesting is the performance of the averaging model. With smaller list sizes it performs on par with the other models, suggesting averaging is a component of the overall wisdom. However, as its performance decreases with larger list sizes, it is evident that averaging does not really capture the distance differences effectively. There is simply more to effective aggregation than simple averaging.

In terms of feedback, it seems there is a clear effect of individual competence that is aggregated into a more competent group answer. However, simple aggregation does not capture the necessary information, whereas computationally intensive methods appear to do so quite effectively, especially with the larger pool of judges. Although judges in the simulations makes independent guesses and subsequent answers are not impacted by the previous inputs, there is nonetheless group wisdom that occurs at the system level once individual answers are aggregated.

5.4.1.1 Application of Findings

These findings have clear impact on the various fields concerned with rank aggregation. It indeed appears that the use of more computationally complex methods is advantageous if the task at hand is complex: the number of independent sources is high, the list size is high and the probability of selecting an optimal list by any one judge is low.

This has direct implications in the field of web search and information retrieval. Based on the modelling, we would suggest that the use of simple models, such as picking the list that occurs most often, or averaging the results based on their relative position would not produce very accurate results beyond very specific conditions. Namely, the number of ranked results would have to be extremely limited (up to 6) and the number of different aggregators quite high. For virtually all other cases, distance based aggregation methods would yield better results. There does not appear to be a computationally optimal shortcut that could be taken here.

5.5 Experiment 9: Rank Revision

Modelling suggests that a system created by multiple answers should become more accurate provided it uses the more computationally intensive aggregation methods. While this is evident from the simulations, these simulations are based on assumptions. They provide a way of creating a baseline against which human behaviour could be tested. Given that there is little empirical studies that have looked at how individuals aggregate ranks, this is a natural extension of this line of inquiry.

There are elements from Experiments 5–8 that can be applied to better understand rank aggregation dynamics. The domain of feedback loop tasks, whereby the same question is asked multiple times and participants are able to see other answers and revise their beliefs accordingly, lends itself quite well to understand individual belief revision dynamics and to see how accuracy enhancing they are. The main research question at this point was: how would human beings perform in a rank aggregation task? What revision methods, if any, would they choose? And would they become more accurate with time?

From the modelling it is clear that different aggregation methods work better under different conditions, it is not clear what strategies human would choose when aggregate ranks. While it may appear that taking the group mean is advantageous from the accuracy point of view, it is also more difficult to calculate quickly and may be the less preferred strategy than simply adopting the majority opinion for example, not to speak of the more computationally complex methods that would be next to impossible to calculate on the fly. To test this, a study was designed that looked at individual rank revision in a group setting, with the emphasis on feedback incorporation and accuracy measurement of the outcome.

The study was a modification of elements of belief revision used in the Lund study and outlined in Chapters 3 and 4. The purpose of the experiment was to test what rules, if any, individuals use to revise their beliefs in light of new information and what methods may explain the revision process as described by the aggregation methods above. Unlike similar studies on the topic that have mostly looked at absolute answers and estimates, we were interested in applying this in the context of rank revision. In other words, the interest was in understanding how participants revise ranked orders when presented with ranked information from their peers. From the modelling exercises above we knew that adopting the group mean is the most beneficial strategy a person can take with smaller list sizes.

5.5.1 Method

5.5.1.1 Participants

Participants for this study were volunteers from the University of London community. Participants were paid £5 for taking part in the study. There were 19 participants who took part, which created three panels of five participants and one panel with four participants (n=19). Each group of participants took part at the same time and were hosted in the same room. No particular exclusion criteria were used and participants were free to self select which of the time slots worked best for them to attend the study. It did not appear that any participants knew each other prior to the study.

5.5.1.2 Materials & Procedure

Participants were seated in a computer lab, spaced apart in a way that prevented them from seeing each others' screens. Each participant had a computer in front of them that contained a NetLogo interface that was connected in a network to other computers in the room. See Figure 5.2 for a sample interface that each participant saw.

Initially, participants were read basic instructions regarding the task. The task involved each participant ranking four cities from the largest to smallest by population size (list size 4). Each city was presented in a text box and contained a number along with the name of the city (see example in Figure 5.2). In the drop down box 'City A' they were instructed to put the number of the city they believed to be the largest, 'City B' were to contain the second largest, and so on. After all four boxes

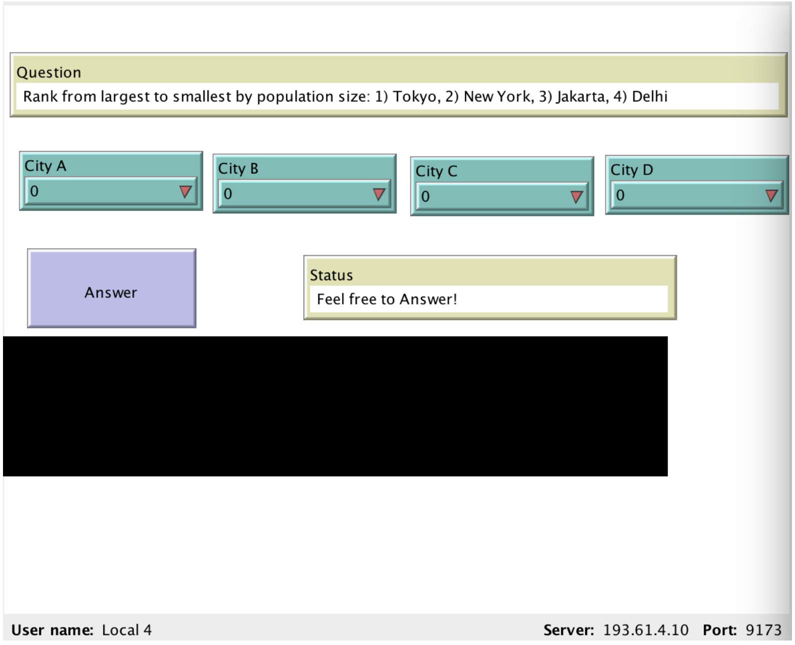


Figure 5.2: An example of a NetLogo interface through which participants communicated with each other.

were filled, participants had to submit their answers and wait for everyone else in the room to finish. Once, all answers were submitted, participants could see how everyone else had ranked the cities. At this point, everyone had an opportunity to revise their answers in light of additional information (see Figure 5.3 and zoomed in view in Figure 5.4). They repeated this process three times for each question, resulting in four rounds - initial round, plus three revision rounds.

In total, each participant answered 21 questions. There was an initial practice question which participants did in a directed manner, followed by 20 other questions, which were done independently and free from any additional instructions. Each question contained a different set of cities and in different order, but the task was the same. There was only a single experimental condition and all participants were treated the same; they were shown the same set of questions, in the same order.

5.5.2 Results

In the first instance we were interested in individual belief revision. We analysed how often individuals changed their answers and what rules they used to do so. Individual revision in this context is a combination of belief revision and rank aggre-

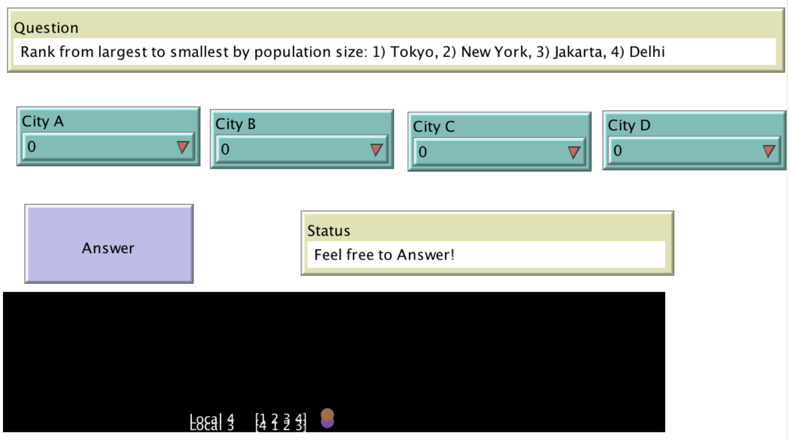


Figure 5.3: An example of a NetLogo interface that a participant would see once other participants submitted their answers.

Figure 5.4: Zoomed in view of the answers of the other participants.

gation as some participants will choose to maintain their own, independent beliefs, while others will choose to revise their answers and adopt those of the group. In this case, participants will, by necessity, need to aggregate group opinions into a single list. Their new answers will reflect at least some of the individual strategies. The findings are outlined below.

5.5.2.1 Individual Revision

Discounting the first question, there were 60 opportunities for each participant to revise their answer (20 questions * 3 revision rounds). On average participants changed their answers 10.3 (SD 7.51) times over the course of the simulation, or about 16% of the time. While some participants changed their answers as little as once, others changed almost a third of their answers. In total, participants revised their answers 196 times.

Most revisions occurred in the first round, where almost as many revisions occurred as in the subsequent two rounds. Table 5.1 breaks down revisions by round. This is quite similar to the previous studies on belief revision conducted in Chapter 4. Revisions occurred unevenly between questions. Seven questions had between 13 and 15 revisions, while remaining 13 questions had between five and nine revisions.

