8-23-2013

Probabilistic and Dynamical Models of Belief Update: Effects of Experimentally Induced Implicit Bias

Stephanie Petrusz

University of Connecticut - Storrs, stephanie.petrusz@gmail.com

Follow this and additional works at: http://digitalcommons.uconn.edu/dissertations

Recommended Citation

http://digitalcommons.uconn.edu/dissertations/229
Organisms undertake actions on the basis of perceptions. Perceptions serve as the basis for what an agent takes to be the case; that is, for the agent’s beliefs. However, the idea of belief is freighted with notions of fallibility and subjectivity, since belief is often considered insofar as it is distinct from knowledge. Here, an attempt is made to link belief more closely with perception and action. This linkage is shown by considering the role belief plays in determining behavior, which is distinct from the role belief plays in language-based philosophical accounts of content and intention. These two notions of belief, and the separation between them, are the subjects of four experiments in which a method of introducing bias in participant responses is employed. In a perceptual task based on the Asch line-length judgment paradigm, participants showed a propensity to respond in keeping with the induced bias rather than with stimulus properties. In a cognitive task based on the Monty Hall Dilemma, participants’ responses were consistent with the induced bias but were less consistent with the best MHD strategy. Experiment 1 established the line-length methodology, which was extended in Experiment 2. Experiment 3 employed the MHD paradigm. Experiment 4 brought the current line-length paradigm into closer contact with the classic Asch paradigm. The overall results were consistent with the claim that belief-as-action and belief-as-assent can, theoretically and methodologically, be treated separately.
Probabilistic and Dynamical Models of Belief Update:
Effects of Experimentally Induced Implicit Bias

Stephanie Petrusz

B.A., University of Pittsburgh, 2005

A Dissertation
Submitted in Partial Fulfillment of the
Requirements of the Degree of
Doctor of Philosophy
At the
University of Connecticut

2013
APPROVAL PAGE

Doctor of Philosophy Dissertation

Probabilistic and Dynamical Models of Belief Update:

Effects of Experimentally-Induced Implicit Bias

Presented by

Stephanie Petrusz, Ph. D.

Major Advisor __________________________________________________________

Michael Turvey

Associate Advisor _______________________________________________________

Claudia Carello

Associate Advisor _______________________________________________________

J. Dixon

Associate Advisor _______________________________________________________

Till Frank

Associate Advisor _______________________________________________________

Michael Lynch

University of Connecticut

2013
ACKNOWLEDGEMENTS

The author thanks the following: My advisor, Michael Turvey, for his guidance, inspiration, and patience. Claudia Carello, without whose facilitation this work would not be possible. My advisory committee, J. Dixon, Till Frank and Michael Lynch, for allowing me to take a flyer on this project and doing their best to keep me from crashing and burning. Those friends, in CESPA and elsewhere, whose conversation and camaraderie gave me much-needed insight and support: Dorothy Baker, Julia Blau, Dobri Dotov, Chris Linder, Lin Nie, Dave Pomper, Alex Reed. Those friends, too many to list, whose influence was less constant but no less meaningful. Those especially who bore the brunt of my failure to smoothly navigate the process of this project: Rob Isenhower, Annelies Rhodes, and Theo Rhodes, who kept me fed, sane, and working, despite my best efforts to stymie theirs. Anna Knuttel and David Rhodes, for their constant warmth and welcome. My parents, Gustav and Sarah Petrusz, who never expected anything less.

This work was supported by a Dissertation Fellowship from the University of Connecticut and, in part, by DARPA Physical Intelligence subcontract HRL 000708-DS. The views, opinions, and/or findings contained in this dissertation are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of DARPA or the Department of Defense.
LIST OF TABLES

Table 1. Strategies for 1\textsuperscript{st} and 2\textsuperscript{nd} door choices in Experiment 3………………..56
Table 2. Significant mean differences between blocks for Experiment 2 $d'$ ………….85
Table 3. Significant mean differences between blocks for Experiment 2 $\beta$……………….87
Table 4. Significant mean differences between blocks for Experiment 2 $c$…………….89
LIST OF FIGURES

Figure 1. Schematic of experimental paradigm in Experiments 1 and 2…………………29
Figure 2. Left clicks as a function of Block, Group and JND in Experiment 1…………32
Figure 3. Mean correct responses and mean confidence as a function of Block, Group and JND in Experiment 1………………………………………………………………………33
Figure 4. Latency of choice as a function of Block and JND in Experiment 1…………33
Figure 5. Accuracy and confidence as a function of Block and JND in Experiment 2…38
Figure 6. Proportion of clicks on the left line as a function of Group and Phase in
Experiment 2……………………………………………………………………………………39
Figure 7. Accuracy as a function of phase and JND (binned) in Experiment 2………40
Figure 8. Confidence as a function of Phase and JND (binned) in Experiment 2……40
Figure 9. Correct 1st and 2nd responses in a MHD trial as a function of Block………50
Figure 10. Confidence and accuracy in the MHD as a function of Block and
Strategy………………………………………………………………………………………51
Figure 11. Accuracy in the MHD as a function of Group and Block…………………52
Figure 12. 1st and 2nd response as a function of Group and Block in Experiment 3……53
Figure 13. Proportion of clicks on the left line as a function of Group and Block in
Experiment 4………………………………………………………………………………54
Figure 14. Absolute value of confidence as a function of Group and Block in Experiment 4……………………………………………………………………………………………67
Figure 15. Signed confidence as a function of Group and Block in Experiment 4……68
Figure 16. Schematic of transition probabilities calculated for Markov chain estimation of attractor strength…………………………………………………………………………………72
Experimentally Induced Bias viii

Figure 17. Individual data and block means of left-click frequency for right- and left-biased groups in Experiment 1……………………………………………………………………73

Figure 18. Comparison of simulated and experimental data for left-click probabilities by block in Experiment 1……………………………………………………………………74

Figure 19. Gamma as a function of Group and Block for Experiments 2 and 3……76

Figure 20. Autoregression parameter $a_1$ as a function of Group and Block for Experiment 2…………………………………………………………………………………………78

Figure 21. Variance in autoregression noise term $q_n$ as a function of Group and Block for Experiment 2……………………………………………………………………79

Figure 22. Autoregression parameter $a_1$ as a function of Group and Block for Experiment 3……………………………………………………………………………………79

Figure 23. Variance in autoregression noise term $q_n$ as a function of Group and Block for Experiment 3……………………………………………………………………80

Figure 24. Mean differences in $d'$ as a function of block and group in Experiment 2…86

Figure 25. Mean differences in $\beta$ as a function of block and group in Experiment 2….88

Figure 26. Mean differences in $c$ as a function of block and group in Experiment 2….90
TABLE OF CONTENTS

ACKNOWLEDGEMENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

CHAPTER TWO: METHODS IN OVERVIEW

CHAPTER THREE: CONSIDERATIONS OF “BELIEF”

CHAPTER FOUR: EXPERIMENT 1

CHAPTER FIVE: EXPERIMENT 2

CHAPTER SIX: EXPERIMENT 3

CHAPTER SEVEN: EXPERIMENT 4

CHAPTER EIGHT: DYNAMICAL ANALYSES

CHAPTER NINE: ANALYSIS IN TERMS OF SIGNAL DETECTION THEORY

CHAPTER TEN: GENERAL DISCUSSION
CHAPTER 1
INTRODUCTION

The conditions under which an agent may be said to believe can be cashed out in
different ways. Belief is essentially an agent’s holding that a proposition or state of
affairs is the case, and as such is to a large extent inaccessible to external assessment. For
this reason some psychologists and philosophers have ignored terms like belief as
meaningless when incorporated into theories. However Fodor (1975) argues that the
etiology and explanation of behavior ultimately rests on propositional attitudes, and can
never be cashed out in strictly behavioral terms—some reference to mental states or
processes is always necessary. In either case such attitudes held by an agent cannot be
directly measured.

There are two related but distinct ways to behaviorally assess belief: an agent may
assent to the truth or falsity of a proposition, or an agent may be disposed to act as though
a proposition is true—or, if one wishes to avoid speaking as though propositions are the
contents of beliefs, to act as though a state of affairs obtains. From an eliminativist or
reductionist perspective these externally-accessible measures are all there is to belief;
however, from a competing perspective these measures are, in some sense, related to, or
determined by, something more.

How can belief, over and above the behaviors it might support, be investigated? A
large literature in contemporary formal epistemology centers on questions of how beliefs
change, rather than the less tractable questions concerning the content or justification of
belief. Existing accounts of belief change share certain features: they concern
propositions, they are framed in terms of state-to-state changes, and they generally
ascribe probabilities as weights to beliefs. Various mathematical models of belief-change (as it might be undertaken by human, artificial, or unspecified agents) have been proposed. Among the arguments leveraged by epistemologists is that a basic species of Bayesian revision is probably inadequate to model belief change. Evaluating the conditionals required by Bayesian update is problematic for reasons including the complexity of agents’ existing commitments (see, for instance, the “web of belief” of Quine and Ullian, 1978) and the difficulty of evaluating the probability of some fact against an unenumerated backdrop of possibilities. Haselager, van Dijk and van Rooij (2008) argue that understanding belief essentially requires understanding the problem of abduction, and doubt whether a traditional brain-based cognitive theory is adequate. Rellihan’s (2009) effort to salvage abduction in the face of Fodor’s (2000) dismissal makes explicit reference to the Quinean belief structure in which agents are enmeshed. Such criticisms and suggested revisions cast serious doubt on the viability of a computational theory of abduction. In Fodor’s (2000, p. 41) words: “Abduction really is a terrible problem for cognitive science, one that is unlikely to be solved by any kind of theory we have heard so far.”

Models of belief update proposed by epistemologists involve a variety of mathematical refinements that are largely foreign to discussions of Bayesian models in psychology. Nonetheless, philosophers friendly to the probabilistic perspective may still entertain doubts as to the adequacy of any computational treatment of abduction (see van Fraassen (1989), cited in Arlo-Costa, unpublished). Though the sophistication of some epistemological accounts should recommend them over models based on toy problems, such models have the luxury of leaving unengaged the question of what the beliefs that
feature in them actually are. But an empirical investigation of belief depends on having some answer to that question.

*The Elio and Pelletier (1997) paradigm*

For a fictional state of affairs X, participants were presented with an initial set of facts and were then given a new fact about X that was incompatible with their initial set of commitments. They were then asked to choose one or more of their initial commitments to disbelieve or regard as uncertain in order to incorporate the new fact. Though the results were informative with respect to belief update, the experiments only involved propositional beliefs – either as statements incorporated into a narrative, or as symbols representative of propositions, as is common in first-order logic. Participants were also asked to treat the information given them as true, and thus were engaged in a kind of suppositional reasoning rather than a direct assessment or manipulation of beliefs actually held. It is probably for logistical reasons that these more direct methods have not been employed in investigating beliefs.

The species of belief dealt with by Elio and Pelletier is closest to the notion of belief as disposition to assent to a proposition. It may be that a disposition to act as though a state of affairs obtains gets at a more deeply rooted notion of belief. However, it may also be that when the cost of an action is low, there is little reason not to go along with what seems to be the case, even if an agent privately maintains reservations. In order to investigate belief empirically it is necessary to attempt to disentangle these notions of belief, since they do not necessarily go together and may lead an agent in different directions. A more fully developed attempt to research belief would have to make sense of these distinct notions, and should permit some conclusions to be drawn regarding
changes in belief, as these may be revelatory of the structure or underpinning of belief. An experimentally based account of belief change may be compared to existing models of belief that are anchored by mathematics rather than behavioral data.

The Baratgin and Politzer (2010) paradigm

Participants in a series of experiments were presented with versions of the Monty Hall scenario, a problem of probability so named because of its similarity to the game show Let’s Make a Deal, hosted by Monty Hall. In the scenario, a player on a fictional game show is offered a choice of three prizes, each one hidden behind a door. One of the doors conceals a car. Behind the other two doors are goats. After the player chooses a door, the host tells the player he will open one of the two remaining doors, invariably revealing a goat. He then allows the player to keep their original door or switch their pick to the third door. The problem consists in the fact that although, mathematically, players always maximize their chance of winning the car by switching, real human agents not only seldom switch, but even upon having the solution explained in detail often resist accepting that switching is the rational decision.

Baratgin and Politzer hypothesize that the widespread difficulty people have in adopting a winning strategy is based on a tendency to incorporate previously unknown facts that are irrelevant to the outcome probabilities as though they were in fact useful. In their experiments the location of a “goat” is presented sometimes in a context that encourages participants to update based on the new fact (as in the Monty Hall scenario) and other times in a way that makes it clear nothing about the problem has changed as a result of the new fact. When presented with the latter case participants do not make the same systematic errors, though their ability to correctly assess the probability of the “car”
being in a certain location collapses quickly with the introduction of more “doors”. The problem of irrelevant facts is not unique to problems of rational agency. Indeed, it may be a challenge for organisms generally. Latty and Beekman (2011) demonstrated that the slime mold (*Physarum polycephalum*) alters its preferences for food sources when its set of alternatives is expanded to include potential sources so unattractive as to be rationally irrelevant to the choice of source. It may be that the *presence* of an additional fact is sufficient for an organism to entertain its *relevance*.

**The Asch (1955) paradigm**

Although often framed in terms of social conformity, the classic Asch line length judgment studies, in our view, constitute a case of conflict between the two sorts of belief introduced above. One of his famous experiments proceeded in this way: a group of participants is shown an image of a line, and then shown a set of three lines. Their task is to choose the line among the three that matches the length of the first line they were shown. Each participant responds individually and publicly, by saying aloud which line they’ve chosen. The trick of the experiment is that only one of the participants is truly participating in the experiment – the others are confederates. When asked, each confederate will choose a line that *does not match* the originally displayed line. The participant, then, has a choice: to allow his actions to be guided by the evidence of his senses, and publicly disagree with his “co-participants”, or to follow the wisdom of the crowd and give an answer he likely believes to be wrong.

