Scaling Function Learning from Individuals to Groups

by

Chrissy M. Chubala

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg, Manitoba

Abstract

Scale invariance, the notion that scientific principles ought to hold over different scales of analysis, is a regularity in the physical and biological sciences but is underappreciated in psychology. Whereas the standard approach in psychology is to explain behaviours at different scales of analysis with different mechanisms, I argue that sophisticated behaviours at any scale are an emergent consequence of simple processes interacting with a structured environment. Changing the scale of analysis, whether temporal, physical, or otherwise, may alter the structure of the environment but need not imply changes to the mechanisms that interact with that environment. To illustrate the centrality of scale invariance and emergence to human cognition, I replicate signature findings from a function learning task after scaling up the unit of analysis, from individuals to groups. In a standard function learning task, individuals learn the relationship between two variables by trial and error, matching one variable (i.e., Y) to a target value of the other variable (i.e., X) and adjusting their responses according to feedback. In an analogous group function learning task, groups of non-communicating individuals learn the relationship between two variables by making individual-level decisions in response to group-level feedback. My experiments with this task demonstrate that groups, like individuals, can learn both simple and complex functions by trial and error, and can generalize their knowledge of a trained function to untrained target values in a transfer test. Groups are, moreover, resilient to disruption of their knowledge, a central feature of distributed representations in biological and artificial neural networks. The results recommend a principled approach to cognition, in which simple processes interact with the structure of the environment to produce sophisticated behaviours, and in which the patterns of behaviour produced at one scale of analysis are reproduced at other scales. Finally, the data show that, when constrained by a collective environment and common goals, individuals self-organize into unique decision roles that support group-level learning. An exploratory analysis of self-reported strategies, individual behaviours, and personality profiles demonstrates how complex social variables can help or hinder the emergence of learning at the level of the group.

Acknowledgements

First and foremost, I want to thank my graduate advisor, Dr. Randy Jamieson, for his invaluable support and mentorship throughout the completion of two degrees. He has generously given me his time and wisdom, often putting my needs ahead of his own to ensure I had the opportunities and experiences necessary to fully develop as a scientist. I could not have hoped for a better mentor, friend, and ally. I would also like to thank the members of my doctoral advisory committee for their guidance and constructive criticism, without which this document would have been inferior in every way: Dr. Murray Singer, Dr. Jason Leboe-McGowan, Dr. Andrea Bunt, and Dr. Valerie Thompson.

Additional gratitude is owed to those brilliant researchers whose kindness, curiosity, collaborative spirit, and advice has contributed to my educational and professional development in countless and immeasurable ways: Dr. Harinder Aujla, Dr. Glen Bodner, Dr. Jim Clark, Dr. Matt Crump, Dr. Chris Eliasmith, Dr. Peter Graf, Dr. Bill Hockley, Dr. Brendan Johns, Dr. Mike Jones, Dr. Harvey Keselman, Dr. Doug Mewhort, Dr. Penny Pexman, Dr. Jean Saint-Aubin, Dr. Steve Smith, Dr. John Vokey, and Dr. Doug Williams. I also owe a debt of gratitude for the generous funding assistance provided by the National Sciences and Engineering Council of Canada (NSERC), the University of Manitoba (Faculty of Arts, Faculty of Graduate Studies, and Department of Psychology), the Bernice D. Lough Psychology Graduate Bursary fund, and the University of Manitoba Graduate Student Association.

I am infinitely thankful for the many friends and peers who shared the late nights, long talks, coffee breaks, and cat .gifs that have kept my mind stimulated and my heart happy over the past few years: Joanna Bhaskaran, Teagan Bisbing, Tom Baril-Bissett, Crystal Boland, Melissa Bulloch, Tiffany Carther-Krone, Paul Cazzorla, Bob and Cindy

Chubala, Dr. Dawson Clary, Matt Cook, Dr. Evan Curtis, Damien Di Fede, Allison Dyck, Andrew Gabel, Heather Girodat, Tim Graham (R.I.P.), Alec Holowka, Eileen Holowka, Stephen Kiazyk, Dr. Cara Kroft, Phil Lavalée, Stephen Lewis, Jeff Lindsay, Sarah McAmmond, Dr. Matt Mechtley, the Pohl family, Kyle Pulver, Dan Reimer, Sarah Rigby, Andreas Rønning, Sal Russo, Adam and Bekah Saltsman, M. Schultz, Jackie Spear, Tamara UnRuh, Rory Waisman, Michelle Wright, and Robert Zetzsche.

Finally, this manuscript would not have existed without the tireless and unconditional support of my partner Dan Pohl, my scalebaby Huxley, my brother Kyle Chubala, and my distant but never absent companion Sara (twobit) Gross. You four keep me sane, happy, and always laughing. For all you've given me, this project is as much your baby as it is mine.

Dedication

This dissertation is dedicated to the memory of my mother, Angela Szczepanski (1958-2016), who taught me by example that a woman could achieve whatever she set her mind to. *Nigdy nie zapomnę o tobie, Anielka*.

Table of Contents

Considerations

List of Tables

List of Figures

Chapter 1: A Principled Approach to Cognition

Humans are capable of a wide array of sophisticated behaviours, from the development of language at a very young age to the use of formal logic when solving complex problems in adulthood. Like language development, many of our most sophisticated abilities seem to develop implicitly; that is, without our intending to learn, without awareness that learning has occurred, and without the ability to easily verbalize the nature of our knowledge (Cleeremans, 2006). Two fundamentally different theoretical approaches to cognition have developed in an effort to explain such instances.

In the first approach, sophisticated behaviours are presumed to be the result of sophisticated mechanisms. For example, our ability to keep a 7-digit telephone number in our minds between the moment we hear it and the moment we are able to write it down was taken as evidence for a sophisticated working memory system (e.g., Atkinson & Shiffrin, 1968; Baddeley & Hitch, 1974), a construct that has gained additional complexity over time (e.g., Baddeley, 2000). This working memory system is in contrast to the episodic memory system (e.g., Tulving, 1983), with which individuals can mentally time travel through previously experienced events, or the semantic memory system (e.g., Squire, 1992), with which individuals can store and recall general world knowledge. Notably, this approach to cognition encourages the partitioning of cognitive systems. When a researcher discovers a phenomenon, they develop a unique system or theory to explain that phenomenon, often without reference to existing theories or phenomena. Although this approach is very common in psychology (see also Knowlton & Squire, 1993; Owen, 1998; Schacter & Tulving, 1994) and undoubtedly valuable, it may overlook commonalities between proposed systems and theories.

In the second approach, sophisticated behaviours are presumed to arise from simple, generalized processing faculties responding to complex or structured environments. This approach was perhaps most eloquently described by Simon's (1969) parable of the ant: if one plotted out on paper the course of an ant's path as it travelled along a beach, from its colony toward a distant food source, one might be tempted to look at the irregular, complex path and conclude that the ant had a sophisticated or intentional route in mind for its walk. Simon argues instead that the apparent complexity of the ant's route is a reflection of complexity in the surface of the beach, rather than complexity in the ant itself. He goes on to argue the same of human behaviours: "Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behaviour over time is largely a reflection of the complexity of the environment in which we find ourselves" (p. 53). Importantly, the explanation for a given behaviour cannot be found in either the human or the environment alone, but in the interaction between them.

In contrast to the first approach, this view of cognition encourages the development of principled accounts: simple mechanisms that support a wide array of behavioural phenomena by virtue of their interaction with different environmental structures. It is the changing structure of the environment that leads to different observable phenomena, for instance in working memory, episodic memory, and semantic memory, rather than different systems altogether.

One such principled approach was given by Surprenant and Neath (2009) and their colleagues, known as the Scale Independent Memory, Perception, and Learning model (SIMPLE; Brown, Neath, & Chater, 2002; Brown, Neath, & Chater, 2007; Chater & Brown, 1999; Neath & Brown, 2006). In the SIMPLE model, items are stored to memory as a point in psychological space, which location varies as a logarithmic function of the item's temporal distance from the current moment. Thus, as time increases and additional items are presented, older items become increasingly compressed in psychological space and thereby less discriminable from their neighbours. This assumption, along with a few others specific to the particular task being modeled, allows the SIMPLE model to account for patterns of retrieval in serial recall and free recall (e.g., serial position curves; Brown, Neath, & Chater, 2007), inter-list and intra-list interference effects (Neath & Brown, 2012), word length effects (Derraugh, Neath, Surprenant, Beaudry, & Saint-Aubin, 2017), retention interval and rehearsal effects typically attributed to a working memory system (Neath & Brown, 2006; Neath, VanWormer, Bireta, & Surprenant, 2014), and list length effects (Quinlan, Neath, & Surprenant, 2015). Critically, the basic principles of the model are held constant over all of these demonstrations; changes in the structure of the task interact with those basic principles to produce behaviours previously believed to be the work of distinct systems.

Scale Invariance

Central to their argument, and to the thrust of this thesis, is the demonstration that these effects are invariant over different time scales. Take, for example, the division of memory systems by temporal scale. Atkinson and Shiffrin's (1968) modal model proposes that distinct systems exist for sensory short-term memory (retention for up to one second), short-term memory (retention for up to thirty seconds), and long-term memory (permanent retention). Different principles, models, and theories have arisen for each of these systems, often overlooking their necessary connectedness.

The SIMPLE model, on the other hand, can account for sensory, short-term, long-term, and working memory with the same principles, because the discriminability of items within psychological space is taken as a ratio rather than an absolute (Neath & Brown, 2006). Moreover, the model can account for semantic memory as a special case of very long-term memory. Kelley, Neath, and Surprenant (2015), for example, demonstrated that semantic memory shares the same serial position curves generally reserved for episodic memory phenomena, further breaking down distinctions between memory systems (see also Neath, Kelley, & Surprenant, 2016).

This notion of scale invariance is an important one in the search for unifying principles of cognition. In the hard sciences, principles and laws tend to hold over a wide range of scales, with certain exceptions at the extremes (for a similar call-to-arms in biology, however, see Castiglione, Pappalardo, Bianca, Russe, & Motta, 2014). Newton's law of gravitation, which states that the force of attraction between two bodies is directly proportional to the product of their masses and inversely proportional to their squared distance, is true at all masses and distances (with exceptions for the extreme of subatomic scales). Scale invariance is, moreover, a feature of nature that can be observed in phenomena as diverse as the fractal patterns of certain plants or the bronchial trees in mammalian lungs (Schlesinger & West, 1991), foraging behaviour (MacIntosh, Pelletier, Chiaradia, Kato, & Ropert-Coudert, 2013; see also Hills, Jones, & Todd, 2012, for patterns of foraging behaviour in human semantic memory search), DNA sequences (Peng et al., 1992), heartbeat rates (Peng, Havlin, Stanley, & Goldberger, 1994), and weather fluctuation patterns (Koscielny-Bunde et al., 1998). Moving beyond the natural sciences, scale invariant patterns have also been demonstrated in perception (Chater & Brown, 1999), urban population growth (Makse, Havlin, & Stanley, 2002), and economics (Mantegna & Stanley, 2002).

In contrast to this rich database of examples in the natural sciences, research in psychology often assigns different explanations to different scales, whether temporal, physical, or spatial. Yet humans are a product of the same evolutionary processes that favoured scale invariance and self-similarity in DNA sequences, bronchial trees, and heartbeat rates. Further, humans live within urban populations that likewise show patterns of scale invariance. It should not therefore be controversial to suggest that our cognition could be similarly described by principles of scale invariance (e.g., Barenblatt, 1996; Surprenant & Neath, 2009). As outlined above, the research group that developed the SIMPLE model has shown that when applied to different time scales of memory, the same principles can account for a range of effects previously assumed to pertain to distinct memory systems (for another computational model of scale-invariant representations of time in memory, see Howard & Eichenbaum, 2013; Howard, Shankar, Aue, & Criss, 2014).

Self-Similarity and Hierarchies

Although I take scale invariance as an ultimate target for psychological science, I acknowledge the difficulty of obtaining this target in light of the stochastic nature of human behaviour. In other sciences, scale invariance is often a mathematically exact argument, as with Newton's law of gravitation: the very same mathematics apply at all scales of analysis. If humans were as well-behaved as elements or planets, however, the problem of psychology would be long since solved. Instead, we must consider less mathematically exact ways in which phenomena and behaviours are self-similar at different scales of analysis.

Indeed, Neath and colleagues' demonstration of scale invariant memory phenomena across time scales (e.g., Neath & Brown, 2006) is more accurately a demonstration of self-similarity across time scales: although they used the mathematics of the SIMPLE model to derive predictions for performance across scales, the predictions were an approximation rather than an exact derivation of recall performance. Throughout this thesis, I follow Neath and colleagues' lead by focusing on this less exact approach to scale invariance. However, I deviate from their work by exploring the approximate self-similarity of cognition across a qualitatively different kind of scale. In examining the effect of time scales on recall memory, Neath and colleagues maintained individual participants as the unit of analysis. Likewise, in examining the effect of object size on gravity, physicists are still interested in the behaviour of the individual object. By contrast, the focus of this thesis will be more in line with studies of fractal self-similarity, in which the scale of interest is a hierarchical one.

Fractals in nature can be described by iterative functions that produce similar, though not identical, shapes at different levels of analysis. On a fern plant, a single leaflet is approximately similar in shape to the leaf of which it is one part; a single leaf is approximately similar in shape to the frond of which it is one part; and so on. In this way, self-similarity is expressed across a hierarchical scale. Within humans, fractal algorithms can capture the shape of physiological growth, such as bronchial trees (e.g., Schlesinger & West, 1991) or the vascular system (Bassingthwaighte, 1992). An interesting question for psychology is whether the behaviour of an individual human can be described as self-similar to the behaviour of a group, of which that human is one part.

The experiments presented in this thesis examine whether the behaviour of groups is similar to the behaviour of individuals, when the two levels of analysis undergo analogous experimental tasks. Whereas an individual is comprised of neural networks cooperating to produce behaviours, a group can be considered a network of individuals that must cooperate to produce behaviours. My hypothesis is that a network of humans, situated within the constraints of a function learning task, will demonstrate similar patterns of behaviour to those demonstrated by a network of neurons under comparable task constraints. Not only would such results be consistent with theories of scale invariance, but they would also support the broader view that the apparent complexity of a system's behaviour is a result of the interaction between simple processes and structured environments (e.g., Simon, 1969).

Group Cognition

In the past, most psychological research on cognition in groups was conducted within the realms of social and organizational psychology (for a review see Levine, Resnick, & Higgins, 1993). Perhaps most famously, Galton (1907) observed that the median estimate of the weight of an ox at a country fair, compiled from a group of individuals' guesses, was more accurate than any one individual guess. He called this the wisdom of crowds, and it is usually explained by the elimination of individual noise, particularly when the diversity of the group is large (Surowiecki, 2004; Yi, Steyvers, Lee, & Dry, 2012). Another well-known phenomenon of group cognition is the aptly named groupthink, whereby a group's decision is negatively influenced by the desire for conformity or harmony within the group (Janis, 1972, 1982). When groups are highly cohesive, insulated from expert opinions, and motivated to reach a consensus, groupthink results in irrational or dysfunctional decision-making. Similarly, the well-known Asch conformity experiments demonstrated that participants would yield to a clearly incorrect opinion when faced with the social pressure of a majority (Asch, 1951).

Although such avenues of study continue to this day, researchers in the cognitive sciences have recently begun examining group cognition by adopting theories of collective structures from the biological and physical sciences (for review see Moussaid, Garnier, Theraulaz, & Helbing, 2009). Such approaches focus on the emergence of global patterns from the interactions of smaller or simpler entities, a theoretical stance that has clear connections to Simon's (1969) parable of the ant. In fact, the complexity of an ant colony is a prime example of an emergent structure. Within a colony, orders are not handed down from the queen to her workers. Instead, each individual ant reacts to the immediate chemosensory stimuli around them, each one in turn adding to the chemical environments of the other ants in the colony. The result is a decentralized system of decision-making, comprised of many individual and autonomous units but nevertheless capable of complex behaviours at the level of the colony (Johnson, 2001).

The indirect communication between ants, and other such agents in similar systems of distributed knowledge, is known as *stigmergy* (Dorigo, Bonabeau, & Theraulaz, 2000). As each agent simultaneously modifies their environment and responds to the collection of modifications made by others, an apparent complexity of behaviour at the level of the group emerges. This concept is entirely in line with Simon's (1969) dictum, except now the interaction between simple processes and structured environments is a recursive one: one agent's behavioural response to an environment changes the structure of that environment and thereby influences the behavioural responses of other agents, whose responses change the structure of the environment, and so on. The complexity of group interactions is itself a suitably structured environment with which to evoke apparently complex behaviours from even the simplest of processing agents. For example, Rubenstein, Cornejo, and Nagpal (2014) demonstrated the emergence of complex structures from simple local behaviours in small autonomous robots. A swarm of 1024 robots were able to assemble into a variety of shapes on command, simply by communicating with and responding to their local neighbours over time.

Although humans are arguably more complex than ants or robots, recall that Simon's (1969) argument parable applies to humans just as well as to ants. A real-world example of emergence in groups of humans comes from the informal online social experiment known as *Twitch Plays Pokemon* ("Twitch plays Pokemon," 2017). The goal of the endeavour was to see whether a large group of humans could collectively play through a videogame by submitting commands as individuals. Live footage of the running game was streamed publically on the popular videogame-focused streaming website *Twitch.tv*. Meanwhile, viewers could type commands to the game in the stream's chat window in order to play (i.e., equivalent to pressing buttons on a gaming console controller). Visual feedback was given in approximately real-time on the stream. The stream's popularity grew rapidly, and soon thousands of viewers were submitting commands at once, peaking at 121,000 simultaneous players.

As might be anticipated, the behaviour of the on-screen avatar was at first erratic and chaotic, plagued by long, unproductive sequences of opening and closing the avatar's equipment bag, turning in circles, repeatedly entering and exiting areas, and so on. Eventually, however, the avatar's behaviour became more organized as individual players developed roles and strategies for their contributions. Knowledge of how to achieve the group's goal (i.e., to successfully play through the game) did not reside in any one player, but was instead represented across all players in a distributed fashion. Ultimately, the game was finally completed after 16 days of continuous gameplay: an extremely large network of humans, each submitting individual responses to the game's current state, was able to complete an extremely complex task at the level of the group. Although the outcome of the experiment is interesting, the absence of experimental control led to a great deal of noise in the form of internet "trolls" who intentionally sabotaged the group's efforts. Combined with the unstructured nature of non-empirical social experiments and the sheer volume of data produced, the results of the Twitch Plays Pokémon experiment are difficult to interpret.

Robert Goldstone and his colleagues have provided numerous demonstrations of emergent group-level cognition in the laboratory (for review see Goldstone & Gureckis, 2009; Goldstone, Roberts, & Gureckis, 2008). In a path formation study, several participants had to independently travel between given points in a two-dimensional space, with the caveat that more well-traveled locations in the space were less costly to traverse (Goldstone, Jones, & Roberts, 2006). Goldstone and colleagues found that participants' movements became a compromise between travelling via the routes they wanted to take and travelling via the routes other participants had taken before. In this way, globally efficient routes emerged from the indirect, stigmergic interactions between individual participants. In an analogous multiplayer foraging game, participants moved toward food-dense regions of a two-dimensional space, but away from other foragers who might do the same, leading to group-level oscillatory migrations between resource patches (Goldstone & Ashpole, 2004; Goldstone, Ashpole, & Roberts, 2005).

Gureckis and Goldstone (2006) offer three properties of distributed systems. First, distributed systems are comprised of individually identifiable units whose functioning can be described independently of the group context. Second, the units in these systems must be loosely coupled and capable of influencing one another. And third, the pattern of connectivity between the units must be dynamic. When these properties are met by groups of humans, group-level behaviour takes on a life of its own and can be evaluated separately from the behaviour of individuals. The groups are said to be distributed systems of knowledge, with behavioural profiles that can be, and usually are, qualitatively different from the behaviours of the individuals in the system.

Whereas Goldstone and his colleagues have been primarily interested in the unique behavioural profiles of distributed cognitive systems, my focus on scale invariance as a principle of nature guides me in a slightly different direction. Just as sophisticated behaviours emerge from the interaction of a brain (i.e., a collection of neurons) with a structured environment, my aim is to show that analogous behaviours emerge from the interaction of a group (i.e., a collection of brains) with a similarly structured environment. In order to make this case, I turn to function learning, a well-studied task in cognitive psychology.

Function Learning

In a standard function learning task, individuals must learn the relationship between two continuous variables, X and Y, by trial and error. The cover story differs by experiment, but one example is the relationship between the dosage amount of a drug (i.e., X) and the level of the drug's effect on a patient (i.e., Y). On every trial of the task, participants give an estimate of the magnitude of the effect that they would expect from a given drug dosage, and then receive feedback as to the correct magnitude of the effect. Over repeated trials, participants' estimates become more accurate, such that the difference between their estimated drug effects and the actual drug effects associated with a given dosage decreases. This improvement in accuracy is taken as evidence of learning. Following a training phase, in which several X-Y mappings are learned, participants are tested on new stimuli that were not seen during training. This transfer test measures how well participants can apply the recently acquired knowledge of the X-Y mapping to new targets. Researchers are interested in participants' ability to both interpolate between previously studied mappings and extrapolate beyond the range of mappings studied.

The X-Y relationship in a given experiment can be formally described by some mathematical function. In a linear function, a one-unit change in X is associated with a constant amount of change in Y. In non-linear functions, the relationship is more complicated. Busemeyer, Byun, DeLosh, & McDaniel (1997) described a number of principles that summarized the basic findings in the function learning literature, a handful of which will be relevant to this thesis. First, monotonic functions are learned faster than non-monotonic functions (e.g., DeLosh, 1995). Second, linear functions are learned faster than nonlinear functions (e.g., Deane, Hammond, & Summers, 1972; DeLosh, Busemeyer, & McDaniel, 1997; Koh & Meyer, 1991). Third, systematically increasing or decreasing training sequences (i.e., as compared to randomly ordered training) facilitates the learning of difficult functions (e.g., DeLosh, 1995). Fourth, performance on test stimuli that interpolated between trained stimuli is almost as accurate as performance on previously trained stimuli themselves (e.g., DeLosh et al., 1997; Koh & Meyer, 1991). Finally, performance on test stimuli that extrapolate beyond the range of trained stimuli is worse than on interpolated test stimuli, but still better than performance before any training (e.g., DeLosh et al., 1997; Wagenaar & Sagaria, 1975).

In order to make sense of these phenomena, researchers have turned to computational models of function learning. The most common approaches are associative or similarity-based models that make use of connectionist architectures (e.g., Busemeyer et al., 1997; DeLosh et al., 1997; McDaniel & Busemeyer, 2005). These models assume that participants learn by forming associations between observed X and Y values and subsequently generalize to interpolated and extrapolated test stimuli as a function of their similarity to trained stimuli. More recently, so-called “hybrid” models, which combine associative mechanisms with more explicit mental representations of the learned functions, have emerged as front-runners due to their ability to predict complications of and abnormalities in extrapolation performance across a range of functions (e.g., Kalish, Lewandowsky, & Kruschke, 2004; Kwantes & Neal, 2006; Kwantes, Neal, & Kalish, 2012).

Throughout this thesis, I take the function learning literature as my benchmark for assessing self-similarity along a hierarchical scale, from individuals to superordinate groups of individuals. The signature findings and computational approaches described above serve as guideposts in my exploration of function learning in groups.

There is at least one comparable example of purely cooperative group learning in the published literature. Roberts and Goldstone (2011) examined groups’ ability to learn in what they called a Group Binary Search (GBS) game. In their web-based task, groups of participants had to match some unknown target number between 51 and 100 by each choosing a response between 0 and 50. The program summed the responses of the participants and gave feedback relative to the target. Groups of varying sizes were able to match targets, though larger groups were slower to do so. Moreover, groups became faster at matching target values over time, indicating group-level learning. Although I was unaware of this work at the outset of developing this thesis, it anticipates some of the results to be presented herein, particularly in Experiment 1. Nevertheless, the group function learning task extends the examination of purely cooperative group tasks beyond the limited purview of the GBS game and builds on the insights provided by Roberts and Goldstone.

Structure of the Thesis

I will examine the ability of groups to learn functions in a series of three experiments. Over the course of the thesis, I aim to satisfy three goals. First, I will demonstrate that signatures of individual function learning generalize to function learning in groups. Specifically, I will show (1) that groups can learn both simple and complex functions; (2) that groups can generalize their learning to new target values in a subsequent transfer test; and (3) that performance on target values that extrapolate beyond those studied during training is worse than performance on target values that interpolate between those studied during training. Second, I will show that groups are resilient to the effects of disruption after learning a function, a result that speaks to the distributed nature of the cognition that emerges during group learning. Third, I will explore how the behavioural characteristics and decision processes of individuals within groups relate to group-level performance in function learning.

The overarching theme of the thesis is that of scale invariance, the notion that scientific principles for explanations of thinking should apply over a wide range of scales. From the perspective of self-similarity, patterns of behaviour observed at one level of a hierarchical system should replicate at other levels of the system. In the context of this thesis, I aim to show that the signatures of function learning in individuals are similar to those demonstrated by groups of individuals in an analogous task. Moreover, I aim to demonstrate that a central characteristic of both biological and artificial neural networks (i.e., resilience to disruption) is similarly expressed in groups of individuals. In both cases, I will argue for a simplified view of cognition whereby a system’s intelligence can be described by simple processes operating within the constraints of a structured environment.

Chapter 2: Function Learning in Groups

In a standard function learning task, individuals must learn the relationship between two variables, for example the dosage amount of a drug (i.e., X) and the level of the drug's effect on a patient (i.e., Y). Participants accomplish this by trial and error: on every trial they provide estimates for the effect of the drug they expect from a given drug dosage, and then receive feedback as to the correct level of the effect. Over a number of repeated blocks, the deviation between participants' estimates of Y for each trained value of X and the correct value of Y for each X, decreases. Evidence of this learning over blocks exists for both linear relationships (i.e., each one unit increase in X is associated with a constant increase in Y) and more complex non-linear relationships (e.g., quadratic, exponential, power).

The goal of Experiment 1 was to replicate signature findings from the individual function learning literature when examining function learning in a group of individuals. There are two particular signatures that I took as the focus of this experiment: the finding that both linear and non-linear functions can be learned, and the finding that linear functions are learned faster than non-linear functions. If principles underlying function learning are self-similar at different scales of analysis, then these signatures should be expected to emerge from group-level function learning as well.

Experiment 1

In this experiment, participants were tested in groups of 10 individuals. On each trial, participants were given a target number that represented the number of them that were to respond with an *i>clicker* remote. In a Simple condition, the target values given to participants were the actual values they were aiming for. In a Complex condition, the true target values were different from (i.e., a cubic-like transformation of) the target values given to participants. After every trial, groups received feedback about both the distance and the direction of their error relative to the true target value (e.g., "On that trial, you were two over the target."), and attempts continued until the target was matched exactly. Note that feedback in the Complex condition pertained to the true (transformed) target values, even though participants were given the untransformed values as goals. For example, all participants were instructed to match the targets 3, 5, 7, and 9. Whereas these were the true targets in the Simple condition, 3, 5, 7, and 9 mapped to the true targets 4, 5, 5, and 8 in the Complex group. For each condition, four target values were trained in this manner for three consecutive blocks of trials.

If groups can learn these functions, we should expect to see an increase in accuracy over the course of training blocks. Because the direction of the error is irrelevant to the research question, the measure of accuracy used was the squared error:

$$Squared\ error = (response\ value - actual\ target\ value)^2$$

This measurement has the added benefit of emphasizing the differences between low-error and high-error trials, thereby increasing the power to detect differences[1]. Thus, learning will be demonstrated by a decrease in squared error over the blocks of training, as determined by a polynomial trend analysis.

A second question of interest is whether the complexity of the function (i.e., simple vs. complex) impacted the shape or extent of learning. In individual function learning experiments, it has often been noted that linear relationships (e.g., my Simple condition) are more readily learned than non-linear relationships (e.g., my Complex condition; see Deane, Hammond, & Summers, 1972; McDaniel & Busemeyer, 2005). DeLosh et al. (1997), for

---

[1] My choice of a more powerful dependent variable was made to compensate for the practical constraints of my experimental protocol that limited my sample sizes.

instance, showed that performance on quadratic functions was initially more prone to error than performance on linear functions, but that the deficit shrank with training. I anticipated a similar trend for groups in my Complex condition, which learned a cubic-like function where the target values 3, 5, 7, and 9 were mapped to the true target values of 4, 5, 5, and 8.

Method

Participants

Two hundred and ten introductory psychology students participated in this experiment in exchange for partial course credit. Participants were run in groups of 10, and each group was randomly assigned to one of two conditions that differed only in terms of the relative complexity of the function they learned: Simple or Complex. Ten groups (i.e., a total of 100 participants) were run in the Simple condition. Eleven groups (i.e., a total of 110 participants) were run in the Complex condition. One group in the Complex Condition was excluded due to intentional sabotage by a student[2], leaving a total of ten groups per condition. Each group of 10 participants served as one subject unit in the analyses below.

Apparatus

All learning data were collected via ten *i>clicker+* remotes in communication with an *i>clicker* USB base system for Windows (v6.41; Macmillan, New York). Each remote was affixed to the center of a halved cardboard box and distributed around the laboratory with all seats facing outward. This setup ensured that participants' actions on the remotes were not visible to any other participant in the room, guaranteeing that each student would make independent choices. A press of the *A* button was taken as a response, whereas a press of the *E* button was taken as a decision to "not respond," thereby ensuring that auditory

---

[2] These instances of sabotage will be described in the Procedures section below, along with my approach to detecting and dealing with them.

feedback in the room was a non-discriminative cue. The buttons were chosen to be as physically distant as possible in order to minimize response errors. Following three blocks of learning, additional information was collected from participants on paper.