The number of revisions made by participants was rather low, but the overall profile of the changes, i.e. mostly in the first round and more for some questions than others, is consistent with the Lund study, along with the UK-based studies in Chapters 3,4.

Table 5.1: Number of revisions made by all participants in each round.

Revision Round 1	Revision Round 2	Revision Round 3
96	58	42

5.5.2.2 Models of Revision

Several models presented in the first part were examined to predict individual belief revision rules that induced the change (such as mean, median, majority and Spearman models). We decided to restrict our fitting to three in particular: mean, majority and Spearman. Mean and majority models had very interesting properties in our earlier modelling exercise and Spearman was chosen as a representative model of the more computationally intensive class of models. Given that all computational models preformed similarly for smaller list sizes, one distance based model was enough to understand participant revision.

Mean and majority models are computationally easy and should have been most likely to be calculable to participants. Since ranked lists were relatively small and were visually represented near each other, identifying the majority list, or calculating the mean list was conceivable and would be something that a participant could engage in prior to revising their answer.

The Spearman model of revision would have been more difficult to calculate. Indeed, it was not expected that participants would consciously calculate the common Spearman list of the different answers. However, this model calculates a distance-based list and we wanted to see how reflective this model may be of the actual rank revision.

In order to test whether participants actually behaved in a way predicted by a model, we generated an answer that a participant would pick if they were guided by a model and then compared the predicted answer with the actual answer in a binary fashion.

Table 5.2 summarises the results. The table is divided into four columns: the total number of revisions for each method, revisions that only occurred when mean and majority models produced different predictions, the average number of revisions per participant corresponding to the model, and the percentage of revisions explained by each model.

Table 5.2 demonstrates that there were significantly more revisions that moved towards the mean than majority. In fact, of the 196 total revisions, 62 or 31.6% were revisions that adopted the group mean, and 44 or 22% that adopted the majority list. On average, the mean model was adopted 3.26 times per participant, while the majority model was adopted 2.32 times. A Fisher's exact test confirms that this is statistically different (1, N = 19) p = 0.025).

Naturally, there were instances where both models predicted the same list and the above numbers include revisions where the mean and majority lists coincide. There were 35 revisions where both models predicted the same result. When removed from the total revision count for each model, there were 27 revisions that adopted the group mean and only 9 revisions that adopted the majority list. This provides strong evidence to suggest that participants in the study adopted the group mean much more readily than the majority list.

A notable finding was how well the Spearman model performed in predicting participant revision. There appears to be a total of 65 or 33% of revisions that could be explained by this model. This is 3 more revisions than the mean model. This is interesting given that Spearman model is complex to calculate. It is, however, an advantageous model to use, as has been demonstrated by our modelling. Although the difference between Spearman and Mean models was not statistically significant

(Fisher's exact test (N=19) p=0.8292), the difference between Spearman and Majority was (p=0.0232).

Table 5.2: Summary of the revisions made by participants which match the predicted revisions of the different models (mean, majority and spearman). The table shows the total number of revisions, revisions that fell into either mean or majority category, average number of revisions explained by the model and the percentage of explained revisions by each model

Model	Total	Model Only	Average	Revision %
Mean	62	27	3.263	31.6
Majority	44	9	2.316	22.4
Spearman	65	-	3.421	33.2

5.5.2.3 Towards a Model of Human Rank Revision

Our findings suggest that human participants are 3 times more likely to adopt the group mean over the majority list in cases where the two do not coincide. This suggests that computational models that emphasise mean ranks may be closer to the way humans make revisions given additional information in a ranked format.

However, it was the more computationally heavy Spearman model that outperformed both, the mean and the majority models, although its performance was only statistically different from the majority model. It must be noted that mean and Spearman were extremely close in terms of the number of revisions made. It is entirely possible that Spearman model outperformed the mean model due to randomness. However, the important fact here is that Spearman was competitive at all. We initially expected that this model would not perform particularity well, perhaps even below the majority, given that there was no indication that participants use sophisticated calculations while revising their answers.

5.5.2.4 Accuracy

Up to this point we did not look at how accuracy enhancing revisions made participants. Table 5.3 summarises the accuracy by round. Unlike previous studies in this thesis, participants did not appear to get more accurate with each revision round. In fact, it appears that collectively participants actually get *less* accurate with revision. Table 5.3 shows that the number of lists that correspond to the correct answer actually decline in the 4^{th} round and the distance (as measured by Spearman score) actually increases. Participants on aggregate appear to be most accurate in their initial estimate, decreasing with each round, and being least accurate after the 3^{rd} revision round.

Individually, none of the participants improved in accuracy over the three revision rounds. In fact, there were only 3 individuals who did improve slightly in accuracy from their initial estimate to the final one. It appears that while revision did occur, it was not as beneficial in the rank situation as with belief revision of absolute values discussed in the previous chapters.

The situation is quite similar in terms of group accuracy. As can be seen from Graph 5.5 only one group becomes slightly more accurate with repeated revision, with three groups becoming less so. Ultimately, none of the groups show statistically significant overall increase in accuracy as a result of repeated group revision.

Table 5.3: Group aggregate accuracy by round.

Round	Absolute Correct	Correct by Distance
1	25	787
2	25	796
3	25	792
4	23	802

5.6 General Discussion

As demonstrated by the modelling at the beginning of the chapter, simple rules perform well under specific conditions. The mean model does well when list size is small, and the majority model does well when individual competence is high. However, it is the more computationally intensive methods, which take into account distance between ranks that are able to capture the complexity of larger lists and perform better under the more complex conditions. These methods also perform better, as the number of judges increases, further enhancing its benefits. To compute an aggregate list using a Spearman model, one would have to calculate distance between every list and then to select a single list that most closely matches all of the

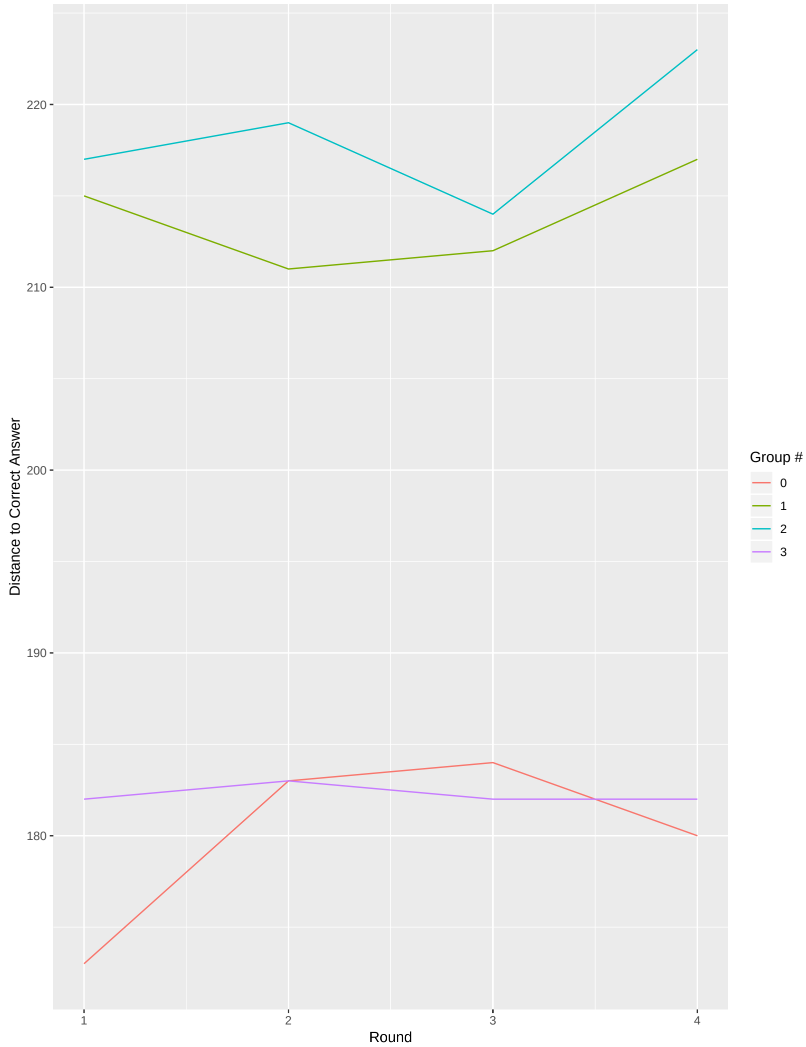


Figure 5.5: Group accuracy over four rounds.

other lists. This is no easy task and computational complexity makes it difficult to extend this beyond relatively small lists.

It was found that human participants in the experiment appeared to use this model in roughly 33% of all revisions that they made. This is notable given that participants spent mere seconds on each revision and did not have any tools at their disposal to do the calculations, or to keep track of the distances. This was done *intuitively*. Participants appeared to revise their answers that we have shown to be advantageous. Given that we did not expect this finding, we did not ask participants to indicate what aggregation methods they may have used, but it would be surprising if they knew anything about distance based measures, or indeed, about Spearman's model.