Asch found that about a third of the time a participant would choose to respond in keeping with the confederates rather than give the obviously correct judgment. He also found that the likelihood of a participant making the conforming response increased with
the number of confederates, up to 7 confederates. These results are typically interpreted as providing evidence of people’s willingness to conform when under social pressure to do so.

Frank (2008) reanalyzed the Asch data using a winner-take-all model. This analysis showed that Asch’s results are not particular either to the task or to the social nature of the pressure brought to bear on the participant – rather, the pattern of participant responses indicates the structure of the underlying dynamics of systems with a limited number of possible states (in this case, a participant’s conforming or not conforming) and pressure to be in one of the possible states.

For present purposes, we take this analysis and combine it with the two species of belief mentioned above. This combination of factors forms the basis of an experimental method for investigating belief in its dual nature as a highly multidimensional, interconnected network, and as the relatively low-dimensional set of possible action responses the dynamics of the network may yield. As described below, in the present set of experiments we take the Asch line length judgment task and extend it further into the realm of dynamical analyses.

Steps towards operationalizing belief

The experiments discussed above leave several issues unaddressed. The Elio and Pelletier case does not address the problem of implicit beliefs and biases. It also raises a particularly thorny issue—the contrast between the description of beliefs and the content of beliefs. This issue will be taken up in the discussion. For the moment it is sufficient to note that, while having the belief that Blorgons are evil is likely to yield both assent to the proposition that “‘Blorgons are evil’ is true” and, in the presence of Blorgons, acting in
accordance with their assumed evil, avoiding Blorgons and ascribing ill intent to them is not necessarily equivalent to having in mind the proposition that Blorgons are evil.

Reflecting on the classical Asch (1951) conformity study raises another central issue: the conflation of belief as assent and belief as the basis for action. As previously argued, the two notions of belief are related but not equivalent. The Asch study seems, intuitively, to be a case where participants were prepared to act as though they believed a proposition they might not have actually believed. Privately held propositional beliefs are difficult to assess, for reasons including experimental demand characteristics and the fact that any response made as the result of probing one’s propositional beliefs is, necessarily, an action in itself. There are psychological methods (name some) for implicitly assessing biases without the participant being aware either of the bias or of the assessment, and some such method is needed to investigate belief.

Addressing these concerns was a guiding principle of the design of the present experiments. A first desideratum was to avoid using propositionally construed beliefs in favor of something more fundamental. Fodor (1975) argued that perception is, in essence, a species of belief—one that, though it may have a propositional description, is not necessarily propositional in character. The experimental task, therefore, was one wherein participants could make perception-based judgments, one that promises to provide insight into the beliefs that ground those judgments. A second desideratum was to offer participants the opportunity (lacking in the Asch study) to vacillate while still making a desired response—in effect to say, “I’m responding this way but I don’t know if I should”. Satisfying this desideratum requires keeping the response and the level of confidence in the response as closely tied as possible. In a task that satisfies this latter
requirement, participants use the spatial layout of a computer screen to respond in two
dimensions with a single action: clicking on the left or right side of the screen provides
the overt action measure of belief, while the click location’s particular distance away
from the center of the screen provides a measure of confidence or willingness to assent to
the same belief.

In a setting satisfying all the above desiderata, the paradigm of judging line-length is
especially attractive because of the existence of the Asch body of literature, including the
analysis of Frank (2008), wherein the Asch data is already treated as a dynamical system.
Since the same analysis was intended in the present research, a task of similar form is to
be preferred.

The challenge of judgments in experiments on perception

Finding reliable and accurate experimental measures of the judgments made in
perception experiments is fraught with difficulty. Judgments in a perception-action task
often change relative to the measures that are imposed. Yet this fact can itself reveal
something of the underlying structure of the perceiving-acting system that confronts a
particular task on a particular occasion. Lee, Lee, Carello and Turvey (2012)
demonstrated that skilled archers’ judgments of the size of an archery target depended
both on the difficulty of the shot and the success of the archer. Easier shots and more
successful shots were judged as being directed towards larger targets. This finding
extends an existing line of research that shows the effects of effort on judgments of
properties like size and distance (Proffitt, 2008; Proffitt & Linkenauger, 2013). It also
indicates, however, that judgments in a perception-action task are a function of more than
the properties of distal objects.
The foregoing lesson is also valuable for experimental considerations of belief. It is commonly accepted that beliefs are not held in a vacuum; however, of the arbitrarily many beliefs that an agent may hold, obtaining any measure of the ones that are local or relevant to a belief that is being perturbed by new facts or by a particular task is a problem that has yet to be resolved. One possible constraint on neighborhoods of belief is the sort of occasion variable that is part of the formal definition of Gibson’s (1979/1986) concept of affordance (Petrusz & Turvey, 2010; Shaw, Turvey & Mace, 1982). Certainly some sort of constraint on neighborhoods of relevant beliefs is desirable. Rellihan’s (2007) analysis of the problem of abduction as presented by Fodor (2000) illustrates why. A computational model of abduction would have three desirable features: it should be modular, it should be reliable, and it should be implementable in polynomial time (i.e., it should be computationally feasible). Any two of these might be gotten together, but at the cost of the third—inferences could be modular and reliable, but then they might take arbitrarily long to determine. Or they could be reliable and relatively quick, but then they could not be modular—they would need to range over the entirety of an agent’s beliefs. Then again, it could be modular and quick, but then there could be no guarantee of reliability.

The present experiments: Basic design and significance

In the research to be reported the experimental design used to study belief was, as presaged above, a conceptual extension of the now-classic Asch (1955) paradigm. In Asch’s experiment participants were swayed by social factors into giving responses on line lengths that were patently not in keeping with the available optical information. Presumably, the participants did not really believe that a physically and noticeably
shorter line was in fact longer, but were pressured to act as though they believed it was. In the present series of experiments, participants must choose a line (or its functional equivalent), thus acting as though they believe that it is the longer line (or the right functional equivalent), but they are also allowed to inform the experimenter about whether they “really” believe in their choice.

To date “beliefs” have functioned as part of computational and representational theories of cognition. Here, beliefs are investigated experimentally as judgment aspects of perception-action tasks. Attempts are made to manipulate beliefs with the goal of understanding belief as an underwriter of action as distinct from ascribing truth to a proposition. Divorcing belief from its propositional and representational character is a first step toward naturalizing belief as a guide for action. Further, the analytical techniques brought to bear on the resulting data constitute an attempt to directly contrast models that are irreducibly probabilistic and computational against those that treat belief as part of the ongoing tasks of a perceiving-acting system.
CHAPTER 2
METHODS IN OVERVIEW

The four experiments presented here are targeted at addressing two categories of questions. First, when there are multiple strategies for dealing with probabilistic stimuli, which will be chosen, and on the basis of what information? Second, when engaged in a task without explicit requirements, is there a separation between participants’ beliefs as the basis for action and propositional beliefs about the task? Addressing these questions created several requirements for the design of the present experiments. Briefly, the stimuli should be probabilistic in character, since the ability of participants to pick up on environmental probability distributions is at issue. The stimulus probabilities should be available to the experimenter but not obviously available to the participant, so that participants can respond to the stimulus distributions on the basis of perceiving rather than cogitating. The experiments must yield a sufficient number of data points to tease apart different strategies that might be used by the participants, and to examine the dynamics of participant strategies over time. Finally, the experiments should provide an opportunity to measure both species of belief while placing the fewest demands on the participant. The experiments, described in detail below, were designed with all these goals in mind.

A Matlab program generated the line-length stimuli presented to participants in Experiments 1 and 2. The experimenter runs the program by means of a script containing five independent parameters that the experimenter may set before each block: the standard length of a line in pixels, the maximum difference (in pixels) between the lines in a pair, the number of pairs to be presented in a block, the proportion of lines that differ
Experimentally Induced Bias  

by less than the JND, and the proportion of pairs in which the left line is longer. This 

design allows for great flexibility in the presentation of the stimuli with little effort on the 

part of the experimenter while the experiment is actually being run. The standard length 

and maximum difference were the same for all participants and all blocks. The number of 

trials differed only by Phase. For example, in Experiment 2 (an elaboration of 

Experiment 1) all participants made judgments on 100 pairs in Block 1 defining Phase 0, 

100 pairs in Blocks 2-5 defining Phase 1, and on 50 pairs in Blocks 6-8 defining Phase 2. 

However, the two probabilistic parameters (proportion of below-JND pairs and 

proportion of left-line-longer pairs) have consequences for the analysis of participant 

data, as described below.

The design of the experiment calls for a certain proportion of line pairs (80% or 20% 

depending on group) to be left-line-longer pairs. Instead of generating the line pairs in 

advance and presenting them randomly, on each trial the program generates a pair of 

lines. In each case, the left line of the pair is longer. Then, based on an experimenter- 

specified “flip percentage” parameter, the pair is displayed on the screen as either right-

line or left-line-longer. The flip parameter makes use of Matlab’s randomization function 

as follows: a number is randomly generated and if it is less than the specified “flip 

percentage” the pair is displayed with the longer line on the right side of the screen. The 

proportion of below-JND pairs is established in the same way: the experimenter specifies 

a proportion of pairs to be below-JND (in this case, .5 for all participants and all blocks), 

and for each line pair generated a random number is also generated. If that number is 

below .5, the difference (in pixels) between lines is (again, randomly) chosen from an 

interval ranging from 0 to the value of the JND. If the number is above .5, the difference
(in pixels) between lines is randomly chosen from an interval ranging from the value of the JND up to the maximum difference between lines.

This program design means that the actual proportions of left-line-longer pairs and below-JND pairs differ slightly from participant to participant and from block to block. For this reason data from each participant are coded as being in either the left-bias or right-bias group, and also as having a certain post hoc presentation proportion. In Experiment 2, the first 100-trial block, alias Phase 0, was unbiased—the probability that a pair would be left-line-longer was .5. Blocks 2 through 5 of Phase 1 were biased as described above, and the final three 50-trial blocks comprising Phase 2 were again unbiased. The basic hypothesis was that in the initial unbiased block participants would click on the right or left line based on a combination of stimulus properties (which line was longer) and guessing (when it was not apparent which line was longer). The introduction of the stimulus bias in Blocks 2 through 5 was predicted to bias responses accordingly, for above-JND stimuli. A central question of the experiments was whether that bias would carry over into the below-JND stimuli. In order to strengthen the bias, the experiments also incorporated an additional bias in the form of a score. Participants were given instructions to choose the longer of each pair of lines, and were told that the score would increase as they performed better on the task, and would decrease if they performed worse. However, the score actually was designed to introduce a bias towards responding that the right line was longer (right-bias condition) or the left line was longer (left-bias condition). This manipulation, heretofore designated as “score”, is a key element of the experimental procedure. To clarify, using Experiment 2 as the example,
“score” was in effect in Blocks 2-5 (Phase 1) but not in effect in Block 1 (Phase 0) and not in effect in Blocks 6-8 (Phase 2).

Experiment 3 used stimuli presented to participants using a webpage designed with Adobe Flash. It consisted of a modified Monty Hall Dilemma (MHD), using a standard 3-door case. Behind each door was a square colored a different shade of red. The red values were randomly generated from a range of RGB values. The “prize” was the brightest of the three shades of red generated on each trial. Although a line-length version of MHD was considered, pilot work indicated that the non-discriminable range around the JND, which was an advantage in Experiments 1 and 2, made it too difficult for participants to find the “prize” lines and discover MHD strategies. Color was therefore chosen as a stimulus property that is both salient to participants and easy to program.

Experiment 4 returned to the method used in Experiments 1 and 2, but more closely approached the original Asch design with the introduction of two new manipulations. First, while the stimuli remained the same as in Experiments 1 and 2, the generation of “score” was altered in the Matlab program to increment when participants responded on the side of the screen with fewer longer lines, and to decrement when participants responded in keeping with the stimulus pair properties. Second, in Experiment 4 participants were told that the “score” variable was a measure of the degree to which their responses matched the responses of a previous participant. This fiction, together with the reversal of the direction of “score”, was intended to generate conflict between stimulus properties (which line is longer) and social pressure (supposed agreement with another participant).
CHAPTER 3

CONSIDERATIONS OF “BELIEF”

The epistemological literature approaches belief from a different standpoint than the psychological literature. Instead of treating belief as a psychological capacity whose explanation is contingent on other psychological capacities, an epistemological inquiry into belief depends on an inquiry into what facts about the world make belief logically and actually possible. Though epistemology as a discipline has traditionally been concerned with problems of knowledge, the definition of knowledge as “justified true belief” has brought focus onto the issue of belief. The central distinction to be drawn between knowledge and belief is that while a belief may be false, knowledge may not. Importantly, though, an agent cannot be said to know something that it does not also believe. The two notions are therefore closely tied. Another important insight following from the close connection between knowledge and belief is that in order for an agent to learn the agent’s beliefs must be able to change. What follows is an attempt to outline the recent history of epistemological approaches to belief, especially efforts to formalize belief and belief update, and what insights such approaches may have for studying belief change from a psychological standpoint. It is important to remember, for the purposes of this discussion, that theory-laden terms such as “learn”, “know”, and “believe” may have different meanings in philosophical and psychological theories, and in some cases may be used more or less colloquially rather than rigorously.