Procedure

Pre-training. After signing consent forms and receiving instructions on the basic nature of the task, two practice trials tested for comprehension and participation, as well as hardware functionality. On the first practice trial, participants were asked to respond by pressing the *A* button on their remotes. On the second practice trial, participants were asked to “not respond” by pressing the *E* button on their remotes. Both practice trials continued until all participants responded correctly; only occasionally did this require any intervention from the experimenter.

Training. Groups participated in three blocks of training. Each block involved learning about four target numbers, always in the same increasing order: 3, 5, 7, and 9. The targets represented the number of participants who were expected to respond with the *A* button on their remote; the remaining participants who chose to “not respond” were required to press the *E* button instead. This ensured a response from every participant. On every trial, the number of participants who responded with the *A* button was taken as the group-level response to be compared against target values. After every trial, groups received directional (i.e., over or under) and numerical (i.e., the distance of error) feedback on their performance, relative to the true target values. Critically, and unbeknownst to participants, the group-level responses in the Complex condition were compared to transformations of the given targets. Target values in that condition were transformed on a cubic function and rounded to whole numbers, so that the feedback compared responses to

4, 5, 5, and 8. In the Simple condition, the group-level responses were compared to the untransformed targets given (e.g., 3, 5, 7, and 9).

Within each block, training on a given target number was repeated until it was achieved on two consecutive trials, or until the group reached a maximum of 19 attempts. The criterion of two consecutive successful trials was set to ensure that participants were actively engaged with the experiment and to demonstrate that a group had in fact learned how to match the target.

Detecting Sabotage. An unexpected consequence of setting a criterion of two consecutive successes was that groups could conceivably achieve a target multiple times, but still not be able to progress through the experiment. Specifically, this requirement provided an opportunity for even a single student to hinder the group's progress by changing his or her response on the trial following a success, behaviour I hereafter refer to as sabotage[3]. Although sabotage was the exception rather than the rule during data collection, its occurrence was prevalent enough to warrant action.

Thankfully, sabotage was easy to detect: a sabotaged group would repeatedly achieve a target on one trial and fail by an error of one (typically in the same direction each time) on the next trial. When this cycle occurred twice in a row, I interrupted the experiment to warn participants that I could tell from the input which remote was responsible for the sabotage, and to remind them that the experiment would take much longer to complete if the saboteur continued. This was usually, but not always, sufficient to deter further sabotage.

---

[3] Of course I cannot prove that the behaviour was intentionally malicious, but the tendency for the behaviour to cease after a warning suggests that this was most often the case.

However, the small sample sizes necessitated by the high time and resource costs of these experiments meant that I could not afford to lose power to artificially noisy data. Sabotage behaviours artificially inflated both the number of trials taken to achieve a target and the average squared error of the trial set (see Chapter 5). Moreover, participant morale was noticeably affected by these behaviours. As such, I opted to exclude from analysis all groups for which sabotage behaviour was noted. Because it could be detected immediately during data collection, exclusion decisions were made and additional experimental sessions were scheduled prior to any analysis of the data. This ensured that the inclusion and exclusion of datasets was impartial, rather than cherry-picked in favour of my hypotheses.

In the present experiment, only one group needed to be excluded. In later experiments, the ratio was higher. Although I excluded them from primary analyses, subsequent exploration of these sabotaged groups nevertheless revealed interesting behaviours that will be discussed later in the document (Chapter 5).

Additional tasks. Between blocks of training, a brief semantic distractor activity was conducted. After successful completion of the first block, participants were each asked to name aloud an animal beginning with a given letter (e.g., G). After successful completion of the second block, participants were each asked to name aloud an animal that might be found in a given environment (e.g., underwater). The purpose of these distractor activities was twofold. First, the semantic nature of the distractor tasks was a strong divergence from the numerical nature of the primary task, potentially inhibiting rehearsal between blocks and limiting ceiling effects. More importantly, the distractor tasks helped to maintain participants' attention, motivation, and engagement. These activities were fun, novel, and engaging where the primary task was repetitive. Whether the more social aspects of these activities (e.g., participants could now look at one another and there were often attempts to make peers laugh with a clever response) had any impact on the group's ability to form a cohesive learner is an interesting question for future inquiry.

Following successful completion of the third and final block of training, participants were asked to write down free-formed thoughts about the strategies they used to make their decisions throughout the experiment. These responses were coded into twenty-six categories according to the strategy or strategies reported. Because this chapter is focused on the group-level analysis of learning, more complete descriptions and analyses of these categorical response codes are reserved for Chapter 5.

Results

The results of Experiment 1 are shown in Figure 2.1. The primary measure of interest in this Experiment is the mean squared error averaged over all trials within a block. A 2 (Condition: Simple vs. Complex) x 3 (Block) mixed measures ANOVA found a main effect of block, $F(1.39, 24.98) = 7.91$, $p = .005$[4], but no main effect of condition, $F(1, 18) = 1.98$, $p = .18$, indicating that performance in both conditions improved over blocks. Polynomial a priori contrasts on the main effect of block confirmed that squared error decreased linearly over blocks of training, $t(36) = 3.97$, $p < .001$. If training performance followed a power (Newell & Rosenbloom, 1981) or exponential (Heathcote, Brown, & Mewhort, 2000) learning curve, with greater benefits in the early portion of training than in the later

---

[4] Throughout this document, tests of sphericity were conducted for all repeated measures analyses. In cases where sphericity was violated, corrections were applied. As per the recommendations of Maxwell & Delaney (2004), the Greenhouse-Geisser correction was applied whenever epsilon was less than 0.75, and the Hyunh-Feldt correction was applied whenever epsilon was greater than 0.75. Thus, when the reported degrees of freedom for repeated measures tests are integer numbers, it can be assumed that the assumption of sphericity was not violated. When the reported degrees of freedom are not integer numbers, the assumption was violated but corrected for with the appropriate correction.

portions, a quadratic contrast might have been expected to fit the data as well; this contrast was not supported, $t(36) = 0.16$, $p = .88$, indicating that learning was best described as a linear decrease in error. Although visual analysis suggests that, consistent with findings in individual function learning (e.g., McDaniel & Busemeyer, 2005), there was an initial deficit for learning complex functions that decreased with additional training blocks, the block by condition interaction was not supported statistically, $F(1.39, 24.98) = 0.57$, $p = .51$. This result suggests that learning may not have differed between the two conditions, but a failure to reject the null hypothesis alone cannot confirm this possibility.

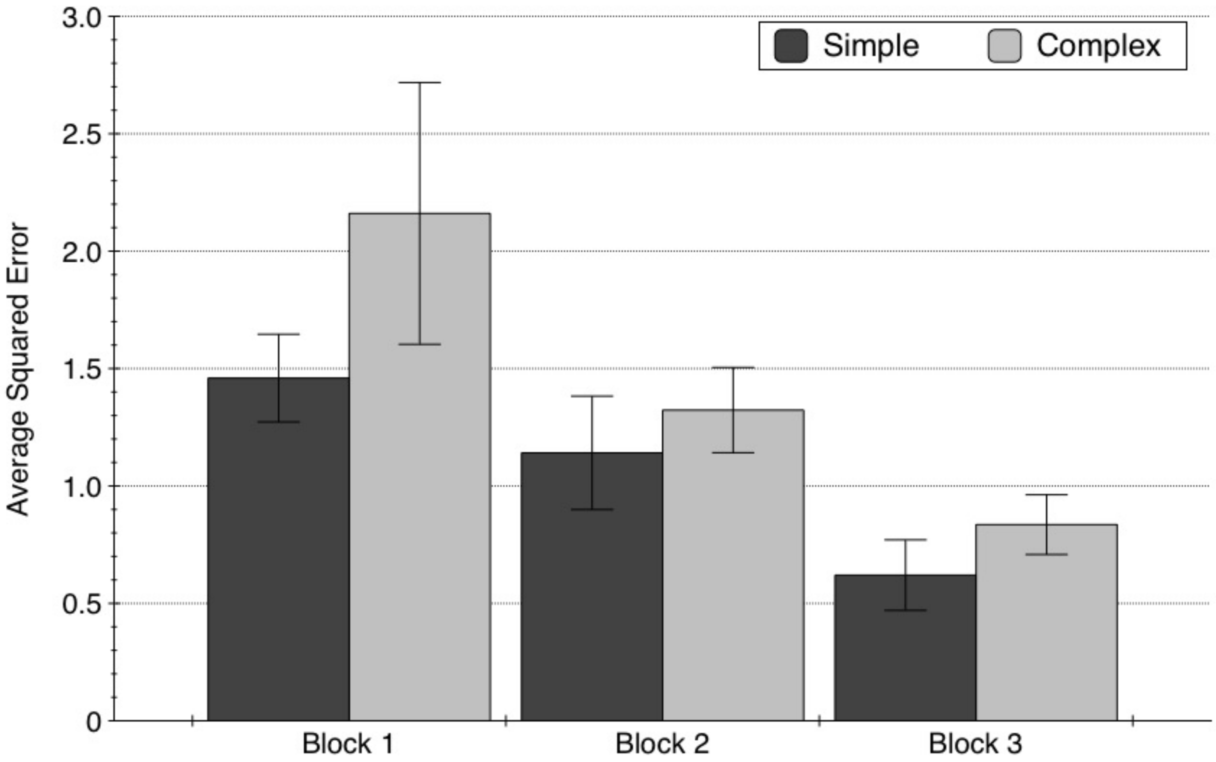

Figure 2.1. Squared error averaged over all attempts in a block in Experiment 1, as a function of block and condition. Error bars represent SEMs.

Although I took as my primary measure the mean squared error averaged over all trials within a block, there are two alternate ways of examining performance in this experiment. First, one could look at the number of trials groups take to match targets. However, this measure does not consider how close a group was to the target on a given attempt, only whether they did or did not exactly match it. Second, one could look at the mean squared error not over *all* trials within a block, but over *only the first attempts* at each target in a block. However, this snapshot of performance may not necessarily reflect how much a group has learned, as disregards performance on all other attempts in a block. By contrast, averaging squared error over all trials within a block captures not only a measure of error (i.e., the numerator in this mean is equal to a sum of all errors within a block) but also a measure of the number of trials taken to achieve a target (i.e., the denominator). In the current experiment, all three measures bore out identical statistical conclusions, and thus only average squared error is reported. A full report of results with the alternative measures can be found in Appendix A.

Discussion

In this experiment, groups of participants learned both simple linear and complex cubic-like functions, as evidenced by decreasing error over the course of training. Learning curves were similar for groups that learned a simple (linear) function and those that learned a complex (non-linear) function. Importantly, groups accomplished this learning without any communication between individuals; simple binary response decisions at the level of the individual resulted in emergent complex behaviours at the level of the aggregate.

The data support my larger argument about self-similarity, that similar behaviours emerge from individuals in a standard function learning tasks as from groups in my group function learning task. In individual function learning, non-linear functions are generally more difficult to learn than linear functions (e.g., Deane, Hammond, & Summers, 1972; McDaniel & Busemeyer, 2005). DeLosh et al. (1997) found an initial deficit in performance while learning quadratic functions as compared to linear functions, but that this deficit shrank with training until performance in the two conditions was identical. Given that the cubic-like function I employed for the Complex condition is mathematically *more* complex than a quadratic function, I expected to easily reproduce this finding. However, although the pattern of results in Figure 2.1 is consistent with that regularity in individual learning, the necessary condition by block interaction was not supported statistically.

There are several possible reasons for this deviation. First, the small sample size of this experiment (N = 20 groups, with n = 10 groups per condition) may have limited the power to detect the interaction. DeLosh et al. (1997), for instance, had 36 participants to my 10 groups per function condition. Indeed, a post-hoc power analysis using G*Power 3.1 (Faul, Erdfelder, Lang. & Buchner, 2007) suggested that the power to detect this interaction was a mere $1-\beta = .21$; with the effect size obtained in the experiment ($\eta_p^2 = .03$), 26 groups, or 260 participants, per condition would have been needed to detect the effect at a power of $1-\beta = .80$. Second, it may be the case that the non-monotonic quadratic function in DeLosh et al.'s experiment was in fact more complex than the cubic-like function learned by my Complex condition. There has been some suggestion in the literature that non-monotonic non-linear functions (e.g., quadratic) are more difficult to learn than monotonic ones (Brehmer, 1974; DeLosh, 1995; Lucas, Griffiths, Williams, & Kalish, 2015). Despite its mathematical complexity, the cubic-like function employed in my Complex condition was nevertheless monotonic; this may have made it easier to learn overall, thereby reducing the magnitude of the observed interaction. This possibility is considered at greater length in Chapter 6.

Finally, it may simply be the case that function learning in groups is fundamentally different from function learning in individuals, such that the two populations are differentially affected by a function's complexity. Although differences between these two scales no doubt exist, I aim to demonstrate over the course of this thesis that their similarities are more remarkable than their differences. Ultimately, both similarities and differences will be informative in understanding how groups of individuals are able to accomplish this learning at all.

This first experiment provided a proof of concept that groups of individuals can indeed learn both simple and complex functions. However, the results have little to say about the nature of this learning. Namely, a group's apparent improvement over training could result from either (1) each individual coming to memorize his or her responses to the four training targets, or (2) each individual gaining a sense of her broader role within the network (e.g., Theiner, Allen, & Goldstone, 2010). The distinction bears similarities to one in the comparative cognition and animal learning literatures, where it is regularly debated whether an apparently complex behaviour is performed by (1) an animal memorizing its responses to particular stimuli (i.e., a simple associative learning account), or (2) an animal gaining a broader understanding of the task, thereby evincing more complex cognition (e.g., Haselgrove, 2010; Le Pelley, 2012; see Shanks, 2010 for review).

Laws of parsimony suggest that more complex explanations should be considered only once simpler explanations have been exhausted (see Morgan, 1903). In studies with animals, simple associative learning explanations are discounted by evidence that the animal’s learned responding is not limited to the stimuli upon which it was trained, but that it can generalize the behaviour to a new set of stimuli. Likewise, evidence that groups of individuals can generalize what they have learned in a function learning task to new target values in a subsequent test phase would discount the simpler explanation that participants merely memorized their trained responses (e.g., DeLosh, McDaniel, & Busemeyer, 1997). In order to better understand the nature of the group learning processes, a second experiment tested whether groups can generalize the knowledge obtained in training to new targets in a subsequent test phase.

Chapter 3: Transfer of Learned Functions in Groups

In standard function learning tasks, training on the relationship between two variables, for example the dosage amount of a drug and the effect of the drug, is followed by a generalization or transfer test. In this example, a transfer test might ask participants to provide estimates of the dosage amount required to achieve some target values of drug effect. Importantly, some of the test values of the drug effect interpolate between those values that participants were trained on, and some of the test values extrapolate beyond those they were trained on. Research has shown that participants are capable of generalizing from both linear and non-linear functions at test, but that performance on extrapolation is consistently worse than performance on interpolation (e.g., DeLosh et al., 1997; Koh & Meyer, 1991; Wagenaar & Sagaria, 1975). Moreover, participants tend to underestimate the value of the Y variable at both the high and low ends of the X variable when extrapolating a linear function (Brown & Lacroix, 2017; DeLosh, 1995; DeLosh et al. 1997; Kwantes & Neal, 2006), but to overestimate the value of the Y variable at both ends of the X variable when extrapolating an exponential or quadratic function (DeLosh et al., 1997).

Experiment 2 sought not only to replicate the learning observed in Experiment 1, but also to examine how groups perform on an analogous generalization test following training. Importantly, the current experiment also increased the size of the group to test whether the learning observed in Experiment 1 would generalize to a larger group. If basic principles of function learning are indeed invariant to scale, then the absolute size of the group should not affect the group-level behaviour (c.f. Roberts & Goldstone, 2011).

In order to better understand how group-level learning emerges from individual-level responding, the current experiment also introduced personality questionnaires to participants' post-experimental tasks. Personality questionnaires are an effective and easily-implemented means of capturing individual differences between participants. In my group function learning experiments, personality variables may capture some of the variability in individual participants' response behaviours and reported strategies. Likewise, the personality profiles of all individuals within a given group may capture variability in groups' success at the task. Although I describe the collection of personality data in the current chapter and subsequently in Chapter 4 (Experiment 3), analyses of these data are reserved for Chapter 5.

Experiment 2

In this experiment, participants were trained and tested in groups of 16 individuals. Training proceeded in much the same way as in Experiment 1, but with new target values (4, 7, 10, and 13) that better captured the range accorded by the larger group size. In a Simple condition, the target values given to participants were the true values they were aiming for. In a Complex condition, the true target values were different from (i.e., a cubic-like transformation of) the target values given to participants; although participants were led to believe the untransformed values were their goals, the feedback provided compared their responses to the transformed values. As in Experiment 1, learning in this Experiment was demonstrated by a minimization of mean squared error over the course of training.

Following three blocks of training conducted in the same manner as in Experiment 1, participants underwent a transfer test, in which they had one attempt at each possible target value (i.e., 1 through 15 inclusive) without any feedback. In addition to the two experimental conditions, a Control condition completed the test phase without any prior training. If groups generalized their learning to the test phase, error on untrained test values should be (1) lower than error achieved on the first block of training in the same condition and (2) lower in the experimental conditions than in the Control condition. Error in the Control condition was assessed against both the untransformed and transformed target values, in order to provide fair comparisons against the Simple (Control-S) and Complex (Control-C) condition, respectively.

Method

Participants

Five hundred and twenty-eight introductory psychology students participated in this experiment in exchange for partial course credit. Participants were run in groups of 16, and each group was randomly assigned to one of three conditions: a Simple condition that was trained and tested on untransformed test targets; a Complex condition that was trained and tested on cubic-like transformations of the target values; and a Control condition that received no training but underwent testing on the same (untransformed) values as the experimental conditions. Ten groups (i.e., a total of 160 participants) were run in each the Simple and Control conditions. Thirteen groups (i.e., a total of 208 participants) were run in the Complex condition. Three groups in the Complex Condition were excluded due to sabotage by students, leaving a total of ten groups per condition. For purposes of data analysis, one group of 16 participants will be referred to as one subject.

Apparatus

Data were collected in a manner identical to that described in Experiment 1.

Procedure

Pre-training & training. For all three conditions, pre-training proceeded in a manner identical to that described in Experiment 1. The Control condition received no training phase whatsoever. In the Simple and Complex conditions, training proceeded identically to the procedure outlined in Experiment 1, except that the target values were altered to accommodate a larger group size. In both conditions, groups received three blocks of training on a set of four target numbers, always in the same increasing order: 4, 7, 10, and 13. Directional (i.e., over or under) and numerical (i.e., the distance of error) feedback was once again provided after each attempt at a target. Critically, and unbeknownst to participants, the group-level responses in the Complex condition were compared to transformations of the given targets. Target values in that condition were transformed on a cubic function and rounded to whole numbers, so that the feedback compared responses to 6, 7, 7, and 9. In the Simple condition, the group-level responses were compared to the untransformed targets given (e.g., 4, 7, 10, and 13). Once again, training on a given target number was repeated until it was achieved on two consecutive trials, or until the group reached a maximum of 19 attempts.

Detecting Sabotage. Sabotage, whether intentional or incidental, was detected in the same manner as described in Experiment 1. In the present experiment, three groups were excluded as a result of this detection. Description and analysis of these sabotaged groups is deferred until Chapter 5.

Test. All three conditions received an identical test phase, in which they were asked for a single attempt at each of the possible target values between one and fifteen, inclusive. Unlike in the training phase, feedback was not provided following an attempt; the obtained number of respondents was simply recorded before the next target was given. Target order was randomized for each group. In the Simple condition, accuracy was measured as the squared error between the obtained number of respondents and the target value. In the Complex condition, accuracy was measured as the squared error between the obtained number of respondents and a cubic-like transformation of the target value, as given in Table 3.1. In the Control condition, accuracy was again measured by squared error, as a function of both the untransformed (Control-S) and transformed (Control-C) target values.

Table 3.1. *Target Values in the Test Phase of Experiment 2, as a Function of Condition.*

Condition	Test Target Value														
Simple	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Complex	1	3	5	6	7	8	8	8	8	8	9	10	11	12	14

Notes. Order of test targets was randomized for each group. Groups in the Complex condition were asked to match the Simple (untransformed) target values but their accuracy was measured against the Complex (transformed) target values.

Additional tasks. As in Experiment 1, a brief semantic distractor activity was conducted between blocks of training. These proceeded exactly as described in Experiment 1. Following completion of the test phase, participants were asked to write down free-formed thoughts about the strategies they used to make their decisions throughout the experiment, as in Experiment 1.

Finally, participants in the Control condition subsequently completed two short personality questionnaires[5]. The Big Five Inventory (BFI) is a brief personality inventory measuring the "big five" traits: openness, conscientiousness, extroversion, agreeableness,

---

[5] These questionnaires were restricted to the Control condition in this experiment due to time constraints.

and neuroticism (John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008). This freely available inventory has participants rate 44 statements about themselves on a scale from 1 (disagree strongly) to 5 (agree strongly). Responses are averaged within the five personality dimensions, resulting in a score out of 5 for each dimension. The Ten-Item Personality Inventory (TIPI) is a very brief measure of the "big five" traits, consisting of only ten items (two items per dimension; Gosling, Rentfrow, & Swann, 2003). Participants rated these statements about themselves on a scale from 1 (disagree strongly) to 7 (agree strongly). I included this shorter inventory largely to test whether it could be reliably used as a quick stand-in for the longer BFI scale. However, within-subject correlations between BFI and TIPI scores varied widely, ranging between $r = -.925$ to $r = .987$, with a standard deviation ($SD = 0.51$) larger than the mean correlation ($M = 0.23$). As such, I opted to discard the TIPI measures in favour of the more robust BFI measures (for a comparison of test-retest reliabilities, external correlates to more comprehensive personality assessments, and self-observer convergence, see Gosling et al., 2003). The results and detailed analysis of all individual-level data, including self-reported strategies and personality measures, can be found in Chapter 5.

Results

Training Phase

The results of the training phase of Experiment 2 are shown in Figure 3.1. Of primary interest is whether the results of Experiment 1 were replicated with a larger group size; that is, whether accuracy improved over the course of training for both experimental conditions. As in Experiment 1, I will first look at the squared error averaged over all trials within a block. A 2 (Condition: Simple vs. Complex) x 3 (Block) mixed measures ANOVA found a main effect of block, $F(2, 36) = 8.43$, $p < .001$, but no main effect of condition, $F(1, 18) = 0.55$, $p = .46$, indicating that learning over blocks was similar across conditions. Polynomial a priori contrasts on the main effect of block confirmed that mean squared error decreased linearly over blocks of training, $t(36) = 4.11$, $p < .001$. The corresponding quadratic contrast was not reliable, $t(36) = 0.05$, $p = .96$. Once again, the block by condition interaction could not support the visual analysis of an initial deficit for learning complex functions that decreased with additional training blocks, $F(2, 36) = 0.71$, $p = .50$. As in Experiment 1, learning of both simple and complex functions was evident, but anticipated differences in the shape of that learning did not bear out statistically.

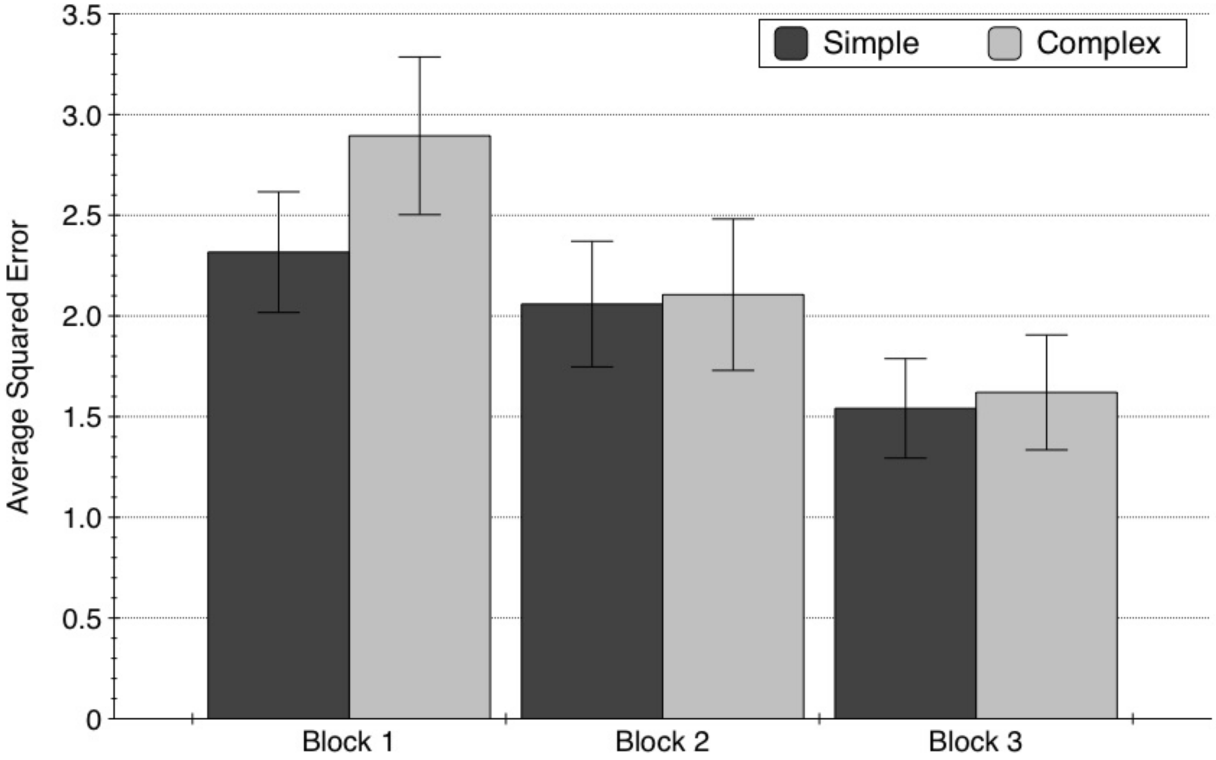

Figure 3.1. Squared error averaged over all attempts in a block in the training phase of Experiment 2, as a function of block and condition. Error bars represent SEMs.

As in Experiment 1, I could have alternately examined performance as a function of the number of trials taken to match targets, or as a function of the squared error achieved on the first attempt at a target. Once again, however, analyses of these alternative measures provided conclusions identical to the primary measure, and are thus not reported (see Appendix A for a full report of these analyses).

Test Phase

To examine how training on a subset of target values influenced group performance on subsequent untrained test target values, analyses could be conducted in two ways: (1) within groups, by comparing performance in the first block of training, prior to any learning, to performance at test; and (2) between groups, by comparing test performance in groups that underwent training (Simple and Complex conditions) relative to those that did not (Control Condition). Finally, an examination of performance across the range of test targets in all conditions allowed a better understanding of how group behaviour compares to that of individuals in standard function learning experiments. Below I delineate each of these approaches in turn.

Within-group analyses. Although mean squared error averaged across all attempts in a block was the most sensible measure for training performance, it does not allow for clear comparisons to test performance, where only one attempt was permitted per target. Instead, Figure 3.2 shows the squared error for first attempts on targets in the three training blocks as well as the test phase. As shown, error on the first attempts during training mirrors the pattern seen when measuring averaged squared error (compare Figure 3.1). A 2 (Condition) x 4 (Time: block 1, block 2, block 3, test[6]) mixed measures ANOVA

---

[6] This analysis only considers the old target values at test.

found a main effect of Block, $F(3, 54) = 4.37$, $p = .01$, but neither an effect of Condition, $F(1, 18) = 2.56$, $p = .13$, nor a two-way interaction, $F(3, 54) = 1.13$, $p = .35$. This pattern of results once again indicates that learning over blocks proceeded similarly in both conditions. Notably, results using the error on first attempts as the measure of interest are consistent with those reported above that used mean squared error over all attempts as the measure of interest.

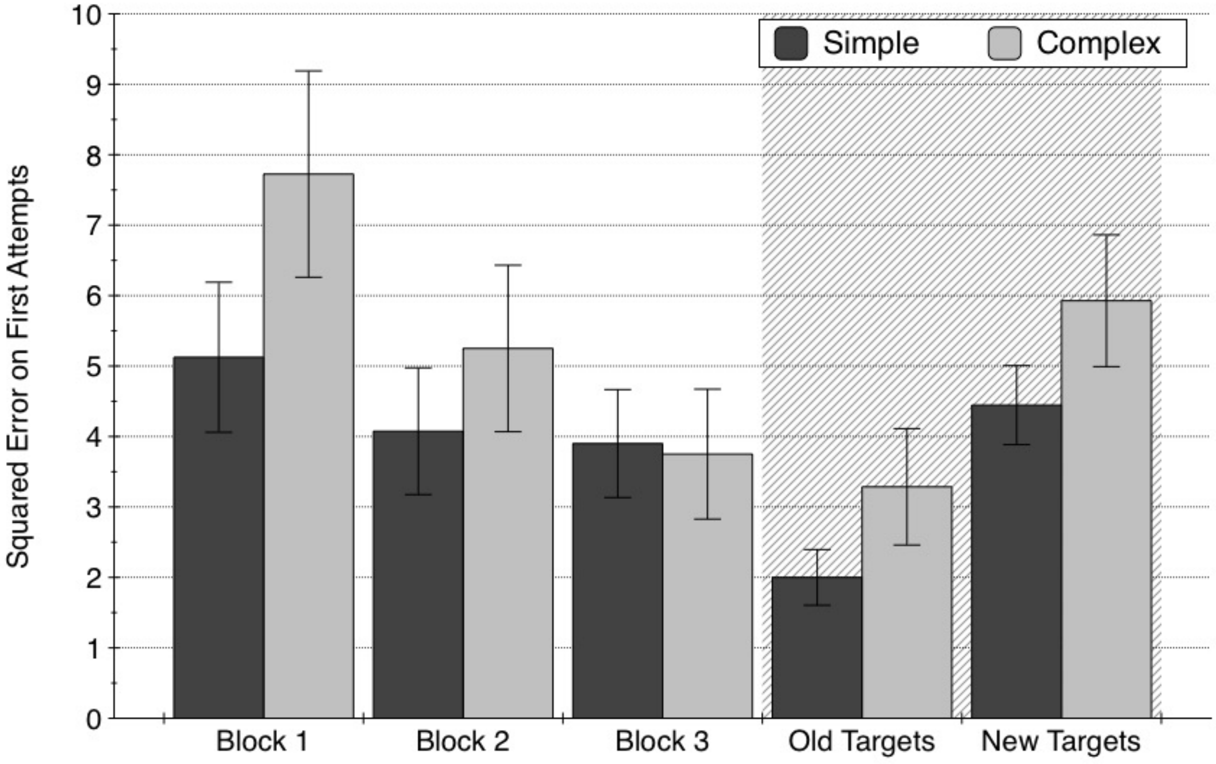

Figure 3.2. Squared error averaged over all first attempts in a block in both the training phase (white background) and the test phase (hashed background) of Experiment 2, as a function of condition. Error bars represent SEMs.