However, this behaviour did not lead to increased accuracy. In fact, collectively all groups got worse with each revision. Participants were most accurate on their initial, independent guess and did not improve their accuracy with more information. Individually, most participants also got worse as a result of the additional information. Although participants may have used models that have shown to be robust in rank aggregation, this did not translate into increase in performance, individual, or otherwise. This is an important finding as it suggests that rank aggregation is a special class of problems in belief revision, where standard findings do not appear to apply. This could be due to the increased complexity of the task, or the need to keep track and incorporate multiple pieces of information, contained in a rank. Regardless, further research is necessary to understand why additional information and opportunity to revise, does not lead to increased accuracy in this domain.

5.7 Conclusion

This chapter addresses two important concepts. In the first instance, our models demonstrate that distance-based methods do not produce significantly better results when list sizes are small, suggesting that the problem of rank aggregation could be satisfactorily solved by simpler methods such as taking the average or majority. This changes markedly as list size grows, suggesting that for most real-world problems, where list sizes can reach hundreds, if not thousands, the more complex, and computationally exhausting aggregation methods need to be employed.

The second finding occurred as a result of the lab experiment, which extended the modelling to better understand how humans aggregate ranks. It was found that participants most often picked the list that reflected the use of distance-based aggregation methods. This is a particularly interesting finding given that distance-based aggregation methods are computationally intense and no previous research has observed such behaviour. Given that these methods have been demonstrated to be robust and performing well as a function of list size, we believe we have identified an important element of adoptive human behaviour suggesting innate positive adaptability in the realm of rank aggregation.

Finally, it was found that participants did not improve in accuracy, and in fact became worse with each revision round. Every group collectively performed worse as a result of being exposed to additional information and the opportunity to revise, while individually only a few participants improved, while the large majority also got worse. This suggests that the standard belief revision dynamics do not appear to apply to the rank aggregation domain.

It would appear that time does not have the same accuracy enhancing effect across all social feedback systems. In the case of rank revision, participants did not become some accuracy, either individually, or as a group, suggesting that time is context depended. This further suggests that there may be other circumstances under which social feedback systems would not produce similar accuracy enhancing characteristics observed in Chapters 3 and 4.

Chapter 6

General Conclusion

Feedback is a fundamental part of human interactions and plays a fundamental role in one's ability to learn from the environment (Orrell, 2011, Powers, 1973, Sterman, 2000, Taleb, 2007). This thesis explored systems, both physical and social that change over time as a result of interactions within them. The goal was to better understand how feedback systems impact individual behaviour and to extend scholarship beyond traditional understanding of systems in classical cognitive literature, as static environments. The research in this thesis builds upon the relatively recent fields of dynamic decision-making and group belief revision by introducing social feedback and observing how human participants interact with and within such systems and how their behaviour changed as a result.

Part I of this thesis focused on one's ability to control physical dynamical systems. Such systems are designed to approximate real-world processes and systems in order to better understand how humans solve problems and make complex decisions (Osman, 2010). Initially designed by Dörner (1975), it was later expanded by Berry and Broadbent (1984) and later extended into a field of dynamic decision-making (for overview see Osman, 2010). Part II of the thesis extended focus onto social dynamical systems. The aim was to better understand how group dynamics influence individual behaviour, which in turn influences the group. The experiments conducted for this purpose involved human participants that could interact with each other by exchanging information. They could learn, while cooperating and competing. Ultimately, this thesis is about understanding and quantifying the impact of positive and negative influences of feedback on performance and the understanding of the different conditions that lead to positive and negative outcomes in social dynamic system.

6.1 Feedback and Performance

So what features of feedback impact performance? Performance was largely defined as the ability to achieve high scores in Part I and the ability to get closer to the correct answer in Part II.

Overall, there were several findings that are important to highlight in regards to feedback loops and performance. Artificial randomness was found to have a substantive effect on performance. When removed from the original sugar factory, participants were able to learn from delayed feedback and achieve mastery over the task. This demonstrates that the delayed nature of feedback does not necessarily lead to the inability to control a feedback simulation, but artificial randomness does.

Natural noise, however, created by participants in group interactions in Part II led to increased to accuracy over time, highlighting an important difference between artificial randomness and naturally occurring noise.

Time was found to be an important consideration in feedback learning. There was a strong practice effect in Part I and participants became more accurate with each revision in Part II. Time allows participants to gain more information and learn to better control the system over time.

The one outlier to this the rank revision experiments performed in Chapter 5, where participants did not improve over time. In the rank revision task, participants did not improve their performance with each round of revision. This suggests that the nature of a social feedback task is quite important and the general view of increased performance does not necessarily extend to all domains. It is possible that rank aggregation is a different class of problems altogether, however, broader claim would require further research.

Finally, Chapter 4 found no statistically significant impact of incentivisation on performance in social feedback loops, or on the revision strategies employed by the participants. It should be noted however, that such findings are not necessarily general in nature and should be interpreted in the context of the task they were found in.

6.2 Feedback Control

Part I, or more specifically, Chapter 2 focused on physical dynamic systems as represented by the sugar factory (Berry and Broadbent, 1984, 1988). The chapter started with a replication study (Experiment 1), which sought to re-evaluate previous findings in light of the time lapse since original research took place. This experiment confirmed the original findings. 30 years since the original sugar factory, human beings are still relatively poor at the task. Experiment 2 changed the original equation and provided participants with a finer control over the number of workers they could assign. Participants were also provided with training materials (see Appendix A for the copy of the materials) designed to help control physical dynamic systems. It was found that participants still struggled to control the output of the sugar factory and materials did not appear to have any impact on performance. In Experiment 3, the cover story was changed to bring the sugar factory into a climate change domain. This was done in order to create a task in a more meaningful domain, where participants might be expected to have more relevant prior real world knowledge. These participants were also provided with training materials to help control the simulation. However, neither bringing the system into a new domain, nor providing the training materials improved performance. Next, Experiment 4 extended the sugar factory beyond the lab, giving participants an ability to play the sugar factory at home. They were free to explore the simulation and play it as many times as they wished. Yet, this did not produce a noticeable increase in performance. These findings are consistent with much of the dynamic control literature. Human participants are not particularly good at controlling such simulations (Osman, 2010).

After repeated failure to improve human performance, a question arose regarding the overall plausibility of controlling such simulations. However, as with much of the literature one unanswered question remained: human performance was poor compared to *what*? There simply were not a practical measure available to compare human performance against. As such, a simulated agent was created, based on a simple reinforcement learning algorithm, which was trained to control the sugar factory.

This led to three main findings. Firstly, human performance was actually on par with the performance of the simulated agent. Secondly, randomness was identified as the main culprit behind poor performance of both human and artificial agents. Once removed the artificial agent's markedly improved, which led the third finding. Experiment 5 confirmed that once randomness is removed human participants are able to learn to control the simulation in a matter of dozen tries; significantly quicker than the simulated agent. Experiment 5 demonstrates that individuals can actually be quite adept at controlling dynamic environments.

This challenges the notion that the temporal nature of feedback is the main reason for poor human performance (Berry and Broadbent, 1988). Indeed, by training a reinforcement learning-based agent – against which human performance could be measured – it became clear that the main problem with the original sugar factory is the randomness factor embedded into the governing equation. Once randomness is removed, participants learned to control the simulation in the matter of a dozen trials. The original rationale for including the randomness was so that, "subjects would exercise continuous control. Subjects might, through chance, hit the target value early in a series of trials. They would be unlikely, however, to remain on this target if they repeatedly entered the same input value" (Berry and Broadbent, 1984, p. 212). Thus, by wishing to ensure continuous control, they inadvertently made the simulation impossible to control.

Randomness, however, is used in many domains without the same adverse consequences in performance. Randomness is used in reaction-time experimentation to display different stimuli (Sternberg, 2010) and in psychophysics randomness is used to delay the display of stimuli at random intervals (Whalen et al., 1999). Even the simulated agent's epsilon greedy algorithm has randomness built in, in order to allow the agent to switch between exploration and exploitation to explore the environment it is in (Gureckis and Love, 2009b, Simonini, 2018). But random noise in a feedback system can have devastating consequences to learning. This is demonstrated by the simulated agent not able to learn. It is also not surprising that Berry and Broadbent (1988) have found that participants achieve fragmented (implicit) learning, without the ability to verbalise what they learned (explicit learning) (Berry and Broadbent, 1988, Cleeremans and Seger, 1994, Sanderson, 1990).

6.3 Belief Revision In Groups

Part II of the thesis focused on a different type of system interaction; that of the group. Unlike the first part, which focused on an individual participant interacting with a physical feedback system, the second part of the thesis focused on social systems. Social systems are created by the interactions of the participants within it. There are two levels of feedback that give rise to such systems. The first level consists of individual-to-individual interactions, which have an impact on behaviour. The second level consists of individual and group accuracy and the interplay between the two. Unlike the physical systems in the first part, the feedback is a function of the these two interactions, which operate within the pre-programmed system, but would not otherwise exist without the participants.

The principal aim of Part II was to better understand the interactions that occur between an individual and the wider group and the two feedback systems that are created by this interaction. Chapter 3 dealt with the strategies that individuals use in group revision tasks. The goal was to better understand and model individual belief revision behaviour and to better understand what strategies individuals use when revising their beliefs, which give rise to the feedback system itself and ultimately has impact on the feedback each participant receives from the system.