Belief as propositional and probabilistic

Belief is commonly construed as a propositional attitude, and beliefs are described as having the form “X believes that P”. That beliefs are taken to feature propositions has led
to beliefs being treated as basically language-like, and sharing characteristics in common with other features of language, such as syntax, truth-value, and reference. Because a belief is a propositional attitude there is a distinction between the truth of the attitude (\(X\) does, in fact, believe that \(P\)) and the truth of the proposition that \(P\) expresses. The meaning of this disconnect is that, while beliefs are about the world or parts of it, beliefs can also be wrong. That is, \(X\) can believe that, for instance, snow is white, and the proposition “snow is white” may be true, but \(X\) could also believe that snow is green.

From this beginning we may address a psychological question: How are belief and action related? We will examine three possible answers.

Armstrong (1973, 2010) argues that perception is essentially a matter of belief. This argument turns on a congruity between features of a propositional account of belief and a representational theory of perception. Representational accounts of perception rely on the assumptions that perceptions can be erroneous, and that the error results from an impoverished mapping between the world and its representation. Perception is the carrier of facts about the world, but it doesn’t do a perfect job. A propositional account of belief will show that belief shares these characteristics—beliefs are about the world, but our evidence about the world is not perfect, and we do not treat all evidence perfectly rationally. So beliefs can be wrong.

Armstrong’s argument that perceptions are beliefs is based on what he terms the “argument from illusion”. He uses perceptual illusions as evidence for the flawed nature of perception and argues that what underlies our reliance on perception is the fact that we believe our perceptions, even in cases where they do not correspond with the world.

Another influential assessment of the degree to which beliefs are (or are like)
propositions is due to Ramsey (1926/1990). In his “General Propositions and Causality” he makes an elegant argument that in many respects beliefs are not really like propositions. He regards beliefs as the sort of things that guide behavior, and says that general propositions – those that are facts about classes of things rather than particulars – are not, because of their very generality, specific enough to serve as guides to action. Further, he claims that, although beliefs may still be sentence-like, they do not behave as truth-carriers in the sense that propositions do; that is, they are not evaluated as true or false, but rather as carrying a certain weight, or may remain unevaluated or even unentertained. He terms sentences of this type “variable hypotheticals” and gives them a status which is particularly revelatory for the present purposes: “Variable hypotheticals…form the system with which the speaker meets the future...[They] are not judgments but rules for judging.” (Ramsey, 1990, p. 241)

Ramsey’s discussion raises several key points. First, beliefs are not necessarily to be understood as propositions. In an important sense their more essential character is as guides for action, and actions are not necessarily undertaken on the basis of agreement or disagreement with a proposition, but on the basis of judgments about what seems likely to the believer to be appropriate on some occasion. Second, these judgments are assessed not in terms of truth or falsity but against a backdrop of likelihood – that is, some kind of probability. Third, these probabilities are subjective – they are determined for individual believers, not (or not exclusively) by external states of affairs.

Ramsey’s arguments show the way in which belief can be subjective and valued by degrees. This is important to bear in mind when examining arguments of the kind offered by Armstrong, in order to maintain the distinctions between a theory of actions
underwritten by the sort of things that can be wrong and a theory of actions underwritten by the sort of things that can be relatively likely or relatively clear.

If beliefs are construed as propositions in the way outlined above, then regarding beliefs as the foundation of perception is something to be avoided by those who take perception to be direct (Shaw, Turvey & Mace, 1982). A theory of direct perception would not allow Armstrong’s argument from illusion: perception is incorrigible, and therefore illusion is by definition not perception. Shaw et al. say that perception is more like knowledge than like belief, but a traditional account of knowledge would say that knowing something also entails believing it. On this reading, perceptions are the sort of things that a perceiver believes, but they are not themselves conditioned on beliefs.

But perceptions are like beliefs in a further sense, which is made clear by Ramsey: perceptions are the sorts of things that guide action, and it is in the sense of belief as a guide for action that belief and perception are closest. Both perceptions and beliefs form the basis for action. Both perception and belief may be more or less thoroughly constrained by the facts of the world. And what is perceived or believed depends strongly on the character of the perceiver/believer, on each one’s particular constitution and history, and on what actions might need to be guided on a given occasion. It is this relation between perception and belief that makes an empirical investigation of belief of value to psychology. We therefore examine what existing theories of belief offer.

What does “probability” mean?

The theory of probability is anchored in Kolmogorov’s axioms, but these do not settle the question of what a probability is. De Finetti (1974) offers a critique of the standard, frequentist interpretation of probabilities. According to this view, the probability of an
Experimentally Induced Bias

event is derived from the observed probability of a certain type of event over a large number of observations or trials. The core assumption is that the frequency of the event over a long enough series of observations is likely to approximate the frequency of the event in the world. This account is problematic for several reasons. As De Finetti points out, the choice of what constitutes a “trial”, as well as what constitutes the event itself, and the conditions under which observations are made, are all indicative of background considerations that preclude the possibility of any probability truly being “the” probability of an event. For this reason De Finetti argues that assessments of probability are necessarily subjective, and in some sense are objects of cognition rather than features of the world.

Popper (1935/2002) has similar grounds to object to the frequentist interpretation, but offers a different type of solution, wherein the probabilistic primitive is a two-place conditional relationship rather than a one-place frequency. This is a critical move for understanding how probabilities might be assigned to beliefs: the probability is not the frequency with which an event occurs, but rather the expected likelihood of that event given some set of circumstances. Arló-Costa (2000; Arló-Costa & Thomason, 2001) offers an account of some of the consequences of construing probability in this way. Building on the work of van Fraassen (1995), he frames the issue in terms of the mathematics required to make update sensible when an agent’s set of beliefs are understood as consisting of conditional probabilities in Popper’s sense taken together with the probability kinematics describing how belief states evolve under new information. The central issues of this approach are how to make mathematical sense of full (or absent) belief, and how to appropriately limit the scope of beliefs whose
Experimentally Induced Bias

associated values are affected by the introduction of new evidence. The latter issue brings this approach into close contact with the frame problem and the problem of abduction, both of which are concerns central to the ecological approach.

Though the effort to systematize probabilistic belief update is compelling, it does not address the concern initially raised in the present body of work. From a psychological perspective, the effect of perturbing particular beliefs, as in the experimental contexts described below, is of more immediate interest than a fully general account of belief update that abstracts away from actual beliefs to unspecified sets of them. This is one point of departure from formal epistemology. Another is reframing the idea of the perturbing influence as pressure, as in the Asch paradigm, rather than as new evidence, as is standard in Bayesian and other probabilistic accounts.

Abduction in theory and practice

One attempt to separate the problem of abduction from the issues of epistemology turns on Popper’s (1932/2002) distinction between the logic of discovery and the logic of justification. Popper claimed that the genesis of scientific theories obeyed a different set of rules than those required to defend theories. Specifically, the defense of theories could be required to conform to the hypothetico-deductive method, whereas discovery could be abductive. If abduction belongs to the realm of discovery, then it is not a necessary feature of a formal account of theory justification or theory change.

Applying the distinction between discovery and justification to accounts of belief revision highlights further divisions among those concerned with belief. Several prominent epistemological accounts of belief revision specifically concern revision of scientific theories (e.g. Alchourrón, Gärdenfors & Makinson, 1985) rather than the belief
states of individual agents, with which psychological theories are generally concerned. The distinctions are blurred, however, by psychological experiments whose focus is the kind of propositional commitments that are also the content of theories (e.g. Elio & Pelletier). Further, there is a tension between the psychological project of describing how agents do reason (and perceive, act, and cognize) and the philosophical goal of prescribing how agents should reason, especially when it comes to the questions of when available evidence warrants giving up prior commitments, which is a central problem in theory justification. For an ecological account of belief these distinctions are crucial, as the commitment to direct realism should entail a collapsing of the prescriptive and the descriptive when it comes to perceptually-based beliefs: what an agent does perceive is what the available information specifies for that agent, and is therefore exactly what the agent should perceive in such a case. This is a strong claim that is implicit in ecological theory but that is absent from other psychological theories that admit the possibility of misperception and seek to overcome the destructive mapping between the distal world of objects and the proximal world of sensations. Although there have been recent moves in philosophy to connect with behavioral data on problems of cognition, which indicates a greater degree of overlap between the goals of philosophy and psychology generally, when comparing accounts of belief revision it is critical to bear in mind the goals and explananda of each specific attempt.

*Do animals reason abductively?*

Constructing an ecological theory of belief revision therefore requires keeping in focus several key points that may be absent from other theories. First, that how agents *do* come by their beliefs and perceptions is a proper explanandum of a psychological theory.
Second, that any such theory must accommodate the behavior of organisms across all phyla—it cannot use any facts of biology or cognition that are particular to humans (e.g., Turvey, in press; Turvey & Carello, 2012). Third, that while deductive reasoning may be profitably applied in many situations, the capacity that fuels perception and action is more akin to abduction than to deduction or induction.

Taken together, these concerns point to the necessity of including data from both animal and human behavior as constraints on theories of any psychological phenomena. In fact, ecological theory wishes to be a theory true by force of existence, not one true by force of argument (cf. Shaw, Turvey & Mace, 1982, Turvey, in press). Several results from animal behavior are of particular interest.

Slime molds (*Physarum polycephalum*) show preferences for food sources based on whether the source is in a safe or risky location (dark vs. illuminated) and on how nutritious the food at a particular location is (percent concentration of oatmeal). They prefer to seek more nutritious food in less illuminated locations, but if they must choose between more nutritious food and a safer location, their behavior depends on their degree of hunger. A well-fed slime mold will choose a safer food source, whereas a starving slime mold will choose the more nutritive source. But slime molds also show a behavioral sensitivity to irrelevant options. Where the preference ordering between food sources is clear, as in the case that there are two equally illuminated food sources, but one is more nutritive (10% oatmeal) than the other (5% oatmeal), a slime mold will choose the more rewarding option. But if a third food source is added which should not disrupt the preference ordering (say, a source with only 1% oatmeal) the slime mold will show an
even stronger preference for the preferred option (Latty & Beekman, 2010). It is not clear why this should be the case.

Pigeons have often featured in experiments on operant conditioning. Two such experiments furnish especially interesting results. Thomas and colleagues (Thomas, 1985; Thomas, Stengel, Sherman, & Woodford, 1987) found that when a pigeon is trained to key peck in response to colored lights, changing the angle of inclination of the floor on which the bird is standing disrupts the conditioned behavior. A seemingly irrelevant variable, which is not apparently related to the targeted behavior, turns out to have specific consequences for the animal. Herbranson and Schroeder (2010), also working with pigeons (*Columba livia*), discovered that pigeons are not susceptible to an error humans make systematically in the Monty Hall paradigm (which is described in detail below). In the Monty Hall scenario, the chance of winning a prize concealed behind one of three doors is always maximized by a willingness to switch from a door initially chosen (when there is a one-in-three chance of choosing the door concealing the prize) to the door remaining after one of the two initially unselected doors is opened to show that it does not conceal the prize. While humans systematically fail to behave optimally in Monty Hall scenarios, Herbranson and Schroeder trained pigeons to respond to a three-door Monty Hall problem. After several days of trials the pigeons reliably switched their door choices, thus maximizing their chances of getting the prize (in this case, access to food). These results call for a reformulation of the problems under discussion. One consideration is the possibility that the difference in performance observed between humans and pigeons is due not to any differences in cognitive machinery but to the differences in the training period. Humans may have fewer than 100
trials to learn MHD strategies (Granberg, 1999), whereas the pigeons in Herbranson and Schroeder’s experiment were trained over a 30-day period. Another possibility is that food, being more biologically salient, provides a more effective framework for discovering MHD strategies. Miller and Matute (1996) found that in conditioning experiments, rats exhibit backward blocking (decrease in response when exposed to a previously conditioned stimulus) when the target stimuli have low biological significance that is not observed with target stimuli of high biological significance. In other words, the significance of the potential prize may change the effectiveness of training the strategy.

A final example of contextual activity is provided by jumping spiders (*Portia labiata*). Tarsitano and Andrew (1999) found that jumping spiders are able to visually distinguish between routes to a target that are complete and routes that have gaps and are impassable to the spiders. When the spiders had a choice of two complete routes they showed no preference in selecting between them, but when one complete route and one incomplete route were available the spiders chose the complete route, regardless of what direction of travel the complete route lay in. Jackson et al. (2002) found that jumping spiders make risk assessments in pursuing spitting spiders, which are both predator and prey for the jumping spiders. The spitting spiders in the experiment (*Scytodes pallidus*) are less likely to spit when they are carrying eggs, which are held in front of the mouth. They are therefore less dangerous when carrying eggs. Jumping spiders attempting to prey upon egg-laden spitting spiders were more likely to take a direct route to reach their intended prey, while jumping spiders pursuing spiders not burdened with eggs were more likely to attempt to approach their prey from behind.
What is of interest in these cases is not just that the behaviors demonstrated are relatively complex. Rather, what is most notable is that in each case the relatively complex behavior is produced by an organism with relatively simple neural machinery. This raises the possibility that instances of what is considered high-level abstract reasoning (abduction, risk assessment) may actually be instances of something based in perceptual, not cognitive, machinery.

*Do humans behave rationally?*

Non-human organisms may perform in surprising fashion when confronted with what seem to be cognitive tasks, but it is also the case that humans often do not live up to expectations in tasks that require certain kinds of reasoning. The philosophical literature is rife with instances of humans’ failure to make rational or logically consistent choices when confronted with certain problems. There are several types of hypothesis one could espouse with respect to an explanation of these failures. One is that precision in such cases is too computationally burdensome, and the heuristics that are employed instead lead to imprecise results under certain conditions. Another, related, explanation is that the contingent history of problems confronted and dealt with on an evolutionary scale has led to inefficiencies that are exploitable but rarely have survival consequences. In either case examining the gap between what human agents do and what they “ought” to do if they behaved rationally may offer insights as to the actual character of cognition. For the purposes of the current discussion two examples will be sufficient illustration.