In each experimental condition, performance on block 1 was compared with test performance for both previously trained (old) targets and untrained (new) targets at test.

In the Simple condition, performance on old test targets was better than performance during the first block of training, *t*(9) = 4.27, *p* = .002, and indeed better than during the third block of training, *t*(9) = 3.27, *p* = .01. In the Complex condition, performance on old test targets was better than performance during the first block of training, *t*(9) = 3.27, *p* = .01, though no better than during the third block of training, *t*(9) = 1.77, *p* = .11. Whereas test performance on old items can be seen as simply an extension of training, generalization of learning would be evinced by better performance on new test targets relative to old targets on the first block of training. Although the differences did not meet statistical significance, they were nevertheless in the correct direction. In the Simple condition, error was greater for old targets in block 1 (*M* = 5.12, *SD* = 2.21) than for new targets at test (*M* = 4.45, *SD* = 1.73), *t*(9) = 0.93, *p* = .38. Likewise in the Complex condition, error was greater for old targets in block 1 (*M* = 7.72, *SD* = 3.35) than for new targets at test (*M* = 6.06, *SD* = 1.85), *t*(9) = 1.27, *p* = .23. It is likely that larger sample sizes would have resolved these differences statistically, a point to which I will return in the discussion.

Accuracy for new test items was further broken down to compare performance on test targets that interpolated between targets studied in training against performance on test targets that extrapolated beyond the range of those studied in training. Research on function learning in individuals consistently shows that participants are better at interpolating than extrapolating, so I expected this result to replicate in group function learning (e.g., DeLosh et al., 1997). At this fine grain of analysis, however, the power to detect effects was severely limited. In the Complex condition, groups were numerically better at interpolated targets (*M* = 4.55, *SD* = 2.49) than extrapolated targets (*M* = 7.86, *SD* = 5.54), but this difference failed to meet statistical reliability, *t*(9) = 1.40, *p* = .20. In the

Simple condition, performance on interpolated and extrapolated targets was nearly identical, $t(9) = 0.34$, $p = .74$.

Between-group analyses. As described in the Methods, accuracy in the Control condition could be computed relative to the untransformed (Control-S) and transformed (Control-C) target values, resulting in a fair empirical comparison of generalization against the Simple and Complex conditions respectively. The results of the following analyses are shown in Figure 3.3. A 4-way ANOVA on average test accuracy revealed that performance did differ over conditions, $F(3, 36) = 16.05$, $p < .001$. A priori contrasts found no difference in overall test accuracy between the Control-S ($M = 14.50$, $SD = 5.45$) and Control-C ($M = 11.71$, $SD = 5.64$) cases, $t(36) = 0.55$, $p = .59$. By contrast, performance was better in the Simple condition ($M = 3.79$, $SD = 1.50$) than in the corresponding Control-S condition, $t(36) = 4.51$, $p < .001$, and better in the Complex condition ($M = 5.19$, $SD = 1.39$) than in the corresponding Control-C condition, $t(36) = 3.25$, $p = .002$. These differences show a benefit of training in the experimental conditions.

It should be noted that the two Control “conditions” were not in fact separate conditions in the experimental design, but instead two different transformations of the same data from one Control condition. Although statistical tests comparing the two cases should thus be interpreted with caution, the comparison is not central to any argument in the thesis. The null difference between the Control-S and Control-C conditions reported above serves as evidence that any differences between the Simple and Complex conditions are not a mere byproduct of how mean squared error was calculated (i.e., against a linear and cubic-like function, respectively).

Additional t-tests were conducted to compare experimental and control conditions on old test items, which experimental conditions would have studied in the training phase, and new test items, with which neither the experimental nor control conditions would have had prior experience. Performance on old test items was better in the Simple condition than in the Control-S condition, *t*(9.80) = 5.04, *p* < .001[7], and better in the Complex condition than in the Control-C condition, *t*(18) = 2.38, *p* = .03.

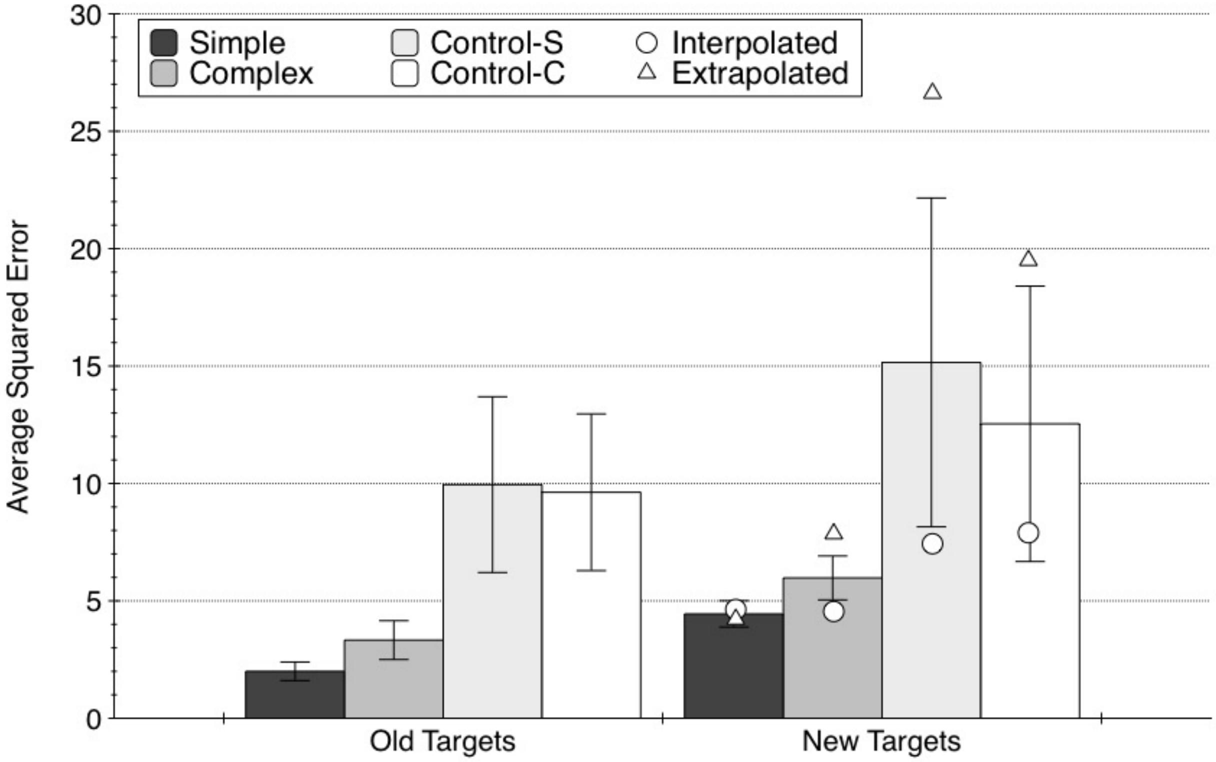

Figure 3.3. Squared error in the test phase of Experiment 2, as a function of target type (Old vs. New). New target values are further separated to indicate performance on both interpolated targets (open circles) and extrapolated targets (open triangles). Error bars represent SEMs.

---

[7] In cases where the assumption of equal variances was violated in independent samples t-tests, as determined by a significant (*p* < .05) Levene's test, the Welch's correction for degrees of freedom was applied. Integer df's indicate the assumption was not violated and no correction was applied, whereas non-integer df's indicate a corrected test.

However, t-tests comparing performance on old test items in the Control condition to performance on those same items in the first block of training in the experimental conditions revealed better performance in the Simple condition than the Control-S condition, $t(12.53) = 2.849$, $p = .01$, suggesting that differences between those two conditions may have in part been present prior to any training; a similar comparison found no difference between the Complex and Control-C condition, $t(18) = 0.65$, $p = .52$. Some caution is thus recommended in the interpretation of differences between the Simple and Control-S conditions.

Performance on new test items was better in the Simple condition than in the Control-S condition, $t(10.20) = 5.37$, $p < .001$, and better in the Complex condition than in the Control-C condition, $t(10.98) = 3.52$, $p = .005$. These results confirm that the experimental conditions benefited from the training phase, relative to a Control condition that received no such training.

As above, performance on new test targets was further broken down to examine differences between the experimental and control conditions on both interpolated and extrapolated test targets. Interpolation performance in the Simple condition ($M = 4.63$, $SD = 3.31$) was numerically better than in the Control-S condition ($M = 7.43$, $SD = 5.21$), and numerically better in the Complex condition ($M = 4.55$, $SD = 2.49$) than in the Control-C condition ($M = 7.90$, $SD = 5.08$), but neither difference was reliable, $t(18) = 1.43$, $p = .17$ and $t(18) = 1.87$, $p = .08$ respectively. By contrast, extrapolation performance was reliably better in the Simple condition than in the Control-S condition, $t(9.32) = 7.40$, $p < .001$, and better in the Complex condition than the Control-C condition, $t(18) = 3.52$, $p = .002$. These results show that performance on both extrapolated test targets and interpolated test targets was improved as a result of training, relative to a control condition that did not undergo training, although only the benefit for extrapolation could be confirmed statistically.

Performance as a function of test target. Figure 3.4 plots the mean value obtained by groups as a function of the target value requested at test, for the Simple condition and Complex condition, top and bottom panel respectively. Mean values obtained by the Control condition are replicated in the two graphs.

A visual analysis of this figure shows that deviations from the trained function are on average larger for Control than experimental conditions. Indeed, the achieved values in the Simple condition follow close to the linear function depicted by the dashed line in the top panel. When compared instead against the cubic-like function, average squared error at test was higher (M = 8.19, SD = 2.37) than when compared against the linear function (M = 3.97, SD = 1.50). Similarly, the achieved values in the Complex condition follow close to the cubic-like function depicted by the dashed line in the bottom panel. When compared instead against the linear function, average squared error at test was higher (M = 7.15, SD = 1.82) than when compared against the cubic-like function (M = 5.19, SD = 1.39). This indicates that the training phase tuned participants' responding to the particular function they studied, and that this tuning generalized to new target values at test.

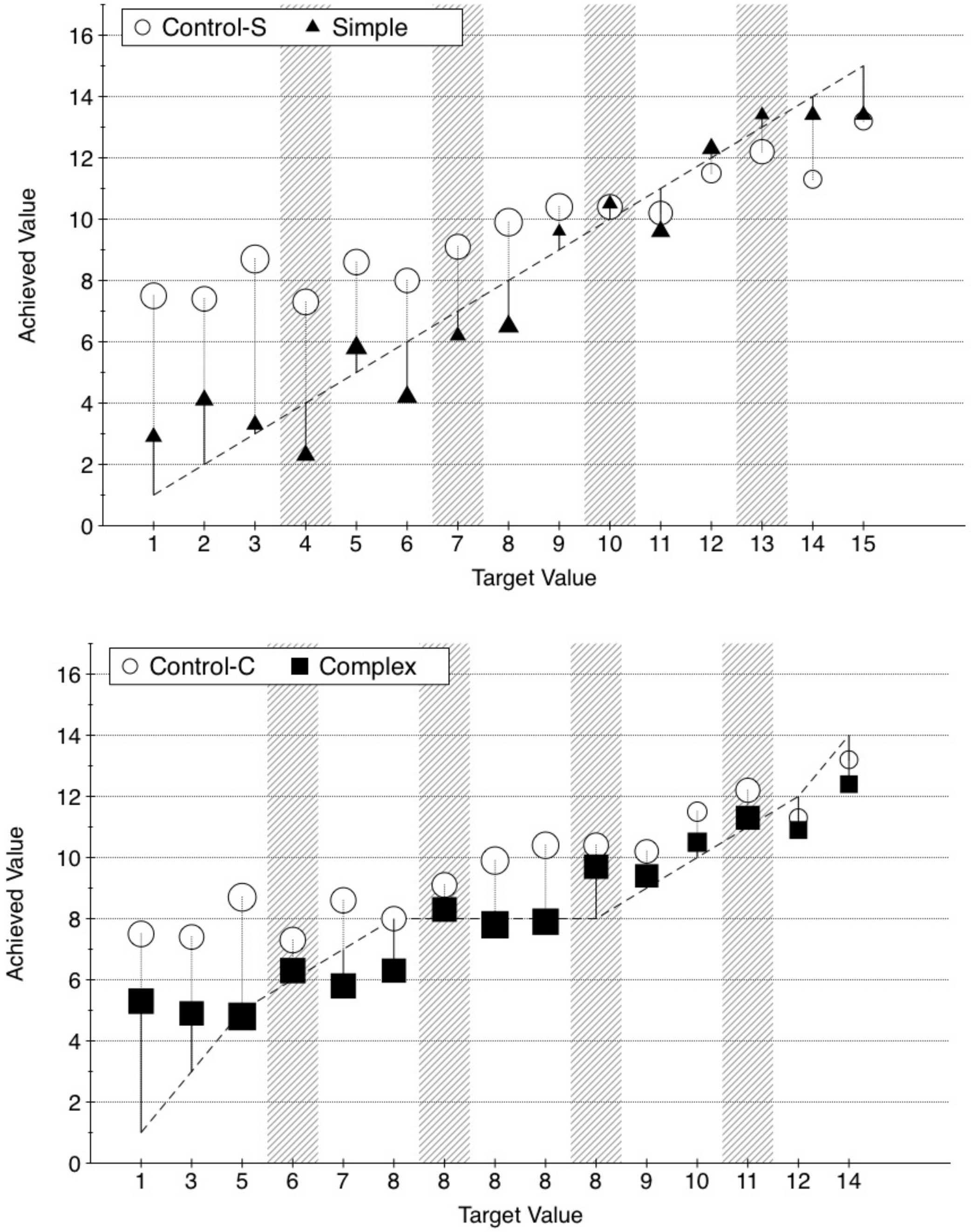

Figure 3.4. Average values achieved in the test phase of Experiment 2, as a function of target value and condition. In the top panel, the Simple (filled triangles) and Control-S (open circles) conditions are compared against the untransformed target values (dashed line). In the bottom panel, the Complex (filled squares) and Control-C (open circles) conditions are

compared against the cubic-transformed target values (dashed line) in Panel B. In both Panels, vertical lines illustrate the deviation of the experimental (continuous lines) and control (dotted lines) conditions from the trained function. The hashed backgrounds indicate old (trained) test targets. All markers are scaled according to their corresponding SEMs.

The Control condition also followed more closely to the cubic-like function (M = 11.71, SD = 5.64) depicted in the bottom panel than the linear function (M = 14.50, SD = 5.45) depicted in the top panel. This finding suggests that the natural bias of groups in a function learning task is toward a cubic-like, rather than a linear, pattern of responding. The possibility is at odds with the fact that linear functions are typically easier to learn than non-linear functions (e.g., McDaniel & Busemeyer, 2005), although I was unable to replicate this standard result with statistical reliability in both Experiments 1 and 2. Nevertheless, I return in the discussion section to the possibility of an inherent bias known in probability estimation and function learning.

The graphs in Figure 3.4 also show that the benefit of training in the experimental conditions, relative to the Control condition, seems to decrease as the test target value increases. That is, the Control condition shows particularly large deviations from the assigned functions in the lower range of target values and for extrapolated target values. To quantify the observation, performance was grouped into low-range extrapolated items (untransformed targets 1, 2, and 3), low-range interpolated items (untransformed targets 5, 6, and 8), high-range interpolated items (untransformed targets 9, 11, and 12) and high-range extrapolated items (untransformed targets 14 and 15).

Table 3.2. *Mean Squared Error During the Test Phase of Experiment 2, Partitioned by Target Range and Type.*

	Low Range		High Range	
	Interpolated	Extrapolated	Interpolated	Extrapolated
Simple	6.23 (4.98)	4.90 (2.37)	3.03 (2.50)	3.20 (1.65)
Control-S	10.83 (8.41)	40.07 (16.33)	4.03 (4.25)	6.45 (3.68)
Complex	5.57 (5.26)	10.60 (8.57)	3.53 (3.05)	3.75 (3.16)
Control-C	6.70 (4.13)	30.53 (13.82)	6.23 (6.06)	1.75 (1.30)

A 2 (Range: low vs. high) x 2 (Target type: interpolated vs. extrapolated) x 2 (Condition: Experimental vs. Control) mixed measures ANOVA confirmed this pattern of results for both the Simple (vs. Control-S) and the Complex (vs. Control-C) cases. A three-way interaction in the Simple, $F(1, 18) = 29.74$, $p < .001$, and Complex, $F(1, 18) = 12.47$, $p = .002$, cases revealed that the majority of the distinction between Experimental and Control conditions was driven by performance on low-range extrapolated target values (see the second column of Table 3.2). At the high-range (see the third and fourth columns of Table 3.2) and for interpolated target values (see the first and third columns of Table 3.2), the differences were comparatively small. This result suggests that one of the major benefits of training in the experimental conditions was improved accuracy for low target values. Relative to the Control condition, the Simple and Complex conditions showed strong generalization of learning when they were tested on small targets, particularly those that extrapolated beyond the range of targets they studied in the training phase.

Discussion

The results of this experiment replicate and expand upon the results of Experiment 1. Although the group size was increased from 10 to 16 participants, groups were once again able to learn both simple linear and complex cubic-like functions, as evidenced by decreasing error over the course of training. The data add to that point by showing a further decrease in error for previously trained target values during a transfer test. As in Experiment 1, anticipated differences in the shape of learning simple (linear) and complex (non-linear) functions were visually apparent in the data but were not supported by statistical analysis. The results of the current experiment also lend support to the hypothesis that groups can generalize their training to untrained target values in a subsequent test phase. Whereas performance on new test values was numerically lower than performance on block 1 of training in both conditions, these findings fell short of statistical reliability. However, comparisons against a Control condition, which did not receive any training, bolster my confidence in the result: performance on new test values was better following training than following no training, suggesting that some generalization did occur. Finally, a closer examination revealed that the effects of training were of most benefit to those target values at the low range, particularly those that extrapolated beyond trained values.

As in Experiment 1, the results shed light on some differences between individual and group function learners. Whereas non-linear functions are generally more difficult for individuals to learn than linear functions (Deane et al., 1982; DeLosh et al., 1997; McDaniel & Busemeyer, 2005), I once again failed to replicate this result in group function learning. As in Experiment 1, group means were consistent with the result but statistical analysis could not support the difference. Once again, the small sample sizes (n = 10 per condition) may have limited the power to detect a real interaction between condition and time. However, repeated failures to find the interaction may be indicative of an alternative explanation. As discussed briefly following Experiment 1, it may be the case that the cubic-like function learned by my Complex condition, due to its monotonicity, was simply not difficult enough to draw out the differences I was hoping to detect. Lucas et al. (2015), for instance, found that non-monotonic non-linear functions (e.g., quadratic) are more difficult to learn than monotonic ones. A closer examination of test performance provides some evidence for this hypothesis.

A typical finding in the individual function learning literature is that participants underestimate the value of the Y variable at both the high and low ends of the X variable when extrapolating a linear function (e.g., Brown & Lacroix, 2017; DeLosh, 1995; DeLosh et al., 1997; Kwantes & Neal, 2006). By contrast, when extrapolating an exponential or quadratic function, participants tend to overestimate the value of the Y variable at both the high and low ends of the X variable (DeLosh et al., 1997). To the best of my knowledge, no examples of cubic function learning in individuals are available for direct comparison. In the current experiment, groups in both the Simple (linear) and Complex (cubic-like) condition showed a tendency to overestimate when extrapolating at the low end and to underestimate when extrapolating at the high end. This pattern of over- and underestimation is unusual, relative to the consistent findings in the individual function learning literature.

Of interest is the fact that the Control condition showed an exaggerated version of this same pattern, suggesting that there is a natural bias for groups to overestimate at the low end and underestimate at the high end; although training in the group function learning task serves to mitigate this bias, it does not ultimately eliminate it (but see Kalish, Griffiths, & Lewandowsky, 2007, for evidence of a different bias in individual function learning).

Precedence for this pattern of results may be found outside the function learning literature. In the field of decision-making, for instance, it has been shown that low probability events are disproportionately weighted in decisions, whereas high-probability events are underweighted (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). When judging physical magnitudes, people overestimate relatively small magnitudes and underestimate relatively large ones (e.g., Poulton, 1982). The bias also appears in frequency judgements, the classical example of which shows that people overestimate low-frequency causes of death, such as tornadoes, and underestimate high-frequency ones, such as cancer (Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978). In sample-size estimation, people overestimate the number of items in small samples and underestimate the number of items in large samples, a phenomenon known as the conservatism bias (e.g., Kaufman, Lord, Reese, & Volkmann, 1949; MacGregor, Lichtenstein, & Slovic, 1988).

Perhaps of most relevance to the bias in groups learning is the classic finding of diffusion of responsibility, whereby people are less likely to act in an urgent event when they believe others will, and more likely to act when they believe others won't (Latané & Darley, 1968). In the context of the present experiment, individual participants may have been more likely to respond to low-range extrapolation values under the assumption that others in the group would not respond, and less likely to respond to high-range extrapolation values under the assumption that others in the group would. Self-report data described in Chapter 5 lend support to this possibility.

An alternate explanation is that the overestimation at the low end and underestimation at the high end are the result of chance. Specifically, if every individual in the group of sixteen were responding (or not) on the basis of chance alone, we would expect a long-term average of eight participants to respond for any test item. The result would be regression to the mean, with overestimation at the low and underestimation at the high end of the target range. However, if chance alone were driving behaviour we would expect performance to be symmetrical, with equivalent error at the low and high ends of the target range and optimal performance for the target value of eight. In all conditions, however, error was not symmetrical around the mean of eight but was instead greater at the lower end of the target range. Even in the Control condition, which did not receive any training and thus might be expected to more closely approximate chance performance, this pattern held. In the Simple condition, mean error was lowest for the test targets nine (*M* = 1.2, *SD* = 2.78) and thirteen (*M* = 1.2, *SD* = 1.55). In the Complex condition, mean error was lowest for the test target thirteen (*M* = 0.9, *SD* = 1.97). In the Control condition, mean error was lowest for the target value of twelve (*M* = 1.9, *SD* = 1.45) when compared to the Simple function and fourteen (*M* = 1.7, *SD* = 1.64) when compared to the Complex function. Although chance may be playing some role in the pattern of test results, it is clearly not capable of fully explaining groups' behaviour.

At least one example within the function learning literature does mirror this pattern of bias in extrapolation, but it comes from a modeling exercise rather than from experimental data. Recall that individuals tend to underestimate when extrapolating from linear functions and to overestimate when extrapolating from quadratic or exponential functions (Brown & Lacroix, 2017; DeLosh et al., 1997; Kwantes & Neal, 2006). In an effort to understand this unexpected pattern, DeLosh et al. compared predictions from four different computational accounts against participant data: the polynomial hypothesis-testing model (Brehmer, 1974), the log-polynomial adaptive-regression model (Koh &

Meyer, 1991), the associative-learning model (ALM; Knapp & Anderson, 1984; Kruschke, 1992), and their own extrapolation-association model (EXAM). I briefly explain each model in turn, with a focus on the ALM.

The polynomial model assumes that participants learn a single regression rule that maps X values to Y values, and subsequently apply that rule in extrapolation. The log-polynomial model instead assumes that participants learn and apply the log transformed relationship between X and Y values, which Koh and Meyer (1991) argue more accurately reflects the non-linear relationship between the physical stimulus and participants' subjective images of the stimulus. Neither of these models does a very good job of capturing extrapolation performance in participants. For linear functions, the log-polynomial model underestimates at the low range but overestimates at the high range, whereas the polynomial model predicts no extrapolation at all. Neither would therefore serve as good models for group extrapolation performance.

In contrast to these strictly mathematical models, the ALM and EXAM models bear similarities to standard associative learning models (e.g., Rescorla & Wagner, 1972) and thus to artificial neural networks (e.g., Rosenblatt, 1958, 1962). In the ALM (Knapp & Anderson, 1984; Kruschke, 1992), participants learn a mapping between stimuli and responses by associating input nodes to output nodes, with one input node corresponding to each possible stimulus magnitude and one output node corresponding to each response magnitude. When a particular stimulus is presented it activates each input node according to its similarity to that stimulus, with a Gaussian function that provides a generalization gradient around activated nodes. Activation is passed from input to output nodes as a function of the strength of their associations and the response is selected as a function of the weighted average of activation over all output nodes. During training, the association strengths between input and output nodes are updated according to a standard delta learning rule.

The EXAM model operates in much the same way as the ALM, except that it uses a more sophisticated matching mechanism to generate responses. The result is that EXAM maps very well onto participant extrapolation performance, underestimating from linear functions and overestimating from exponential and quadratic functions. By contrast, ALM maps very well onto group extrapolation performance: for both linear and exponential functions, ALM overpredicts at the low range and underpredicts at the high range. Whereas DeLosh et al. (1997) took these predictions as a failure of the ALM model, their match to group function learning performance may provide insight into how groups learn in this task. That is, in order to understand the ability of groups to learn simple and complex functions, as demonstrated in Experiments 1 and 2, it may be useful to think of these groups as artificial neural networks.

Chapter 4: Function Learning Groups as Networks

As illustrated by the patterns of extrapolation performance described in the previous chapter, groups of individuals tested on a function learning task behave similarly to predictions of an associative-learning model described by DeLosh, Busemeyer, and McDaneil (1997). Insofar as principles of that model, such as the delta learning rule, can be re-expressed as principles of artificial neural networks (ANN's), the relationship between groups of individuals in a function learning task and groups of nodes in a network ought to be explored.

In both arrangements, individual units activate (or not) in response to input; the summed input is compared against a target; and each unit readjusts to feedback in order to minimize error between the previous group-level outcome and the target. To the extent that ANNs are representative of biological networks, it could likewise be argued that groups in my function learning task share some correspondence to biological neural networks. Of course, the correspondence is necessarily superficial: neither artificial networks nor groups of humans communicate via synapses, ion channels, or neurotransmitter release. Moreover, groups of humans in the function learning task cannot communicate directly with one another, in contrast to the lateral connections known to exist in neural networks and many ANNs. Bearing in mind that caveat, it is nevertheless useful within the current theoretical framework to consider the similarities between networks and groups of individuals that successfully learn and generalize. Below, I briefly outline the history of research on artificial neural networks and consider an important property of modern ANNs that makes contact with the notion of emergent cognition in groups. I then revisit self-similarity as the guiding force behind an experimental analysis of network-like behaviour in the group function learning task.

Artificial Neural Networks

The first formal model of a neural network associated simple binary units by means of propositional logic (McCulloch & Pitts, 1943), effectively applying mathematical descriptions to the conceptual diagrams of connected nodes proposed by Spencer (1855) and James (1890). With the perceptron, Rosenblatt (1958, 1962) formalized the processes by which those associations were formed and the means by which they changed in response to their environments (i.e., learning). Originally designed as a tool for visual recognition, the perceptron posits weights between input and output units that can be altered by experience. Specifically, a network aims to reduce the error between its predicted outcome (i.e., the weighted sum of input activations, transformed by an output function) and the target outcome supplied by the modeler (i.e., supervised learning) for a given input set. The perceptron is able to converge on suitable weight values in order to recognize a number of patterns. Once it was shown that there were several classes of patterns that the perceptron was incapable of learning, however, ANNs fell out of favour (Minsky & Papert, 1969).

The focus of cognitive psychology then shifted toward (classical) computationalism, in which the brain was conceptualized as information processor that operates on symbols, not unlike a computer (e.g., Newell & Simon, 1972, 1976). In this view, meaningful symbolic representations are the base unit of information in the brain, which uses rules to manipulate linguistic or sentence-like symbols in a serial manner. Thus, the classicist view postulated a “language of thought” with syntactic and semantic structure, which need not necessarily make any direct contact with the neurological level of analysis (Fodor & Pylyshyn, 1988). An important implication of classicism is that information is localized, in the sense that the base units of thought are fully realizable as meaningful representations, even when considered outside of the context of the network.

McClelland and Rumelhart's seminal two-volume introduction of parallel distributed processing networks (PDP; McClelland, Rumelhart, & The PDP Research Group, 1986) marked the birth of connectionism as an alternative computationalist approach, and a return to the belief that theories of mind ought to make contact with the brain. Instead of symbolic representations, connectionism took neurons to be the base unit of thought and reinstated the practice of modeling neurons as nodes within an interconnected network. Like Rosenblatt's (1958) perceptrons, PDP and other connectionist networks are characterized by input nodes, output nodes, weighted connections between the two, and a learning rule by which experience can modify patterns of connectivity. Unlike perceptrons, however, the new ANNs included at least one additional (hidden) layer of nodes between the input and output layers and a rule by which patterns of activity can be propagated both forward (i.e., through hidden layers, which transform the activation in some way) and backward (i.e., backpropagation of error; Rumelhart, Hinton, & Williams, 1986) through the multiple layers to accommodate learning.