Initial work focused on the Lund study conducted by Jönsson et al. (2015). Unlike the paper, however, the focus was on the strategies individuals use to revise their answers in light of the information received from the group. Several different belief revision models taken from the literature were applied. However, it was found that the ‘no-change’ model performed the best in predicting actual revision. Experiment 6 focused on individual belief revision. Mean, standard deviation and distance to participant’s answers were manipulated in order to better understand what impact each of these variables would have on revision. It was found that the mean distance to the other answers had the most impact on revision. Standard deviation and distance to the closest answer had a moderate mediating impact on revision. And participant’s confidence had very little correlation with revision. Interestingly, Experiment 6 also showed that confidence did not appear to have significant impact on belief revision on its own, but it did when combined with the distance to the mean. In order words, if individuals were less confidence and were further away form the group them they did move closer to it.

Experiment 7 focused on a single experimental manipulation, sometimes leaving participant’s answer enclosed by other answers and sometimes making it the most extreme answer. This was done to measure the impact of being on the ‘inside’ or ‘outside’ of the other answers would have on belief revision. The main finding of this experiment was that individuals were quite sensitive to where they were in relation to the group. Individuals whose answers are outside of the group tend to become more conformist, by moving closer to the mean of the group.

Over the course of several revisions, each individual acts as both a positive and a negative feedback loop. In the first round of revision, participants tend to readily move towards the group mean, especially those whose views are more extreme. This produces a positive loop in terms of the amount of revision, whereby new information is introduced into the system, further encouraging participants to revise their answers on subsequent rounds. At the same time, participants act as negative feedback loops in subsequent rounds, where the probability of each person revising drops precipitously, which leads to the lack of new information being introduced, which leads to further reduction in revision.

In the context of feedback loops, change has an impact on accuracy. Given that revision by necessity is the only way for the group to get more accurate, it provides an observable link between feedback and real world outcomes. Consequentially, when participants did change their answers they moved in the right direction, improving group accuracy (see Jönsson et al., 2015, for detailed discussion on group accuracy).

Research in this chapter culminated in the creation of a new model of human belief revision that incorporated much of the behaviours that participants displayed in the Lund study and Experiments 6 and 7 into a framework that could be used to approximate human behaviour more accurately than the 'no-change' model. When applied to the Lund dataset, the *new* model outperformed the previously best performing 'no-change' model for all of the groups. The new model also performed well Experiment 8 discussed in Chapter 4, which applied the model to a different set of participants in an entirely new country (the UK). Given that this model's performance was not statistically different from the 'no-change' model, it certainly does not capture the whole range of revision behaviours that participants have exhibited and further research with larger groups is required.

6.4 Individual vs Collective Incentivisation

Whereas Chapter 3 focused mainly on the feedback loop between the participants. The focus was on the individual, their revision strategies and how they influences other participants. Chapter 4 and specifically Experiment 8 focused on the second-level feedback loop, that of the interplay between individual and group accuracy by manipulating the nature of incentivisation. The second loop is not as mysterious as it may appear at first and through various incentivisation it can be influenced and its effects studied.

Many of the individuals in the real world systems are incentivised in some way. Two types of incentivisation were used in Chapter 4: individual and group. In the individual condition, bonus was paid to participants who achieved the highest individual accuracy, regardless of how well the group did. In the group condition, a bonus is paid to participants of the group that achieved the highest accuracy as the group, compared to other groups. By contrast, a single participant per group was paid the bonus in the Lund study.

It was found that participants in Experiment 8 sought much greater conformity than in the Lund study. However, when compared to each other, the two types of incentivisation (individual vs group) had modest impact on revision. Participants in the two conditions did revise differently. Participants in the group incentivisation condition were much more likely to revise towards the group mean, suggesting that incentivisation leads to greater conformity. This, however, did not lead to increased accuracy. This created to a unique phenomena of decreased accuracy in the group incentivisation condition in the last round. These differences were minor, however and further studies are required to corroborate these effects.

As demonstrated by Hertwig and Ortmann (2001) in their large review of performance-based monetary reward studies, there is a real mix of findings in this area. Many studies show the impact of monetary incentivisation and yet fail to reach statistical significance (for summary discussion on group accuracy see Hertwig and Ortmann, 2001, p.392). Our findings suggest that performance-based monetary rewards have an effect of producing greater conformity, which is not accuracy enhancing, however.

6.5 Rank Revision

Chapter 5 expanded belief revision to a new domain of rank aggregation. The purpose of this chapter was to better understand how humans aggregate ranks and whether their methods correspond to those that have been shown to produce more accurate results in modelling exercises.

First, different models of rank aggregation were modelled and examined on their properties to produce accurate lists under different conditions. It was found that under smaller list sizes of 4-5 items, simpler aggregation methods, such as averaging produced results that were on-par, or better than the more complicated, computationally intensive methods, such as Kemmeny-Snell distance based aggregation. Furthermore, it was found that depending on the individual competence of the judges, majority aggregation would either produce very accurate results, or fail completely, regardless of the number of judges. When list size was increased to 6 and beyond, the computationally intensive methods began to outperform the simpler methods, suggesting that in the limit they would be the only reliable way of aggregating large ranks. Furthermore, the individual accuracy required for the majority method also increases, making this a less desirable aggregation method with larger list sizes.

Next, we ran a study actually looking at human performance. In Experiment 9 participants were asked to rank different cities in the order of population size from largest to smallest. They were then given a chance to revise their rankings in light of the other answers they saw from the participants around them. This study used the existing belief revision experimental design employed in the Lund study, as well as Experiment 8, applied to the new domain of rank revision. Participant behaviour obtained in Experiment 9 showed that distance-based aggregation methods better accounted for the revision. This is a notable finding given that distance-based aggregation methods are computationally intense and no previous research has observed such behaviour. Given that these methods have been demonstrated to be robust and performing well as a function of list size, this finding shows participants demonstrating a high level of sophistication in ranked belief revision.

6.6 Research Limitations

This thesis makes a number of claims and findings that are backed by nine experiments and a number of modelling exercises. However, as with any research, there are a number of considerations that limit the wider applicability of this research.

Naturally, experiments in Chapter 2 approximate some of the tasks that individuals may face in the real world. However, real-world tasks that contain feedback loops are simply too complex to model fully. Thus, the findings in the chapter regarding randomness should be considered in a specific context from which it is derived.

Similarly, the findings in Chapter 3 around the lack of impact of confidence on belief revision may have been caused to the nature of the experiment itself, where the task was not one where confidence plays an important role. Given that these were general knowledge questions, it is possible that confidence, or the accuracy of the self-reported confidence measures, was simply not important to the participants and that in a different context, such measure would indeed correlate with belief revision.

All lab experiments are time-bound and the findings are limited by the number of trials one could do. The sugar factory was bound by 30 trials, while belief revision tasks asked participants to revise their answers up to 3 times. The number of measures such tasks generate are too few to draw general conclusions about human behaviour in the real world. It is certainly an indication, but the limited nature of such experiments precludes us from being about to draw wider conclusions.

6.7 Future Research

This thesis introduces several lines of inquiry and models that require further testing if they are to be incorporated into the wider literature on human decision-making and dynamic systems. There findings can be divided into three sub-sections and are outlined in greater detail below.

6.7.1 Dynamic Control

Given that the common wisdom of the difficulties of the sugar factory due to the delayed feedback nature of the system, may not be correct, it is worthwhile to explore the approach presented in this paper to other virtual physical systems that have been employed in the literature (Cleeremans and Seger, 1994, Funke, 1988, Gonzalez et al., 2005, Osman, 2010). In particular, it would be important to test different simulations with an artificial agent to see if such tasks have a solution in the first place. The second step would be to identify where artificial agents have difficulties and to rework the tasks in such a way that they become controllable. The last step would be to have human participants conduct the new simulations and to see if they performance improved as a result.

The second research strand that is worthwhile in this space is to reduce the reliance on randomness in these simulations. The artificial randomness introduced into these simulations, which is usually intended to make the task more engaging, actually makes it quite difficult for participants to do. This somewhat negates the original intent behind the experiment, as it becomes a case of human against artificial randomness. As has been demonstrated in this paper, randomness ultimately prevails, but such an outcome does not necessarily provide a useful insight into the real world dynamics, as real world systems are not governed by artificially created randomness, but by the complexity of the interactions of the agents with in.

6.7.2 Models of Human Belief Revision

Chapter 3 introduced an improved model of human belief revision in a dynamic system. This comprehensive model of belief revision appears to closer reflect true human behaviour on belief revision tasks. The hope is that this research leads to more accurate models of belief revisions than are currently used across multiple disciplines to model, or incorporate human decision-making in a group setting. This research seems particularly important due to the proliferation of computational models that incorporate human decision-making in order to explain a range of real-world phenomena, from opinion dynamics to network science.