MHD is a classic example of flagrant and persistent failure to act in accordance with the rules of probability. On the game show *Let’s Make a Deal*, the host (Monty Hall) would offer a contestant the choice of opening one of three doors on stage. Behind one of
the door lay a desirable prize such as a car, and behind the other two were (presumably less-desirable) goats. Upon the selection of a door, Monty Hall would show the contestant what was behind one of the two unselected doors. Importantly, he would always choose to open a “goat” door. He would then offer the contestant the chance to switch their pick from their originally chosen door to the remaining unselected door. Probabilistically, a contestant always maximizes their chances of winning the car by switching their pick, yet in experiments that follow this paradigm, people seldom choose to switch. In fact, many people resist the logic of switching even after it is demonstrated or explained to them that switching is always the better bet. This unintuitive result requires an explanation in an account of human reasoning.

Abduction and ecological reliability

What can be gleaned from the discussions above? One possible conclusion that can be drawn centers on the evidence available for addressing a situation. The expectation of what a rational agent will do on some occasion may require certain types of information not usually present, or may hinge on the way in which information is presented. This is the essence of Baratgin and Politzer’s discussion of MHD – the context of the situation makes a great deal of difference to the approach one takes. The same lesson is present in the findings of Latty and Beekman. In these cases the context that confronts the organism results in different patterns of behavior.

Problems like MHD hinge essentially on discovering the correct description of a situation so that an agent confronted with the situation will behave rationally: that features which should be distinct are not conflated and vice versa, that precise strategies and heuristics are not employed erroneously. The conviction that lies behind much of the
literature seems to be that a whole-truth-and-nothing-but-the-truth description will yield rational behavior, and that where rational behavior is not observed it is as a result of a failure to achieve the correct understanding of the matter at hand.

These findings are in keeping with the ecological standpoint. Action is underwritten by perception, and perception is founded on available information and the specifying relations that obtain in a state of affairs. Nothing else is possible. If an additional fact or set of facts would change an organism’s behavior, but those facts aren’t available in that situation, or a distinction might be made on some basis, but that distinction isn’t specified for the organism, it’s fruitless to say the organism isn’t behaving rationally. The neighborhood of ecologically reliable relations may not be the only description available, but it is the only neighborhood an organism’s perceptions will allow it to navigate.
CHAPTER 4

EXPERIMENT 1

The considerations identified in Chapters 1-3 form the basis for Experiment 1. Those considerations, as they bear specifically on Experiments 1 and 2, are schematized in Figure 1. Participants were presented with a forced-choice line judgment task, administered in two parts. In Phase 1, 400 pairs of lines oriented horizontally were presented on a laptop monitor, one pair at a time. Participants were asked to choose which line in the pair was longer, and to register their confidence in their choice. They responded by using the laptop trackpad to click on the screen. Both responses were given simultaneously—the choice of line by which side of the screen participants clicked on, and the confidence by the click location’s horizontal distance from the center of the screen. In this way participants’ responses allow for an assessment of both species of belief under discussion (see Chapter 2): (a) the choice of longer line constitutes a participant’s disposition to behave as though something is the case, e.g. to act in accordance with the belief that the left line is longer; (b) the degree of confidence in the choice provides insight into the extent of the disconnect between the disposition to act and the willingness to assent to a proposition, e.g. “I believe that the left line is longer”. Importantly, a participant’s response on one measure is independent of their response on the other—it is possible for the participant to make a choice and have very little confidence in the choice.
Intuitively, choosing which of a pair of lines is longer is not a task about which there may be much doubt, if the lines differ sufficiently. Therefore, half of the presented pairs differed by less than a JND, so as to provide participants with opportunities for doubt. The JND for line length had been well established experimentally as being approximately three-hundredths (.03) of the length of a reference line. However, initial attempts using that value yielded results indicating that participants performed better than expected, indicating that they were able to perceive smaller differences in line lengths. Different values were tried, until pre-pilot volunteers’ performance was equal to chance, a value of .019 times the reference length.
As noted, the experimental design is a conceptual extension of the Asch (1955) paradigm. Participants in Experiment 1 experienced pressure on their beliefs, in the form of “score” applied on the less-than-JND trials during Phase 1 of the experiment (see Chapter 2). To reiterate, “score” was designed to introduce a disposition to respond that the left line was longer or that the right line was longer (see Figure 1).

Participants

Participants were 15 graduate students from the University of Connecticut who volunteered to participate. Participants were randomly assigned to one of two groups, left-biased (Group L, 8 participants) or right-biased (Group R, 7 participants).

Stimuli and procedure

The line pairs were of two kinds, less than a JND difference and more than a JND difference. There were 280 instances of the former kind and 220 instances of the latter kind. Participants viewed the line-pair stimuli (generated by a Matlab program) on a MacBook laptop. They used the laptop trackpad to guide an onscreen cursor in order to make their responses (see Chapter 2 for details).

Design

Following the basic methodology introduced in Chapter 2, Phase 1 and Phase 2 (there was no Phase 0) were defined over five blocks of 100 trials: Blocks 1-4 with bias and “score” (Phase 1), and Block 5 without “score” and with bias reversed (Phase 2). For Group L in Phase 1, 80% of the line pairs were left-line-longer. For Group R in Phase 1, 80% of the line pairs were right-line-longer. “Score” incremented or decremented according to the following rule: For Group L, whenever a participant clicked on the left side of the screen (indicating a left-line-longer choice), “score” for the participant, with
probability .8, would increase by 10 points on that trial, *regardless of whether the left line was actually longer on that trial*. If the participant clicked on the right (indicating a right-line-longer choice), then with a probability of .2 the participant’s score would increase by 10, again *regardless of which line was longer*. Mutatis mutandis, for Group R, “score” for the participant incremented by 10 with probability .8 for right clicks and probability .2 for left clicks.

“Score” decremented whenever it did not increment. Crucially for the experiment, “score” only changed on less-than-JND trials. On greater-than-JND trials “score” remained unchanged from the most recent less-than-JND trial. This was done to maintain the plausibility of “score”, which was targeted only at those line pairs where the fact of the matter could be in doubt.

Phase 2 of the experiment, Block 5, consisted of an additional 100 horizontally oriented line pairs. It was distinguished from Phase 1 by the absence of “score”. It was further distinguished by a different bias, a different proportion of left-line longer pairs. For Group L, the pairs on Block 5 were 30% left-line-longer. For Group R, the pairs in Block 5 were 30% right-line-longer. The proportions were changed in order to determine whether the bias introduced would persist in the absence of “score”. A reversal of proportion was chosen, instead of presenting an unbiased block, in order to make distinguishable whether participants responded in keeping with the stimulus properties, in keeping with the bias introduced in Phase I, or in keeping with a guessing strategy. Finally, the proportion of less-than-JND pairs was increased on Block 5 from 50% to 80%, in order to provide more of the trials of interest.
Results and Discussion

*Left-line clicking.* This measure indexes the influence of the bias that distinguished Groups L and R. Figure 2 shows that the number of left-line clicks varied systematically as a function of Group, Block, and JND, $F(4, 52) = 23.47, p < .0001, \eta^2 = .08$.

![Figure 2. Left clicks as a function of Block, Group (squares Group L, circles Group R) and JND (open < JND, closed > JND) in Experiment 1.](image)

*Accuracy and confidence.* Figure 3 compares performance accuracy with performance confidence (measured in pixels; see Chapter 2). Inspection of the accuracy data suggests indifference to both Group and Block ($F$s < 1) for both levels of JND. In comparison, inspection of the unsigned confidence data (lower panel) suggests a decline with Block, $F(4, 52) = 3.47, p < .014, \eta^2 = .046$, a decline that was at the same rate for both Group L and Group R.

![Figure 3. (Left) Mean Correct Responses and (Right) Mean Confidence Measure as functions of Block, Group and JND in Experiment 1.](image)
Latency. Figure 4 shows that latency of responding declined with Block, \( F(4, 13) = 9.46, p < .001, \eta^2 = .16, \) and did so at a faster rate for JND > 1, \( F(4, 52) = 10.76, p < .0001, \eta^2 = .028. \)

*Figure 4.* Latency of choice as a function of Block and JND in Experiment 1.

If participants respond based on stimulus properties alone, one would expect to see the present distributions reflected in the greater-than-JND trials (participants correctly choosing the longer lines) and responding with equal probability on the left and right in the less-than-JND trials (guessing). If “score” affected participants’ disposition to act (choice of longer line), one would expect that participants would respond in accordance with “score” on the less-than-JND trials. If “score” affected participants’ assent to a belief that a particular line is longer, one would expect participants’ confidence in the judgments made to vary with whether or not the response agreed with the disposition introduced by “score” (higher confidence in responses concurring with “score”, lower confidence in responses opposing “score”).
Generally, results indicated that the introduction of “score” was sufficient to alter participants’ responses over the course of Phase 1, and that the effect persisted somewhat in Phase 2. Though the Phase 2 left-click frequencies differed, the particular aftereffect that the bias introduced in Phase 1 might have had could not be revealed using ANOVA.

Experiment 1 was aimed at fulfilling several criteria for an empirical investigation of belief. First, it attempted to elicit responses on two separate construals of belief—belief as assent and belief as action. Second, it incorporated stimuli with probabilistic properties, as a means of assessing the empirical adequacy of probabilistic models of belief and belief change, which are common in both psychology and epistemology. Third, it constituted a step toward drawing a connection between the Asch study, which has been primarily discussed in the context of social influence on behavior, and a perception-action perspective whereby behavior may be influenced by many types of conditioning, learning, and more-or-less reliable information.
Experimentally Induced Bias 35

CHAPTER 5

EXPERIMENT 2

Experiment 2 extended Experiment 1 to include Phase 0 in the form of a single block of 100-pairs presented without a bias and in the absence of “score”. Its purpose was to establish participants’ baseline rates of clicking on the left and right sides of the screen. Experiment 2 also expanded Phase 2 of Experiment 1 from one block to three blocks of 50 trials. Additional modifications included a three bins classification of JND (less-than-JND, greater-than-JND, and a range from 0.5JND to 2.5JND identified in the figures as approximately 1JND) and changing the Phase 2 frequencies of Experiment 1 to 50% right- and left-line-longer pairs for both groups. The reclassification of trials into three JND bins was motivated by the method of determining the value of the JND for line length used throughout the experiment. The JND value established in pilot work was the value where participants performed equal to chance, rather than a hard cut-off. This raised the question of whether some of the variance in the effect of the line length difference might be missed by treating the line pairs as strictly above- or below-JND, and a slightly finer-grained analysis was therefore employed. The Phase 2 frequencies were changed from the frequencies used in Experiment 1 because the analysis of Experiment 1 showed no convincing reason why the two groups should be treated separately in Phase 2. Experiment 2 was conducted using the same Matlab program that was used for Experiment 1, with modifications as described below.

Participants and procedure

Participants were 30 undergraduate students from the University of Connecticut who received course credit for participation. All participants completed the Edinburgh
Handedness Inventory before beginning the experiment. This constituted the only
demographic information collected from the participants. Participants were randomly
assigned to either the left-biased (Group L) or right-biased (Group R) condition.

Apparatus and procedure were the same as in Experiment 1.

**Design**

There were eight blocks, one in Phase 0, four in Phase 1, and three in Phase 2. In
Block 1 100 line-pairs were presented, unbiased (50% of pairs were left-line-longer, 50% of pairs differed by less than the JND) and without “score”. Blocks 2-5 were the same as in Experiment 1, with each block containing 100 line pairs, of which 50% were below-JND and either 20% (right-biased group) or 70% (left-biased group) of line pairs were left-line-longer, and in which participants received “score”. In contrast, Blocks 6-8 consisted of 50 line-pairs without bias and without score.

**Results**

Several analyses with somewhat different foci were applied to the data. The initial analysis took the form of four parallel 2 (Group) × 8 (Block) × 3 (Well-above, Just-above- or Below-JND) ANOVAs conducted on the dependent measures of latency, confidence, correct responses, and number of responses on the left side of the screen (a measure of the efficacy of the introduced bias).

**Latency.** Neither the main effect of group nor any interaction with group was significant, $F s \approx 1$. The main effect of JND, $F(2, 56) = 22.65$, $p < .0001$, $\eta^2 = .06$, indicates that latencies were shorter when the two lines were obviously different in length ($M = 1.63$ s) than when they were close in length ($M = 2.02$ s and 1.99 s for just below and just above the JND). The main effect of block, $F(7, 196) = 23.00$, $p < .0001$, $\eta^2 = .23$,
indicates that latencies became shorter with experience in the task but with a slight lengthening when “score” was no longer provided. The Block × JND interaction, $F(14, 392) = 3.15, p < .0001, \eta^2 = .008$, indicates that the latency difference across differing JND trials diminished with experience in the task, again with a slight increase when “score” was first eliminated.