Since McClelland and Rumelhart's (1986) PDP volumes, interest in connectionist networks has proliferated such that their complexities and their consequent capabilities have advanced at a dramatic pace, and in several different directions. ANNs of various kinds have been used to model phenomena as diverse as language development (Elman, 1993; Howell & Becker, 2001), serial memory (Botvinick & Plaut, 2006), consciousness (Dehaene

& Naccache, 2001), implicit learning (Becker, 1999; Kinder, 2000; Kinder & Lotz, 2009), mathematical cognition (Verguts & Fias, 2005), associative learning (Delamater, 2012; Gluck & Bower, 1988; Turnock & Becker, 2008), and several forms of brain damage (Hinton, Plaut, & Shallice, 1993; Montague, Dolan, Friston, & Dayan, 2012; Turnock & Becker, 2008). Recently, much media attention and government funding has been directed toward collaborative projects to simulate entire brains with advanced ANNs bearing a brain-like number of nodes and connections combined with more biologically plausible mechanisms of activation (e.g., the Blue Brain Project, Markram, 2006; the BRAIN initiative, Jorgenson et al., 2015; NENGO, Eliasmith et al., 2012).

Distributed Representation in Networks

Of most relevance to this thesis is the fact that, unlike the serial processing and localized representations espoused by classicism, connectionist networks are inherently parallel in their operations and distributed in their representations. Input activates all input nodes in parallel, and activation is propagated to or from all nodes in a given layer in parallel. The representations are distributed and subsymbolic, in that the state of any given node is not symbolically meaningful; meaning emerges only once the full pattern of activity across the network is considered.

Research conducted by Karl Lashley in the early 20th century first established the notion of distributed representation as a viable possibility for biological networks. Lashley was interested in localizing the means by which memories are stored and represented in the brain, a theoretical construct that he called the *engram*. Decades of research examining the effect of lesions on previously learned behaviours, however, led him to conclude that the engram could not be attributed to a given brain area; rather, he concluded that representations must be distributed throughout the cortex (see Lashley, 1950 for review). For example, rats were trained on a maze-learning task prior to having varying proportions of their bilateral cortices surgically removed. Upon reintroducing the rats to the maze task, he found that the number of errors produced, and the number of trials taken to relearn the task, were an exponential function of the percentage of cortex excised (Lashley & Wiley, 1933). That is, there was a *graceful degradation* of learned behaviour as the amount of damage increased. Ultimately, Lashley concluded that the resistance of the learned behaviours in the neural network to lesions could be attributed to a distributed representation scheme: insofar as representations for the learned behaviours were not localized to a given area of cortex, excising even a sizeable chunk of cortex could only hinder or slow performance, not destroy it altogether.

Although there remains some debate to this day as to the biological plausibility of fully distributed representations (see e.g., Bowers, 2009, 2010; Page, 2000, for arguments in favour of localist representations; see e.g., Farah, 1994; Plaut & McClelland, 2010, for arguments in favour of distributed representation), the distributed assumption of modern connectionism makes useful contact with the theories of emergence that are central to my examination of function learning in groups.

As described in the introduction to this thesis, when a unique group-level pattern of behaviour emerges from the local interaction of individuals, the group is taken to be an example of a distributed system of knowledge (e.g., Goldstone & Ashpole, 2004; Goldstone, Ashpole, & Roberts, 2005; Goldstone & Gureckis, 2009; Goldstone, Jones, & Roberts, 2006). Recall that Gureckis and Goldstone (2006) described distributed systems according to three related properties: that they are comprised of individually identifiable units whose functioning can be described independently of the group context, that those units are loosely coupled and capable of influencing one another, and that the pattern of connectivity between them is dynamic. This description can be applied to neural networks as well as to groups in my function learning task. In a biological network, a given neuron can be described independently yet it both influences and is influenced by other neurons in the network dynamically over time. The same can be said of nodes in an artificial neural network. In my function-learning groups, each individual can undoubtedly be described without reference to others in their group. Individuals in my groups are loosely coupled by their connection to the group-level responses and feedback, and the relationships between individuals change as each one develops unique response patterns and strategies for matching target values.

To the extent that networks and groups of individuals can be described by properties common to all distributed systems, phenomena expressed in networks as a function of those properties should be replicable in groups. That is, the patterns of findings observed in distributed collections of neurons, real or artificial, ought to scale up to distributed collections of individuals. To test this hypothesis of scale invariant principles, I adapted Lashley and Wiley's (1933) cortical lesioning procedure in rats to disruption in groups of participants in a function learning task.

Groups in Experiment 3 underwent the same three blocks of training as those in Experiments 1 and 2, and afterward had small (10%), medium (30%), or large (50%) amounts of disruption applied to the groups' responding for a subsequent relearning phase. To the extent that similar principles are operating at these two scales, groups' performance during the relearning phase should not be significantly impacted unless the amount of disruption is great; Lashley found that at least some of the original learning was preserved until approximately 50 per cent or more of the cortex was lesioned (see Lashley, 1950). I anticipated the same pattern of graceful degradation in my function-learning groups.

Experiment 3

After confirming in Experiment 2 that the basic effects of learning scale up to groups of 16 individuals, the present experiment returns to training and testing groups of 10 individuals. Training proceeded identically to Experiment 1, returning once again to the target values 3, 5, 7, and 9. In all conditions, the target values given to participants were the true values they were aiming for. As in the previous experiments, learning in this Experiment was demonstrated by a decrease in mean squared error over the course of training.

After three blocks of training, participants were deceived. Unbeknownst to them, a set percentage of participants in each group was "lesioned"[8] such that their responses would no longer be counted in a subsequent relearning phase. The experimental conditions differed only in the amount of disruption applied: 1 of 10 participants was excluded in the Small disruption condition, 3 of 10 were excluded in the Medium disruption condition, and 5 of 10 were excluded in the Large disruption condition. Relearning proceeded identically to the training phase, with the exception of the deception: groups attempted to achieve target values and received feedback after every attempt.

If the behaviour of groups in this task mimics the behaviour of real or artificial neural networks, then degradation of learning following disruption should be *graceful.* That is, behaviour should be relatively spared for Small (10%) and Medium (30%) disruptions

---

[8] No undergraduates were harmed in the conducting of these experiments.

but not for Large (50%) disruptions. The detrimental effect of disruption will be determined by (1) the change in squared error within groups before and after disruption, and (2) differences in those changes between conditions.

Method

Participants

Three hundred and ten undergraduate students participated in this experiment in exchange for partial course credit. Participants were run in groups of 10, each of which was randomly assigned to one of three conditions: groups in the Small condition underwent a 10% disruption (i.e., one participant) during the relearning phase; groups in the Medium condition underwent a 30% disruption (i.e., three participants) during the relearning phase; and groups in the Large condition underwent a 50% disruption (i.e., five participants) during the relearning phase. The different conditions received identical training on four untransformed target values (i.e., 3, 5, 7, and 9), and differed only on their treatment during the relearning phase. Ten groups (i.e., a total of 100 participants) were run in each the Small and Large conditions. Eleven groups (i.e., a total of 110 participants) were run in the Medium condition, but one group was excluded due to sabotage by students, leaving a total of ten groups per condition. For purposes of data analysis, each group of 10 participants served as one subject.

Apparatus

Training data were collected in a manner identical to that described in Experiment 1.

Procedure

Pre-training & training. For all three conditions, pre-training was identical to that described in Experiment 1 with groups of 10 participants. Training in all three conditions proceeded identically to the Simple condition of Experiment 1. As a reminder, groups were trained on each of four target values (i.e., 3, 5, 7, and 9) for three blocks with numeric and directional feedback after every attempt. Once again, training on a given target value was repeated until it was matched on two consecutive trials, or until the group reached a maximum of 19 attempts. All participants' responses were counted throughout the training phase.

Relearning phase. Following the third block of training, participants each selected one of ten slips of paper. Each slip had either a letter *A* or a letter *B* printed on it, and participants were instructed to push the corresponding button on their *i>clicker* remote. In the Small condition, the pool of slips consisted of nine *A*'s and one *B*; in the Medium condition, seven *A*'s and three *B*'s; and in the Large condition, five *A*'s and five *B*'s. These slips were used to randomly assign students to have their responses included (i.e., *A*'s) or excluded (i.e., *B*'s) from the retraining phase. The exclusion of responses was accomplished without participants' knowledge using a demographic filtering function in the *i>clicker* software.

In all three conditions, relearning proceeded as though it were a fourth block of training. In the Small condition, groups were asked to once again match the targets 3, 5, 7, and 9, with feedback provided after every attempt, excluding responses from the one disrupted participant; with nine remaining participants, all of the original training target values could be matched. In the Medium condition, only seven undisrupted participants remained, so these groups were asked to once again match the targets 3, 5, and 7. Responses were excluded from the three disrupted participants and feedback was provided after every attempt. In the Large condition, only five undisrupted participants remained, so these groups were asked to once again match the targets 3 and 5. Responses were excluded from the five disrupted participants and feedback was provided after every attempt. In the Medium and Large conditions, no explanation was given for the exclusion of some target values until debriefing. This procedure was intended to mimic demonstrations of lesioning in brains (i.e., Lashley & Wiley, 1933) or ANNs (e.g., Hinton & Shallice, 1991; Plaut & Shallice, 1993).

Detecting sabotage. Sabotage, whether intentional or incidental, was detected in the same manner as described in Experiment 1. In the present experiment, one group of participants in the Medium condition was excluded as a result of this detection. Description and analysis of these sabotaged groups is deferred until Chapter 5.

Additional tasks. As in Experiments 1, a brief semantic distractor activity was conducted between blocks of training. These proceeded exactly as described in the methods section of Experiment 1. Between the third block of training and the retraining phase, a similar distractor activity was conducted, in which participants were each asked to name aloud an animal that begins with the same letter as their first name.

Following completion of the retraining phase, participants were asked to write down free-formed thoughts about the strategies they used to make their decisions throughout the experiment, as in Experiments 1 and 2. Finally, all participants completed two short personality questionnaires, as described in Experiment 2: the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008), and the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003). As in Experiment 2, within-subject correlations between BFI and TIPI scores varied widely, ranging from r = -.89 to r = .97 in the Small condition (M = 0. 35, SD = 0.52), from r = -.98 to r = .99 in the Medium condition

($M = 0.19$, $SD = 0.49$), and from $r = -.89$ to $r = .98$ in the Large condition ($M = 0.27$, $SD = 0.46$). As such, I opted to discard the TIPI measures in favour of the more robust BFI measures. Analysis of all individual-level data, including self-reported strategies and personality measures, are found in Chapter 5.

Results

Training Phase

The results of the training phase of Experiment 3 are shown in the top panel of Figure 4.1. As in the previous experiments, my primary interest in the training phase was whether accuracy improved over the course of training. Once again, I first looked at the mean squared error averaged over all trials within a block. A 3 (Condition: Small, Medium, or Large) x 3 (Block) mixed measures ANOVA found a main effect of block, $F(2, 54) = 13.08$, $p < .001$. Polynomial a priori contrasts on the main effect of block confirmed that squared error decreased linearly over blocks of training, $t(54) = 5.11$, $p < .001$. The corresponding quadratic contrast was not reliable, $t(54) = 0.27$, $p = .79$. In this experiment, all three conditions were identically trained, so no effect of condition or block by condition interaction were expected; both predictions were confirmed, $F(2, 27) = 1.19$, $p = .32$ and $F(4, 54) = 0.88$, $p = .48$, respectively. As in Experiments 1 and 2, I could have alternately examined performance as a function of the number of trials taken to achieve targets, or as a function of the squared error achieved on the first attempt at a target. Once again, however, analyses of these alternative measures provided conclusions identical to the primary measure, and are thus not reported in the text (see Appendix A).

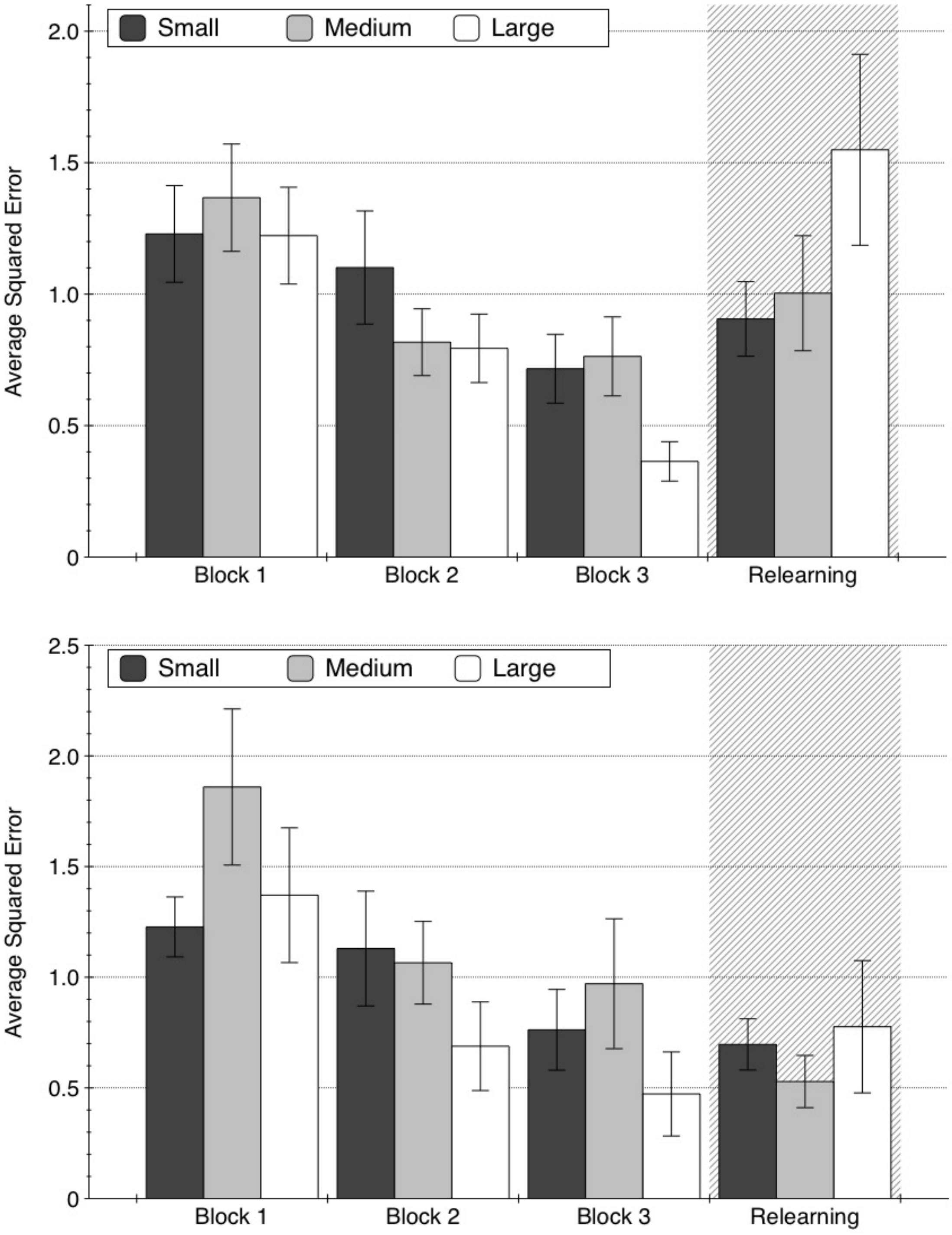

Figure 4.1. Squared error averaged over all attempts in a block in both the training phase (white background) and the relearning phase (hashed background) of Experiment 3, as a function of condition. Error bars represent SEMs. Top panel: all target values are included

in the averages for all blocks. Bottom panel: the averages for all blocks exclude those target values not presented during the relearning phase, as well as those target values at the boundary of what each condition could achieve following disruption at relearning. Thus, averages include the target values 3, 5, and 7 for the Small condition; target values 3 and 5 for the Medium condition; and only target value 3 for the Large condition.

Examined individually, the Small condition, $t(18) = 2.13$, $p = .047$, the Medium condition, $t(18) = 2.329$, $p = .03$, and the Large condition, $t(18) = 5.0$, $p < .001$, each showed the same pattern of linear decrease in mean squared error averaged across all attempts in a block.

Relearning Phase

In order to examine how disruption affected group performance on subsequent relearning of previously trained target values, two types of analyses were conducted: (1) within groups analyses that compared performance in the first block of training, prior to any learning, to performance during relearning, and (2) between groups analyses that compared re-learning performance as a function of the disruption size. I delineate both of these approaches in turn.

Within-group analyses. The analysis within each group can be conducted in a number of different ways, owing to differences between the training and re-learning phases. Below, analysis of the Small condition is used to highlight the different approaches to the within-group comparisons, which will then be applied to the remaining conditions in turn. The results of these analyses are compiled and presented in Table 4.1 for ease of comprehension.

Small condition. In the Small condition, groups relearned the full range (i.e., 3, 5, 7, 9) of training targets, so performance on all four targets could be directly compared between phases. When comparing performance on the full range of targets between block 1 and relearning (hereafter referred to as the Full-B1 vs. Full-RL comparison), error was lower during relearning (M = 0.91, SD = 0.40) than during block 1 (M = 1.23, SD = 0.54), but this difference was not reliable, $t(9) = 1.61$, $p = .14$. However, with only nine undisrupted participants in each group, the target value of 9 demanded that *all* participants respond in order to succeed. Including the boundary target (e.g., target 9 in the Small condition; target 7 in the Medium condition; target 5 in the Large condition) in the calculation of mean relearning performance may skew the error too far for such small sample sizes and thus render the comparison overly conservative.

A more liberal comparison excludes the boundary target from the calculation of mean relearning performance, while either including (i.e., Full-B1 vs. Exclusion-RL) or excluding (i.e., Exclusion-B1 vs. Exclusion-RL; see the bottom panel of Figure 4.1) the boundary target from the calculation of mean performance during block 1. In the Small condition, both comparisons showed better performance during relearning than during block 1 of training, although the difference was only reliable for the Exclusion-B1 vs. Exclusion-RL comparison, $t(9) = 2.35$, $p = .04$; the Full-B1 vs. Exclusion-RL comparison fell just short of statistical significance, $t(9) = 2.25$, $p = .051$. Finally, performance on the boundary target value was compared between block 1 and relearning (i.e., Boundary-B1 vs. Boundary-RL). Performance during relearning (M = 1.53, SD = 0.90) was slightly worse than during block 1 of training (M = 1.23, SD = 1.59), but this difference was not reliable, $t(9) = 0.88$, $p = .40$.

To summarize, groups were relatively robust to a Small amount of disruption. When the boundary target value (i.e., 9) was excluded from consideration, performance during relearning was comparable to or better than performance during block 1 of training. Groups in this condition did struggle to match the boundary target as a result of the disruption, but performance on the remaining targets was minimally affected.

Medium condition. In the Medium condition, groups relearned a restricted range (i.e., 3, 5, 7) of the training targets. When comparing Full-B1 vs. Full-RL performance between phases, error was lower during relearning (M = 0.86, SD = 0.72) than during block 1 (M = 1.32, SD = 0.46), but this difference was not reliable, $t(9)$ = 1.76, p = .11. Under more liberal comparisons that exclude the boundary target from the calculation of mean relearning performance, groups performed better during relearning than during block 1 whether the boundary target was included (i.e., Full-B1 vs. Exclusion-RL) or excluded (i.e., Exclusion-B1 vs. Exclusion-RL) from the calculation for block 1, $t(9)$ = 4.29, p = .002 and $t(9)$ = 3.62, p = .006, respectively. Finally, the Boundary-B1 vs. Boundary-RL comparison showed better performance during block 1 (M = 0.77, SD = 0.71) relative to relearning (M = 1.70, SD = 2.83), but this difference was not reliable, $t(9)$ = 1.13, p = .29.

As with a Small amount of disruption, groups were relatively robust to a Medium amount of disruption. When the boundary target value (i.e., 7) was excluded from consideration, performance during relearning was consistently better than performance during block 1 of training. Groups did struggle to match the boundary target as a result of the disruption, but performance on the remaining targets was unaffected.

Table 4.1. *Descriptive Statistics and Hypothesis Tests for Within-Group Comparisons in Experiment 3.*

Condition	Phase	Comparison	Targets	M (*SD*)	t-value	*p*
Small	*Block 1 (B1)*	Full	3, 5, 7, 9	1.23 (*0.54*)		
	Relearning (RL)	Full	3, 5, 7, 9	0.91 (*0.40*)	$t(9) = 1.61$	.14
	Block 1 (B1)	Full	3, 5, 7, 9	1.23 (*0.54*)		
	Relearning (RL)	Exclusion	3, 5, 7	0.70 (*0.37*)	$t(9) = 2.25$	.05*
	Block 1 (B1)	Exclusion	3, 5, 7	1.23 (*0.43*)		
	Relearning (RL)	Exclusion	3, 5, 7	0.70 (*0.37*)	$t(9) = 2.35$	.04**
	Block 1 (B1)	Boundary	9	1.23 (*1.59*)		
	Relearning (RL)	Boundary	9	1.53 (*0.90*)	$t(9) = 0.88$	.40
Medium	*Block 1 (B1)*	Full	3, 5, 7, 9	1.32 (*0.46*)		
	Relearning (RL)	Full	3, 5, 7	0.86 (*0.72*)	$t(9) = 1.76$	.11
	Block 1 (B1)	Full	3, 5, 7, 9	1.32 (*0.46*)		
	Relearning (RL)	Exclusion	3, 5	0.53 (*0.37*)	$t(9) = 4.29$	.002**
	Block 1 (B1)	Exclusion	3, 5	1.86 (*1.11*)		
	Relearning (RL)	Exclusion	3, 5	0.53 (*0.37*)	$t(9) = 3.62$	.006**
	Block 1 (B1)	Boundary	7	0.77 (*0.71*)		
	Relearning (RL)	Boundary	7	1.70 (*2.83*)	$t(9) = 1.13$	.29
Large	*Block 1 (B1)*	Full	3, 5, 7, 9	3.00 (*1.39*)		
	Relearning (RL)	Full	3, 5	4.20 (*3.63*)	$t(9) = 0.91$	.39
	Block 1 (B1)	Full	3, 5, 7, 9	1.22 (*0.56*)		
	Relearning (RL)	Exclusion	3	0.78 (*0.94*)	$t(9) = 1.07$	.31
	Block 1 (B1)	Exclusion	3	1.37 (*0.96*)		
	Relearning (RL)	Exclusion	3	0.78 (*0.94*)	$t(9) = 1.52$	.16
	Block 1 (B1)	Boundary	5	1.90 (*1.53*)		
	Relearning (RL)	Boundary	5	2.32 (*1.83*)	$t(9) = 0.50$	.63

Note. * $p < .10$, ** $p < .05$.

Large condition. In the Large condition, groups relearned a restricted range (i.e., 3, 5) of the training targets. When comparing Full-B1 vs. Full-RL performance between phases, error was higher during relearning ($M = 4.20$, $SD = 3.63$) than during block 1 ($M = 3.00$, $SD = 1.39$), but this difference was not reliable, $t(9) = 0.91$, $p = .39$. Even under the more liberal Full-B1 vs. Exclusion-RL and Exclusion-B1 vs. Exclusion-RL comparisons, performance during relearning was not reliably better than during block 1, $t(9) = 1.07$, $p = .31$ and $t(9) = 1.52$, $p = .16$ respectively. Finally, the Boundary-B1 vs. Boundary-RL comparison showed slightly better performance during block 1 ($M = 1.90$, $SD = 1.53$) relative to relearning ($M = 2.32$, $SD = 1.83$), but this difference was not reliable, $t(9) = 0.50$, $p = .63$.

In contrast to the Small and Medium conditions, groups in the Large condition were detrimentally affected by the disruption regardless of how liberal a comparison was made. In all cases, performance during relearning was either no different than or worse than performance during block 1 of training, indicating that their training was not robust to such widespread disruption.

Between-group analyses. A 2 (Block: block 3 vs. relearning) x 3 (Condition) ANOVA conducted on the mean squared error averaged over all attempts in a block showed a main effect of block, $F(1, 27) = 7.23$, $p = .01$, indicating that disruption affected groups' relearning in some way. More importantly, a significant block by condition interaction, $F(2, 27) = 4.08$, $p = .03$, shows that groups were differentially affected as a function of the size of the disruption applied. To examine this result more thoroughly, I conducted additional analyses on difference scores (performance during relearning - performance on block 3) that capture the detrimental effect of disruption within each group following block 3. This approach moreover minimizes the influence of differing baselines among the three conditions, increasing the power to detect differences among them. As with the within-group analyses, the nature of the task leaves the between-group analyses open to a number of approaches.

First, a 3-way ANOVA demonstrated significant differences across conditions in the overall difference scores, *F*(2, 27) = 3.68, *p* = .04. A priori contrasts found no difference between the Small (*M* = 0.19, *SD* = 0.45) and Medium (*M* = 0.24, *SD* = 0.98) conditions, *t*(27) = 0.12, *p* = .90, but difference scores in the Large condition (*M* = 1.18, *SD* = 1.19) were reliably greater than scores collapsed over the Small and Medium conditions, *t*(27) = 2.71, *p* = .01. Calculated in this way, the difference scores were no greater than zero in either the Small condition, *t*(9) = 1.33, *p* = .22, or the Medium condition, *t*(9) = 0.78, *p* = .46, but were greater than zero in the Large condition, *t*(9) = 3.16, *p* = .01.

Although overall difference scores are informative, this measure averages across 4 target values for the Small condition (i.e., 3, 5, 7, 9), 3 target values for the Medium condition (i.e., 3, 5, 7), and 2 target values for the Large condition (i.e., 3, 5); the differences in the number of observations in each group may skew the results. A second approach is to take difference scores for only those target values that all conditions have in common (i.e., 3 and 5). A one-way ANOVA (i.e., 3 levels of condition) on these scores shows the same pattern of results. There was a significant effect of condition, *F*(2, 27) = 6.99, *p* = .004. A priori contrasts found no difference between the Small (*M* = -0.14, *SD* = 0.78) and Medium (*M* = -0.44, *SD* = 0.97) conditions, *t*(27) = 0.71, *p* = .48, but difference scores collapsed over the Small and Medium conditions were significantly smaller than difference scores in the Large condition (*M* = 1.08, *SD* = 1.11), *t*(27) = 3.67, *p* = .001. Calculated in this way, the difference scores were no greater than zero in either the Small condition, *t*(9) = 0.55, *p* = .60, or the Medium condition, *t*(9) = 1.45, *p* = 0.18, but were greater than zero in the Large condition, *t*(9) = 3.08, *p* = .01.

As in the within-group comparisons, including the boundary target (i.e., target 9 in the Small condition; target 7 in the Medium condition; target 5 in the Large condition) in the calculation of difference scores may inflate estimates. A more conservative approach compares the conditions only on the non-boundary target value that they all have in common (i.e., 3). A 3-way ANOVA did not find any difference between the three conditions, $F(2, 27) = 2.16$, $p = .14$, suggesting that even the Large condition was not greatly disadvantaged in relearning the target value 3. Further to this interpretation, difference scores for the target value of 3 were no greater than zero in any condition, $t(9) = 0.14$, $p = .89$ in the Small condition, $t(9) = 1.74$, $p = .12$ in the Medium condition, and $t(9) = 1.01$, $p = .34$ in the Large condition.

Similarly, difference scores for just the boundary target in each condition can be compared (i.e., comparing target 9 in the Small condition, target 7 in the Medium condition, and target 5 in the Large condition). A 3-way ANOVA found no differences between the three conditions, $F(2, 27) = 0.97$, $p = .39$. The difference score for this target was reliably different from zero in each the Small, $t(9) = 5.00$, $p < .001$, Medium, $t(9) = 2.74$, $p = .02$, and Large conditions, $t(9) = 3.01$, $p = .01$.

To summarize, the amount of disruption applied had a differential impact on how well groups retained learning between block 3 of training and the relearning phase. Overall, the Small and Medium conditions were less detrimentally affected than the Large condition. However, performance on the lowest target value (i.e., 3) was spared even following a Large amount of disruption, whereas performance on the boundary target value was impaired relative to block 3 performance following any amount of disruption.

Discussion

The results of this experiment provide not only a replication of the results of Experiment 1, but also a test of the robustness of function learning in groups. As in the previous experiments, groups in the current experiment were able to learn a linear function, as evidenced by decreasing error over the course of training. The results from the relearning phase provide evidence that groups are relatively resistant to the effects of disruption. Taken together, within-group and between-group analyses suggested that groups were minimally impacted by a Small or Medium amount of disruption. With a Large amount of disruption, however, relearning behaviour suffered. Of course, this pattern of results was not uniform for all approaches to the analysis. For example, performance in all conditions was most apparently affected for boundary target conditions, even with a Small or Medium amount of disruption. By contrast, performance on the non-boundary target value 3 was comparable between block 3 and relearning even when with a Large amount of disruption.

Altogether, the results provide preliminary support for my hypothesis that the behaviour of groups in a function learning task mimics at least some of the behaviours of neural networks, despite their differences in scale. Just as Lashley and Wiley (1933) showed an exponential increase in learning deficits as a function of increasing area of cortical lesion, a similar relationship was found when I excluded different proportions individual respondents from groups. Critically, disruption didn't prevent groups from relearning the task altogether. Indeed, relearning performance following Small or Medium amounts of disruption was largely unimpaired, except for boundary target values. Following a Large amount of disruption, however, relearning performance was generally slower and more error-prone than initial learning. To Lashley (e.g., 1950), this pattern of results was evidence of two interconnected properties of distributed representation in the brain: *mass action*, the notion that all areas of cortex are involved in a given behaviour, and *equipotentiality*, by which the function of any cortical area is not predetermined, but is flexible according to need. Applying these properties to groups in a function learning task, all participants in a group are involved in matching target values, and the role a given participant is not predetermined, but is flexible according to the particular needs of the group.