Undoubtedly, further research is required in this area. The new model shows promise, although much remains to be understood. In particular, the relationship between source reliability, incentivisation, group size and willingness to change should be explored, along with role of individual difference, such as personality, intelligence and prior knowledge. The dynamics of constant updating also need to be better understood. The final model assumes no revision beyond the first round, and empirical data suggests a big drop off, however, some revision does occur. The nature and dynamics of this revision is beyond the scope of this thesis, but presents an important avenue for future research.

Future research of individual revision will need to focus on expanding on this model and evaluating the conditions that influence the alpha factor. The seminal feature of the new model includes the so-called stickiness (alpha) factor. The factor was first introduced in Chapter 3 and further expanded in Chapter 4. This factor refers to groups opinion elasticity, or the propensity to revise belief. The greater the factor, the more revision occurs in the system. Experiment 8 showed that this this factor can vary based on incentivisation and geography. As a follow up it would be quite important to understand individual and group dynamics that lead to its increase or decrease. One such factor appears to be different incentivisation, another may be geography and culture, and there may be others. This line of research would further extend the understanding of group cohesion by specifically focusing on factors that make individuals revise their answers more, or less depending on their context.

Rank revision introduced in Chapter 5 also requires further research. Experiment 9 dealt with list size 4 and future experiments should extend the list size to better understand how humans revise larger ranks. Furthermore, modelling can be extended to list sizes of 10 or greater, although that would require new computational strategies to model. The larger list size modelling would show which aggregation methods are optimal as list sizes get bigger and bigger and come to approximate many of the real world problems where rank aggregation is important.

6.8 Terminus

This thesis and the research contained therein was compiled over four years and represents an effort to understand a complex phenomena of the impact of feedback loops on individual decision-making and the systems that these interactions give rise to. As more and more interactions take place online, social systems are coming to increasingly resemble feedback properties discussed in these chapters.

There are several lessons that I have learned while doing research in this space. Firstly, overall, individuals are uncomfortable with being on the margins. Whenever possible, they will move to conform, adopting just enough of the opinions of others to fit into the group. Secondly, we will listen to opinions even if they are extreme. Being exposed to even unbelievably crazy opinions still makes us wonder and revise, even if slightly towards the mean. Third, two individuals is a complex system, but three people is a truly complex world. It becomes increasingly difficult to truly discern patterns and desires of individuals in a complex system and eventually the two concepts become one. It becomes impossible to understand individual actions without understanding the context and vice versa. Finally, not all revision leads to more accuracy. When aggregating ranks, having more information did not lead to greater accuracy.

It is my hope that this research may be a call to action to further bridge the gap between theoretical understanding and practical evaluation of actual human performance. A bridge that psychology and computer science is bound to build as cognitive science becomes more prevalent as a discipline.

Appendix A

Sugar Factory Experimental Materials

PILOT Testing (n=30)

Condition 1: Cover Story – Sugar Factory (Exact) (n=15)

You are in charge of running a sugar production factory in an underdeveloped country. You control the rate of production by simply changing the size of the work force, ignoring all other factors. You start with 600 workers that produced 6000 tonnes of sugar in the previous month.

Your task is to reach and maintain a target output of 9,000 tons per month. To help with the task, the maximum output of the factory has been set at 12,000 and the minimum to 1,000.

You will have to run the factory for 30 months. Each month you will assign a number between 1-12 representing the number of workers that would work in the factory that month. The computer will multiply your number by 100 to get the actual number of workers. Example: 6 is 600 workers.

Condition 2: Cover Story – Sugar Factory (Range) (n=15)

You are in charge of running a sugar production factory in an underdeveloped country. You control the rate of production by simply changing the size of the work force, ignoring all other factors. You start with 600 workers that produced 6000 tonnes of sugar in the previous month.

Your task is to reach and maintain a target output of 9,000 tons per month. However, we will also score any output between 8000 and 10,000 as being on target.

To help with the task, the maximum output of the factory has been set at 12,000 and the minimum to 1,000.

You will have to run the factory for 30 months. Each month you will assign a number between 1-12 representing the number of workers that would work in the factory that month. The computer will multiply your number by 100 to get the actual number of workers. Example: 6 is 600 workers.

PART II

Condition 1: Cover Story – Sugar Factory (Target) (n=20)

You are in charge of running a sugar production factory in an underdeveloped country. You control the rate of production by simply changing the size of the work force, ignoring all other factors. You start with 600 workers that produced 6000 tonnes of sugar in the previous month.

Your task is to reach and maintain a target output of 9,000 tons per month. To help with the task, the maximum output of the factory has been set at 12,000 and the minimum to 1,000.

You will have to run the factory for 30 months. Each month you will assign a number between 1-12 representing the number of workers that would work in the factory that month. The computer will multiply your number by 100 to get the actual number of workers. Example: 6 is 600 workers.

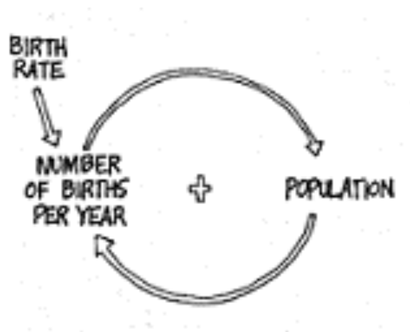
Condition 2: Cover Story – Sugar Factory (Non-Linear) (n=20)

You are now in charge of running a factory in a developed country with a well-integrated supply chain management and dynamic output control that is more responsive to the changes to the workforce. Your goal has increased and you are to reach and maintain a target output of 100,000 tons of sugar per month. You start with 6000 tons that were produced by 60 workers. To help with this task, the maximum output has been set at 240,000 and the minimum at 1,000.

You will run this factory for 30 months, please note that since this factory is much more advanced, you will now have the ability to fine-tune the number of workers you assign each month. You may assign a number from 0-1200 representing the number of workers that would work in the factory that month. The computer will no longer multiply your number by 100.

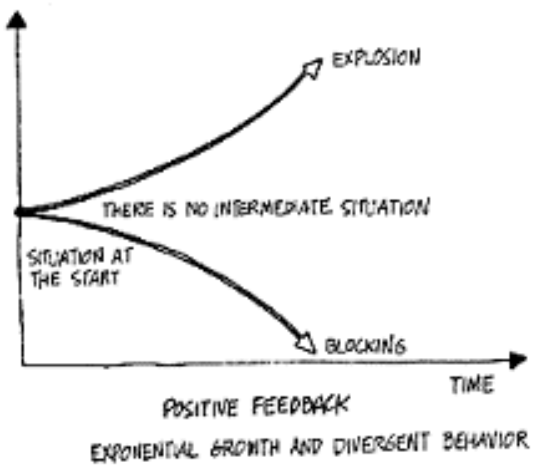
Basics of 'feedback' loops/Controlling Feedback loops

Feedback loop is a phenomenon where previous output becomes part of future input. For example, think of population growth, more births result in a larger population and larger population results in more births. The relationship between births (input) and population (output) is known as a **positive feedback loop**. A figure 1 below illustrates this characteristic:



(Figure 1: Birth Rate Feedback Loop) (de Rosnay, 1979) http://pespmc1.vub.ac.be/feedback.html

Positive feedback loops, if left alone can lead to significant, uncontrollable, growth or decay. As can be seen from Figure 2, the growth or decay can become exponential in nature, which makes reversal more difficult, the longer the loop is allowed to operate.



(Figure 2: Positive Feedback Loops) (de Rosnay, 1979) http://pespmc1.vub.ac.be/feedback.html

Positive feedback loops

The sugar factory you have just controlled was governed by a positive feedback loop. The formula for production took into account previous month's output in determining how much the factory produced in the subsequent month. When dealing with such systems, it is important to keep the following strategies in mind:

- Attempt to understand the underlying processes: What is the governing equation of the system? Are there any patterns to the output? What is the impact of the different variables? How does the system grow and decline?
- Test various theories: Try to vary your answers and test different patterns. Switch between different hypothesises and test them in a methodical way. Seek and accept evidence that may disprove the hypothesis you are testing. Do not only seek and accept answers that confirm your hypothesis.
- Make changes gradually when testing your theories. Sudden and large changes in input can have adverse reactions within the system and taint any feedback. This is particularly true for systems governed by non-linear growth, or decline.
- Account for noise. Some variation and randomness is to be expected.
- Use your answers to create convergence on the goal. Try a strategy that instead of focusing on hitting the target attempts to converge on it over several turns.

Appendix B

Climate Change Study Experimental Materials

PART III

Cover Story – Sugar Factory (Non-Linear) (n=20)

You are in charge of running a sugar factory in a developed country with a well-integrated supply chain management system that is very responsive to any changes in the workforce you make. Your goal is to assign workers to work in the factory, and reach and maintain a target output of 100,000 tons of sugar per month. In the previous month 60 workers produced 6000 tons. To help with this task, the maximum output of the factory has been set at 240,000 and the minimum at 1,000.

As you progress, you will see a black line with a number above it indicating the factory output, as well as dots with a number of workers that you assigned.

You will run this factory for 30 months, each month you will assign a number from 0-1200 representing the number of workers that would work in the factory that month.

Cover Story – Climate Change

In front of you is a simulation of the Kaya identity, which economists use to express the relationship between several social and economic factors and CO_2 emissions.