*Signed Confidence.* Confidence was measured in screen pixels, with sign indicating left (–) or right (+) and size indicating confidence in that judgment. Because of the sign difference, the main effects of group, $F(1, 28) = 6.33, p = .018, \eta^2 = .11$, and JND, $F(2, 56) = 5.49, p = .007, \eta^2 = .024$, should be interpreted in light of their interaction, $F(2, 56) = 14.46, p < .0001, \eta^2 = .064$. In particular, confidence was highest—clicks were farther left or farther right—when line lengths were most distinct. The main effect of block was significant, $F(7, 196) = 4.63, p < .0001, \eta^2 = .05$, as were its interactions with group, $F(7, 196) = 11.89, p < .0001, \eta^2 = .13$, and with JND, $F(14, 392) = 5.24, p < .0001, \eta^2 = .058$. The Block × Group × JND interaction, $F(14, 392) = 6.16, p < .0001, \eta^2 = .07$, indicates that confidence was diminished for blocks without “score”, especially when lines were close in length.

*Correct Responses.* The main effects for group and block were not significant. The main effect of JND, $F(2, 28) = 291.5, p < .0001, \eta^2 = .9$, indicates that participants made more correct responses on the well-above-JND trials ($M = .9, SD = .12$) than on the below- ($M = .54, SD = .13$) and just-above-JND trials ($M = .6, SD = .11$). The JND × Block interaction, $F(14, 392) = 2.54, p < .01, \eta^2 = .017$, indicates that while the proportion of correct responses remains higher for the well-above-JND trials throughout the experiment, the proportions of correct responses for the just-above-JND and below-
JND trials are initially different but converge over the course of Phase 1 of the experiment before once again diverging in Phase 2. Figure 5 summarizes the results for accuracy and (absolute) confidence.

![Figure 5. Accuracy (Top) and Confidence (bottom) as a function of Block and JND in Experiment 2.](image)

**Left-line clicking.** The main effect of group, $F(1, 28) = 9.54, p = .0045, \eta^2 = .37$, indicates a higher proportion of left-line clicks for Group L ($M = .57$) than Group R ($M = .41$). The main effect of JND, $F(2, 56) = 4.68, p = .013, \eta^2 = .032$, indicates a higher proportion of left responses when line lengths were less distinct ($M = .51$ and $M = .50$ vs. $M = .46$). The main effect of block, $F(7, 196) = 5.83, p < .0001, \eta^2 = .106$, indicates that the proportion of left responses diminished with experience in the task but with a resurgence when “score” was no longer provided. The Block × JND interaction, $F(14, 392) = 4.70, p < .0001, \eta^2 = .066$, indicates that the resurgence was due to different patterns as a function of JND level: The proportion of left responses shot up for distinct lines pairs and then diminished again but for the close-to-JND pairs it gradually increased. Finally, the Block × Group × JND interaction, $F(14, 392) = 6.46, p < .0001, \eta^2 = .09$, indicates that the diminishing of left responses was due to Group R when a score was provided, and to the farthest from JND levels when the score was no longer provided.
Figure 6. Proportion of clicks on the left line as a function of Group and Phase in Experiment 2. (Phase 0 – Block 1; Phase 1 – Blocks 2-5; Phase 2 – Blocks 6-8)

All of the above outcomes can be conveniently communicated by a focus on Phase. A minimalist analysis takes Block 1 as representative (obviously) of Phase 0, Block 2 as representative of Phase 1, and Block 6 as representative of Phase 2. Figures 6-8 present the results in terms of the foregoing phases.

Figure 7. Accuracy (proportion correct) as a function of Phase and JND (binned) in Experiment 2.

The important interactions are (a) Group × Phase, $F(2, 56) = 15.6, p < .001$, for Left-side clicking (demonstrating that the manipulations of bias and “score” worked, see
Figure 6), (b) JND × Phase, $F(4, 112) = 3.9, p < .005$, for proportion correct (Figure 7), suggesting that accuracy depended on phase (alias context), and (c) JND × Phase, $F(4, 112) = 9.05, p < .0001$, for confidence, suggesting that confidence declined from Phase 0 to Phase 2 in a JND-dependent way (Figure 8).

![Figure 8](image)

**Figure 8.** Confidence as a function of Phase and JND (binned) in Experiment 2.

Discussion

Each of the behavioral aspects reported (latency, confidence, left-side clicking, and correct responses) carries a different part of the experimental hypothesis. The proportion of correct responses is informative about the extent to which participants are successful with respect to the stimuli: that is, whether they accurately perceived the difference in line lengths. The proportion of left responses is informative about whether the manipulation was successful: that is, whether participants responded in keeping with the bias introduced in Phase 1. (For the two preceding forms of successful, refer to “Key elements” in Figure 1.) If the manipulation were not successful, the expectation would be that participants would respond on the side with the longer line in the above-JND trials but would respond with equal likelihood on the left and right in the below-JND trials.
(essentially, they would guess). A successful manipulation, however, should bias the participants toward responding on either the left (Group L) or the right (Group R), especially in the below-JND cases. Taken together, the proportion correct and the proportion of left responses represent the species of belief that is more overt: the disposition to act as though a state of affairs obtains.

Confidence is, naturally enough, informative of the participants’ confidence in their responses: how sure they are that the action they undertake on each trial is the appropriate one. Latency provides a more subtle measure of confidence, one that is not generated consciously by the participants. The latency measure also goes some way towards teasing apart participants whose average confidence is low because their confidence is low (in which case latency would be expected to be large) and participants whose average confidence is low because they are using a relatively narrow range of the confidence measure (in which case latency would be expected to be smaller). Taken together, confidence and latency represent the species of belief that is propositional: holding that a proposition is true, or being prepared to assent to a proposition such as “I believe that the left line in this pair is longer”. Because confidence is a continuous measure it is in keeping with the idea of propositional beliefs as being the sorts of things that carry weights.

Throughout the experiment Group L made significantly more left responses than Group R. This demonstrates that the primary stimulus manipulation is effective: participants are capable of responding in keeping with the stimulus properties. Further, the significant difference between groups of the number of left responses on the below-
JND and near-JND trials indicates that the introduced bias affects responses in the cases where what response is correct is not obvious.

A less straightforward pattern of left responses is observed when taking Block into account. Throughout Phase 1 (Blocks 2 through 5) Group L participants responded more on the left, while Group R participants responded more on the right. In Phase 2 (Blocks 6 through 8), however, Group L’s rate of responding on the left decreased smoothly, while Group R showed a different pattern: in Block 6, Group R began to respond on the left. In Blocks 7 and 8 Group R’s responses on the right again increased. The effect of the manipulation was therefore successful (participants responded on the side to which they were being biased), but more complex than expected. The pattern of responses made by Group R bears a similarity to responses observed in prism experiments (see Blau, Stephen, Frank, & Turvey, 2009), wherein an effect of an adaptation to context is observed while the context is present, and a rebound effect is observed when the context is removed. While this phenomenon may be of like kind, no explanation for why the effect should be observed in Group R but not Group L immediately suggests itself. One possibility, briefly considered in Experiment 2A (below), is that handedness plays a role in differentiating responses on the right side of the screen from responses on the left, and may constitute an additional effect of context.

In keeping with the main conceptual argument—that belief as the basis for action and belief as a proposition to which assent or a weighted assent might be assigned—confidence and the proportion of correct responses are best interpreted together. Whether a correct response is given is the experimental measure of belief as the basis for action, while the confidence rating is the measure of the degree of assent to the proposition
Experimentally Induced Bias 43

represented by the line length judgment given on a trial. Rates of correct responses are highest for the above-JND condition, though they are not perfect. This indicates that even for obviously different line-pairs errors are still made. The lack of an effect of block indicates that rates of correct response also remained fairly constant throughout the experiment, though the significant interaction of JND and block shows that, while the below-JND and near-JND trials were initially treated as different by the participants, as the experiment went on they became more similar.

Unsurprisingly, confidence ratings were highest for the above-JND trials, where the line differences were distinct. However, confidence ratings decreased for all JND bins throughout the experiment. In particular, confidence decreased in Phase 2, when score was removed. This is particularly interesting when considered in tandem with the fact that the proportion of correct responses remained fairly constant across the entire experiment. These two results support several of the intuitions whose investigation motivated the present experiments. Most importantly, these results indicate that it is possible to experimentally distinguish between an action and the actor’s confidence. It is this component that was not part of Asch’s (1951, 1956) data, but which on our reading of the dual nature of belief is critical to understanding the way in which propositional beliefs are not fully determinative of action.

Finally, the results for latency indicate straightforwardly that participants are quicker to respond to more distinct stimuli, and their responses are given more quickly as the experiment progresses, presumably as a result of increased facility with the experiment.
EXPERIMENT 2A

Experiment 2 revealed a clear distinction in consistency of observed data between the left-biased and right-biased groups. Participants in the left-biasing group showed less within-group variation in terms of bias-side click frequency than participants in the right-biasing group. Because the preponderance of participants in Experiment 2 was right-handed, a first hypothesis was that the left-biasing condition constituted greater pressure on participants, perhaps moving them through a phase transition. Moving across the midline of the body is more effortful (Rosenbaum, van Heugten, & Caldwell, 1996), and right-handed participants using their right hand to interface with the computer via the touchpad experience less of this effort in moving the cursor to or on the right side of the screen than moving the cursor to or on the left side of the screen. To test this hypothesis five left-handed participants were recruited. Each participant in this left-handed subgroup was instructed to use their left hand to interface with the touchpad regardless of which hand they would ordinarily use. All of the left-handed subgroup completed the right-biasing condition of the experiment. Their results serve as a contrast to the largely right-handed participants from Experiment 2. Because this subgroup was being biased to move across the midline of the body, the frequencies of their bias-side clicks were conjectured to more closely resemble the left-biased group from Experiment 2. The data did not offer any substantive support for the conjecture. It is noteworthy that some participants reported spontaneously that they ordinarily used their right hands for computer mice. On questioning, all concurred that right hand usage was the norm and was apparently so in this small group even for those who scored strongly left-handed on the Edinburgh
Handedness Inventory. The reason for commonplace right hand usage at the computer is that this is a standard setup in many school and work environments.

EXPERIMENT 2B

In an additional attempt to clarify the potential handedness factor in Experiment 2, a version of the experiment was conducted with vertically-oriented lines stacked on top of each other, wherein participants responded using the top and bottom of the screen rather than the left and right sides. Five participants were tested. The results did not suggest any obvious difference between performance in this up-down arrangement and the performance in the left-right arrangement of the main experiment. Both performances might be regarded as reflections of the kind of phenomena studied under the label “S-R compatibility”.
CHAPTER 6
EXPERIMENT 3

Experiment 3 was a version of the Monty Hall dilemma (MHD). In the classic Monty Hall, players are confronted with a scenario in which they must choose between three doors, two of which conceal “goats” (undesirable prizes), and one of which conceals a “car” (a desirable prize). Though they do not see what’s behind the door they choose at the outset, they are shown what’s behind one of the remaining two doors. Per the rules of the game, the door that is opened is always shown to be a “goat” door. The player then has the opportunity to switch their pick to the remaining unopened door (“switching”) or to stay with their original door choice (“sticking”).

Famously, the “winningest” strategy to resolve MHD is to switch. Equally famously, sticking is by far the more popular strategy (see e.g. Granberg, 1999). In fact, the sticking strategy is so entrenched that even when the optimal strategy is explained people resist believing it. The explanation of the winning strategy is as follows: Initially, the probability that the prize lies behind any given door is 1/3. Therefore the initial choice of door may as well be made at random. The probability that the prize lies behind one of the unchosen doors is 2/3. In the second stage of the scenario, as the “goat” door is opened, the probability that the remaining unchosen door is the “car” door has essentially doubled due to the removal of the other unchosen door as a possibility. But the probability of the first door remains unchanged, as no new information about it has been introduced by the opening of the “goat” door. Therefore the remaining unchosen door has a higher probability at the end of the scenario of being the “car” door, even though the probabilities are equal at the outset.
Baratgin and Politzer (2010) advanced the hypothesis that both the dilemma and the resistance to the winning strategy are rooted in the manner in which information is presented throughout the scenario. They argue that the opening of the “goat” door is taken to be relevant to the overall probabilities, and that players do not effectively isolate the new information to the real space of outcomes, which includes only the unchosen doors. They presented several versions of MHD and concluded that when information is presented as being relevant only to the unchosen doors, people do not commit the typical MHD error. Granberg (1999) tested a 4-door version of MHD with unequal probabilities, where the optimal strategy is to choose the lowest probability door initially and then switch to the highest probability door remaining after the reveal stage. Participants in Granberg’s study were explicitly informed of the door probabilities on each trial, yet even so no participants reliably used the optimal strategy.

MHD presents a classic problem in biased reasoning. Because it has straightforward probabilities and a well-understood set of strategies it provides a good starting point for an experiment investigating the effect of introducing bias, thereby altering the probability distribution of stimuli. In that sense it provides an extension of Experiments 1 and 2. In order to more closely tie MHD to the first two experiments, we constructed a version conditioned on a perceptual task—in this case, color discrimination.

**Method**

**Participants.** Ten undergraduate students (mean age = 22.4, 5 male, 5 female) from the University of Connecticut participated in the experiment. They received either $10 or course credit for participating. All participants completed the Edinburgh Handedness Inventory, which categorized all participants as right-handed.
Materials. Participants viewed MHD stimuli on a MacBook laptop. They used the laptop trackpad to guide an onscreen pointer in order to make their responses.

Design. Experiment 3 used stimuli presented to participants using a webpage designed with Adobe Flash. It consisted of a modified MHD, using a standard 3-door case. Behind each door was a square colored a different shade of red. The red values were randomly generated from a range of RGB values. The “prize” was the brightest of the three shades of red generated on each trial. Although a line-length version of MHD was considered, pilot work indicated that the range of indiscriminability around the JND, which was an advantage in Experiments 1 and 2, made it too difficult for participants to find the “prize” lines and discover MHD strategies. Color was therefore chosen as a stimulus property that is both salient to participants and easy to program. While the standard MHD has a unique “prize” and two identical “goats”, i.e. losing choices, the version of the MHD employed for this experiment had 3 distinct stimuli. This was done because participants were making comparative rather than categorical judgments, though it may have affected the outcome of the experiment (discussed below).