One alternate explanation for the results of Experiment 3 is that that probability structure of the task differed between conditions, making it more likely that participants would have to change their responses in the Large condition than in the Small or Medium conditions. Thus, the appearance of graceful degradation would be a function of probability rather than any feature of the hypothesized distributed representation of the task. In order to address this possibility, the probability structure of the task facing each condition in the Relearning phase is given in Appendix B. As shown, the conjunctive probability of any given participant being excluded from their group's response and of that participant having been a responder to a given target (equivalent to the probability of at least one participant needing to change her response for the group to succeed) does indeed favour the Small condition. This fact is not inconsistent with neurological lesioning studies, in which a neural network would have to re-arrange more extensively in order to accomplish a task following a larger amount of cortical excision than a smaller amount of excision. However, the corresponding results on the right-hand side of the table in Appendix B show that this fact cannot fully account for performance during relearning in Experiment 3. For example, the conjunctive probability for both the Small condition relearning the target 9 and the Medium condition relearning the target 3 is 9/100 or 9% of cases. Yet performance is dramatically better in the Medium condition in these instances. Likewise, the conjunctive probability for both the Medium condition relearning the target 5 and the Large condition relearning the target 3 is 15/100 or 15% of cases. Yet performance in these instances favours the Large condition. Although it is undoubtedly a factor in performance, the probability structure of the task alone cannot explain the results of Experiment 3. Instead, the results are also consistent with Lashley and Wiley's (1933) finding of graceful degradation.

It is worth reiterating that I intend the analogy between groups and networks to be taken superficially, applying to behaviour rather than structure. There are undeniable differences in the composition of biological networks, artificial neural networks, and groups of individuals. However, the *behavioural* correspondence between them, as demonstrated in this experiment, speaks to the overarching theory of self-similarity and a novel application of network theory. Just as Lashley and Wiley (1933) showed graceful degradation of a learned behaviour in rats following an increasing amount of cortical lesioning, learned function behaviours in groups was largely spared following a Small or Medium amount of disruption but severely hindered following a Large amount of disruption.

Summary of Chapters 2–4

The series of experiments reported in this thesis served as the first presentation of a novel experimental design and the first steps in a new line of inquiry. Although the ability of individuals to learn the relationship between two variables has long been appreciated, function learning in groups has heretofore not been studied. The question of whether groups, like individuals, can learn functions is of both empirical and theoretical importance.

In Experiment 1, I demonstrated that groups could learn both linear and nonlinear functions. In standard individual function learning tasks, linear relationships are generally learned faster than non-linear relationships (e.g., McDaniel & Busemeyer, 2005). In partial replication of this finding, groups in my experiment learned a Simple (linear) function numerically better than a Complex (cubic-like) function, but this difference between conditions could not be supported statistically.

Experiment 2 replicated the findings of Experiment 1 with a larger group size, thereby showing that the effect was not limited to groups of 10 participants but instead could scale up to 16 participants. In a subsequent transfer test, groups demonstrated a benefit of prior training on matching new target values that hadn't been previously trained: in both the Simple and Complex conditions, performance for new target values at test was numerically better than for old target values in the first block of training (i.e., prior to any learning), and significantly better relative to a Control condition that did not undergo any training. A typical finding in the individual function learning literature is that participants underestimate the value of the Y variable at both the high and low ends of the X variable when extrapolating a linear function (Brown & Lacroix, 2017; DeLosh, 1995; DeLosh et al., 1997; Kwantes & Neal, 2006). In contrast to this literature, groups in both the Simple (linear) and Complex (cubic-like) condition overestimated when extrapolating at the low end and underestimated when extrapolating at the high end. Moreover, this pattern held for the Control condition, suggesting a natural bias that can be mitigated but not eliminated by training in a group function learning task.

In Experiment 3, groups demonstrated that learned functions are relatively resistant to the effects of disruption following training. Generally speaking, small (10%) to medium (30%) amounts of disruption resulted in little to no degradation of performance in a relearning phase, whereas a large (50%) amount of disruption was quite detrimental. This pattern of findings is similar to the graceful degradation demonstrated by Lashley and Wiley (1933) when an increasing proportion of rat cortex was excised: performance on previously learned maze suffered as an exponential function of excision size.

The results of these experiments have important implications. First, they show that findings from function learning tasks are self-similar when scaling from individuals to groups of individuals. This is an important discovery as psychology tries to establish itself among the natural sciences, where scale invariance and self-similarity are wide-ranging principles (e.g., Koscielny-Bunde et al., 1998; MacIntosh, Pelletier, Chiaradia, Kato, & Ropert-Coudert, 2013; Peng et al., 1992; Peng, Havlin, Stanley, & Goldberger, 1994; Schlesinger & West, 1991). The data presented in this thesis add to the growing database of demonstrations of scale invariance within psychology (e.g., Howard & Eichenbaum, 2013; Howard, Shankar, Aue, & Criss, 2014; Neath & Brown, 2006).

Second, these results support a principled account of cognition wherein sophisticated behaviours are the result of simple processes operating within the constraints of a structured environment (e.g., Simon, 1969). Whereas our awareness of our own cognition gives us the unshakeable intuition that the mechanisms underlying our behaviours must themselves be sophisticated, it is difficult to argue the same for the group-level behaviours of a decentralized collection of individuals. Because a group has no frontal cortex within which an executive control could reside, the group function learning task removes that default from our explanatory toolbox and leaves only emergence. The ability of a group to learn a function is the result of simple individual responses to the constraints of the dynamic and structured task environment.

Finally, and relatedly, the results point to the distributed representation of a collaborative task among individuals in groups. As Theiner, Allen, and Goldstone (2010) argue, group cognition is not a simple aggregation of individual cognitions, but rather the emergent outcome of a division of labour (see also Goldstone & Gureckis, 2009). Knowledge of how to succeed in the group-level task does not reside in any individual, but is rather distributed across them dynamically as the group learns. This division of labour typically occurs without intent and as a result of interactions (often indirect) among agents. That is, group cognition is self-organizing (see Moussaid, Garnier, Theraulaz, & Helbing, 1993). This point moreover reinforces the theory of self-similarity, because self-organization is a fundamental component of biological neural networks, as well some of their artificial counterparts (e.g., Becker, 1996; Becker & Hinton, 1992; Kitzbichler, Smith, Christensen, & Bullmore, 2009; Kohonen, 1982; Ultsch, 2007). That features of individual brains can be replicated in collections of brains only amplifies the importance of self-similarity as a guiding principle.

I turn next to a more in depth analysis of how groups are able to accomplish this learning. By examining individuals' self-reported strategies, histories of responding in the task, and personality profiles, the means by which groups adaptively solve a complex problem become clearer.

Chapter 5: Understanding How Groups Learn

The experimental analyses presented in this thesis have demonstrated that groups can learn both simple and complex functions (Experiment 1), that groups can generalize their learning to unstudied regions of a function in a transfer test (Experiment 2), and that the learned behaviours of groups are relatively resistant to the detrimental effects of disruption (Experiment 3). Having shown that groups *can* learn, a natural follow-up question asks *how* they learn.

Groups as Artificial Neural Networks

In earlier chapters, I drew analogies between learning in groups of individuals and learning in networks. In both groups and networks, individual units activate in response to input and feedback alters patterns of activation in order to minimize error between the resulting behaviour and the target behaviour. Within a network, individual cells or nodes operate in concert to achieve some common goal. In order to contribute to the network-level behaviour, an individual cell need not have knowledge of the higher-level goal and need not have any insight into the activation tendencies of its neighbours; its sole obligation is to respond locally to feedback. Likewise, in my experiments with groups, individual participants are operating in concert to achieve some common goal (i.e., matching target values and finishing the experiment). In order to contribute to the group-level behaviour, each participant makes local decisions in response to feedback, without any insight into the decisions of his or her peers. In the case of groups learning nonlinear functions (Complex condition; Experiments 1-2), groups can achieve the higher-level goal despite that participants' knowledge of that goal is inaccurate. Though the analogy is useful, it bears natural limitations. Specifically, an adult human is neither a node in an ANN nor a cell in a biological network. The aim of this thesis is not to argue otherwise, but instead to uncover similarities and differences in the way principles of function learning operate at different scales of analysis.

I conducted simulations of Experiment 1 using a basic single-layer feedforward neural network, in which each participant was represented by a unique node. The network was given targets (e.g., 3, 5, 7 ,9), and individual nodes would respond (or not) on the basis of their individual connection weights with the output node. The combined activation over the network would represent the number of group members responding on a given trial. The difference between the target and the network's actual output would result in changes to activation weights, as a function of (1) the size of the mismatch between the group's target and its actual output, and (2) an individual's learning rate. Under this analogy, each node in the network operated as a threshold function. As long as the target value was less than or equal to a node's individual connection weight, that node would respond to the target. If a target value was greater than a node's connection weight, that node would not respond. For each simulated group, each node in the network was randomly seeded with a starting connection weight between 0 and 1, and a learning rate between 0.01 and 0.5. Randomized learning rates were introduced under the assumption that different individuals would have a different propensity to change their responses.

The results of this simulation were not as impressive as anticipated. Although simulated groups could learn to match target values, each group was restricted to a single set of weights as a solution for a given target value. To the extent that the solutions for the target value 3 and 9, for example, relied on different connection weights, catastrophic interference was inevitable and savings between blocks was limited. Because they were unsuccessful, the results of the simulations are not reported in this document.

A common computational approach to individual function learning is a different kind of artificial neural network, wherein each input node corresponds to one possible stimulus magnitude and each output node corresponds to one response magnitude (e.g., ALM, Busemeyer et al., 1997; EXAM, DeLosh, Busemeyer, & McDaniel, 1997). When a given stimulus is presented to the network, each node is activated to a degree dependent on its similarity to the stimulus. The model chooses the best response by comparing the similarity of a stimulus target and all input nodes, using a Gaussian retrieval gradient. The weighted sum of this activation is then passed to some decision process that selects the optimal output, and the association strengths between input and output nodes are updated via some learning rule.

As reported in Chapter 3, patterns of overestimation and underestimation on a transfer test matched the predictions of the ALM model (e.g., Busemeyer et al., 1997). As such, I conducted simulations of Experiment 2 using this model ($\alpha = .25$, $\gamma = 1$). Although more successful than simulations with the simple feedforward network, the results of these simulations nevertheless failed to fully capture the behaviour of groups in my group function learning task. Proper learning curves emerged in both the Simple and Complex conditions, with savings between blocks. Moreover, both simulated conditions experienced savings at test relative to an untrained Control condition and performance was best for old items, followed by interpolated items, followed by extrapolated items. However, in all measures simulated performance was better in the Complex condition than in the Simple condition; in the ALM model, there is no way to simulate participants' belief that the target is an untransformed value (e.g., 4, 7, 10, 13) when in fact it is a transformed value (e.g., 6, 8, 8, 11). Because the transformed targets in the Complex condition cover a smaller range than the untransformed targets in the Simple condition, the Gaussian retrieval process will always have more success with the transformed targets.

A second problem with this type of network model is that the mappings to the group function learning task are less clear. Unlike an input node in the ALM model, a given participant does not correspond to one target value in the range of possible values. Further, whereas the ALM model selects the best response from the activation of a few input nodes, the group-level response in my experiments involves the individual descisions of all participants. Solutions in the ALM model are represented much more locally than the results of my experiments, particularly Experiment 3, would suggest they are represented in groups. It may be the case that these computational models likewise oversimplify how neural networks accomplish individual function learning, but as long as science is unable to query individual neurons during a function learning task, the simplifying assumption is a useful one for making progress.

Alternate Approaches

Ideally, neural network simulations would have fully captured performance in my experiments. Throughout this thesis I have argued that individual and group function learning are examples of self-similar behaviours along a hierarchical scale; I would have liked to be able to demonstrate that the two levels of analysis were supported by self-similar mechanisms as well, using the same mathematics to explain function learning across the hierarchy. As it stands, however, the data limit the analogy. Whereas the results of my experiments offer some support for self-similarity of behaviour (e.g., at the level of the group), the results of the simulations call into question whether similar mechanisms are employed across the hierarchical scale to accomplish the task. Approaches beyond the neural network analogy are required to fully understand how groups learn functions.

Thankfully, the structure of the experiments in this thesis allows direct inquiry into the decision profiles of the individuals that constitute a given learning group. Three different but complementary types of data were collected during the group function learning task that can shed light on the behaviour of individual participants. First, self-report data were collected on the strategies that participants used to make their decisions throughout the experiment. Second, a record of participants' response patterns throughout the experiment was collected by the *i<clicker* software. Third, personality data were collected from participants in Experiment 3 and in the Control condition of Experiment 2. An analysis of these three types of data is an important addition to the results already discussed in Chapters 2 through 4. Understanding how groups learn must necessarily go beyond a group-level analysis, because each member of the group is herself a complex whole. These analyses will help to uncover the conscious and unconscious use of strategies, as well as the influences of pre-existing individual traits on both individual-level and group-level behaviours.

A Preemptive Summary

Due to the nature of these experiments, wherein each group of ten or sixteen individuals serves as one subject for the group-level analyses, the sheer volume of individual-level data collected is immense. The post-hoc and exploratory analyses to be reported in this chapter are numerous as a result. I here want to anticipate the general pattern of results in order to aid comprehension moving forward.

First, analyses of self-reported strategies will demonstrate that a homogenous set of strategies within a group was associated with better performance in the function learning task, but that the inclusion of some flexible strategies (i.e., strategies for altering responses when the group failed to match a training target value) was a critical predictor of a group's success. Second, analyses of participants' actual response patterns in the task will show that participants' self-reported strategies generally matched their behaviour, although there were high levels of usage of certain strategies even among those participants who did not report using them. These results indicate that self-report does not fully capture what participants are doing in the group function learning task, consistent with findings in the implicit learning literature (e.g., Neil & Higham, 2012; Reber, 1967). Finally, analyses of personality variables will suggest that, like self-reported strategies, more homogeneity of personality profiles was associated with better group performance. There were moreover some relationships between participants' personality scores and their tendency to self-report certain strategies, particularly for the Extroversion dimension. I expand on each of these findings below.

Self-Reported Strategies

At the end of each experiment, self-report data were collected from every participant about the strategies he or she used in order for the group to succeed. Responses were hand-coded to determine which of twenty-six categories of strategy each participant reported. The list of possible categories was initially limited to those I predicted participants might report. On a first pass of all participants' self-report data, the list of categories changed as new strategies were encountered and added, and unused categories were discarded or merged with similar but more common categories. On a second pass of the data, participants were re-categorized according to the finalized list of useful categories, which are listed and described in Table 5.1. As shown, these categories were not mutually exclusive; a given participant could be coded to have used any number of strategies[9]. Some categories pertained only to training blocks in all experiments (categories A through J in Table 5.1). For example, Extreme Adjust Only (category I) was assigned when a participant indicated that they changed their response during training if the group was extremely far from the target value. Other categories pertained only to the test phase of Experiment 2 (categories X through Z). For example, Anchor For New (category Y) was assigned when a participant indicated that they based their responses for New test target values on their responses to Old target values during training. Finally, some categories could pertain to either training or test, or both. Single Threshold (Category M), for instance, was applied when a participant indicated that they would respond one way when the target was over some set value and the other way when the target was below that value.

The right-hand side of Table 5.1 gives the percentage of participants in Experiments 1 through 3 who self-reported each strategy. As shown, some strategies were reported more frequently than others. However, strategy distributions were consistent across experiments, with correlations of $r = .97$ between Experiments 1 and 2, $r = .95$ between Experiments 1 and 3, and $r = .96$ between Experiments 2 and 3[10]. The most commonly reported strategies were Recall and Repeat (category D; 28.5% overall), Let Others Adjust (category E; 23%), and Majority (category P; 20.3%), with at least one fifth of all participants reporting using each.

---

[9] The one exception to this was category W (None), which criteria states that the participant "did not answer or did not understand the question," because it was only assigned when no other category fit the participant's self-report.

[10] Correlations with Experiment 2 exclude the Control condition and Test Only categories.

Table 5.1. *Coding Guidelines for Self-Reported Strategies, and Percentage of Participants Reporting Each Strategy Per Experiment.*

				Percentage Reported		
		Category	Description	E1	E2	E3
Training Only	A.	Adjust after one	Often changed response when group was over/under once	5.0	4.1	6.3
	B.	Adjust after two	Often changed response when group was over/under twice in a row	4.5	1.9	0.7
	C.	Adjust after many	Changed response only when group was over/under repeatedly	4.5	5.3	6.3
	D.	Recall and repeat	Recalled and repeated responses from previous block(s)	27.0	27.5	31.0
	E.	Let others adjust	Maintained response when group was over/under, expecting others to change	23.5	22.5	23.0
	F.	Directional adjust	Adjusted if over but not under, or vise versa	5.0	5.3	8.3
	G.	Specific adjust	Adjusted if over/under for particular targets but not others	1.5	1.3	2.3
	H.	Occasional adjust	Adjusted sometimes (no clear pattern described)	10.0	12.2	7.0
	I.	Extreme adjust only	Adjusted only when group was over/under by large amounts	6.5	5.9	2.7
	J.	Compensate	Compensated for others adjusting by adjusting in opposite direction	1.5	4.7	3.3
Either Training or Test	K.	Respond tendency	Tended to respond	11	11.9	15.3
	L.	Not respond tendency	Tended to not respond	11	13.1	11.7
	M.	Single threshold	Responded if target was above/below a given threshold, otherwise not	15	17.8	16.3
	N.	Complex threshold	Responded to particular targets but not others, no single threshold	5.0	1.6	1.3
	O.	Preferred numbers	Responded based on some personal preference for certain numbers	0	0.6	0
	P.	Majority	Tendency to respond to higher numbers	17.0	22.8	21.0
	Q.	Minority	Tendency to respond to lower numbers	5.5	8.1	2.7
	R.	Probability	Decided based on probability/chance/odds, given group size	3.5	2.8	2.3
	S.	Social decisions	Decided based on some judgement about the peers in the room	0.5	1.6	4.3
	T.	Intuition	Decided based on intuition/gut feeling	3.5	1.6	1.7
	U.	Random responding	Claimed to respond randomly	2.5	1.6	0.3
	V.	Uninformative	Reported some nonsensical/unhelpful/idiosyncratic strategy	4.5	2.8	4.3
	W.	None	Did not answer or did not understand the question	4.0	2.5	4.3
Test Only	X.	Maintain for old	Gave same responses for Old targets at test as during training	--	3.1	--
	Y.	Anchor for new	Anchored responses for New targets at test around responses for Old targets	--	1.6	--
	Z.	Unpatterned	Alternated between respond/not respond throughout test	--	1.9	--

Notes. E1 = Experiment 1 (200 participants total); E2 = Experiment 2 (320 participants total); E3 = Experiment 3 (300 participants total). Response categories are not mutually exclusive. E2 percentages exclude Control condition.

Given this degree of overlap between participants' reported strategies, it is of interest to know whether overlap aided or hindered performance in the group function learning task. Specifically, does group performance fare better when participants use the same set of strategies, or when participants diversify their approaches to the task? There was no a priori reason to anticipate either effect, though one could imagine reasons why either might be the case. If success relied upon cohesion among group members, for instance, we would expect a benefit for groups with redundant strategies. Bettencourt (2009), for instance, argued that coordination in collective cognitive tasks could be understood as the minimization of uncertainty and the maximization of predictability. Alternately, if success in group function learning relied upon a distributed representation of task knowledge, we would expect to see a benefit for more diverse group makeups. With their Group Binary Search (GBS) game, Roberts and Goldstone (2011) emphasized the importance of complementary responses and the differentiation of individual roles, showing that greater variability in response patterns among participants predicted better group performance. Given the differences between the GBS game and my function learning task, it was not clear what effect variability in strategies would play in group success.

To answer this question, I applied a measure of redundancy from information theory to the self-report data from each group. In information theory, redundancy is a statistic, ranging from 0 to 1, which quantifies the degree to which a sequence predicts itself (e.g., Garner, 1962; McGill, 1954; Shannon & Weaver, 1949). In a perfectly redundant sequence (i.e., redundancy is 1), any part of the sequence perfectly predicts any other part. In a perfectly non-redundant sequence (i.e., redundancy is 0), no part of the sequence can be predicted from any other part because the sequence is random. Most sequences fall between the two extremes, indicating some amount of structure with some randomness.

In the context of the self-report data, the statistic is calculated from a vector of probabilities of each category being reported, rather than a "sequence" per se. For each of g groups, R_g is the measure of redundancy within the group, which indexes the degree to which any participant's strategies can be predicted by the strategies reported by any other participant in the same group,

$$R_g = 1 - \frac{-\sum_{i=1}^{n} p_i log_2(p_i)}{log_2(n)},$$

where n is the number of possible categories (i.e., 23 in Experiments 1 and 3, 26 in the experimental conditions of Experiment 2, and 16 in the Control condition of Experiment 2) and p_i is the probability of the category i being reported. The resulting redundancy measures can then by converted into percentages. Table 5.2 shows the percent redundant measures for each group. As shown, reported strategies within groups were about 30% redundant on average, with a reasonable amount of variability between groups.

Correlation analyses assessed whether more or less redundancy in self-reported strategies was associated with better performance in group function learning, looking first at the training phase only. Because all three experiments involved a similar training phase, the only difference being more participants per group in Experiment 2, I collapsed over all experiments to increase the power to detect significant correlations. Note that the Control condition of Experiment 2 was excluded from these analyses, because groups in that condition did not complete a training phase.

Table 5.2. *Percent Redundant Measures for Each Group in Experiments 1-3.*

	Condition	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Mean (*SD*)
Exp. 1	*Simple*	41.3	27.7	30.8	26.4	37.9	31.4	40.7	26.1	32.4	55.7	35.0 (*9.13*)
	Complex	33.5	33.5	17.6	39.2	26.0	21.0	20.0	19.9	28.5	27.9	26.7 (*7.14*)
Exp. 2	*Simple*	20.5	16.7	32.4	24.7	19.6	25.4	41.0	30.9	25.0	22.5	25.9 (*7.16*)
	Complex	25.3	21.8	24.0	28.0	36.8	26.3	31.2	24.0	29.9	34.3	28.2 (*4.85*)
	Control	39.4	28.8	28.9	20.6	27.9	23.2	30.4	24.3	38.1	29.4	29.1 (*5.98*)
Exp 3.	*Small*	23.0	39.2	37.1	39.2	33.1	22.6	37.8	28.8	22.6	20.8	30.4 (*7.70*)
	Medium	25.6	31.8	41.2	31.4	30.2	29.6	37.6	29.5	22.6	30.4	31.0 (*5.29*)
	Large	33.2	53.2	32.7	32.6	33.2	29.3	25.6	33.3	34.4	45.1	35.3 (*7.98*)

Redundancy was negatively correlated with mean squared error averaged across all trials in a block, for each block 1 (r = -.24, p = .04), block 2 (r = -.27, p = .03), and block 3 (r = -.40, p < .001). That is, in each block of training, higher redundancy within a group was associated with better performance, as indicated by lower error scores. Likewise, redundancy was negatively correlated with the number of attempts taken to match a target value in block 2 (r = -.27, p = .03) and block 3 (r = -.43, p < .001), and also with mean squared error on first attempts at targets for each block 1 (r = -.26, p = .03), block 2 (r = -.249, p = .04), and block 3 (r = -.34, p = .004). Overall, agreement between participants in a group was associated with better performance in the training phase.

In Experiment 1 alone, redundancy was associated only with better performance in block 3, as measured by the number of attempts taken to match targets (r = -.49, p = .03), mean squared error averaged over all attempts (r = -.44, p = .05), and mean squared error averaged over first attempts (r = -.50, p = .03). In Experiment 2 alone, redundancy was not significantly associated with any measure of performance, in either training (maximum r =

-.19, $p = .42$ for mean square error averaged over attempts in block 3) or test (maximum $r = .27$, $p = .15$ for mean squared error on Old target values). In Experiment 3 alone, redundancy was significantly associated only with the number of attempts taken to match targets in block 3 ($r = -.44$, $p = .02$) and marginally associated with the mean squared error averaged over attempts in block 3 ($r = -.342$, $p = .07$). During the relearning phase, redundancy was not associated with any measure of performance (maximum $r = .18$, $p = .34$ for mean squared error averaged over attempts).

Because mean squared error averaged over all attempts in block 3 repeatedly showed correlations with redundancy, this seemed like a good measure to examine more closely. Specifically, I asked which of the 23 training strategies best predicted mean squared error performance in block 3, collapsed over experiments. When the strategy categories were regressed on mean squared error in block 3 in a stepwise fashion, with an entry criterion of $p < .05$ and a removal criterion of $p > .10$, four categories emerged as reliable predictors: Majority (category P), $b = 0.52$, $t(65) = 5.82$, $p < .001$; Adjust After Two (category B), $b = 0.36$, $t(65) = 4.02$, $p < .001$; Adjust After Several (category C), $b = 0.25$, $t(65) = 2.85$, $p = .006$; and Adjust After One (category A), $b = 0.23$, $t(65) = 2.59$, $p = .01$[11]. This model accounted for nearly half of all variance in performance in block 3, $R^2_{adj} = .46$, $F(4, 65) = 15.52$, $p < .001$.

Thus, performance in the training phase was aided by a combination of strategies. Groups performed better when a sizeable proportion of participants intentionally responded with the majority (i.e., respond for high numbers, do not respond for low numbers) but when the group nevertheless remained flexible, with some participants

---

[11] All reported regression coefficients represent standardized beta weights.

altering responses when the group failed to match the target for one, two, or several consecutive attempts. The inclusion in the predictive model of three different strategies for altering responding suggests some degree of variability in approaches is beneficial for success. If too many participants changed their responses after the group failed to match a target once, the group's behaviour would vary too liberally and optimal matching behaviour would not be achieved. However, if too many participants refused to change their responses until the group had failed to match a target for several attempts, the group would be too inflexible for optimal matching behaviour to be achieved. Instead, the combination of different strategies for altering responses optimized matching behaviour during training.

Nevertheless higher levels of redundancy in reported strategies, rather than more distributed strategy sets, were correlated with better performance overall. There are a number of possible explanations for the discrepancy, but I focus on three below.

First, recall that the strategy categories were not mutually exclusive. A given participant could report responding with the majority (category P), recalling and repeating previous responses (category D), adjusting after several mismatches (category C), responding to preferred numbers (category O), *and* responding according to a single threshold (category M), as one example participant in Experiment 2 did. Such instances could lead to reduced redundancy (i.e., more variability in strategy) within a group, without necessarily implying an increase in strategy variability *across* participants in the group—because much of the variability was the result of one participant reporting a variety of strategies.

Second, some categories, such as majority responding, were clearly most adaptive when many participants in a group employed them, whereas others, such as compensating for others' adjusting (category J), were most adaptive when few participants in a group employed them. This fact necessarily inflates redundancy within a group, without necessarily diminishing the importance of some variability (e.g., the one participant compensating for others) across participants in that group.

Finally, the redundancy statistics could only be calculated for those strategies participants actually reported. Some strategies, such as recalling and repeating previous responses, are easily verbalized by participants. Others, such as a Complex Threshold (category N), were perhaps more difficult for participants to verbalize even if they made use of such a strategy. Moreover, participants may not have explicitly recognized their use of certain strategies, even if their behaviour implicitly suggested their use. For example, participants might have adjusted to extreme mismatches (category I) without appreciating that this was a strategy, instead considering it a natural response to the current situation. In light of these possibilities, self-reported strategies should not be taken as evidence in isolation. Instead, the reports can be complemented by examining actual participant behaviour in the group function learning task, as recorded by the *i<clicker* software.

Individual Response Patterns

On every attempt at a target value, participants' responses were recorded by the *i<clicker* software and saved to a .csv file. These data complement the self-report data because they not only provide insight into what participants were doing throughout the experiment and how their behaviour changed over training, but also because they allow us to compare self-reported strategies with actual behaviour.

With hundreds of individual participants responding to three blocks of training on four target values, each with up to 19 possible attempts, the amount of data involved is enormous. Furthermore, it is not immediately clear how to quantify behaviour over every attempt, because each is dependent upon the pattern of responding from previous attempts and previous blocks. Thus, the approach taken here was to use the self-reported strategies as a benchmark for analyzing specific individual behaviour patterns. I look specifically at the three most commonly reported strategies, which are easily quantifiable: Recall and Repeat (category D), Let Others Adjust (category E), and Majority responding (category P).

Recall and repeat. The most frequently self-reported strategy in all Experiments was recalling and repeating previous responses to training targets (28.5% of participants overall). If every participant in a group employed this strategy during training, performance would be optimized, a fact that was articulated several times in participants' self-reports. Indeed, it may be that participants reported this strategy because they recognized it as the optimal solution in an optimal group, not because they actually employed it throughout training.

In order to assess the degree to which this strategy was followed, I examined whether the pattern of responding for the four training target values was consistent between the final response (i.e., the response given when the target was matched) in one block and the initial response (i.e., the response given on the first attempt at a target) in the subsequent block. For each participant in a group, I calculated the proportion of times (ranging from 0 to 1) the responses matched between blocks (e.g., final response in block 1 vs. initial response in block 2; final response in block 2 vs. initial response in block 3).

Thus, for Experiment 1, the proportions were calculated as the number of matches observed divided by the total number of matches possible (i.e., 2 comparisons x 4 target values = 8 possible matches). Overall, those participants who reported using the recall and repeat strategy followed this strategy more often (*M* = 0.87, *SD* = 0.15) than those participants who did not report using this strategy (*M* = 0.77, *SD* = 0.21), *t*(198) = 3.44, *p* = .001.

For Experiment 2, the proportions were calculated in the same way as Experiment 1. Responses during the test phase were excluded from this analysis, because a separate strategy category (category X, Maintain for old) was assigned when participants indicated repeating responses from training during the test phase. Thus, proportions were once again calculated as the number of matches observed divided by the total number of matches possible (i.e., 2 comparisons x 4 target values = 8 possible matches). Overall, those participants who reported using the recall and repeat strategy followed this strategy more often (*M* = 0.76, *SD* = 0.21) than those participants who did not report using this strategy (*M* = 0.75, *SD* = 0.20), but this difference was not reliable, *t*(318) = 0.79, *p* = .43. Interestingly, participants in Experiment 2 followed this strategy less often than those in Experiment 1; this was true whether or not they reported using the strategy, although the decrease was more precipitous for those who did report it. Despite this decrease, a comparable percentage of participants reported this strategy in Experiments 1 and 2, 27% and 27.5% respectively. It could be that the larger group sizes of Experiment 2 made this particular strategy less successful, thus preventing participants from following it even when they believed overall that they were. If this is the case, we should expect to see a return to greater usage of the strategy in the smaller group sizes of Experiment 3.