This simulation starts in 2000. You have just been appointed as the Prime Minister of a developed country and your job is to stay in power by carefully balancing economic growth with the rise in CO_2 emissions. Climate change is about to become a major political issue as it is becoming clear that if carbon emissions are not curbed, the global temperatures will rise, with potentially unpredictable consequences. As such, you are to navigate a path between economic growth and carbon emissions. Your task is to maintain healthy GDP growth, while ensuring that your country's emissions stay below the critical threshold of 6 billion tons of CO_2 a year.

You are to run your country for the next 40 turns, controlling quarterly GDP growth. Each turn you may enter any number between 10 and -10 indicating GDP growth or decline that you set in the budget. Keep in mind that consistent GDP decline will lead to popular unrest and ultimately will result in your dismissal.

Your approval rating is to your right and will change color to reflect your popularity. If it becomes too red, you will be voted out of office and the simulation will end prematurely. The black line on the graph shows current CO_2 emission levels. The red line represents the limit of 6000 as mentioned before.

Appendix C

Climate Change Study Questionnaire

Climate Change Questionnaire – PART I

Below are 23 questions about your views on climate change. Please indicate your answer by marking the appropriate item on the scale.

1. Claims that human activities are changing the climate are exaggerated

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

2. Climate change is just a natural fluctuation in earth's temperatures

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

3. I do not believe climate change is a real problem

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

4. I am uncertain about whether climate change is really happening

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

5. The evidence for climate change is unreliable

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

6. There is too much conflicting evidence about climate change to know whether it is actually happening

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

7. Too much fuss is made about climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

8. Floods and heat-waves are not increasing, there is just more reporting of it in the media these days

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

9. Many leading experts still question if human activity is contributing to climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

10. The media is often too alarmist about issues like climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

11. I believe climate change is reversible

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

12. I often set my thermostat to 20 degrees or cooler in wintertime to reduce my contributions to climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

13. It is my responsibility to do something about climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

14. I support government action to tackle climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

15. I use public transportation or car-pooling to reduce my contributions to climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

16. I believe government action is necessary to moderate climate change

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

17. Dealing with climate change requires drastic collective action

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

18. I believe government action will be effective in moderating climate change

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

19. I <u>intend</u> in the future to set my thermostat to 20 degrees or cooler in wintertime to reduce my contributions to climate change

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

20. I <u>intend</u> to write a letter, email, or phone a government official about climate change over the next 12 months

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

21. I <u>intend</u> to in the future use public transportation or car-pooling to reduce my contributions to climate change

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

22. Dealing with climate change requires <u>immediate</u> collective action

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

23. I believe government should cut CO2 emissions even if that leads to increased unemployment

Strongly Disagree								Strongly Agree	
1	☐	2	☐	3	☐	4	☐	5	☐

23. I support carbon tax (tax on companies emitting carbon into the atmosphere)

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

23. I support drastic economic measures (ex. closing down factories) in order to cut CO2 levels

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

23. I believe UK government is doing enough to combat climate change

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

23. I suppose cleaner energy sources (ex. wind, solar) even if it means paying more for electricity

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

23. Climate change is too complex and uncertain for scientists to make useful forecasts

Strongly Disagree — Strongly Agree

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

Appendix D

UK Census Questions

Answer	Question
2.6	"What is the percent of adults (over the age of 16) who have been unemployed for over 12 months?"
21	"What is the percentage of adults in the UK who report not drinking alcohol at all?"
86	"What is the proportion of individuals living in the UK that have been born in the UK?"
19	"What is the proportion of children in the UK living in absolute poverty?"
78.4	"What is the average male life expectancy in the UK?"
72	"What is the proportion of people in the UK engaging in any volunteering activity at least once a year?"
98	"What is the proportion of people in the UK who report having a partner, family member, or friend to rely on if they have a serious problem?"
41	"What is the proportion of people agreeing that people in their neighbourhood can be trusted?"
58	"What percentage of workers are part of a workplace pension scheme?"
19.5	"What proportion of the population in the UK smokes?"
28.5	"What proportion of adults does less than 30 minutes of exercise a week?"
4.1	"What percentage of energy is consumed from Renewable Sources?"
43.2	"What percentage of households recycle?"
25	"What percentage of total land in the UK is used for agriculture?"
53	"What percentage of the food consumed in the UK comes from the UK?"
30	"What percentage of the food consumed in the UK comes from the EU and Europe?"
20	"What percentage of rivers in the UK failed chemical testing in 2012?"
47	"What percentage of fish is harvested sustainably in the UK?"
15	"What percentage of the population in the UK reported binge drinking (8+ drinks on a single night) at least once in the previous week?"
15	"What is the percentage of self-employed workers in the UK?"
41	"What is the average age of a worker in the UK?"
42	"What percentage of marriages end in divorse in the UK"
34.4	"What percentage of adults(18+) in the UK have a unversity degree"
64	"What percentage of adults(18+) in the UK are classified as overweight or obese?"
10	"What percentage of the population is lefthanded?"

Bibliography

Akerlof, G. A. and Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.

Aronson, E., Turner, J. A., and Carlsmith, J. M. (1963). Communicator credibility and communication discrepancy as determinants of opinion change. *The Journal of Abnormal and Social Psychology*, 67(1):31–36.

Bakshy, E., Messing, S., and Adamic, L. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Socience Express*, pages 1–5.

Baltrunas, L., Makcinskas, T., and Ricci, F. (2010). Group Recommendations with Rank Aggregation and Collaborative Filtering. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, pages 119–126, Barcelona, Spain.

Barto, A. G. and Singh, S. P. (1991). On the Computational Economics of Reinforcement Learning. *Connectionist Models*, pages 35–44.

Becker, J., Brackbill, D., and Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences of the United States of America*, 114(26):E5070–E5076.

Bernheim, B. D. and Rangel, A. (2008). *Choice-theoretic foundations for behavioral welfare economics*. NY: Oxford University Press.

Berry, D. and Broadbent, D. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, 36(2):209–231.

Berry, D. C. (1991). The Role of Action in Implicit Learning. *The Quarterly Journal of Experimental Psychology*, 43(4):881–906.

Berry, D. C. and Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological Research*, 49(1):7–15.

Berry, D. C. and Broadbent, D. E. (1988). Interactive taks and the implicit-explicit distinction. *British Journal of Psychology*, 79(79):251–271.

Bojarski, M., Yeres, P., Choromanska, A., Choromanski, K., Firner, B., Jackel, L., and Muller, U. (2017). Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car. *arXiv preprint*, pages 1–8.

Bonner, S. E. and Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society*, 27(4-5):303–345.

Bovens, L. and Hartmann, S. (2002). Bayesian Networks and the Problem of Unreliable Instruments. *Philosophy of Science*, 69(1):29–72.

Bovens, L. and Hartmann, S. (2003). *Bayesian epistemology*. Oxford University Press.

Bowen, H. J. and Kensinger, E. A. (2017). Cash or Credit? Compensation in Psychology Studies: Motivation Matters. *Collabra: Psychology*, 3(1).

Brancotte, B., Yang, B., Blin, G., and Cohen-boulakia, S. (2015). Rank aggregation with ties : Experiments and Analysis Rank Aggregation problem. *Proceedings of the 41st International Conference on Very Large Data Bases*, pages 1202–1213.

Brandes, U., Robins, G., McCranie, A., and Wasserman, S. (2013). What is network science? *Network Science*, 1(01):1–15.

Brase, G. L. (2009). How different types of participant payments alter task performance. *Judgment and Decision Making*, 4(5):419–428.

Brase, G. L., Fiddick, L., and Harries, C. (2006). Participant recruitment methods and statistical reasoning performance. *Quarterly Journal of Experimental Psychology*, 59(5):965–976.

Brehmer, B. (1990). Strategies in real-time, dynamic decision making. *Insights in Decision Making*, pages 262–279.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta psychologica*, 81:211–241.

Brehmer, B. (1995). Feedback delays in complex dynamic decision tasks. In Frensch, P. and Funke, J., editors, *Compex Problem Solving: European Perspective*, chapter 4, pages 103–130. Lawrence Erlbaum Associates Publishers, NJ.

Buchner, A., Funke, J., and Berry, D. C. (1995). Negative Correlations between Control Performance and Verbalizable Knowledge: Indicators for Implicit Learning in Process Control Tasks? *The Quarterly Journal of Experimental Psychology Section A*, 48(1):166–187.

Cadsby, C. B., Song, F., and Tapon, F. (2016). The Impact of Risk-Aversion and Stress on the Incentive Effect of Performance-Pay. In *Experiments in Organizational Economics*, pages 189–227. Emerald Group Publishing Limited.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions. *The Leadership Quarterly*, 13(1):69–93.

Centola, D. (2018). *How Behavior Spreads: The Science of Complex Contagions*. Princeton University Press.

Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3):1247–1250.

Clark, D., Ranney, M. A., and Felipe, J. (2013). Knowledge Helps: Mechanistic Information and Numeric Evidence as Cognitive Levers to Overcome Stasis and

Build Public Consensus on Climate Change. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 35(2070):2070–2075.