Experiment 3 had the same structure as Experiment 2: an initial 100-trial block with an equal chance of the “prize” being behind each of three doors, followed by 4 100-trial blocks with unequal probabilities of the “prize” being behind each of the doors, and 3 50-trial blocks with equal probabilities. In the unequal-probabilities blocks (Blocks 2 through 5, composing Phase 1), either the leftmost or rightmost door had a higher chance of concealing the “prize”. Participants were randomly assigned to either the left-biased (Group L) or right-biased (Group R) condition. Participants were asked both to select a door by clicking on it, and also to register their confidence in their choice, measured in
how far from the center of the chosen square in any direction they clicked, while still making the response within the confines of the door. This lack of directionality was included specifically to avoid the sort of laterality effects that were observed in Experiments 1 and 2, and also to simplify the verbal instructions to the participants. The center represented the highest possible confidence and the inside edge of the door represented the lowest confidence. In all there were 650 trials for each participant.

Procedure. Before beginning the experiment participants saw an instruction screen, and were told that they could begin the experiment by clicking anywhere on the screen. Each individual trial consisted of an instance of MHD, as follows: three “doors” (square boxes) were displayed on screen, along with the text prompt “Select a door”. A participant would click on one box, whereupon the color behind one of the remaining doors (never the prize door) would be revealed. The participant would then be asked to stick or switch, via the text prompt: “Switch your pick?” They would then either switch (by clicking on the unopened door they did not originally choose) or stick (by once again clicking on the originally-chosen door). The color behind the door they ultimately chose would then be revealed, along with text telling them either “You win” or “You lose” depending on the outcome of their final choice. At the end of the experiment participants were asked to briefly describe any strategy they might have employed in the course of the experiment.

Results

The focus was on the pattern of the 1st and 2nd responses comprising a trial. A 2 (Response) × 2 (Group) × 8 (Block) ANOVA was conducted on the confidence measure (derived as a distance from the center of the screen). This distance was a shallow U-
Experimentally Induced Bias

shaped function of block for both responses, $F(7, 56) = 3.54, p = .003, \eta^2 = .073$. No other main effects or interactions reached significance, all $Fs \approx 1$ except the Block × Group interaction, $F(7, 56) = 1.61, p > .15, \eta^2 = .033$.

A parallel ANOVA was conducted on the proportion of correct 1st and 2nd responses. As shown in Figure 9, the main effect of Response, $F(1, 8) = 12.30, p = .008, \eta^2 = .41$, indicates a higher proportion of correct 2nd responses ($M = .66$) than 1st responses ($M = .42$). The main effect of block on proportion correct, $F(7, 56) = 10.09, p < .0001, \eta^2 = .4$, was realized as an inverted U-shaped function of block. No other main effects or interactions reached significance.

![Figure 9. Correct 1st and 2nd responses in a MHD trial as a function of Block.](image)

In a trial the 2nd response may repeat the 1st response (a case of *hold*) or not (a case of *switch*). A $2 \times 2 \times 8$ ANOVA was conducted on the distance of response from center, as an indicator of confidence. The distance measure associated with the 2nd response (hold or switch) was a shallow U-shaped function of Block, $F(7, 49) = 2.43, p = .032, \eta^2 = .081$ (Figure 10 left). No other main effects or interactions reached
significance. A parallel ANOVA was conducted on the proportion of correct holds and switches. Of the two strategies, switch ($M = .68$) was associated with a higher proportion of correct responses to MHD than hold ($M = .44$), $F(1, 7) = 5.62, p = .05, \eta^2 = .37$ (Figure 10 right).

![Figure 10. Confidence (left) and accuracy (right) in the MHD task as a function of Block and Strategy (hold where 2nd response = 1st response, or switch where 2nd response ≠ 1st response.]

A Block × Hold interaction, $F(7, 49) = 2.71, p = .019, \eta^2 = .172$, suggests that the Block effect characterized only those trials when participants remained with their first choice ($F = 9.70, p < .0001$). The proportion of correct 2nd responses on those trials in which participants switched from the 1st response was unaffected by Block ($F < 1$) remaining high across Blocks 1-8. The main effect of group, Group L versus Group R, was not significant, $F = 1$. The effect of group was seen in the interaction with block, $F(7, 49) = 2.91, p = .013, \eta^2 = .12$, as Figure 11 makes evident: the inverted U-function was more characteristic of Group R. No other interactions reached significance, $Fs < 1$. 
Experimentally Induced Bias

Figure 11. Accuracy in the MHD task as a function of Group and Block.

A 2 (1\textsuperscript{st} and 2\textsuperscript{nd} response) × 2 (Group) × 8 (Block) ANOVA was conducted on the choice of door (1 = left, 2 = middle, 3 = right). The results are shown in Figure 12. The main effect of group, \(F(1, 8) = 21.39, p = .0017\), indicates that Group L was more likely than Group R to choose the left door (\(M = 1.58\) vs. \(M = 2.20\)). The main effects of Trial and Block were not significant, \(F_{s} \approx 1\), but their effects were seen in interactions. The Block × Group interaction, \(F(7, 56) = 16.56, p < .0001\), indicates a strong directional bias to the unequal-probabilities blocks. The marginal Response × Block interaction, \(F(1, 8) = 4.59, p = .065\), and the significant Response × Group × Block interaction, \(F(7, 56) = 2.90, p = .012\), indicate that the directional bias during the unequal-probabilities blocks is stronger on the second response. This characterization was verified with post hoc \(t\)-tests (\(df = 8, p = .05\) requires \(t > 2.3\)). There was no directional bias for the first responses of Blocks 1, 7, and 8, \(|t| < 1\); the first responses of Blocks 2 through 6 showed a directional bias (\(t = -2.81, -2.42, -2.43, p < .05; -1.78, -1.33, p > .05\)). For the second responses, Block 1 showed a directional bias in the opposite direction, \(t = 3.23, p = .012\). The second responses of Blocks 2 through 7 showed a directional bias (\(t = -7.38, -8.75, -7.04, -16.64, \text{ etc.} \).
Experimentally Induced Bias 53

$p = .0001; -1.31, p > .05; -3.12, p = .014)$. The directional bias was eliminated on Block 8, $|r| < 1$.

Figure 12. 1st and 2nd response as a function of Group and Block in Experiment 3.

Participants were characterized as ‘holders’ or ‘switchers’ by dividing them into quartiles based on the percentage of trials per block on which they switched. This was used as the basis for an ANOVA comparing the quartiles on confidence and proportion of correct responses. For the purposes of this ANOVA, only 2nd responses were included. Switch quartile significantly affected performance, $F(3, 76) = 4.698, p < .01, \eta^2 = .156$, but did not significantly affect confidence. Participant blocks falling into the lowest quartile (fewest switch trials) had the highest performance ($M = 67.4, SD = 29.3$). These values did not significantly differ from the performance of those in the highest quartile (most switch trials), which had the next highest performance ($M = 63.7, SD = 27.3$). Post hoc analyses showed that these blocks differed significantly only from the second quartile, $p < .01$.

Discussion

The winning MHD strategy is switching, but as described above, a preponderance of players not only fail to discover or employ this strategy, but resist acknowledging that it
Experimentally Induced Bias

is better even after having it explained. As Granberg (1999) reported, in an unequal-probabilities MHD the best strategy is even less obvious: to optimize the chances of winning a player should initially select a low-probability door and then switch to the highest remaining probability door after a “goat” has been revealed. In the equal-probabilities case, participants may only adopt two strategies: holding or switching. In the unequal-probabilities case, there are four potential strategies that a participant might be expected to adopt (see Table 1). The strategies shown in the top two rows of the table are parallel to those available in the standard MHD (hold and switch), but the outcomes are particular to the unequal-probabilities variant. Those in the bottom two rows of the table are unique to the unequal-probabilities MHD. The principal difference in strategy between the two versions of the dilemma is that in the standard case, there is no advantage to initially selecting one door over another. Players therefore begin with a random door selection, and the win probabilities associated with each strategy are not dependent on which door is selected. In the unequal-probabilities case, if the player chooses at random (that is, does not choose on the basis of an assessment of the door win probabilities), then a given turn, from the player’s perspective, unfolds in the same way as a turn of the standard MHD. But as the top half of the table illustrates, from the perspective of the game (or the experimenter), the outcome depends on whether the door initially selected by the player is, in fact, a high-probability or low-probability door.

The breaking of the outcome symmetry among doors in the unequal-probabilities MHD further results in the availability of the two strategies described in the bottom half of the table. In these strategies players choose a particular door based on their assessment of the door probabilities. The data from Experiment 3 indicate that by the end of Phase 1
Experimentally Induced Bias

Players are able to capitalize on the presence of the high-probability door, which suggests that players in the unequal-probabilities MHD of Experiment 3 can be usefully characterized as holders (who will select the high-P door and stick to it) and switchers (who will select a low-P door and switch away from it) on the first choice of the trial. This is for two reasons. First, as stated above, although it is possible for players to choose a door at random, it appears as though they are able to coordinate with the unequal-

<table>
<thead>
<tr>
<th>Door choice</th>
<th>P(win)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Definitional</strong></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>Hold</td>
<td>random low-P</td>
</tr>
<tr>
<td></td>
<td>random high-P</td>
</tr>
<tr>
<td>Switch</td>
<td>random low-P</td>
</tr>
<tr>
<td></td>
<td>random high-P</td>
</tr>
<tr>
<td>Max door probability</td>
<td>intentional high-P</td>
</tr>
<tr>
<td>Optimal</td>
<td>intentional low-P</td>
</tr>
</tbody>
</table>

*aBold text shows standard/general MHD logic; italicized text shows logic with the biasing introduced in Experiment 3. Note that biasing allows a best guess for choosing (or avoiding) the high-P door with the first choice.*
probabilities structure of the game, and therefore it is unlikely that they are making their initial choice at random. Second, because their overall strategy is dependent on the initial selection, so long as the assertion that players know there is a high-probability door holds, the initial selection is sufficient in the unequal-doors case to characterize a player as a holder or a switcher.

In the present experiment, the win probability associated with the third strategy (choose high-probability door and stick) is .8, while the win probability associated with the fourth strategy (choose low-probability door and switch to high-probability door if it remains) is .9. It is therefore clear that the fourth strategy remains the optimal one, but the difference in win percentages is small.

In Granberg’s 4-door unequal-probabilities study participants were informed of the probabilities of finding the prize behind each door. In the present experiment participants were not informed of the probabilities at any point. The fact that both groups more frequently selected the highest-probability doors (leftmost for Group L, rightmost for Group R) is therefore indicative of participants’ ability to coordinate with the structure of the experimental task, but is not in itself indicative of the strategy a participant may have adopted.

Only block significantly affected confidence. In this experiment confidence is measured as distance from the center of the chosen door: The further from the center a response is made, the less confident the participant was in that response. Therefore the greater the distance, the lower the confidence. Distance decreased after the first block before increasing again over the final blocks (Figure 10 left). This indicates that confidence increased initially before decreasing at the end of the experiment. This can be
considered an indicator of participants’ level of comfort with the task: After having initially been exposed to an equal-probabilities MHD in Block 1 (Phase 0), participants grew more comfortable with the task in Blocks 2 through 5 (Phase 1). However, confidence decreases at the end of Phase 1 and through Phase 2. Here participants found themselves confronted with new task constraints, as the unequal-probabilities MHD they had been solving reverted to the equal-probabilities case and its attendant strategies.

The results of the performance measure (whether the correct door was chosen) show a similar pattern to the confidence results (Figure 10 right). Initially, there is no difference in performance between groups. Over the course of Phase 1 (the unequal-probabilities MHD) performance increases before decreasing again in the final three equal-probabilities blocks (Phase 2). However, the higher proportion of correct responses on 2nd responses over 1st responses evident in Figure 10 (right) shows that switching remains the most effective strategy even in the unequal-probabilities case. This fact is also indicated by the fact that an overall higher proportion of correct responses are seen in 2nd responses that are switches versus 2nd responses that are holds.

The analysis in terms of hold versus switch trials also makes sense of the overall higher performance in the unequal-probabilities blocks. The significant Block × Response interaction may be attributed to the fact that having a high-probability door reduces the difference in success between hold and switch trials, leading to an overall increase in the likelihood of selecting the winning door on a hold trial in Phase 1.

Importantly, confidence does not vary with performance (Figure 10 left). “ HOLDERS” and “SWITCHERS” show almost identical patterns of average confidence, even though switching is the more effective strategy throughout the experiment, and even though
participants are aware of their performance (by being told “You win” or “You lose” after every trial). This finding was further supported by the analysis in terms of switch quartiles, which showed that how often a participant switched in a given block significantly affected performance but not confidence. This independence of confidence and performance lends support to the hypothesis that the aspects of the task about which a participant might have propositionally-construed beliefs are not necessarily the aspects that are used to guide behavior (Anderson & Runeson, 2008; Runeson, Juslin, & Olsson, 2000). It is important to note that performance did not significantly differ between the most and least frequent switchers. As noted above, in this particular version of the MHD the win percentage of choosing the highest probability door is very close to the win percentage of adopting the “winningest” strategy. Therefore in this particular instance it is not altogether surprising that the overall benefit of adopting the “switch” strategy does not manifest clearly. However it should be emphasized that as a direct consequence of the unequal probabilities MHD, adopting either strategy, which would lead to a participant being in the highest switch quartile (switch strategy adopters) or lowest switch quartile (hold strategy adopters), yields higher performance than adopting neither strategy (as we may characterize those who fall into the middle quartiles).