For Experiment 3, behaviour during the relearning phase was included in the calculation of proportions, because participants believed it to be a fourth block of training. Thus, three comparisons were made for each target value. However, the nature of the relearning phase meant that proportions had to be calculated in slightly different ways for each condition, because the number of matches possible varied. In the Small condition, the number of matches observed was divided by 3 comparisons x 4 target values = 12 possible matches. In the Medium condition, the number of matches observed was divided by 3 comparisons x 4 target values – 1 comparison (i.e., because the target value 9 did not reappear in the relearning phase) = 11 possible matches. In the Large condition, the number of matches observed was divided by 3 comparisons x 4 target values – 2 comparisons (i.e., because the target values 7 and 9 did not reappear in the relearning phase) = 10 possible matches. Collapsed across conditions, those participants who reported using the recall and repeat strategy followed this strategy more often (M = 0.86, SD = 0.16) than those participants who did not report using this strategy (M = 0.79, SD = 0.18), t(298) = 3.27, p = .001. As hypothesized above, participants in Experiment 3 used this strategy more often than those in Experiment 2, whether or not they reported using it. This suggests that the recall and repeat strategy may not be as adaptive in larger group sizes, though the reasons why this might be so are not immediately clear.

Finally, collapsing across Experiments, those participants who reported using the strategy did in fact follow it more often (M = 0.82, SD = 0.18) than those who did not report using it (M = 0.76, SD = 0.20), t(818) = 4.07, p < .001. Notably, however, even those participants who did not report using this strategy nevertheless behaved as though they did 76.5% of the time on average. This finding suggests that the self-report data may be underestimating the extent to which participants did recall and repeat their responses on consecutive blocks of training. It is not clear from the data whether the non-reporting participants were aware they were using this strategy but simply failed to report it, or whether they were unaware that their memory for their responses in previous blocks was influencing their behaviour.

Let others adjust. The next most common strategy reported was to maintain one's response when the group failed to match the target, under the expectation that other participants would alter their responses (23% of participants overall). Using this strategy has the effect of minimizing the degree of variability between attempts, although too many participants in a group employing this strategy would lead to inflexibility that may hinder target-matching success. In order to assess the degree to which this strategy was followed, I examined whether participants' responses remained consistent between the initial and final response of each block, for each target value. For each participant in a group, I calculated the proportion of times the initial response given to a target value matched the final response given to the same target value in the same block.

For Experiment 1, this proportion was calculated as the number of matches between initial and final responses, divided by the total number of possible matches (i.e., 4 target values x 3 blocks = 12 possible matches). Overall, those participants who reported letting others adjust followed the strategy more often ($M = 0.87$, $SD = 0.14$) than those participants who did not report this strategy ($M = 0.83$, $SD = 0.16$), but this difference was not reliable, $t(198) = 1.51$, $p = .13$.

For Experiment 2, proportions for this strategy were calculated identically to Experiment 1. Overall, those participants who reported letting others adjust followed the strategy more often ($M = 0.83$, $SD = 0.15$) than those participants who did not report this strategy ($M = 0.81$, $SD = 0.15$), but this difference was not reliable, $t(318) = 0.97$, $p = .33$.

For Experiment 3, behaviour during the relearning phase was once again included in the calculation of proportions, leading to four comparisons for each target value. However, the nature of the relearning phase meant that proportions had to be calculated in slightly different ways for each condition, because the number of matches possible varied. In the Small condition, the number of matches observed was divided by 4 comparisons x 4 target values = 16 possible matches. In the Medium condition, the number of matches observed was divided by 4 comparisons x 4 target values – 1 comparison (i.e., because the target value 9 did not reappear in the relearning phase) = 15 possible matches. In the Large condition, the number of matches observed was divided by 4 comparisons x 4 target values – 2 comparisons (i.e., because the target values 7 and 9 did not reappear in the relearning phase) = 14 possible matches. Collapsed across conditions, those participants who reported using the recall and repeat strategy followed this strategy more often (M = 0.86, SD = 0.12) than those participants who did not report using this strategy (M = 0.85, SD = 0.12), but this difference was not reliable, $t(298)$ = 0.68, p = .50.

Finally, collapsing across Experiments, those participants who reported using the strategy did in fact follow it more often (M = 0.86, SD = 0.16) than those who did not report using it (M = 0.83, SD = 0.14), but this difference too fell short of statistical reliability, $t(818)$ = 1.80, p = .07. Not even the extremely large sample size, when collapsed over all Experiments, could resolve the difference between those participants who reported maintaining their responses and those who did not. Indeed, it is remarkable that participants who did not report this strategy nevertheless behaved as though they were following it 82.7% of the time. Regardless of their self-reports, participants were overall unlikely to alter their responses between the first attempt and the final attempt at a target[12].

As was concluded for the recall and repeat strategy, the results here suggest that the self-report data may be underestimating the extent to which participants maintained their responses when the group failed to match targets. Those who did report this strategy often described it explicitly as a strategic choice. For example, one participant in Experiment 1 wrote:

> Throughout the experiment, my general strategy was to stick with the response I initially chose, as I knew if the response was over/under the target, someone else would very likely change their response and if I did as well every time, it would take longer to get the desired target.

It may be that most participants simply did not see their unwillingness to alter their responses as a strategy, per se, and thus did not report it as such. Despite this, having most participants maintain their responses most of the time is clearly beneficial to a group's success. If too many participants were to alter their responses too often, the group would vary too liberally around a target and would be less likely to match it. Yet recall that mean squared error in block 3, averaged over all Experiments, was best predicted by a model that included three different strategies for adjusting: Adjust After One (category A), Adjust After Two (category B), and Adjust After Several (category C). Despite that fewer participants reported these strategies (5.13%, 2.37%, and 5.37%, respectively) than the Let Others

---

[12] It should be noted that the calculation of proportions included those instances where groups matched a target on their first attempt; in such cases, the initial and final responses were necessarily identical. Although these cases represented 23.8% of all possible matches over all Experiments, proportions would be similarly inflated for those participants who reported this strategy and those participants who did not and thus would not influence the finding of no difference.

Adjust strategy (23%), the willingness of some participants to do so was critical for a group's success. Evidently a balanced combination, wherein most participants let others adjust most of the time while a minority opted to adjust as needed, was a sensible approach to the problem of group function learning.

Majority responding. The third most common strategy reported was to respond with A when the target value was more than half of the group size, and to respond with E when the target value was less than half of the group size. This strategy is wise from a probabilistic standpoint but, as with the previous strategy discussed (i.e., let others adjust), success could be hindered if too many participants within a group adopted it. In order to assess the degree to which this strategy was followed, I examined the proportion of times participants responded A to those target values smaller than half the group size (i.e., 3 in Experiments 1 and 3; 4 and 7 in Experiment 2) and E to those target values larger than half the group size (i.e., 7 and 9 in Experiments 1 and 3; 10 and 13 in Experiment 2). For each participant, distinct proportions were calculated for initial responses (i.e., the response given on first attempt at a target per block) and final responses (i.e., the response given when the target was matched per block). However, as might be expected from the analysis of the Let Others Adjust category above, initial and final proportions were very strongly correlated both for participants who reported the majority strategy, $r = .80$, $t(169) = 17.11$, $p < .001$, and those who did not, $r = .75$, $t(647) = 29.23$, $p < .001$. As such, I report only the results for final responses below.

Thus, for Experiment 1, the proportions were calculated as the number of times a given participant responded *A* to the low target value (i.e., 3)[13], plus the number of times he or she responded *E* to the higher target values (i.e., 7 and 9), divided by the total number of possible instances to abide by this rule (i.e., 3 target values x 3 blocks = 9 possible instances). Overall, those participants who reported using this strategy followed it more often (M = 0.82, SD = 0.23) than those who did not report the strategy (M = 0.67, SD = 0.20), $t(198) = 3.84$, $p < .001$.

For the training phase of Experiment 2, proportions were calculated as the number of times a given participant responded *A* to the low training target values (i.e., 4 and 7), plus the number of times he or she responded *E* to the higher training target values (i.e., 10 and 13), divided by the total number of possible instances to abide by this rule (i.e., 4 target values x 3 blocks = 12 possible instances). Furthermore, proportions were calculated for the test phase as the number of times a given participant responded *A* to the low test target values (i.e., 1 through 7), plus the number of times he or she responded *E* to the higher test target values (i.e., 9 through 15)[14], divided by the total number of possible instances to abide by this rule (i.e., 14 test target values). Overall, those participants who responding in the majority did so more often during training (M = 0.69, SD = 0.19) than those who did not report the strategy (M = 0.62, SD = 0.18), $t(318) = 3.05$, $p = .002$, and more often during test (M = 0.76, SD = 0.16) than those who did not report the strategy (M = 0.66, SD = 0.18), $t(318) = 4.64$, $p < .001$.

---

[13] The target value 5 was excluded because it is the exact halfway point of the group size, and thus does not count as either a majority or minority target.

[14] The target value 8 was excluded because it is the exact halfway point of the group size, and thus does not count as either a majority or minority target.

For Experiment 3, behaviour during the relearning phase was once again included in the calculation of proportions. As in Experiment 1, the numerator for the proportions was calculated as the number of times a given participant responded A to the low target value (i.e., 3), plus the number of times he or she responded E to the higher target values (i.e., 7 and 9). However, the nature of the relearning phase meant that denominator for the proportions differed slightly across conditions. In the Small condition, the instances of majority responding observed was divided by 4 blocks x 3 target values = 12 possible instances. In the Medium condition, instances of majority responding observed was divided by 4 blocks x 3 target values – 1 instance (i.e., because the target value 9 did not reappear in the relearning phase) = 11 possible instances. In the Large condition, the instances of majority responding observed was divided by 4 blocks x 3 target values – 2 instances (i.e., because the target values 7 and 9 did not reappear in the relearning phase) = 10 possible instances. Collapsed across conditions, those participants who reported responding with the majority in fact did so more often (M = 0.85, SD = 0.14) than those participants who did not report using this strategy (M = 0.73, SD = 0.18), $t(198) = 5.15$, $p < .001$.

Finally, collapsing across Experiments, those participants who reported responding with the majority followed this strategy more often (M = 0.77, SD = 0.19) than those who did not report using it (M = 0.67, SD = 0.19), $t(818) = 6.30$, $p < .001$. It is notable, though not surprising, that participants who did not report responding with the majority nevertheless did so 67.1% of the time. This finding suggests that the self-report data may be underestimating the extent to which participants’ responses were influenced by the size of the target relative to the size of the group. It may be the case that most participants didn’t see majority responding as a strategy, per se, but rather as a natural reaction to the probability structure of the task—though only 2.87% of participants on average used the keywords “probability,” “odds,” or “chance” in their self-reports. Recall, moreover, that use of the Majority response strategy emerged as a reliable predictor of mean squared error performance in block 3, collapsed across experiments. Indeed, in the first step of the stepwise regression analysis, the frequency with which this strategy was reported predicted 20.5% of variance in block 3 performance by itself. Intuitively, this makes sense: as more participants respond A to high target values and E to low target values, the group is more likely to match the target, particularly as the target value approaches the extremes of the target range.

Summary of individual response patterns. To summarize, I examined the extent to which participants’ self-reported strategies matched their behaviour, focusing on the three most commonly reported strategies: Recall and Repeat (category D), Let Others Adjust (category E), and Majority responding (category P). Overall, participants who reported recalling and repeating previous responses (category D) and responding with the majority (category P) were more likely than their non-reporting peers to behave consistently with these strategies. This was not the case for the Let Others Adjust strategy (category E), with which behaviour was generally consistent even among those participants who did not report it as a strategy. Notably, participants in Experiment 2 showed lower levels of adherence to the Recall and Repeat strategy (category D), whether they reported using it or not, suggesting that recalling and repeating prior responses was less adaptive in larger group sizes. It is not immediately clear why this would be the case; at the group level, training behaviour in the larger groups of Experiment 2 (Chapter 3) was largely consistent with behaviour in the smaller groups of Experiments 1 (Chapter 2) and 3 (Chapter 4).

Future research should more closely examine the effects of group size on both group- and individual-level behaviour in group function learning tasks.

Personality Variables

As briefly described in Chapters 3 and 4, personality variables were measured for participants in the Control condition of Experiment 2 and all conditions of Experiment 3[15] using the Big Five Inventory (BFI), a brief personality inventory measuring the "big five" traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism (John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008). Descriptive statistics of the BFI results from all four conditions are given in Table 5.3. As shown, mean scores on all five personality variables were generally stable across conditions.

Table 5.3. *Mean (SD) BFI Results for Experiment 2 (Control Condition) and Experiment 3 (All Conditions).*

	Experiment 2	Experiment 3			
	Control	Small	Medium	Large	Average
Openness	3.48 (*0.55*)	3.54 (*0.55*)	3.52 (*0.48*)	3.48 (*0.55*)	**3.51**
Conscientiousness	3.64 (*0.64*)	3.50 (*0.66*)	3.59 (*0.58*)	3.45 (*0.65*)	**3.55**
Extroversion	3.16 (*0.71*)	3.27 (*0.78*)	3.32 (*0.80*)	3.27 (*0.78*)	**3.25**
Agreeableness	3.83 (*0.59*)	3.93 (*0.53*)	3.90 (*0.58*)	3.85 (*0.55*)	**3.88**
Neuroticism	2.94 (*0.77*)	2.97 (*0.76*)	2.81 (*0.84*)	2.97 (*0.73*)	**2.92**

Note. All personality variables are scored out of a maximum of 5.

The relationship between personality variables and behaviour in the group function learning task can be examined either (1) at the group level, by examining whether performance during training or relearning was correlated with the spread of personality

---

[15] Personality data were unable to be collected for one group in the Small condition of Experiment 3. As such, this group is excluded from the analyses below.

variables within a group, or (2) at the individual level, by correlating individuals' scores on the five personality variables with their self-reported strategies.

Group-level analyses. Much like the analysis of self-reported strategies presented above, this approach to the question examines the relationship between group performance and personality variability within groups. However, with each participant having continuous scores on each of five personality variables, the information theory approach described above was not appropriate for this question.

Instead, I borrowed a technique from mathematics that examines the spread of values within a matrix of numbers. For each group in Experiment 3, the matrix of personality variables by individual participants (i.e., a 5 x 10 matrix) can be conceived of as a high dimensional space. In this space, each participant is represented by a vector that defines his or her position on the five personality dimensions. In order to simplify this high dimensional space, each group's matrix was converted into a Euclidian distance matrix using the R function *dist*. Classical multidimensional scaling (MDS; Mardia, 1978) was then applied to each group's distance matrix using the R function *cmdscale.* Both R functions used can be found in the default R package *stats*. Multidimensional scaling is a useful tool for examining these data because it not only measures the variance in personality profiles within each group, but also captures the relationships between individuals on all five personality dimensions (i.e., by finding the principal components in each group's "meta-personality"). The MDS procedure returns eigenvalues for each matrix, which capture the major axes of variance in the space. MDS moreover allows us to visualize the personality variable space of a group in two dimensions, as illustrated in Figure 5.1.

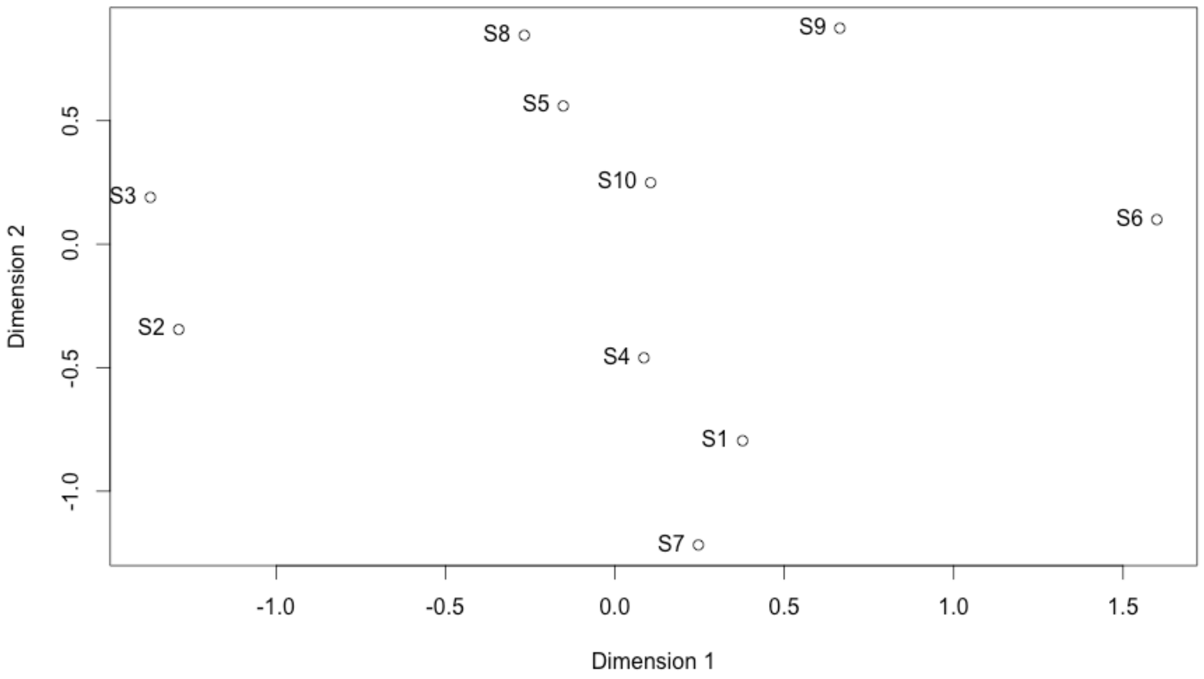

Figure 5.1. A two-dimensional plot of the five-dimensional personality space for a group in the Small condition of Experiment 3, following Multidimensional Scaling (MDS). Each participant's personality profile results in a unique location in the space, relative to the other participants in the group. The two dimensions shown are not meaningful, per se, but instead capture the two most important axes of variability in the matrix.

As shown, each participant in this example group has a unique position in the space relative to the other participants in their group; their relative positions capture the similarity structure between participants' personality profiles. Importantly, a different configuration of personality profiles and a different set of eigenvalues emerge for each group. A question of interest is whether the degree of within-group variability in personality profiles was related to performance on the function learning task. One hypothesis is that a more diverse set of pre-existing personality traits would support a naturally occurring (i.e., without communication or planning) division of behavioural roles within a group. An alternative hypothesis is that groups with converging personality profiles would operate more cohesively. A useful measure of within-group variability is the spread of the matrix, defined as the largest absolute distance between any two eigenvalues of a matrix (Mirsky, 1956). Spread was calculated for each group and subsequently entered into a correlational analysis with metrics of performance in the function learning task.

For the Control condition of Experiment 2, spread was not correlated with overall mean squared error at test, $r = -.33$, $p = .35$, nor with performance on any subset of test items, maximum $r = -.38$, $p = .28$ for extrapolated test items. Collapsing across conditions in Experiment 3, spread was not correlated with any measure of performance (i.e., number of trials taken to match target, mean squared error averaged over all attempts in a block, and mean squared error averaged over first attempts in a block) either during training, maximum $r = .35$, $p = .06$ for mean squared error averaged over all attempts in block 1, or during relearning, maximum $r = .14$, $p = .47$ for mean squared error averaged over all attempts. Moreover, the spread of personality variables was not significantly correlated with the redundancy of strategies self-reported, $r = -.22$, $p = .24$, suggesting that variability in personality variables was not associated with variability in strategies.

Despite the statistical non-significance of the correlations in Experiment 3, it is worth noting that the relationships between a group's spread of personalities and its error were almost uniformly positive, whereas the relationship between spread and redundancy was negative. Recall that redundancy was negatively associated with error in block 3. Taken together, these results hint at a coherent pattern of relationships: groups with more variability in personality among participants reported less redundant (i.e., more varied) strategies, and showed worse performance (i.e., more error) on the group function learning task. Of course, this pattern must be considered with caution until a replication with larger sample sizes can resolve the statistical uncertainties in some of the relationships. Research in personality psychology typically relies on very large sample sizes to pull out modest correlations between personality and behaviour. Fraley & Vazire (2014), for instance, showed that the average sample sizes in social and personality psychology is 104 participants; in order to achieve such a sample size in a group learning task, I would require at least 104 x 10 = 1040 participants.

Individual-level analyses. Participants were assigned a value of 1 for each strategy category that they reported, and a value of 0 for each strategy category they did not report. Subsequently, participants' scores on each of the five personality dimensions were entered into a correlation analysis with the binary scores for strategy use, collapsing across groups and experiments.

Openness was positively correlated with Social Decisions (category S), $r = .12$, $p = .03$, and negatively correlated with Recall and Repeat (category D), $r = -.12$, $p = .04$. Thus, according to self-report, more open participants were more likely to make decisions based on some social aspect of other participants in the room, but less likely to recall and repeat their previous responses to target values. Conscientiousness, by contrast, was not significantly associated with any self-reported strategy categories.

Extroversion was negatively correlated with Specific Adjust (category G), $r = -.20$, $p < .001$, Single Threshold (category M), $r = -.15$, $r = .01$, and Majority responding (category P), $r = -.22$, $p < .001$, and positively correlated with Respond Tendency (category K), $r = .15$, $p = .01$. Thus, according to self-reports, more extroverted participants tended to respond more often, while less extroverted participants more often adjusted when the group failed to match specific target values, more often used a single threshold for response decisions, and more often responded with the majority.

Agreeableness was associated only with Random Responding (category U), $r = .12$, $p = .04$, suggesting that more agreeable participants reported responding randomly more often than less agreeable participants. Finally, neuroticism was positively correlated with Majority responding (category P), $r = .15$, $p = .01$, and negatively correlated with Recall and Repeat (category D), $r = -.12$, $p = .047$. According to self-reports, more neurotic participants were therefore more likely to respond with the majority, and less likely to recall and repeat their previous successful responses to targets.

Although the lack of a priori hypotheses about role of personality variables in behaviour during a group function learning task recommends caution when interpreting these results, a few of these findings are generally consistent with what might be expected from the definitions of personality dimensions in the BFI scale[16]. All conceptual definitions are taken from John, Naumann, and Soto's (2008) book on personality.

Openness, for example, is defined as the breadth and depth of originality and complexity of an individual's mental and experiential life. Insofar as the Openness dimension tracks originality, it is perhaps not surprising that individuals scoring high on this dimension eschewed the most commonly reported strategy (Recall and Repeat, 28.5% of participants overall; see Table 5.1) in favour of the more obscure strategy of socially-

---

[16] Note that Bonferroni corrections applied to the correlations, with a family-wise error rate of $\alpha = .05/25 = .002$, suggest that most of the relationships should be considered with caution. Those still reliable after correction are negative correlations between Extroversion and the Specific Adjust strategy, $p < .001$, and Extroversion and the Majority Responding strategy, $p < .001$.

based decisions (category S, 2.1% overall). Moreover, those participants reporting the social decision strategy gave inventive descriptions of their decision processes that did not necessarily map onto reality. For instance, one participant in Experiment 2 wrote**,** **"**I took note of the people in the room who appeared to be more timid and maybe more reluctant to hit *A*."

Extroversion is described as an energetic approach toward the social and material world, including assertiveness and activity. The finding that extroversion was associated with a greater tendency to respond (category K) is consistent with expectations based on this definition; high levels of energy and activity might make participants want to take what they perceived as an active role in the group's success. Likewise, a tendency to not describe their behaviour as responding with the majority (category P), relative to less extroverted participants, is consistent with expectations for individuals wishing to assert themselves.

The finding of no relationship between conscientiousness and strategy self-report cannot be accounted for by the conceptual definition, which describes conscientiousness a socially prescribed impulse control that facilitates goal-directed behaviour. We might have expected, for instance, a tendency for more conscientious individuals to let other participants adjust when the group failed to match a target. Likewise, there was no reason to expect more random responding from more agreeable participants, described conceptually as having a pro-social or communal orientation.

The relationships between neuroticism and self-reported strategies are similarly difficult to interpret in light of the BFI definition of neuroticism, which is associated with negative emotionality including anxiety and sadness. Participants who scored higher on the neuroticism dimension tended to report responding with the majority, and tended not to report recalling and repeating their previous responses to targets; there is no clear reason why negative emotionality would encourage such self-report patterns.

Summary of Individual Behaviour

Ideally behaviour in the group function learning task would have been fully captured by neural network models, in line with the argument for self-similarity across the hierarchical scale from individuals to groups. However, neural network simulations were unsatisfying and demanded consideration of additional and alternative explanations. In order to better understand the behaviour of individual participants in the group function learning task, three different types of data were analyzed. First, I examined the self-reported strategies that participants provided at the end of their experiments. After categorizing written responses, I used information theory to assess whether more or less diversity of reported strategies within a group was associated with better performance in the function learning task. Overall, less diversity, or more redundancy, was associated with better performance. However, a regression analysis found that a subset of four strategy categories could predict almost half of all variance in mean squared error during block 3 of training. This analysis showed that a combination of stability in responding and various forms of flexibility to feedback were most predictive of a group's success.

Next, I examined the extent to which participants' self-reported strategies matched their actual behaviour in the task, focusing on the three most commonly reported strategies: Recall and Repeat (category D), Let Others Adjust (category E), and Majority responding (category P). Overall, participants who reported using Recall and Repeat (category D) and Majority responding (category P) were more likely than their non-reporting peers to behave consistently with these strategies. This was not the case for the

Let Others Adjust strategy (category E), with which behaviour was generally consistent even among those participants who did not report it as a strategy. This sort of approach to the data bears some resemblance to protocol analysis, as described by Newell and Simon (1972), in which participants verbally reported their thinking as they completed some cognitive task.

Finally, I examined the relationship between performance on the group function learning task and individual participants' personality variables, as assessed in Experiment 3 and the Control condition of Experiment 2. A mathematical technique was used to measure, separately for each group, the spread along the major axes of variability on the five dimensions of personality. This measure of spread failed to correlate significantly with any measure of performance on the group function learning task, but the pattern of findings was nevertheless consistent with those of the self-report data: more variability seemed to be associated with poorer performance. Individuals' scores on the five personality dimensions were also correlated with their self-reported strategies. The uncovered relationships were consistent with the conceptual definitions of Extroversion and Openness, but no reasonable interpretation could be found for the relationships between strategy reports and Agreeableness, Conscientiousness, or Neuroticism.

Overall, the results of these analyses paint a complex picture of what participants were doing in order for their groups to succeed in the function learning task. Just as each individual in a more typical cognitive psychology task takes a slightly different approach to the problem, each group in my function learning task was comprised of unique individuals that necessarily brought different approaches to the task. Despite this variability, however, groups succeeded in learning both simple and complex functions. Replication with larger sample sizes will be necessary to resolve some of the statistical uncertainties of analyses at the level of the individual, but the results presented in this chapter provide a rough sketch of how individuals contribute to group-level success in this task.

Sabotaged Groups

As described in the Methods section of Experiment 1, a small proportion of groups in every experiment was discarded due to what I termed sabotage. That is, a single participant could hinder their group's progress by changing his or her response on the trial immediately following a successful match, thereby preventing the group from reaching the criterion of two consecutive target matches. These groups were immediately detected during data collection, and their data were excluded from all analyses reported in the document so far. In Experiment 1, there was one sabotaged group in the Complex condition. In Experiment 2, there were three sabotaged groups in the Complex condition. And in Experiment 3, there was one sabotaged group in the Medium condition. In this section, I describe the performance of these groups and the behaviour of the individuals within them.

Table 5.4 provides a comparison of performance between the sabotaged groups and those analyzed for the experimental results in Chapters 2 through 4. As shown, performance was generally worse on all metrics for groups in which sabotage occurred. Notably, the deficit for sabotaged groups was not restricted to the block(s) in which the sabotage behaviour occurred. In most cases, the sabotage behaviour was restricted to the first block of training; as stated in the Methods section of Experiment 1, a warning was usually sufficient to deter further sabotage behaviours. Nevertheless, performance in sabotaged groups was almost uniformly worse on every block of training than the average performance for non-sabotaged groups in the same experimental conditions.

Although the large differences in sample sizes (e.g., 3 sabotaged groups vs. 20 non-sabotaged groups in Experiment 2) preclude any statistical analysis to confirm this pattern, it suggests that group cohesion was detrimentally affected for the duration of the experiment once sabotage occurred. This finding is interesting in its own right, and supports my decision to exclude these sabotaged groups.

Table 5.4. *Comparison of Mean Scores Between Sabotaged and Analyzed Groups During Training.*

		Sabotaged Groups			Analyzed Groups			Average
		Block 1	Block 2	Block 3	Block 1	Block 2	Block 3	Difference
Exp. 1	*Trials*	16.50	5.75	6.75	6.20	4.38	3.98	**4.82**
	Average Error	2.24	1.51	2.04	2.16	1.32	0.83	**0.49**
	First Error	9.50	6.25	6.75	4.35	3.65	2.40	**4.03**
Exp. 2	*Trials*	11.33	8.33	7.08	6.32	5.32	5.40	**3.23**
	Average Error	3.90	2.14	2.09	2.89	2.11	1.62	**0.50**
	First Error	9.75	4.42	6.50	7.72	5.25	3.75	**1.31**
Exp. 3	*Trials*	10.75	4.25	4.50	6.45	4.42	3.98	**1.54**
	Average Error	1.48	1.41	1.38	1.32	0.85	0.81	**0.43**
	First Error	4.00	5.00	3.25	3.52	2.45	1.72	**1.52**

Notes. For Experiments 1 and 2, all Sabotaged groups were in the complex condition; therefore, comparisons are made against means in the Complex condition only. For Experiment 3, the Sabotaged group was in the Medium condition; therefore, comparisons are made against means in the Medium Condition only.