Cleeremans, A. and Seger, C. (1994). Implicit learning. *Psychological bulletin*, 115(Article 97):1–27.

Condorcet, M. d. (1785). *Essai sur l'application de l'analyse la probability des decisions rendues la plurality des voix*. Imprimerie Royale, Paris, 1 edition.

Currie, D. J., Smith, C., and Jagals, P. (2018). The application of system dynamics modelling to environmental health decision-making and policy - A scoping review. *BMC Public Health*, 18(1):1–11.

Dalal, R. S. and Bonaccio, S. (2010). What types of advice do decision-makers prefer? *Organizational Behavior and Human Decision Processes*, 112(1):11–23.

Das, a., Gollapudi, S., and Munagala, K. (2014a). Modeling Opinion Dynamics in Social Networks. *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, pages 403–412.

Das, A., Gollapudi, S., and Munagala, K. (2014b). Modeling opinion dynamics in social networks. *Proceedings of the 7th ACM international conference on Web search and data mining - WSDM '14*, pages 403–412.

Dayan, P. and Niv, Y. (2008). Reinforcement learning: The Good, The Bad and The Ugly. *Current Opinion in Neurobiology*, 18(2):185–196.

Deci, E. L., Koestner, R., and Ryan, R. M. (2005). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6):627–668.

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional options. *Quarterly Journal of Economics*, 118(3):909–968.

Diaconis, P. and Graham, R. L. (1977). Spearman's Footrule as a Measure of Disarray. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(2):262–268.

Dienes, Z. (1990). *Implicit Concept Formation*. PhD thesis, University of Oxford.

Dienes, Z. and Fahey, R. (1995). Role of Specific Instances in Controlling a Dynamic System. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4):848–862.

Doer, B., Fouz, M., and Friedrich, T. (2012). Why rumors spread so quickly in social networks. *Communications of the ACM*, 55(6):70.

Dörner, D. (1975). Wie Menschen eine Welt verbessern wollten [How people wanted to improve the world]. *Bild der Wissenschaft*, 12(2):48–53.

Dörner, D. (1997). *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations*. Perseus Books, Reading, MA, new editio edition.

Dubois, D. and Prade, H. (1997). Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge. In *Qualitative and quantitative practical reasoning*, pages 96–107. Springer.

Dwork, C., Kumar, R., Naor, M., and Sivakumar, D. (2001). Rank aggregation methods for the Web. In *Proceedings of the tenth international conference on World Wide Web - WWW'01*, pages 613–622, Hong Kong, China.

Farrell, S. (2011). Social influence benefits the wisdom of individuals in the crowd. *Proceedings of the National Academy of Sciences*, 108(36):E625–E625.

Fernández, M., Vallet, D., and Castells, P. (2006). Probabilistic Score Normalization for Rank Aggregation. In *Proceedings of the 28th European Conference on {IR} Research (ECIR)*, pages 553–556, London, UK.

Flynn, B. (2011). Maurice Merleau-Ponty.

Folia, V., Udd, J., Forkstam, C., and Petersson, K. M. (2010). Artificial Language Learning in Adults and Children. *Lenguage Learning*, 60:Suppl.(December):188–220.

Forrester, J. (1971). Counterintuitive behavior of social systems. *Theory and Decision*, pages 1–29.

Funke, J. (1988). Using simulation to study complex problem solving. *Simulation & Games*, 13(3):277–303.

Funke, J. (1992). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *German Journal of Psychology*, 16(1):24–43.

Galton, F. (1907). Vox populi. *Nature*, 75:450–451.

Gaskett, C., Wettergreen, D., and Zelinsky, A. (1999). Q -Learning in Continuous State and Action Spaces. *Australian Joint Conference on Artificial Intelligence*, 1747:417–428.

Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel*. Cambridge University Press.

Gibson, F. P. (1996). The sugar production factory - A dynamic decision task. *Computational and Mathematical Organization Theory*, 2(1).

Gibson, F. P., Fichman, M., and Plaut, D. C. (1997). Learning in Dynamic Decision Tasks: Computational Model and Empirical Evidence. *Organizational Behavior and Human Decision Processes*, 71(1):1–35.

Gigerenzer, G. (2001). The Adaptive Toolbox. In Selten, R. and Gigerenzer, G., editors, *Bounded Rationality: The Adaptive Toolbox*, pages 37–50. The MIT Press, Cambridge, MA.

Gigone, D. and Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin*, 121(1):149.

Goldman, A. I. (1999). *Knowledge in a Social World*. Oxford University Press.

Gonzalez, C., Lerch, J., and Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4):591–635.

Gonzalez, C., Vanyukov, P., and Martin, M. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, 21(2):273–286.

Granovskiy, B., Gold, J. M., Sumpter, D. J. T., and Goldstone, R. L. (2015). Integration of social information by human groups. *Topics in Cognitive Science*, 7(3):469–493.

Guala, F. (2005). *The Methodology of Experimental Economics*. Cambridge University Press, New York, NY, USA.

Gureckis, T. M. and Love, B. (2009a). Learning in Noise: Dynamic Decision-Making in a Variable Environment. *Journal of Mathematical Psychology*, 53(3):180–193.

Gureckis, T. M. and Love, B. C. (2009b). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*, 113(3):293–313.

Hahn, U., Hansen, J. U., and Olsson, E. J. (2018). *Truth tracking performance of social networks: how connectivity and clustering can make groups less competent*. Springer Netherlands.

Hahn, U., Harris, A. J. L., and Corner, A. J. (2009). Argument content and argument source: An exploration. *Informal Logic*, 29:337–367.

Harvey, N. and Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, 70(2):117–133.

Hayes, N. A. and Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition*, 28(3):249–276.

Hegselmann, R. and Krause, U. (2002). Opinion Dynamics and Bounded Confidence. *Simulation*, 5(3):2.

Heiser, W. and D'Ambrosio, A. (2013). Clustering and Prediction of Rankings Within a Kemeny Distance Framework. In Lausen, B., Poel, D. V. d., and Ultsch, A., editors, *Algorithms from and for Nature and Life*, chapter 1, page 547. Springer International Publishing, Switzerland, 1 edition.

Hertwig, R. and Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, 24:383–451.

Hill, G. W. (1982). Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin*, 91(3):517.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., Group, C. S. R., and others (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, 19(1):46–54.

Hogarth, R. M. (2001). *Educating intuition*, volume 17. University of Chicago Press.

Homer, J. B. and Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health*, 96(3):452–458.

Horne, Z., Powell, D., and Hummel, J. (2013). A single counterexample leads to moral belief revision. In *35th Annual Conference of the Cognitive Science Society*, pages 1–23.

Howes, A., Lewis, R. L., and Vera, A. (2009). Rational Adaptation Under Task and Processing Constraints: Implications for Testing Theories of Cognition and Action. *Psychological Review*, 116(4):717–751.

Hu, H. (2017). Competing opinion diffusion on social networks. *Royal Society Open Science*, 4.

Ibrahim Shire, M., Jun, G. T., and Robinson, S. (2018). The application of system dynamics modelling to system safety improvement: Present use and future potential. *Safety Science*, 106:104–120.

IMF (2019). Real GDP Growth.

Jackson, M. O. (2009). Social Structure, Segregation, and Economic Behavior. *SSRN Electronic Journal*, pages 1–19.

Jackson, M. O. (2010). *Social and Economic Networks*. Princeton University Press.

Jönsson, M. L., Hahn, U., and Olsson, E. J. (2015). The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition*, 142(191-204):191–204.

Kaya, Y. (1990). Impact of Carbon Dioxide Emission Control on GNP Growth: Interpretation of Proposed Scenarios. Technical report, IPCC Energy and Industry Subgroup, Response Strategies Working Group, Paris.

Keerthi, S. S. and Ravindran, B. (1994). A tutorial survey of reinforcement learning. *Sadhana*, 19(6):851–889.

Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, 30(1/2):81.

Kieras, D. E. and Meyer, D. E. (1998). The Role of Cognitive Task Analysis in the Application of Predictive Models of Human Performance. Technical report, University of Michigan.

Kolde, R., Laur, S., Adler, P., and Vilo, J. (2012). Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics*, 28(4):573–580.

Krackhardt, D. and Hanson, J. R. (1993). Informal networks.

Kretzschmar, M. and Morris, M. (1996). Measures of concurrency in networks and the spread of infectious disease. *Mathematical biosciences*, 133(2):165–195.

Kumar, R. and Vassilvitskii, S. (2010). Generalized Distances Between Rankings. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 571–580, New York, NY, USA. ACM.

Langley, P. and Simon, H. A. (1981). The Central Role of Learning in Cognition. In Anderson, J., editor, *Cognitive skills and their acquisition*, chapter 12, pages 361–380. Lawrence Erlbaum Associates, Hillsdale,NJ.

Laureiro-Martinez, D., Brusoni, S., Canessa, N., and Zollo, M. (2015). Understanding the exploration exploitation dilemma:: An fMRI study of attention control and decision-making performance. *Strategic management journal*, 36(3):319–338.

Lazer, D. and Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4):667–694.

Leiserowitz, A., Maibach, E., Roser-Renouf, C., and Hmielowski, J. (2012). Global Warming's Six Americas. Technical Report november 2011, Yale Project on Climate Change Communication, New Haven, CT.