The existence of the directional bias over the unequal-probabilities blocks (Blocks 2-5 composing Phase 1) showed that participants have a general ability to select the highest probability door. What is most of interest is the persistence of the directional bias in Block 6, which was the first equal-probabilities block following Phase 1 of the experiment. This indicates a carryover of strategy from the unequal-probabilities blocks, until participants once again coordinate with the new task requirements.
When categorizing the data on the basis of “hold” versus “switch” trials, the success of the switching strategy is apparent. The other noteworthy finding is that Group L switched less often than Group R. No explanation consistent with the body of work on MHD immediately suggests itself, though it may be in keeping with other laterality effects noted in the present set of experiments. In other words, one hypothesis would be that participants found it effortful to make successive responses on the left side.
Experiment 4 was designed with the idea of bringing the current paradigm into closer contact with the original Asch study. Discussions of Asch (1955) in the years since publication have focused almost exclusively on the social component of the experiment, attributing a majority of the explanatory power to the idea of conformity to a group. Asch noted that his participants’ error rate in line matching (that is, their rate of trials on which they chose to conform to the group opinion) achieved a maximum when participants were confronted with between three to seven people who asserted that a non-matching line was the correct choice. This suggests that it is not just the presence of social pressure but the degree brought to bear that makes a difference. Even in this case the error rate topped out at about 1 in 3 trials.

Although Experiments 1 and 2 used a paradigm heavily inspired by Asch, neither experiment incorporated a social dimension. This was done in Experiment 4. Once again the manipulation was intended to be as minimal as possible while still being effective. In the present case, rather than putting people into a “truly” social context (e.g. having several participants in a room together), a degree of subterfuge was added to the original paradigm. Score was identified as reflecting the participant’s performance as relative to a prior participant. The key manipulation of Experiment 4 was a conflict between two biases, the stimulus bias of Experiments 1-3 and the social bias conveyed by the aforementioned score manipulation.

Methods
Participants. Participants were 6 volunteers (ages 25 to 56, 3 males, 3 females) from the community. Four participants were randomly assigned to the two conditions. Two left-handed participants (personal report) were assigned to separate conditions.

Design. Participants followed exactly the same procedure as in Experiment 2. There were only two changes to the design of the previous experiment. First, participants in were told that the “score” manipulation in Phase I (Blocks 2 through 5) was a measure of the degree to which their responses correlated with a previous participant in the study. Crucially, in order to test whether this subterfuge would influence participants’ responses, the direction of bias of “score” was reversed. For below-JND trials in Phase I, with probability .8, when participants clicked on the side of the screen towards which the line pairs were biased, their score decremented by 10, and increased by 10 when they clicked on the opposite side of the screen. The latter was done to put the influence of the stimulus properties into conflict with the “social” manipulation.

Results

Preliminary examination revealed that the ANOVA method used for Experiments 1-3 was inadequate for the small number of participants. The analysis that follows was conducted with SPSS. It allows the incorporation of all degrees of freedom in the experimental design.

A 2 (Group) × 8 (Block) × 3 (JND bin) ANOVA was conducted for the dependent measures of correct responses, left-click frequency, confidence, and latency.

Correct responses. The only main effect was JND bin, $F(2, 3852) = 315.828, p < .001$. Post-hoc tests showed that all bins significantly differed from each other, $p < .001$. The smallest-difference bin (Bin 1, comprising line pairs differing by less than half the JND
of 3.4 pixels) had the lowest proportion of correct responses ($M = .56, SD = .5$), the middle-difference bin (Bin 2, comprising line pairs differing by half the JND to 1.5JND) had more correct responses ($M = .65, SD = .476$), and the greatest-difference bin (Bin 3, comprising line pairs differing by greater than 1.5JND) had the most correct responses ($M = .98, SD = .144$).

*Left-click frequency.* Group, Block, and JND bin were significant: Group, $F(1, 3852) = 38.114, p < .001$, Block, $F(7, 3852) = 2.47, p < .05$, and JND bin, $F(2, 3852) = 12.17, p < .001$. Group 1 clicked on the left less ($M = .46, SD = .498$) than Group 2 ($M = .53, SD = .499$). Post-hoc tests on Block showed that Block 6 differed significantly from Blocks 3 and 4 ($p < .05$). Post-hoc tests on JND bin showed that Bin 3 differed significantly from Bin 1 and Bin 2 ($p < .001$). Group × Block × JND bin was significant, $F(14, 3852) = 4.768, p < .001$. The effect of Group depended on Block, $F(7, 3852) = 7.742, p < .001$, and JND $F(2, 3852) = 6.052, p < .01$, and the effect of Block depended on JND bin, $F(14, 3852) = 2.398$.

*Confidence.* Group, Block, and JND bin were significant: Group, $F(1, 3852) = 147.674, p < .001$, Block, $F(7, 3852) = 6.164, p < .001$, and JND bin, $F(2, 3852) = 147.336, p < .001$. Group 1 was more confident overall ($M = 360.42, SD = 201.594$) than Group 2 ($M = 326.57, SD = 181.772$). Post-hoc tests on JND bin showed that Bin 3 differed from Bin 1 and Bin 2 ($p < .001$). The Group effect was Block dependent, $F(7, 3852) = 147.336, p < .001$, and JND dependent, $F(2, 3852) = 12.393, p < .001$. Block × JND bin was also significant, $F(14, 3852) = 2.095, p < .05$.

*Latency.* There were significant effects of Block, $F(7, 3852) = 6.379, p < .001$, and JND bin, $F(2, 3852) = 43.913, p < .001$. Post-hoc tests showed that Block 1 differed
Experimentally Induced Bias

significantly from Blocks 3 ($p < .05$) and 5 ($p < .001$). Post-hoc tests on JND bin showed that the greatest-difference bin (Bin 3) differed significantly from both Bin 1 and Bin 2 ($p < .001$). The effect of Group depended on Block, $F(7, 3852) = 5.489, p < .001$.

**Discussion**

That the difference between line lengths, coded as JND bin, significantly affected the number of correct responses made by participants is not surprising. That Block did not affect the number of correct responses is surprising.

*Left-click frequency.* As in Experiments 1 and 2, the left-click measure is of central importance. To reiterate, the proportion of responses made on each side of the screen is a measure of the extent to which the experimentally-induced biases work. While both the line-length and score biases in Experiments 1 and 2 were intended to influence participants’ responses in the same direction, in the present experiment these same biases were intended to influence participants’ responses in opposite directions. Because of this conflict, the proportion of responses made on each side of the screen may be taken as indicative of which bias, stimulus/perceptual or social, comes to dominate in driving responses to different ends, and not as indicative of the effectiveness of the biases. Left-click frequencies for Group R and Group L are displayed in Figure 13. As expected, Group R clicked on the left less frequently overall than Group L. While there was an overall difference in left-click frequency among several blocks, averaged across groups, and between the greatest-difference JND bin and the other two bins, these results are not of immediate interest. The significant Group × Block interaction, however, is of
immediate interest. Phase 1 (Blocks 2 through 5) is the portion of the experiment in which the biases are present, with the social bias in conflict with the stimulus/perceptual bias. As Figure 14 shows, at the beginning of Phase 1 (Blocks 2 and 3), participants responded primarily on the side of the screen towards which they were being biased perceptually. Then, in Block 4, both groups responded predominantly on the side of the screen towards which they were being biased socially. The implication is that participants were in fact responding on the basis of the social pressure. The pattern of responses evident in the present experiment was not seen in Experiments 1 and 2. To reiterate, Experiment 4 differs from them only in the directionality and social character of “score”. Block 4 is also the locus of the lowest average confidence in both groups, indicating that although participants had begun to coordinate successfully with the social pressure, as was the case in Experiments 1-3 their success was not concomitant with confidence in their performance.
A final comment on left-click frequency is in respect to Phase 2 (Blocks 6 through 8). When both types of bias were removed, Group L showed a carryover tendency to respond on the left. This tendency decayed smoothly throughout Phase 2. Group R, on the other hand, exhibited a quick rebound to responding almost equally on both sides of the screen, followed by a slight uptick in left-side responses in Block 8. This asymmetry of responses in Phase 2 is consistent with the laterality effects observed in Experiments 1 and 2, but there is still no immediately compelling explanation. One candidate explanation remains the possibility that the bias towards responding on the left, which requires a crossing of the midline of the body when using the right hand to interface with the trackpad, constitutes a stronger pressure, which then has a stronger aftereffect.

**Confidence.** While all the independent variables (Group, Block, and JND bin) significantly affected confidence, only JND bin affected the number of correct responses, with larger length differences between lines straightforwardly yielding more correct responses. This points towards the same dissociation between confidence and correctness noticed in Experiments 1 and 2.

Further unpacking the results of the ANOVA on confidence requires some stage setting. In this experiment, as in the three previous experiments, confidence was measured as number of pixels away from the center of the screen a participant clicked. The center-line of the screen is treated as the zero point of horizontal distance across the screen. Responses on the right side of the screen are positive numbers, while responses on the left side of the screen are negatively valued. Because of this convention, and because at its most basic confidence is a measure of distance-from-center, we have thus far found it more useful to conduct analyses on the *absolute value* of confidence. This is
so because using the signed confidence measure, which is useful for distinguishing responses on the left side of the screen from responses on the right side of the screen, will produce statistically significant differences where, in fact, no difference in confidence exists. That is, a participant who clicks 100 pixels to the left of center while making a judgment that the left line is longer is exactly as confident as a participant who clicks 100 pixels to the right of center while making a judgment that the right line is longer — but analyzing such responses in their signed forms will show that they differ.

The present experiment, however, introduces another lateral symmetry break that required a different handling of confidence. Whereas in Experiments 1 and 2 the bias induced by the stimulus properties and the bias induced by score drove participant responses in the same direction (to either the right or left side of the screen), in Experiment 4 the two conflicted. The group designated left-biased saw a set of line pairs containing a left-line-longer bias but, crucially, the “score” bias (which was ostensibly a measure of coordination with a previous participant, and therefore a “social” bias) would bias participants to respond on the other side of the screen—in this case, on the right. As described above, this conflict allowed us to assess the relative strength of the perceptual bias and the social bias, and constituted the heart of Experiment 4. However, this added manipulation made it important to be able to distinguish right from left responses on a block-by-block basis.
To see why the latter is so, we can compare the graphs of confidence (Figure 14) and signed confidence (Figure 15). The graph of confidence shows that responses for Group R (right-biased) increase in value from Block 1 to Block 2, while responses for Group L (left-biased) decrease in value from Block 1 to Block 2. This could be interpreted as indicating that Group R on average is more confident in Block 2. However, the graph of signed confidence indicates that Group L’s absolute confidence decreases in Block 2 in virtue of becoming more negative—that is, both more confident and more leftward. Similarly, there is an overlap in confidence for both groups in Block 4, which occurs at a positive value. In absolute terms it appears as though Group R exhibits a decreasing trend in confidence from Block 2 through Block 4, while Group L exhibits a sudden increase in confidence in Block 4 compared to Block 3 and Block 5. The same data represented in signed terms tells a different story: on average, both Group R and Group L decrease in confidence in Block 4, with the average confidence for both groups sitting nearly at zero.
Confidence and signed confidence, therefore, characterize participants’ responses slightly differently. Once again, because signed confidence conflates magnitude of confidence with direction of response, until this point the absolute value of confidence was the more useful measure. With the introduction of the conflicting perceptual and social biases in Experiment 4 it became desirable to have some way of distinguishing changes in the magnitude of confidence from changes in the direction of response.

Special contributions of Experiment 4. There were three findings that are important for the overall picture of the present series of experiments. First, the dissociation between confidence and task performance, which had been seen in Experiments 1-3, is also seen in Experiment 4. The presence of this dissociation most clearly supports our central argument, that belief-as-action and belief-as-assent to a proposition are neither the same thing nor straightforwardly related. Second, the line-length bias that is perceptual in nature, and the score bias that in this case takes on the character of a social manipulation,
were both effective. Third, while the social bias is effective, with sufficient exposure the perceptual bias comes to dominate.

This last finding is the one that allows us to consider Experiment 4 in the context of the original Asch studies. As previously discussed, Asch did not find that participants committed errors wholesale on the basis of social pressure. What is remarkable about his findings is that participants rendered false judgments at all, because the social bias in Asch’s studies flew in the face of the perceptual evidence. The distress that participants evinced, as well as their willingness to offer conforming rather than correct answers, has been taken on the whole to indicate that participants value social considerations above considerations of truth. But this is not how we take the results. Asch’s work was not framed in the terms we find compelling—the distinction between belief-as-action and belief-as-assent. The present experiments achieve some separation between these notions, through the introduction of the confidence response as well as the line-length-judgment response. But in this final experiment, we also created a distinction between the two types of pressure on participant responses: the perceptual pressure and the social pressure. We can therefore see how these two types of pressure manifest differently for participants in Experiment 4. In keeping with Asch’s findings, when participants believe their responses are being considered against another person’s responses, they exhibit a tendency to respond in keeping with the social pressure, rather than the perceptual fact of the matter (or, in the present case, the perceptual bias). But this tendency does not wholly override participants’ tendency to respond in keeping with perceptual states of affairs. Asch found that the majority of participant responses were correct on the basis of the perceptual information—again, what was surprising in his experiment was that participants ever
responded according to social rather than perceptual factors. We found that participants begin responding in keeping with perceptual factors, then demonstrate a pattern of responses consistent with the social pressure. This change coincides with an overall decrease in confidence. Then, in the final block of Phase 1, participants once again begin responding consistent with the perceptual bias. Taken as a whole the evidence from Asch’s study and the present Asch-inspired work suggests that social pressure affects responses but never completely overrides perceptual factors.