Despite that learning of the group function was disrupted in sabotaged groups, there was evidence of a different kind of learning not observed in the non-sabotaged groups analyzed in earlier chapters. Specifically, certain individuals in these groups developed the unique role of compensating for the behaviour of the saboteur, attempting to anticipate what the saboteur's response would change to after the group matched the target once and then adjusting their own response to make up the difference. One participant in Experiment 2 explicitly admitted to doing so on their self-report, saying, "during the confirmation part [i.e., the trial immediately following a target match] where someone kept giving the 11th response, I stopped so we made it to 10. It was annoying." Unfortunately, the right balance of sabotage behaviour and compensation behaviour was difficult to achieve. In several groups, more than one person attempted to compensate for the sabotage, thereby negating the good intentions of the compensators. To a compensator, another compensator's good intentions might appear to be more sabotage, further eroding trust and cohesion within the group.

Personality variables were collected for the one sabotaged group in Experiment 3, and the saboteur's scores were examined relative to the average scores given in Table 5.3. One might expect that a saboteur would score low on Agreeableness, for example, since their behaviour betrayed antagonism. Fascinatingly, however, this participant in fact scored higher than would be expected on average, achieving a 4.11 out of 5 as compared to the average 3.88 out of 5. Whether this was a true reflection of their personality or some conscious or unconscious attempt to mask their bad behaviour, the collected data cannot say. It is worth noting, however, that saboteurs were very likely to include in their self-reports some assertion that part of their strategy was to always repeat the same response after the group matched the target once. A more detailed analysis of the social psychological aspects of group cooperation, or lack thereof, is beyond the scope of this dissertation (for a review of the social aspects of cognition, see Levine, Resnick, & Higgins, 1993). Ultimately, the observations presented here serve to reinforce my decision to exclude sabotaged groups from analysis, as the effects of the antisocial act on a group's ability to cooperate and collaboratively solve a problem were many.

Although the small sample sizes of the present experiments make it difficult to rule out chance, it is potentially of interest that sabotaged groups in Experiments 1 and 2 were uniformly in the Complex condition, learning a more difficult function than those in the corresponding Simple condition. Further research should examine whether the difficulty of a group's task has any effect on the rate of saboteur or otherwise uncooperative behaviours, as such findings would undoubtedly carry important implications for real-world group learning situations. As briefly discussed in Chapter 1, for example, the collaborative online videogame experiment "Twitch plays Pokémon" also resulted in incidents of sabotage that the group learned to overcome. If such instances are indeed a universal feature of group learning phenomena, then removing sabotaged groups from my analyses may have limited the ecological validity of the experiments. However, in light of the exploratory nature of the experiments and the high resource costs to run them, I feel that any reduction in external validity was justified by the internal validity gained by controlling extraneous noise from sabotage behaviours.

Chapter 6: General Discussion

The experiments presented in this thesis examined function learning in groups of individuals, exploring both group-level behaviours and the individual-level variables that supported group-level learning. In Chapter 2, I showed that groups could learn both simple and complex functions, though no statistical difference could be detected in the rate of learning between the two levels of function complexity.

In Chapter 3, I demonstrated that larger groups could likewise learn both simple and complex functions. Moreover, learning in both conditions generalized to untrained areas of the functions, both extrapolating beyond the range of trained target values and interpolating between the trained target values, relative to a Control condition that did not undergo any training.

In Chapter 4, I showed that groups in the function learning task were relatively resistant to a small (10%) or medium (30%) amount of disruption following training. Performance was degraded, however, when a large (50%) amount of disruption was applied, mirroring Lashley and Wiley's (1933; see also Lashley, 1950) classic finding of graceful degradation in neurological lesioning.

Chapter 5 presented an exploratory analysis of the relationships between group-level performance in function learning and the traits or behaviours of individuals within the group. Overall, the results showed that better group performance was associated with more homogenous groups, in terms of both self-reported strategies and, to a lesser extent, personality profiles. Success in the group function learning task was nevertheless predicted by a combination of strategies that balanced rigidity (e.g., responding with the majority) with flexibility (e.g., adjusting one's response when the group failed to match a target after one, two, or several attempts). Self-reported strategies tended to reflect actual behaviour within the task, but certain strategies were regularly employed regardless of whether participants reported using them or not. Finally, some relationships existed between a participant's personality scores and her self-reported strategies, particularly with regard to the Extroversion dimension.

Function Learning

In a typical function learning experiment, individual participants learn the relationship between two variables by trial and error: given a target value of X, participants provide an estimate for the associated value of Y before receiving feedback as to the correct Y value. In this manner, participants have been able to learn linear (e.g., Brehmer, 1974), quadratic (e.g., DeLosh, 1995; DeLosh, Busemeyer, & McDaniel, 1997), exponential (e.g., DeLosh et al., 1997), power (e.g., DeLosh et al., 1997), logarithmic (e.g., Koh & Myer, 1991), and cyclic (e.g., Bott & Heit, 2004) relationships.

My demonstrations of group function learning replicate some of the basic results observed in individual function learning tasks. Namely, that groups can learn both simple and complex functions (Experiment 1), and that groups can generalize their learning to areas of the function that were not previously trained (Experiment 2). Like individual learners, groups performed better on new targets that interpolated between trained target values than on new targets that extrapolated beyond trained target values (e.g., DeLosh et al., 1997; Koh & Meyer, 1991; Wagenaar & Sagaria, 1975).

Nevertheless, there were several notable differences between individual and group function learners. First, non-linear functions are generally more difficult for individuals to learn than linear functions (e.g., Deane, Hammond, & Summers; DeLosh et al., 1997; Koh &

Meyer, 1991; McDaniel & Busemeyer, 2005). Although this result replicated numerically in both Experiments 1 and 2, it was not supported statistically. Aside from issues of sample size, which will be considered later, these results may reflect a difference between a function's statistical complexity and its psychological complexity. As Koh and Myer (1991) pointed out, most demonstrations comparing performance on linear and nonlinear functions could be re-described as comparing monotonic and non-monotonic functions; several demonstrations have shown that non-monotonic functions (e.g., quadratic) are more difficult to learn than monotonic ones (e.g., DeLosh, 1995; DeLosh et al., 1997). Although statistically complex, the cubic-like function that defines my Complex condition may not carry the same psychological complexity as, for instance, a less statistically complex but non-monotonic quadratic function. Moreover, DeLosh (1995) showed that a systematic presentation of targets from small to large (as in my experiments) facilitated the learning of difficult functions, compared to a randomly ordered training sequence. In my experiments, learning in the Complex condition may have been aided by the order of the presentation of target values.

Second, whereas participants in an individual function learning task tend to underestimate the value of the Y variable when extrapolating a linear function (e.g., Brown & Lacroix, 2017; DeLosh, 1995; DeLosh et al., 1997; Kwantes & Neal, 2006), groups in all conditions of my function learning task overestimated when extrapolating at the low end and underestimated when extrapolating at the high end. This pattern was especially true of the Control condition, which did not receive any training. That result suggests a natural response bias in groups that is mitigated but not eliminated by training on functions (see Kalish, Griffiths, & Lewandowsky, 2007, for a different bias in individual function learners).

Interestingly, the pattern of extrapolation performance observed in Experiment 2 is consistent with the predictions of an associative learning model (ALM) of function. DeLosh, Busemeyer, and McDaniel (1997) took this prediction as a failure of the ALM model. In contrast, I took the match between their model and my data as evidence that groups in my function learning task may mirror some of the properties of artificial neural networks.

Third, whereas extrapolation performance was notably worse than interpolation performance in the Complex and Control conditions of Experiment 2, there was no difference between the two in the Simple condition. This lies in contrast to the individual function learning literature, where extrapolation performance is always found to be worse than interpolation performance (see Busemeyer et al., 1997). Indeed, a number of the within-group differences anticipated in Experiment 2 did not reach statistical reliability. This was most likely due to small sample sizes, a fact that will be discussed later, but it must nevertheless be acknowledged. Instead, my argument for generalization of learning relied upon comparisons between the experimental conditions and their respective controls. To my knowledge, no examples of individual function learning have included a test-only control condition, making it difficult to make direct comparisons to the individual learning literature for those conclusions.

Considerations

Differences in behaviour. These differences in behaviour between groups and individuals recommend some caution when interpreting the results of the experiments from the perspective of scale invariant processes and self-similar behaviours. Although I would like to argue that function learning operates invariantly across different levels of a hierarchy, the data warrant a more conservative conclusion.

There are indeed ways in which groups of individuals in a function learning task behave like individuals in a function learning task. Given analogous methodologies, approximately analogous behaviours emerged: groups learned both simple and complex functions, generalized that learning to new areas of the function relative to an untrained Control condition, and showed resistance to disruption of their learning. However, as discussed above, not all details of the results replicated when scaling from individuals to groups. Moreover, artificial neural network models were unable to capture a complete picture of the individual-level responding that supported group-level learning.

The differences are nevertheless instructive. Despite the general consistency when scaling a function learning task from individuals to groups, individuals within the group are not directly comparable to neurons in a brain. Individuals in the group task are not limited to the to basic learning behaviours encapsulated by artificial neural network models, but instead incorporate more complex decision processes: as shown in Chapter 5, participants' decisions were also guided by memory (i.e., recalling and repeating previous responses to targets) and simple heuristics (e.g., responding with the majority, letting others adjust when the group mismatched a target). It would therefore be misleading to suggest that individuals within a group are directly comparable to neurons or nodes within a network, despite some similarities in behaviour.

Differences in methodology. The differences in behaviour may also be attributable to differences in the way the task was conducted, between individuals and groups. While some changes in methodology were an unavoidable product of the way group behaviour is measured and collected, other changes could, and in retrospect should, have been avoided. I hereafter refer to the experiments conducted by DeLosh, McDaniel, & Busemeyer (1997) as an illustration of how individual function learning tasks are typically run. In their experiments, for example, training targets are randomly ordered within a block. By contrast, in the current experiments trials were sequentially ordered (e.g., 3, 5, 7, 9) on every block. DeLosh (1995) showed that the learning of difficult functions was facilitated by systematically ordering training items. Assuming that functions would be more difficult for groups to learn than for individuals to learn, I decided to take advantage of this facilitation. In retrospect, however, this systematic ordering may have changed the nature of the task and minimized the anticipated differences between the Simple and Complex conditions.

Second, my experiments required that groups matched a given target before progressing through a block of training. In individual function learning, successful matching is not a requirement during training. Once again, this decision was made to simplify the task for groups, but may have altered the nature of the task. Moreover, it may have introduced total number of training trials as a potential confound. Whereas DeLosh et al. (1997) provided all participants with 200 training trials, learning either 8 (over 25 blocks), 20 (over 10 blocks), or 50 (over 4 blocks) unique targets, groups in my experiments learned 4 unique targets over 3 blocks of training with a variable number of total trials. Groups in my Experiments 1 and 3 underwent an average of 53.6 training trials, while those in Experiment 2 underwent an average of 70.5 training trials, both of which fall quite short of the standard 200 trials in individual function learning. This shortened training may have limited learning and impeded performance on the transfer test phase; it is possible, for example, that within-group differences between old items, interpolated items, and extrapolated items, may have been more pronounced had training been more extensive.

Future research into group function learning ought to more closely equate methodological details to those of individual function learning tasks, in order to minimize opportunities for divergent behaviours to emerge. Whereas it is presently unclear the extent to which differences in function learning behaviours are the result of differing mechanisms or differing methodologies, the differences nevertheless limit the force of my argument for self-similarity across a hierarchical scale, from individuals to groups.

Limitations of analogy. Lashley (1950) famously summarized his search for the engram with the conclusion that the many parts of a neural network are not assigned specialized and localized representations, but are rather flexible parts of an integral whole across which representations are distributed. Likewise, the results of my group function learning task suggests a representation of the function that is distributed across the individuals in a group. Throughout this thesis, I have drawn analogies between neural systems and groups in the function learning task. In both cases, behaviour at the larger scale of analysis can be described as an emergent outcome of distributed knowledge at the smaller scale (Goldstone & Gureckis, 2009; Moussaid, Garnier, & Theraulaz, 2009; Roberts & Goldstone, 2011; Theiner, Allen, & Goldstone, 2010). I must, however, acknowledge two limitations of this analogy.

First, debate about whether localized or distributed representations are more biological plausible has endured, unresolved, for decades. The argument for localization of function view has been bolstered by a long history of case studies on subjects with neurological damage, in whom certain functions are spared while others remain intact (e.g., Broca, 1861). Reliance on this dissociation logic is perhaps most pervasive in the study of memory, where functionally and structurally distinct memory systems have been proposed on the basis of selective impairments following excision of or lesion to brain areas (e.g., Patient H.M.; Scoville & Milner, 1957; Patient K.C.; Tulving, Schacter, McLachlan & Moscovitch, 1988). The subtractive method commonly used in functional neuroimaging, whereby blood oxygenation changes during a control task are subtracted from blood oxygenation changes during a task of interest, likewise seems to support the localization of functional representations in the brain (c.f. Poldrack, 2006 for criticisms of this inference).

Despite their popularity, both dissociation logic and the subtractive method have been criticized as encouraging fallacious or otherwise unsupported interpretations (e.g., Chater, 2003; Dunn & Kirsner, 1988, 2003; Farah, 1994; Medler, Dawson, & Kingstone, 2005; Prince, Brown, & Heathcote, 2012; Yeates, Willis, Jones, & McLaren, 2014). Namely, neither single case studies of neurological damage nor fMRI data showing subtracted activation (e.g., *which brain areas are active for function A minus function B?*) can speak to the *sufficiency* of a given area for a function. They can correctly be taken as evidence for *necessity*, but necessity alone is not conclusive evidence for localized representation. By contrast, distributed representations have been uncovered for a wide range of cognitive processes, including visual category recognition in the ventral temporal cortex (Haxby et al., 2001) and visual cortex (Cox & Savoy, 2003), the focus of visual attention at the level of the primary visual cortex (Kamitani & Tong, 2005), representations of actions (Grafton & de C. Hamilton, 2007), gustatory stimuli (Simon, de Araujo, Gutierrez, & Nicolelis, 2006), conceptual processing (Kiefer & Pulvermuller, 2012), and semantic memory storage and retrieval (Martin & Chao, 2001).

Increasingly, researchers have tempered their stances to acknowledge that the brain may employ a mixture of both local and distributed representations (Fair et al., 2009;

Medler et al., 2005; Quiroga & Kreiman, 2010a, 2010b; Quiroga, Kreiman, Koch, & Fried, 2008; see also the CALM model, Happel & Murre, 1994; Murre, Phaf, & Wolters, 1992). On the basis of self-reported strategies (see Chapter 5), an argument may be made the group function learning task also employs both local (i.e., individual strategies) and distributed (i.e., group-level learning) representations.

I take this opportunity to reiterate that I intend the analogy between groups of individuals and neural networks to be taken superficially. The data point to self-similarity of behaviour, not necessarily of process. Individual humans are ostensibly more complex than individual neurons and undoubtedly cannot be represented by synaptic weights, neurotransmitters, and spiking electrical signals. That being said, the constraints of the group function learning task seem to evoke from individuals behaviours that are in many ways consistent with neural networks in more standard function learning tasks. In my view, it would be an error to overlook principles of behaviour because different systems, or scales of analysis within a system, implemented those principles in different ways. Behavioural principles can help to identify and distinguish the different ways in which the behaviours of interest are computed and implemented at the level of mechanism.

Statistical uncertainties. The group function learning task is an expensive experimental protocol to conduct, in terms of both participant requirements and time demands. In order to collect data for the 80 groups across the three experiments presented in this thesis, a total of 980 participants were run. Moreover, arranging for groups of 10 or 16 participants to attend an experiment at the same time required many additional hours of organizational effort. Because of the relative expense of these studies, the sample sizes (n = 10 groups per condition in every experiment) were smaller than the oft-cited ideal of n = 30 per cell, a fact that may have limited my ability to detect some effects. In this sense, groups in this task can be considered alongside more costly experiments with animals or studies of clinical populations, where a practical balance must be struck between the amount of data collected and the costs of collecting those data. In such cases, as in this thesis, broad conclusions drawn from small sample sizes need to be considered with an appropriate amount of caution.

I want to focus particularly on the standard finding in function learning, that non-linear functions are generally more difficult to learn than linear functions (Deane et al., 1972; DeLosh et al., 1997; Koh & Meyer, 1991; McDaniel & Busemeyer, 2005). Although the average mean squared errors in my Simple (linear) and Complex (non-linear) conditions were in line with this expectation in both Experiments 1 and 2, the anticipated block by condition interaction could not be statistically supported in either experiment. As discussed in Chapter 2, a total of 26 groups (520 participants) would have been needed in each condition to detect this interaction in Experiment 1 ($\eta_p^2 = .03$) at a power of $1\text{-}\beta = .8$. To detect the interaction at the same power in Experiment 2 ($\eta_p^2 = .04$), a total of 22 groups (352 participants) would have been needed in each condition. Thus, had I followed the guidelines of $n = 30$ observations per cell, power analyses suggest I would have been likely to find the interaction effects in both Experiments—assuming those differences are in fact true.

Nevertheless, the statistical uncertainty of this result may also have been due to the monotonic nature of the non-linear function chosen. It is possible that a cubic-like function, although mathematically complex, lacked the psychological complexity necessary to draw out the differences between conditions. I anticipate that, like individuals, a quadratic function would have proven more difficult for groups to learn than a linear function (DeLosh et al., 1997; Lucas, Griffiths, Williams, & Kalish, 2015). In retrospect, it would have been a better strategy to examine groups' ability to learn the psychologically more complex quadratic function than the mathematically more complex cubic function.

Future Work

Variations on group function learning. The work presented in this thesis constitutes first steps in an exploration of the group function learning task. Although I have shown that groups of individuals can learn both simple and complex functions (Experiment 1), that groups can generalize their knowledge to untrained areas of the function in a subsequent transfer test (Experiment 2), and that groups are relatively resistant to the detrimental effects of disruption to their knowledge (Experiment 3), many questions remain unanswered.

First, the sizes of the groups studied in these experiments represent a fairly restricted range. Future work should examine whether the results would scale down to smaller groups, or scale up to larger groups. This question is particularly pertinent in light of my theoretical focus on scale invariance. Briefly described in Chapter 1, Roberts and Goldstone (2011) provided a comparison of small (2-3 participants), medium (4-9 participants) and large (10 or more participants) groups in a Group Binary Search (GBS) game. In the game, groups of participants had to match some unknown target number between 51 and 100 by each choosing a response between 0 and 50. The program summed the responses of the participants and gave group-level feedback relative to the target. Small groups learned to match targets faster than medium groups, which in turn learned faster than large groups. They moreover showed that decreases in individual reactivity to feedback and higher within-group variations in reactivity were both predictive of better performance in large groups, but not in small groups. The group function learning task would likewise benefit from comparisons between different group sizes (see Individual Response Patterns in Chapter 5 for an example of possible differences as a function of group size).

Moving beyond the range of group sizes studied by either myself or Roberts and Goldstone would likewise be a useful avenue of future research, because artificial neural networks are not known to scale up reliably (Feldman-Stewart & Mewhort, 1994; c.f. Elman, 2005; Ghahramani, Korenberg, & Hinton, 1999; see also recent efforts with extremely large-scale networks: Markram, 2006; Jorgenson et al., 2015; Eliasmith et al., 2012). To the extent that groups in this task mirror the behaviour of artificial networks, similar limitations may be observed in larger groups.

A second branch of enquiry for future exploration involves the degree of communication permitted between group members. In my function learning task, individuals could only communicate their decisions indirectly, through the group-level feedback provided after every attempt to match a target. It is of interest to know whether additional communication between individuals would help or hinder group-level performance. Miller and Steyvers (2011), for example, found that knowledge of another participant’s response series in an order reconstruction task improved the accuracy of the group’s aggregate response series. In a point estimation task, Mason, Jones, and Goldstone (2008) had participants try to find the highest point in an unknown multimodal distribution. On every attempt, participants guessed at the location of the highest point on the X-axis. Participants received feedback not only regarding their guesses, but also the guesses made by the other participants with whom they were connected in a network. By varying the degree of connectivity in the networks, they showed that individuals' convergence on the actual highest point was fastest in in networks that were fully connected locally (i.e., each participant was connected to the neighbour on either side of him or her) with a small number of non-local connections. Networks that were entirely connected (i.e., every participant was connected to every other participant), networks with a small number of random connections, and networks with strictly local connections, all took longer to converge—particularly on more difficult problems (see also Wisdom, Song, & Goldstone, 2013).

A related question is whether existing social relationships between individuals in a group would impact performance. Brennan and Enns (2015) found facilitative effects for both communication between individuals and existing social affiliations in a collaborative pair task, suggesting that two heads were better than one for reasons much more complex than mere statistics. In light of my finding that group success in the function learning task was associated with lower amounts of variability in personality profiles and higher amounts of redundancy in self-reported strategy use (see Chapter 5), it is reasonable to hypothesize that affiliated groups in the function learning task would show a similar advantage over unaffiliated groups.

An alternate approach. Because, as I argued earlier, the time and resource costs of running group my group function learning experiments are high, it is worth considering a different experimental approach to collaborative learning. One alternative that I am currently exploring makes use of Arduino, an inexpensive and open-sourced platform for prototyping on simple electronics hardware (i.e., single-board microcontrollers). Arduino boards can be flexibly configured with a variety of input sensors and output interfaces to serve as an ideal engineering solution for unique experimental designs.

To simplify the group function learning task, I might create a hardware device that takes in pressure-sensitive touch input from a small group of individuals (e.g., 3 or 4 participants) and performs some output function on the aggregate to power a coloured LED lamp. Groups would be tasked with matching the colour and intensity of a target LED lamp by estimating the appropriate amount of pressure to apply. After every attempt, the group would receive feedback comparing the LED output they created to the target LED lamp. In a group of three participants, one participant's input might determine the Red value, one might determine the Blue value, and one might determine the Green value, thereby recreating a target LED output according to the RGB colour model. A group of four participants might instead recreate a target LED output according to the CMYK colour model (i.e., cyan, magenta, yellow, black). Groups would have to learn the function that transforms their aggregate input into an LED output without communication. Critically, it is easy to generate a machine that translates a person's pressure response according to any mathematical function—thus allowing the experiment to measure learning in the group as a function of the input-to-output relationship specified in the machine. Unlike my current function learning task, where an individual's decision is binary and the feedback is one-dimensional (e.g., the group was over, under, or matched the target), the continuous nature of pressure-sensitive input would limit the degree to which a simple conscious strategy like those reported in Chapter 5 will be useful. Thus, it may prove to be a purer method of examining emergence in group cognition.

Modeling group cognition. In order to better understand how groups of people learn in the function learning task, future research should focus on computational modeling of the empirical results presented in this thesis. As discussed in Chapter 5, my original goal was to show that artificial neural network models could capture behaviour in both individual and group function learning tasks. To that end, I ran simulations with two ANN models: a basic single layer feedforward network (e.g., Rosenblatt 1958), and the ALM model (Knapp & Anderson, 1984), which was designed to explain findings from the individual function learning literature. However, both ANN models proved insufficient to capture to details of how groups were learning and how individual decisions were contributing to that learning. In light of the shortcomings of network models, I briefly outline alternative approaches from two different traditions below: (1) Minerva-2 (Hintzman, 1984, 1986), an instance model originally developed to explain episodic memory, and (2) agent-based models (e.g., Epstein & Axtell, 1996; Goldstone & Janssen, 2005), which are increasingly used in cognitive science to model emergent behaviours.

Minerva-2. The ALM approach bears some similarities to the Minerva-2 model (Hintzman, 1984, 1986), an instance model of episodic memory that has been successfully adapted to a great many problems, including decision-making (Dougherty, Gettys, & Ogden, 1999; Neil Bearden & Wallsten, 2004), implicit memory (Chubala & Jamieson, 2013; Chubala, Johns, Jamieson, & Mewhort, 2016; Jamieson, Holmes, & Mewhort, 2010; Jamieson & Mewhort, 2009, 2010), false memory (Arndt & Hirshman, 1998), semantic memory (Kwantes, 2005), and associative learning (Jamieson, Crump & Hannah, 2012; Jamieson, Hannah, & Crump, 2010). The model stores each experience as separate vector in a memory matrix. A retrieval probe presented to the memory matrix activates all vectors in proportion to their similarity to the probe. The sum of the weighted activations results in a measure of the probe's familiarity, while the vector created from the pattern of activation in the matrix is itself a representation of the content retrieved. Minerva-2 combines localized storage of individual events with distributed retrieval process (Hintzman, 1990). As previously discussed, an argument may be made that the group function learning task also combines local (i.e., individual strategies) and distributed (i.e., group-level learning) processes. Because a large proportion of participants reported recalling and repeating their previous responses to given target values (see Chapter 5), a memory-based model that makes contact with neural networks (Kelly, Mewhort, & West, 2014) may also be useful in modeling group function learning behaviour.

Agent-based models. Instead of attempting to directly capture group-level behaviours with computational models intended to describe individual cognition, agent-based models (ABMs) describe the interactions among individual agents and their environments that lead to emergence at a higher level of analysis (e.g., Epstein & Axtell, 1996; Goldstone & Janssen, 2005). These models simulate individuals as virtual agents, providing each with basic decision-making processes, a structured environment, and period of time over which interactions can occur. The computational description is at the level of the agents, but the outcome of a simulation is considered at the level of the emergent organization.

As an example, Roberts and Goldstone (2011) used an ABM to model their Group Binary Search (GBS) game. Agents were seeded with varying levels of reactivity to feedback following a trial, which value could change probabilistically over time. The group-level results of their GBS game were best fit by a model of flexible strategies, such that agents decreased their reactivity to feedback when the group feedback changed from one round to the next (e.g., from "too high" to "too low" or vise versa). Because of this adaptive individual-level strategy, groups avoided getting stuck in an oscillation between overshooting and undershooting a target value. It is not difficult to see how such a model could be applied to my group function learning task.

Although the majority of current ABMs tout oversimplified representations of agents, increasing awareness of these models within psychology ought to improve them. For example, imagine an ABM in which the individual agents were described by a model of memory like Minerva-2 (Hintzman, 1984, 1986). Depending on the individual agent's immediate surroundings in the model, it would develop a unique memory matrix to represent its experience, and would subsequently respond to the structure in the environment in ways that are inherently different from other agents with other histories of experience. Modern technology has effectively removed the computational barriers to implementing such rich simulations. I have begun using Repast Simphony, an open-sourced ABM platform (North et al., 2013), and I fully intend to pursue this line of research.

Implications

Scale invariance and self-similarity. In addition to the empirical demonstrations provided by my experiments, the results of this thesis make contributions to the broader theoretical landscape of cognitive psychology. In demonstrating that findings from individual function learning tasks can scale up to groups of individuals, the data help to express human behaviour according to general principles of behaviour and structure in the natural sciences. One such regularity is that of self-similarity, whereby regularities at one scale of a hierarchical system repeat at other scales of the system. In scaling the function learning task from individuals to groups, similar methodologies led to similar, though not identical, behaviours. Whereas numerous phenomena in the physical and biological sciences have proven to be invariant to scale (e.g., Koscielny-Bunde et al., 1998; MacIntosh, Pelletier, Chiaradia, Kato, & Ropert-Coudert, 2013; Peng et al., 1992; Peng, Havlin, Stanley, & Goldberger, 1994; Schlesinger & West, 1991), psychology has generally opted to provide different theories and models for different scales of analysis (e.g., Atkinson & Shiffrin, 1968; Baddeley & Hitch, 197; Knowlton & Squire, 1993; Owen, 1998; Schacter & Tulving, 1994).

The data presented in this thesis are consistent with a database of scale invariant phenomena within psychology (e.g., Chater & Brown, 1999; Howard & Eichenbaum, 2013; Howard et al., 2014; Neath & Brown, 2006) and the social sciences more generally (e.g., Maske, Havlin, & Stanley, 2002; Mantegna & Stanley, 2002). Such a database is critical if psychology is to situate itself among the other branches of science. As a regularity of both natural organic phenomena and large-scale societal structures, scale invariance should be an uncontroversial principle of human behaviour—which finds itself between the micro lens of biology and the macro lens of society. Indeed, in the broader scope of science as a whole, it is the suggestion that humans are somehow the exception to this wide-ranging rule that is controversial.

Whereas other research groups have demonstrated scale invariance on the dimension of time (i.e., the SIMPLE model; Brown, Neath, & Chater, 2007; Chater & Brown, 1999; Neath & Brown, 2006, 2012), my own demonstrations show a less exact form of the principle: self-similarity on different levels of a hierarchy, from individual learners to groups of learners. Unlike the linear unfolding of time, the dimension of interest in the current demonstrations is hierarchical. Just as neurons are parts that comprise a larger cognitive system, individual humans are parts that comprise a larger cognitive system in my function learning task.

Hierarchical systems. The writer and journalist Arthur Koestler (1967) gave the name *holons* to organisms that are both parts and wholes at different levels of the same hierarchical system. Holons encapsulate the notion that the wholeness or part-ness of an organism do not exist as independent absolutes, but are instead determined by both the organism's place in a hierarchy of organization and the current level of analysis being considered. This idea was inspired by Simon's (1962) parable of the two watchmakers, in which one watchmaker, whose assembly process involves building upon stable intermediate states, fares better than one whose assembly process is not stable until it is fully complete. In this parable, the stable intermediate states are wholes in and of themselves, but are also parts of the finished watch; thus, they are holons. To the extent that individual humans exist as parts in the larger whole of society, they too are examples of holons. Indeed, Koestler argues that holons are the key to evolutionary complexity, and are therefore ubiquitous in nature.

In the context of this thesis, individual humans are the stable intermediate states that connect the hierarchy between biological neural networks and groups of humans (see also Goldstone & Gureckis, 2009). Individuals in my group learning tasks are simultaneously autonomous units that are capable of making local decisions independent of other individuals (see e.g., Chapter 5 of this thesis), and subordinate units that answer to higher levels of the hierarchy (i.e., group-level learning; see Chapters 2 through 4).