Levi, I. (1983). *The enterprise of knowledge: An essay on knowledge, credal probability, and chance*. MIT press.

Lim, J. S. and O'Connor, M. (1995). Judgemental adjustment of initial forecasts: its effectiveness and biases. *Journal of Behavioral Decision Making*, 8(3):149–168.

Lorenz, J., Rauhut, H., Schweitzer, F., and Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22):9020–9025.

Lorge, I. and Brenner, M. (1958). A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance: 1920-1957. *Psychological Bulletin*, 55:337–372.

Ma, J., Liu, W., Dubois, D., and Prade, H. (2010). Revision rules in the theory of evidence. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 1:295–302.

Mann, M. and Gaudet, B. (2018). The "Kaya Identity".

March, J. G. (2008). Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1):71–87.

McGeorge, P. and Burton, A. M. (1989). The effects of concurrent verbalization on performance in a dynamic systems task. *British Journal of Psychology*, 80(4):455–465.

Melville, N. P. (2010). Information Systems Innovation for Environmental Sustainability. *MIS Q.*, 34(1):1–21.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. *Nips*, pages 1–9.

Moussaïd, M., Herzog, S. M., Kämmer, J. E., and Hertwig, R. (2017). Reach and speed of judgment propagation in the laboratory. *Proceedings of the National Academy of Sciences*, 114(16):4117–4122.

Moussaid, M., Kammer, J. E., Analytis, P. P., and Neth, H. (2013). Social influence and the collective dynamics of opinion formation. *PLoS ONE*, 8(11).

Neutel, A.-M., Bodini, A., and Berlow, E. (2014). Feedback loops: How nature gets its rhythms.

Nickerson, R. S., Butler, S. F., and Barch, D. H. (2017). Set size, assertion form, thematic content and sampling in the selection task. *Thinking and Reasoning*, 23(2):134–157.

Niu, S., Lan, Y., Guo, J., Cheng, X., Yu, L., and Long, G. (2015). Listwise Approach for Rank Aggregation in Crowdsourcing. In *Proceedings of the Eighth ACM*

International Conference on Web Search and Data Mining - WSDM'15, pages 253–262, Shanghai, China.

Orrell, D. (2011). Economyths: Ten Ways That Economics Gets It Wrong. *International Journal of Social Economics*, 38(9):821–822.

Osman, M. (2010). Controlling uncertainty: a review of human behavior in complex dynamic environments. *Psychological bulletin*, 136(1):65–86.

Over, D. E., Hadjichristidis, C., Evans, J. S. B., Handley, S. J., and Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, 54(1):62–97.

Pearl, J. and Russell, S. (2011). Bayesian Networks.

Pentland, A. (2014). *Social Physics*. Penguin Press.

Pihur, V., Datta, S., and Datta, S. (2007). Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach. *Bioinformatics*, 23(13):1607–1615.

Powers, W. T. (1973). Feedback: Beyond Behaviorism: Stimulus-response laws are wholly predictable within a control-system model of behavioral organization. *Science*, 179(4071):351–356.

Ranney, M. A., Clark, D., Reinholz, D. L., and Cohen, S. (2012). Changing Global Warming Beliefs with Scientific Information: Knowledge, Attitudes, and RTMD. *Proceedings of 34th Annual Meeting of the Cognitive Science Society*, pages 2228–2233.

Renda, M. E. and Straccia, U. (2003). Web Metasearch: Rank vs. Score Based Rank Aggregation Methods. In *Proceedings of the 2003 ACM symposium on Applied computing - SAC '03*, pages 841–846, Melbourne, FL, USA.

Revkin, A. and Seelye, K. (2003). Report by E.P.A. Leaves Out Data on Climate Change.

Ritchie, H. and Roser, M. (2019). Fossil Fuels.

Rosenzweig, C., Karoly, D., Vicarelli, M., Neofotis, P., Wu, Q., Casassa, G., Menzel, A., Root, T. L., Estrella, N., Seguin, B., Tryjanowski, P., Liu, C., Rawlins, S., and Imeson, A. (2008). Attributing physical and biological impacts to anthropogenic climate change. *Nature*, 453(May):353–357.

Sanderson, P. M. (1990). Implicit and explicit control of a dynamic task: Empirical and conceptual issues. Technical report, University of Illinois at Urbana-Champaign, Urbana, IL.

Schoenberg, T., Daw, N., Joel, D., and O'Doherty, J. (2007). Reinforcement Learning Signals in the Human Striatum Distinguish Learners from Nonlearners during Reward- Based Decision Making. *Applied and Environmental Microbiology*, 27(47):12860–12867.

Shanks, D. R. and St John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17(3):367–395.

Sherif, M. and Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change.* Yale Univer. Press, Oxford, England.

Simon, H. A. (1956). Rational Choice and the Structure of the Environment. *Psychological Review*, 63(2):129–138.

Simonini, T. (2018). Diving deeper into Reinforcement Learning with Q-Learning.

Singmann, H., Klauer, K. C., and Over, D. (2014). New normative standards of conditional reasoning and the dual-source model. *Frontiers in Psychology*, 5(APR):1–14.

Spence, A., Venables, D., Pidgeon, N., Poortinga, W., and Demski, C. (2010). Public Perceptions of Climate Change and Energy Futures in Britain: Summary Findings of a Survey Conducted from January to March 2010.

Stanley, W. B., Mathews, R. C., Buss, R. R., and Kotler-Cope, S. (1989). Insight Without Awareness: On the Interaction of Verbalization, Instruction and Practice in a Simulated Process Control Task. *The Quarterly Journal of Experimental Psychology Section A*, 41(3):553–577.

Sterman, J. (1994). Learning in and about complex systems. *System Dynamics Review*, 10(2-3):291–330.

Sterman, J. D. (1989). Misperceptions of Feedback in Dynamic Decision Making. *Organizational Behavior and Human Decision Processes*, 43(3):301–335.

Sterman, J. D. (2000). *Business dynamics: systems thinking and modeling for a complex world*. McGraw-Hill Education.

Sterman, J. D. (2002). System Dynamics: Systems Thinking and Modeling for a Complex World.

Sternberg, S. (2010). Psychology 600-301 Proseminar in Psychological Methods, Spring Semester 2004 Reaction-Time Experimentation. Technical report, UPenn.

Stewart, N., Chater, N., and Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, 53(1):1–26.

Sutton, R. and Barto, A. (1998). *Reinforcement Learning: An Introduction*. Bradford Books, MIT.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*, volume 2. Random House.

Thomas, M. S. C. and Van Heuven, W. J. B. (2005). Computational models of bilingual comprehension. In *Handbook of bilingualism: Psycholinguistic approaches*, pages 202–225. Oxford University Press New York, NY.

Uddén, J., Ingvar, M., Hagoort, P., and Petersson, K. M. (2012). Implicit Acquisition of Grammars With Crossed and Nested Non-Adjacent Dependencies: Investigating the Push-Down Stack Model. *Cognitive Science*, 36(6):1078–1101.

Wanzel, S. K., Schultze, T., and Schulz-Hardt, S. (2017). Disentangling the effects of advisor consensus and advice proximity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(10):1669.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20(3):273–281.

Watkins, C. (1989). *Learning from Delayed Reward*. PhD thesis, University of Cambridge.

Watkins, C. J. C. H. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4):279–292.

Watts, D. J. (1999). Networks, Dynamics, and the Small World Phenomenon. *American Journal of Sociology*, 105(2):493–527.

Weiner, I. B., Tennen, H. A., and Suls, J. M. (2012). *Handbook of Psychology, Personality and Social Psychology*. John Wiley & Sons.

Whalen, J., Gallistel, C. R., and Gelman, R. (1999). Nonverbal Counting in Humans: The Psychophysics of Number Representation. *Psychological Science*, 10(2):130–137.

Whitmarsh, L. (2011). Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. *Global Environmental Change*, 21(2):690–700.

Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1):79–82.

Yaniv, I. (1997). Weighting and Trimming: Heuristics for Aggregating Judgments under Uncertainty. *Organizational Behavior and Human Decision Processes*, 69(3):237–249.

Yaniv, I. (2004a). *Receiving other people's advice: Influence and benefit*, volume 93. Organizational Behavior and Human Decision Processes.

Yaniv, I. (2004b). The Benefit of Additional Opinions. *Current Directions in Psychological Science*, 13:75–78.

Yaniv, I. and Choshen-Hillel, S. (2012). Exploiting the wisdom of others to make better decisions: Suspending judgment reduces egocentrism and increases accuracy. *Journal of Behavioral Decision Making*, 25(5):427–434.

Yaniv, I., Choshen-Hillel, S., and Milyavsky, M. (2011). Receiving advice on matters of taste: Similarity, majority influence, and taste discrimination. *Organizational Behavior and Human Decision Processes*, 115(1):111–120.

Yaniv, I. and Milyavsky, M. (2007). Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes*, 103(1):104–120.

Zeron, E. S. (2008). Positive and Negative Feedback in Engineering and Biology. *Mathematical Modelling of Natural Phenomena*, 3(2):67–84.