In their roles as intermediate states, individuals allow information to travel bidirectionally between the level of the group and the level of neurons supporting individual behaviour. Koestler (1967) emphasized that disruption of the bidirectional flow of information would result in a breakdown of the hierarchical system, such that parts would no longer recognize their role in the superordinate whole. Whereas he focused his argument on societal breakdowns, my group function learning task provides a straightforward empirical demonstration. In the test phase of Experiment 2, average squared error on the target value of 1 was substantially greater (*M* = 14.30, *SD* = 17.28) than the next highest errors (*M* = 7.10, *SD* = 7.82 for the target value of 2; *M* = 6.40, *SD* = 7.92 for the target value of 5). I argue that the target value of 1 primed the sense of an autonomous whole in the individual group members, thereby disrupting the bidirectional flow of information and reducing individuals' attention to their roles as integral parts of the group. Most of the time, however, individuals' behaviours appropriately balanced their personal inclinations with the group-level feedback. The result was self-organization of individual roles and the expression of group-level learning.

Self-organization and emergence. Indeed, a key feature of Koestler's (1967) holons is that they are self-organizing, open (i.e., dissipative) systems. Such systems move from chaos to organization over time as a result of bidirectional interactions with the environment. Life itself, for example, is generally considered a self-organizing system: living organisms maintain a bidirectional relationship with their environments, and have evolved over time into stable intermediate states reciprocal to their unique environmental niches. An individual brain can also be considered a self-organizing system, as neural development proceeds by shaping connections in response to the particular experiences of the individual over time. Indeed, some computational modelers have taken a self-organizing approach to the brain with some success (e.g., Becker, 1996; Becker & Hinton, 1992; Kitzbichler, Smith, Christensen, & Bullmore, 2009; Kohonen, 1982).

Likewise, individuals in the group function learning task self-organized into unique response roles as a result of group-level feedback. There was no direct communication between individuals and no executive control faculty to facilitate this division of labour. Instead, it emerged from the structure of the task environment. This hearkens back to Simon's (1969) parable of the ant: the apparent complexity of the group's behaviour is a reflection of the complexity of the group function learning task. Simple decision mechanisms operating at the lowest level of the hierarchy (i.e., the neural level) interacted with the constraints of the structured task environment to produce impressive learning capabilities at the highest level of the hierarchy (i.e., the group-level). Importantly, the structure in this task goes beyond the mere target values given to groups, but instead includes the entire feedback history of a group and the a priori traits or tendencies that individuals bring with them to the experiment (see Chapter 5). Moreover, while the task environment (e.g., feedback) is altering the behaviour of individuals within the group, the structure of that environment is changing in response to individuals' behaviours. As Koestler (1967) argued, bidirectional relationships between levels of a hierarchy are critical for the stability of the system and, I argue, for successful group-level learning in my function learning task.

Summary

In one view of cognition, sophisticated behaviours are presumed to be the result of sophisticated mechanisms, where each phenomenon can be explained by its own mechanism (e.g., Atkinson & Hitch, 1968; Schacter & Tulving, 1994). In a more parsimonious view, sophisticated behaviours are an emergent consequence of simple mechanisms that operate over time in a structured environment (e.g., Simon, 1969). The explanation for a given behaviour cannot be found in either the human or the environment alone, but instead emerges from the interaction between them. Critically, this latter view can be applied to cognition at any scale. That is, whether one is considering different time scales in memory (e.g., Neath & Brown, 2006) or different levels of a hierarchical system of learners (Experiments 1 through 3 of this thesis), behaviours at any scale can be explained as a function of simple mechanisms interacting with structured environments. The principle of scale invariance is important insofar as it brings psychology in line with a host of discoveries in the physical and biological sciences that suggest the principle is a fundamental feature of nature.

Throughout this thesis, I have approached the principle of scale invariance in a simplified and approximated laboratory model. By applying a standard task in cognitive psychology to groups rather than individuals, I have examined the self-similarity of behaviours when a similar task is applied across two levels of a hierarchy. Recent advances in the cognitive sciences have focused on how group cognition emerges from the local interaction of individuals in a social environment (Goldstone & Ashpole, 2004; Goldstone, Ashpole, & Roberts, 2005; Goldstone, Jones, & Roberts, 2006; Goldstone & Gureckis, 2009; Gureckis & Goldstone, 2006; Moussaid et al., 2009), a theoretical stance that is clearly in line with Simon's (1969) dictum. I sought to join this literature by examining function learning in groups of individuals.

In a series of three experiments, I demonstrated that groups can learn both simple and complex functions (Chapter 2: Experiment 1), that groups can generalize their learning to untrained areas of a learned function (Chapter 3: Experiment 2), and that this group learning is resilient to the effects of disruption (Chapter 4: Experiment 3). An extensive evaluation of individual behaviour within the groups (Chapter 5) showed a distributed representation of the learned functions, wherein individuals self-organize into differentiated behavioural roles in order for the group to succeed. Whereas an ideal expression of scale invariance would have shown individuals and groups to accomplish function learning by the same mathematical processes (i.e., error reduction learning in an artificial neural network), the behaviours of both groups and the individuals within them were not fully captured by this computational approach. Instead, individuals incorporated elements of memory-based models and simple heuristics into their decisions, the combination of which across individuals led to behaviour consistent with network-like learning. That is, despite differing mechanisms, the environmental constraints of a function learning task nevertheless led individuals and groups to produce many similar behaviours.

Taken together, the results presented in this thesis help to emphasize the centrality of scale invariance, emergence, and distributed representations to cognition. Apparently complex behaviours, whether displayed in individuals or groups of individuals, need not be the result of sophisticated mechanisms. As introspective beings, we hesitate to discount the phenomenological sensation of our own thoughts and behaviours as mere functions of our environment, but the notion becomes less intrinsically unpalatable when applied to groups (see Cooley et al., 2017 for perceptions of cognition in groups). Groups do not have a phenomenological consciousness, yet they can behave in ways that mimic the behaviour of conscious individuals. Groups do not have a frontal cortex, yet they can behave in ways that nobody would dispute as signs of intelligence, were those behaviours were demonstrated by individuals instead.

This understanding allows us to reconceive human intelligence, which at its most remarkable is traditionally taken as evidence for human uniqueness. Borrowing from both Simon (1962, 1969) and Koestler (1967), I argue instead that intelligence is an emergent property that is (1) a function of the parallel interaction of the parts of a system with a structured environment, and (2) visible at the level of the whole, which whole will necessarily depend upon the scale of the analysis under consideration. Intelligence is thus not a property of the system itself, but an attribution made about the behaviour that emerges from the system. By this metric, I argue that the group-level learning demonstrated throughout this thesis represents a form of intelligence, albeit a form of intelligence that is functionally distinct from that observed when considering individual humans as wholes instead of parts.

As I hope to have demonstrated with my empirical analyses, a group of individuals can be taken as a meaningfully intelligent unit, of as much interest to cognitive science as intelligence at other levels of analysis. I join comparative cognition researchers in arguing against the conceit of uniqueness assumed by humans (e.g., De Waal & Ferrari, 2010; Shettleworth, 2010) and the inimitable Stephen J. Gould in arguing against the tendency to reify intelligence as an entity (1981/1996). I propose instead that intelligence is an attribution we make about adaptive behaviours, and that its emergence is observable in all self-organizing systems situated within structured environments: from ants on beaches, to humans moving through the world, to groups of undergraduates made to complete a collaborative function learning task for course credit.

References

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory and Language*, *39*, 371–391.

Asch, S. E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men; research in human relations* (pp. 177–190). Oxford, England: Carnegie Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*, 89-195.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation, 8*, 47-89.

Barenblatt (1996). *Scaling, self-similarity, and intermediate asymptotics*. Cambridge, UK: Cambridge University Press.

Bassingthwaighte, J. B. (1992). Fractal growth patterns. *Acta Stereologica, 11* (Suppl. 1), 305-319.

Becker, S. (1996). Mutual information maximization: Models of cortical self-organization. *Network: Computation in Neural Systems*, *7*, 7–31.

Becker, S. (1999). Implicit learning in 3D object recognition: The importance of temporal context. *Neural Computation*, *11*, 347–374.

Becker, S., & Hinton, G. E. (1992). Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature*, *355*, 161–163.

Bettencourt, L. M. A. (2009). The rules of information aggregation and emergence of collective intelligent behavior. *Topics in Cognitive Science, 1*, 598–620.

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Decision Processes, 11*, 1– 27.

Brennan, A. A., & Enns, J. T. (2015). When two heads are better than one: Interactive versus independent benefits of collaborative cognition. *Psychonomic Bulletin & Review, 22*, 1076–1082.

Bott, L., & Heit, E. (2004). Nonmonotonic extrapolation in function learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 38–50.

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*, 201–233.

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: Implications for neural network theories in psychology and neuroscience. *Psychological Review*, *116*, 220–251.

Bowers, J. S. (2010). More on grandmother cells and the biological implausibility of PDP models of cognition: A reply to Plaut and McClelland (2010) and Quian Quiroga and Kreiman (2010). *Psychological Review*, 117, 300–306.

Broca, P. (1861) Remarks on the Seat of the Faculty of articulated language, following an observation of aphemia (C. D. Green, Trans.). *Bulletin de la Société Anatomique*, *6*, 330–357.

Brown, G. D. A., Neath, I., & Chater, N. (2002). *A ratio model of scale invariant memory and identification* (Memory Lab Technical Report 2002-01). West Lafayette, IN: Purdue University.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*, 539–576.

Brown, M. A., & Lacroix, G. (2017). Underestimation in linear function learning: Anchoring to zero or X-Y similarity? *Canadian Journal of Experimental Psychology*. doi: 10.1037/cep0000129.

Busemeyer, J. R., Byun, E., DeLosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts & D. Shanks (Eds.), *Knowledge concepts and categories* (pp. 405– 437). Cambridge, MA: MIT Press.

Castiglione, F., Pappalardo, F., Bianca, C., Russo, G., & Motta, S. (2014). Modeling biology spanning different scales: An open challenge. *BioMed Research International*, e902545.

Chater, N. (2003). How much can we learn from double dissociations? *Cortex*, *39*, 167–169.

Chater, N., & Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition, 69*, B17-B24.

Chubala, C. M., & Jamieson, R. K. (2013). Recoding and representation in artificial grammar learning. *Behavior Research Methods*, *45*, 470–479.

Chubala, C. M., Johns, B. T., Jamieson, R. K., & Mewhort, D. J. K. (2016). Applying an exemplar model to an implicit rule-learning task: Implicit learning of semantic structure. *Quarterly Journal of Experimental Psychology, 69*, 1049-1055.

Cleeremans, A. (2006). Implicit learning models. In *Encyclopedia of Cognitive Science*. (Article 97). Hoboken, NJ: John Wiley & Sons.

Cooley, E., Payne, B. K., Cipolli III, W., Cameron, C. D., Berger, A., & Gray, K. (2017). The paradox of mind: “People in a group” have more mind than “a group of people”. *Journal of Experimental Psychology: General,* online first publication*.*

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, *19*, 261–270.

De Waal, F. B. M., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, *14*, 201–207.

Deane, D. H., Hammond, K. R., & Summers, D. A. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology, 92*, 20–26.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.

Delamater, A. R. (2012). On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*, *40*, 1–23.

DeLosh, E. L. (1995). *Hypothesis testing in the learning of functional concepts*. Masters thesis, Psychology Department, Purdue University.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 968–986.

Derraugh, L. S., Neath, I., Surprenant, A. M., Beaudry, O., & Saint-Aubin, J. (2016). The effect of lexical factors on recall from working memory: Generalizing the neighborhood size effect. *Canadian Journal of Experimental Psychology, 71*, 23-31.

Dorigo, M., Bonabeau, E., & Theraulaz, G. (2000). Ant algorithms and stigmergy. *Future Generation Computer Systems, 16*, 851–871.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*, 180–209.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review*, *95*, 91–101.

Dunn, J. C., & Kirsner, K. (2003). What can we infer from double dissociations? *Cortex, 39*, 1–7.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, *338*, 1202–1205.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.

Elman, J. L. (2005). Connectionist models of cognitive development: where next? *Trends in Cognitive Sciences*, *9*, 111–117.

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, D. C.: Brookings Institution.

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., … Petersen, S. E. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Computational Biology*, *5*, e1000381.

Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the "locality" assumption. *Behavioral & Brain Sciences*, *17*, 43–61.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Feldman-Stewart, D., & Mewhort, D. J. K. (1994). Learning in small connectionist networks does not generalize to large networks. *Psychological Research*, *56*, 99–103.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Fraley R. C., & Vazire, S. (2014). The N-Pact factor: Evaluating the quality of empirical journals with respect to sample size and statistical power. *PLoS ONE, 9*, e109019.

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

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. Oxford, England: Wiley.

Ghahramani Z., Korenberg A. T., & Hinton G. E. (1999). Scaling in a hierarchical unsupervised network. *Proceedings of the 9th International Conference on Artificial Neural Networks—ICANN* (pp. 13-18). Edinburgh: Springer.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 227-247.

Goldstone, R. L., & Ashpole, B. C. (2004). Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review, 11*, 508-514.

Goldstone, R. L., Ashpole, B. C., & Roberts, M. E. (2005). Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review, 12*, 81-87.

Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science, 1*, 412–438.

Goldstone, R. L., & Janssen, M. A. (2005). Computational models of collective behavior. *Trends in Cognitive Sciences, 9*, 424–430.

Goldstone, R. L., Jones, A., & Roberts, M. E. (2006). Group path formation. *Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 36*, 611–620.

Goldstone, R. L., Roberts, M. E., & Gureckis, T. M. (2008). Emergent Processes in Group Behavior. Current Directions in Psychological Science, 17(1), 10–15.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the big five personality domains*. Journal of Research in Personality, 37*, 504-528.

Gould, S. J. (1981/1996). *The Mismeasure of Man* (Rev. ed.). New York, NY: Norton.

Grafton, S. T., & de C. Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, *26*, 590–616.

Gureckis, T. M., & Goldstone, R. L. (2006). Thinking in groups. *Pragmatics and Cognition, 14*, 293–311.

Happel, B. L. M., & Murre, J. M. J. (1994). Design and evolution of modular neural network architectures. *Neural Networks*, *7*, 985–1004.

Haselgrove, M. (2010). Reasoning rats or associative animals? A common-element analysis of the effects of additive and subadditive pretraining on blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 296-306.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*, 185–207.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119*, 431–440.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98*, 74-95.

Hinton, G. E., Plaut, D., & Shallice, T. (1993). Simulating brain damage. *Scientific American*, *269*, 76–82.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*, 96–101.

Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hintzman, D. L. (1990). Human learning and memory: Connections and dissociations. *Annual Review of Psychology*, *41*, 109–139.

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General, 142*, 1211–1230.

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2014). A distributed representation of internal time. *Psychological Review, 122*, 24-53.

Howell, S. R., & Becker, S. (2001). Modelling language acquisition: Grammar from the lexicon. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (429-434). Hillsdale, NJ: Erlbaum.

James, W. (1890). *The Principles of Psychology*. New York, NY: H. Holt & Company.

Jamieson, R. K., Crump, M. J. C., & Hannah, S. D. (2012). An instance theory of associative learning. *Learning & Behavior*, *40*, 61–82.

Jamieson, R. K., Hannah, S. D., & Crump, M. J. C. (2010). A memory-based account of retrospective revaluation. *Canadian Journal of Experimental Psychology*, *64*, 153–164.

Jamieson, R. K., Holmes, S., & Mewhort, D. J. K. (2010). Global similarity predicts dissociation of classification and recognition: evidence questioning the implicit-explicit learning distinction in amnesia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 1529–1535.

Jamieson, R. K., & Mewhort, D. J. K. (2009). Applying an exemplar model to the artificial-grammar task: inferring grammaticality from similarity. *Quarterly Journal of Experimental Psychology (2006)*, *62*, 550–575.

Jamieson, R. K., & Mewhort, D. J. K. (2010). Applying an exemplar model to the artificial-grammar task: String completion and performance on individual items. *Quarterly Journal of Experimental Psychology, 63*, 1014–1039.

Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*. Boston, MA: Houghton Mifflin.

Janis, I. L. (1982). *Groupthink* (2nd ed.). Boston MA: Houghton Mifflin.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory--Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). Amsterdam, NL: Elsevier.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York, NY: Scribner.

Jorgenson, L. A., Newsome, W. T., Anderson, D. J., Bargmann, C. I., Brown, E. N., Deisseroth, K., … Wingfield, J. C. (2015). The BRAIN Initiative: Developing technology to catalyse neuroscience discovery. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *370*, 20140164.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review, 14*, 288–294.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111*, 1072–1099.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*, 679–685.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.

Kelley, M. R., Neath, I., & Surprenant, A. M. (2015). Serial position functions in general knowledge*. Journal of Experimental Psychology: Learning, Memory, & Cognition, 41*, 1715–1727.

Kelly, M. A., Mewhort, D. J., & West, R. L. (2014). The memory tesseract: Distributed MINERVA and the unification of memory. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Cognitive Science Society*, Quebec City, QC.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*, 805–825.

Kinder, A. (2000). The knowledge acquired during artificial grammar learning: Testing the predictions of two connectionist models. *Psychological Research*, *63*, 95–105.

Kinder, A., & Lotz, A. (2009). Connectionist models of artificial grammar learning: What type of knowledge is acquired? *Psychological Research*, *73*, 659–673.

Kitzbichler, M. G., Smith, M. L., Christensen, S. R., & Bullmore, E. (2009). Broadband criticality of human brain network synchronization. *PLoS Computational Biology*, *5*, e1000314.

Knapp, A. & Anderson, J. A. (1984). A signal-averaging model for concept formation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 616-37.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science, 262*, 1747-1749.

Koestler, A. (1967). *The Ghost in the Machine* (1990 reprint ed.). London, UK: Penguin.

Koh, K., & Meyer, D. E. (1991). Function learning: Induction of continuous stimulus-response relations. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 811–836.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*, 59–69.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H. E., Goldreich, Y., & Schellnhuber, H.-J. (1998). Indication of a universal persistence law governing atmospheric variability. *Physical Review Letters, 81*, 729.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44.

Kwantes, P. J. (2005). Using context to build semantics. *Psychonomic Bulletin & Review*, *12*, 703–710.

Kwantes, P. J., & Neal, A. (2006). Why people underestimate y when extrapolating in linear functions. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32*, 1019–1030.

Kwantes, P. J., Neal, A., & Kalish, M. (2012). Item order matters in a function learning task. *Canadian Journal of Experimental Psychology, 66*, 90–97.

Lashley, K. S. (1950). In search of the engram. In *Physiological mechanisms in animal behavior. (Society's Symposium IV.* pp. 454–482). Oxford, UK: Academic Press.

Lashley, K. S., & Wiley, L. E. (1933). Studies of cerebral function in learning IX: Mass action in relation to the number of elements in the problem to be learned. *Journal of Comparative Neurology, 57*, 3-55.

Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, *10*, 215–221.

Le Pelley, M. E. (2012). Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 686–708.

Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology, 44*, 585–612.

Lichtenstein S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning & Memory, 4*, 551-578.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review, 22*, 1193–1215.

MacGregor, D., Lichtenstein, S., & Slovic, P. (1988). Structuring knowledge retrieval: An analysis of decomposed quantitative judgments. *Organizational Behavior & Human Decision Processes, 42*, 303-323.

MacIntosh, A. J. J., Pelletier, L., Chiaradia, A., Kato, A., & Ropert-Coudert, Y. (2013). Temporal fractals in seabird foraging behaviour: Diving through the scales of time. *Scientific Reports, 3*, Article #1884.

Makse, H. A., Havlin, S., & Stanley, H. E. (2002). Modelling urban population growths. *Nature, 377*, 608-612.

Mantegna, R. N., & Stanley, H. E. (2002). Scaling behaviour in the dynamics of an economic index. *Nature, 376*, 46-49.

Mardia, K.V. (1978). Some properties of classical multidimensional scaling. *Communications on Statistics: Theory and Methods, A7*, 1233–1241.

Markram, H. (2006). The blue brain project. *Nature Reviews: Neuroscience*, *7*, 153–160.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, *11*, 194–201.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General, 137*, 422–433.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*. (2nd ed.). New York, NY: Psychology Press.

McClelland, J. L., Rumelhart, D. E., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*, 115–133.

McDaniel, M. A., & Busemeyer, J. R. (2005). The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychonomic Bulletin & Review, 12*, 24–42.

McGill, W. J. (1954). Multivariate information transmission. *Psychometrika, 19*, 97–116.

Medler, D. A., Dawson, M. R. W., & Kingstone, A. (2005). Functional localization and double dissociations: The relationship between internal structure and behavior. *Brain & Cognition*, *57*, 146–150.

Miller, B. J., & Steyvers, M. (2011). The Wisdom of Crowds with Communication. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Cognitive Science Society, Austin, TX.

Minsky, M., & Papert, S. (1969). *Perceptrons*. Oxford, UK: MIT Press.

Mirsky, L. (1956). The spread of a matrix. *Mathematika, 3*, 127-130.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*, 72–80.

Morgan, C. L. (1903). *An introduction to comparative psychology*. Newcastle, UK: Walter Scott.

Moussaid, M., Garnier, S., Theraulaz, G., & Helbing, D. (2009). Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science, 1*, 469–497.

Murre, J. M. J., Phaf, R. H., & Wolters, G. (1992). CALM: Categorizing and learning module. *Neural Networks*, *5*, 55–82.

Neath, I., & Brown, G. D. A. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. *Psychology of Learning and Motivation, 46*, 201-243.

Neath, I., & Brown, G. D. A. (2012). Arguments against memory trace decay: A SIMPLE account of Baddeley and Scott. *Frontiers in Cognition*, *3*, 1–3.

Neath, I., Kelley, M. R., & Surprenant, A. M. (2016). Three semantic serial position functions at the same time. *Experimental Psychology, 63*, 351–360.

Neath, I., VanWormer, L. A., Bireta, T. J., & Surprenant, A. M. (2014). From Brown-Peterson to continual distractor via operation span: A SIMPLE account of complex span. *Canadian Journal of Experimental Psychology, 68*, 204–211.

Neil, G. J., & Higham, P. A. (2012). Implicit learning of conjunctive rule sets: An alternative to artificial grammars. *Consciousness & Cognition, 21*, 1393–1400.

Neil Bearden, J., & Wallsten, T. S. (2004). MINERVA-DM and subadditive frequency judgments. *Journal of Behavioral Decision Making*, *17*, 349–363.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive Skills and their Acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, *19*, 113–126.

North, M. J., Collier, N. T., Ozik, J., Tatara, E., Altaweel, M., Macal, C. M., Bragen, M., & Sydelko, P. (2013). Complex adaptive systems modeling with Repast Simphony. *Complex Adaptive Systems Modeling, 1*, 3.

Owen, A. M. (1998). Memory: Dissociating multiple memory processes. *Current Biology, 8*, R850-R852.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *The Behavioral and Brain Sciences*, *23*, 443–467; discussion 467–512.

Peng, C.-K., Buldryev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., & Stanley, H. E. (1992). Long-range correlations in nucleotide sequences. *Nature, 356*, 168-170.

Peng, C.-K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos, 5*, 82-87.

Plaut, D. C., & McClelland, J. L. (2010). Locating object knowledge in the brain: Comment on Bowers's (2009) attempt to revive the grandmother cell hypothesis. *Psychological Review*, *117*, 284–288.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*, 377-500.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*, 59–63.

Poulton, E. C. (1982). Biases in quantitative judgments. *Applied Ergonomics, 13*, 31-42.

Prince, M., Brown, S., & Heathcote, A. (2012). The design and analysis of state-trace experiments. *Psychological Methods*, *17*, 78–99.

Quinlan, J. A., Neath, I., & Surprenant, A. M. (2015). Positional uncertainty in the Brown-Peterson paradigm. *Canadian Journal of Experimental Psychology, 69*, 64–71.

Quiroga, Q. R., & Kreiman, G. (2010a). Measuring sparseness in the brain: Comment on Bowers (2009). *Psychological Review*, *117*, 291–297.

Quiroga, Q. R., & Kreiman, G. (2010b). Postscript: About grandmother cells and Jennifer Aniston neurons. *Psychological Review*, *117*, 297–299.

Quiroga, R. Q., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not "Grandmother-cell" coding in the medial temporal lobe. *Trends in Cognitive Sciences*, *12*, 87–91.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior, 6*, 855-863.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64-99). New York, NY: Appleton-Century-Crofts.

Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology*, *63*, 101–128.

Roberts, M. E., & Goldstone, R. L. (2011). Adaptive group coordination and role differentiation. *PLOS ONE, 6*, e22377.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386–408.

Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, DC: Spartan.

Rubenstein, M., Cornejo, A., & Nagpal, R. (2014). Programmable self-assembly in a thousand-robot swarm. *Science*, *345*, 795–799.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Schacter, D. L., & Tulving, E. (1994). *Memory Systems*. Cambridge, MA: MIT Press.

Schlesinger, M. F., & West, B. J. (1991). *Physical Review Letters, 67*, 2106.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, & Psychiatry*, *20*, 11–21.

Shanks, D. R. (2010). Learning: From Association to Cognition. *Annual Review of Psychology*, *61*, 273–301.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shettleworth, S. J. (2010). Clever animals and killjoy explanations in comparative psychology. *Trends in Cognitive Sciences*, *14*, 477–481.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*, 467-82.

Simon, H. A. (1969). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Simon, S. A., de Araujo, I. E., Gutierrez, R., & Nicolelis, M. A. L. (2006). The neural mechanisms of gustation: a distributed processing code. *Nature Reviews: Neuroscience*, *7*, 890–901.

Spencer H. (1855). *The Principles of Psychology*. London, UK: Longman, Brown, Green, & Longman.

Squire, L. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4*, 232–243.

Surowiecki, J. (2004). *The Wisdom of Crowds.* New York, NY: Doubleday.

Surprenant, A. M., & Neath, I. (2009). *Principles of Memory*. New York, NY: Psychology Press.

Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research, 11*, 378–395.

Tulving E. (1983). *Elements of episodic memory*. Oxford, UK: Clarendon Press.

Tulving, E., Schacter, D. L., McLachlan, D. R., & Moscovitch, M. (1988). Priming of semantic autobiographical knowledge: A case study of retrograde amnesia. *Brain and Cognition*, *8*, 3–20.

Turnock, M., & Becker, S. (2008). A neural network model of hippocampal–striatal–prefrontal interactions in contextual conditioning. *Brain Research*, *1202*, 87–98.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Twitch Plays Pokémon. (n. d.). In *Wikipedia*. Retrieved March 30, 2017, from http://en.wikipedia.org/wiki/Twitch_Plays_Pokémon

Ultsch, A. (2007). Emergence in self organizing feature maps. In H. Ritter & R. Haschke (Eds.), *Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)*. Bielefeld, Germany: Neuroinformatics Group.

Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition*, *33*, 1–16.

Wagenaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics, 18*, 416–422.

Wisdom, T. N., Song, X., & Goldstone, R. L. (2013). Social learning strategies in networked groups. *Cognitive Science, 37*, 1383–1425.

Yeates, F., Wills, A. J., Jones, F. W., & McLaren, I. P. L. (in press). State-trace analysis: Dissociable processes in a connectionist network? *Cognitive Science*.

Yi, S. K. M., Steyvers, M., Lee, M. D. and Dry, M. J. (2012), The wisdom of the crowd in combinatorial problems. *Cognitive Science, 36*, 452–470.

Appendix A

Results of analyses on training phase data in Experiments 1-3, using alternative measures of performance.

Experiment 1.	Trials taken to match target.			First attempt squared error.		
ANOVA	*df*	*F*	*p*	*df*	*F*	*p*
Block	1.49, 26.83	16.19	< .001*	1.47, 26.47	6.56	.01*
Condition	1, 18	1.11	.31	1, 18	5.56	.03*
Block x Condition	1.49, 26.83	0.01	.98	1.47, 26.47	0.02	.95
Contrasts: Block	*df*	*t*	*p*	*df*	*t*	*p*
Linear	36	5.37	< .001*	36	3.59	<.001*
Quadratic	36	1.87	.07	36	0.51	.61

Experiment 2.	Trials taken to match target.			First attempt squared error.		
ANOVA	*df*	*F*	*p*	*df*	*F*	*p*
Block	2, 36	5.78	.01*	2, 36	5.36	.01*
Condition	1, 18	0.41	.53	1, 18	1.57	.23
Block x Condition	2, 36	1.30	.28	2, 36	1.44	.25
Contrasts: Block	*df*	*t*	*p*	*df*	*t*	*p*
Linear	36	3.09	.004*	36	3.21	.003*
Quadratic	36	1.43	.16	36	0.66	0.51

Experiment 3.	Trials taken to match target.			First attempt squared error.		
ANOVA	*df*	*F*	*p*	*df*	*F*	*p*
Block	2, 54	10.97	< .001*	2, 54	11.30	< .001*
Condition	2, 27	3.17	.06	2, 27	1.09	.35
Block x Condition	4, 54	0.89	.48	4, 54	0.71	.59
Contrasts: Block	*df*	*t*	*p*	*df*	*t*	*p*
Linear	54	4.58	< .001*	54	4.75	< .001*
Quadratic	54	0.96	.34	54	0.08	.93

Note. * $p < .05$.

Appendix B

The probability structure of the Relearning phase of Experiment 3 (Chapter 4). P(excluded) is the probability that any one participant is excluded from their group's response during the relearning phase. P(responder) is the probability that any one participant was a responder for a particular target prior to the relearning phase. Conjunctive Prob. represents the probability that a responder is excluded, and thus the probability that at least one person needs to change their response in order for the group to obtain the target.

Condition	P(excl.)	Target	P(resp.)	Conjunctive Prob.	Trials	Average Error	First Trial Error
Small	1/10	3	3/10	3/100	5.5	0.99	1.6
		5	5/10	5/100	3.8	0.74	1.0
		7	7/10	7/100	3.7	0.36	0.8
		9	9/10	9/100	14.3	1.53	2.2
Medium	3/10	3	3/10	9/100	2.9	0.25	0.6
		5	5/10	15/100	5.5	0.81	2.3
		7	7/10	21/100	8.5	1.95	5.9
Large	5/10	3	3/10	15/100	4.2	0.78	1.9
		5	5/10	25/100	10.2	2.32	6.5