**Coda.** There are three aspects of this experiment which must be considered distinct: belief-as-action (the responses a participant makes), belief-as-assent (a participant’s confidence in their responses), and which type of pressure (social or perceptual) dominates a participant’s responses at a given point in time. Experiment 4 builds on the findings of Experiments 1, 2, and 3, which all demonstrate the dissociation of the two notions of belief. It also distinguishes between the two types of pressure, showing that social pressure affects responses over and above confidence, but that perceptual pressure tends to dominate over the course of the experiment. From an ecological perspective this is perfectly reasonable. The social environment is one to which a social organism must successfully coordinate, and social information should therefore be expected to drive behavior. But social information is neither as stable nor, at bottom, as consequential, as perceptual information, and is unlikely in a setting such as this to ever be the sole driver of behavior.
CHAPTER 8
DYNAMICAL ANALYSES

Beliefs are the sorts of things that occur in large numbers, with linkages between them that may not be directly observable (Quine & Ullian, 1978). How best to understand the dynamics of belief has been a matter of some debate. One constraint on models of belief revision is that they must accommodate arbitrarily large numbers of beliefs. Another constraint arises because when one belief is updated or perturbed it will, in turn, perturb those beliefs that are somehow connected to it. Some way of assessing these changes, and the neighborhoods over which they range, is therefore also a needed component.

So far as we are aware there has been no other attempt to construe the dynamics of belief using the tools of nonlinear dynamics. We began from the standpoint that the actions that arise from an underlying belief network might constitute attractors in a relatively high-dimensional space. We therefore assessed the response measures in the present experiments as attractors, and the degree to which participants made the responses as a measure of attractor strength.

We proceeded in keeping with the methods described in Isenhower, Frank, Kay and Carello (2012). In this method, attractor strength is calculated as the average value of the transition probabilities towards the response of interest—in this case, left clicks. These transition probabilities are calculated by treating the data as a Markov process. To calculate transition probabilities left clicks were represented by ones and right clicks were represented by zeros. We focused only on below-JND stimuli, and did not incorporate responses to above-JND stimuli.
Markov Chain Modeling Experiment 1

A Markov Chain Model in close analogy to that presented by Isenhower, Frank, Kay and Carello (2012) was developed for the results of Experiment 1 (Chapter 4). As a first step, a two state model for less-than-JND stimuli was examined. In what follows, the two-state model for less-than-JND stimuli will be sketched. The two-state model for less than-JND pairs aims to understand the progression of less-than-JND response trajectories with the states “left click” and “right click” as shown in Figure 16.

![Diagram of transition probabilities](image)

*Figure 16.* Schematic of transition probabilities calculated for Markov chain estimation of attractor strength.

The model describes switches from left-click to right-click responses as well as subsequent responses of the same type. The model has four parameters, only two of which are independent. The two independent parameters are the transition probability $p(L \text{ to } R)$ of switching from a left-click response to a right-click response and the transition probability $p(R \text{ to } L)$ of switching from a right-click response to a left-click response.

The transition probabilities were determined for each of Blocks 1-5 (see Chapter 4) and for each participant using the parameter estimation method discussed in Isenhower et al. (2012). From the transition probabilities, the attractor strength of the right-click and
left-click states can be estimated. Note that there is only one independent measure for attractor strength. We used the strength of the attractor of the left-click responses.

The primary expectation was a main effect of Group: The left-click attractor should be weaker for Group R (right-biased) than for Group L (left-biased). That is, \( p(L \text{ to } R) \) should be higher for Group R than Group L. In contrast, the speculation was that \( p(R \text{ to } L) \) is lower for Group R than for Group L. One issue to consider is whether or not the transition probabilities change across blocks. This would be a hint that the biased-structure that emerges during the experiment affects the response-making dynamics. That is, if there is a block-group interaction then this could support the hypothesis of a circular causality in opinion-making processes as suggested by current models of opinion-making (see e.g. Frank, 2008 and references therein).

Results of an analysis of the Markov chain transition probabilities are shown in Figure 17. Black dots are participant data. Gray dots indicate sample means. The top panels refer to participants of Group R and the bottom panels refer to participants of Group L. Across Blocks 1-4, the left-click probability for less-than-JND stimuli decays for Group L and increases for Group R. In line with the prediction made above, the transition probability \( p(L \text{ to } R) \) was higher for Group R than for Group L, whereas the

Figure 17. Individual data and block means of left-click frequency for right- and left-biased groups in Experiment 1.
transition probability \( p(\text{R to L}) \) was lower for Group R than for Group L. Interestingly, transition probabilities showed a trend across Blocks 1-4. For Group R the transition probability \( p(\text{L to R}) \) increased, whereas \( p(\text{R to L}) \) decreased, which accounts for the increase of the right-click probability (i.e., decrease in left-click probability) across Blocks 1-4. The opposite relationship can be observed for Group L. Simulation of the Markov model yields results consistent with the data. Figure 18 shows simulation results (crosses) of the model for the right-biased group using the estimated transition probabilities as model parameters and starting with an initial left-click probability of 0.5. Experimental data are shown as circles.

![Figure 18. Comparison of simulated and experimental data for left-click probabilities by block in Experiment 1.](image)

Hypothesis testing was conducted for the transition probabilities \( p(\text{right to left}) \) for which a clear pattern is also visible on the basis of the raw data. A two-way ANOVA for the effect of block and group on transition probability showed a significant main effect of group \( (F(1,17) = 13, \ p < 0.005) \), indicating that \( p(\text{right to left}) \) is lower for Group R than for Group L. There was a significant interaction \( (F(3, 51) = 3.197, \ p < 0.05) \), indicating that \( p(\text{right to left}) \) decreased for Group R, whereas it increased for Group L.
Markov Chain Modeling Experiment 2

The transition probabilities calculated in the foregoing fashion were used to calculate attractor strength, $\gamma$, for left click for each participant and block. Four participants’ data were removed; one because the participant responded only on the bias side (in this case, the right side), which called into question the legitimacy of the responses, and an additional 3 participants whose data were not analyzable with this method. A repeated measures ANOVA was conducted on the remaining data. The effect of block was significant, $F(7, 168) = 3.749, p < .01$. However, neither group nor the interaction of group and block was significant. Means for block indicate that $\gamma$ decreased for both groups throughout the experiment. As an overall trend $\gamma$ was lower for Group R than Group L (see Figure 19, left panel).

Markov Chain Modeling Experiment 3

The same method was again applied to the data from Experiment 3. Our focus in this experiment was different: in the previous experiments the behavior of interest was whether participants clicked more often on the right or on the left, and whether the side they favored was also the side towards which they were being biased. In the Monty Hall-style experiment, the behavior of interest was what strategies participants adopted under the different conditions of the experiment. At its most general the MHD turns on whether players are willing to adopt the switching strategy, and the question we wish to answer with the current method is under what conditions the likelihood of switching changes. For Experiment 3, therefore, the Markov model of attractor strength was calculated for hold and switch rather than for responding on the left- or rightmost door.
A 2 (group) × 8 (block) repeated measures ANOVA was performed for the attractor strength $\gamma$ of switching. There was no significant main effect of either group or block, and no significant interaction (see Figure 19, right panel).

**Figure 19.** Gamma as a function of Group and Block for (left) Experiment 2 and (right) Experiment 3.

**Discussion of Markov Chain Modeling**

The Markov chain models of attractor strength for Experiments 2 and 3 did not yield the expected results. In Experiment 2 Group — that is, in which direction participants were being biased — did not have a significant effect. This goes directly against the hypothesis that the attractor for right-side responses would increase for the right-biased group and decrease for the left-biased group. It is perhaps less surprising that there was no significant effect of Group for Experiment 3, since the attractor strength was calculated for the hold-switch variable rather than for right- or left-side responding. There is no reason to suspect that the strategy adopted by participants would systematically vary with Group. However, we expected to see an effect of Block on attractor strength consistent with participants learning or honing in on the winning MHD strategy. We did not see this. Reasons why this might be the case are covered more thoroughly in the general discussion; for the moment, we simply observe that a participant would need to
appreciate that one strategy was superior to the other as a precursor to consistently adopting it.

**Autoregression Modeling Experiment 2**

Each of the present experiments yielded two measures of core interest: judgment and confidence. Because the judgments in each experiment were binary, the above Markov method was appropriate for estimating the strength of the attractors of the different possible judgments. The confidence response, however, was continuously-valued, and therefore not treatable using this method. A related method was therefore employed. The time series of confidence responses were analyzed as order 1 autoregressive (AR(1)) processes. Such an analysis allows a time series to be categorized as a stable or an unstable process and, if stable, whether the process is oscillatory or nonoscillatory – that is, whether it tends toward a specific value, or not, and if so, whether the process moves toward its fixed point monotonically. (Chatfield, 1991). The data here were treated only as a first-order autoregressive process, with a model of the form

\[ X_{n+1} = a_1 \cdot X_n + q_n \]

where \( X \) is the value of the process at each time step \( n \), \( a_1 \) is the first-order autoregression parameter, and \( q_n \) is a noise term (Diggle, 1990). It is the value of the autoregression parameter \( a_1 \) that describes the stability of the process, with values falling in the interval \((-1, 1)\) describing stable process, and values outside this interval describing unstable ones. Negative values falling within the interval describe processes that oscillate about a fixed point, while positive values within the interval describe processes that converge monotonically on a fixed point. Using the time series of the unsigned confidence responses, we calculated \( a_1 \) and the variance of the noise term \( q_n \) for each
participant and block in Experiment 2. That is, in this analysis we considered only the magnitude of the confidence responses and not their sign (which indicates on what side of the screen a response was made). Therefore the variable $X_n$ in the first-order autoregression model above reflects deviations from the mean unsigned confidence. More precisely, for each participant the mean value of the unsigned confidence across all blocks was calculated. Next the deviations from this time-averaged mean value were computed. In summary, the variable $X_n$ of the model is related to the observed signed confidence responses $S_n$ in the following fashion:

$$S_n \rightarrow U_n = |S_n|$$

$$X_n = U_n - \text{mean}(U_n)$$

We conducted 2 (group) $\times$ 8 (block) repeated-measures ANOVAs on the resulting values of $a_1$ and $\text{Var}(q_n)$. A 2 (group) $\times$ 8 (block) repeated-measures ANOVA for $a_1$ showed no significant main effects or interactions. Means are displayed in Figure 20.

![Figure 20: Autoregression parameter $a_1$ as a function of Group and Block for Experiment 2.](image)

A 2 (group) $\times$ 8 (block) repeated-measures ANOVA for the variance in $q_n$ was significant for Block, $F(2.376, 61.778) = 12.729, p < .001, \eta^2 = .32$, with the
Greenhouse-Geisser correction for a significant Mauchly’s test of sphericity applied.
Neither the main effect of Group nor the Group × Block interaction was significant.
Within-subjects contrasts indicate that the data were consistent with polynomials through
order 4, but visual inspection suggests that the data are probably best fit by a quadratic,
\( F(1, 26) = 14.5, p < .01 \) (see Figure 21).

![Figure 21. Variance in autoregression noise term \( a_n \) as a function of Group and Block for Experiment 2.](image)

**Autoregression Modeling Experiment 3**

The same technique was applied to the confidence data from Experiment 3. Again, 2
(group) × 8 (block) repeated-measures ANOVAs were performed for the values of \( a_i \) and
the variance of \( q_n \) calculated for each participant and block.

![Figure 22. Autoregression parameter \( a_1 \) as a function of Group and Block for Experiment 3.](image)
A 2 (group) × 8 (block) repeated-measures ANOVA for $a_1$ yielded no significant main effects or interactions (see Figure 22). A 2 (group) × 8 (block) repeated-measures ANOVA for the variance of $q_n$ was significant for Block, $F(3.415, 27.318) = 3.523, p < .05, \eta^2 = .3$, with the Greenhouse-Geisser correction for a violation of the sphericity assumption applied. Within-subjects contrast show that the data are best fit by a quadratic function, $F(1, 8) = 9.109, p < .05$. This pattern is visible in Figure 23.

![Figure 23](image)

**Figure 23.** Variance in autoregression noise term $q_n$ as a function of Group and Block for Experiment 3.

**Discussion of Autoregression Modeling**

In both Experiments 2 and 3 the values of $a_1$ calculated per participant and block on the confidence responses fell into the (-1, 1) interval. This indicates that the confidence measure represents a stable process; that is, one which approaches a single value. Since the values are negative, we can conclude that the process represented by the confidence measure approaches its stable value in an oscillatory fashion: responses will be both above and below the fixed point as the process approaches it. Values of $a_1$ did not change significantly in either Experiment 2 or Experiment 3, which is consistent with the claim that confidence in these experiments tracks a stable aspect of the behavior under investigation.
This pattern of results for $a_i$ on the confidence measure is, in our view, consistent with the findings for confidence from the standard inferential statistics described earlier. We have already shown that confidence is independent of performance, where performance can be understood as the measure more closely tied to the facts of the task. The process that underlies the confidence remains stable throughout the experiment, and converges on a fixed point, rather than monotonically increasing or decreasing towards some value. As a first pass at interpreting these results, we argue that this is further indicative of the extent to which confidence is independent of the task and of performance on the task. It may be that confidence is somehow endogenous to the responding system (i.e., the participant), perhaps constituting a relatively more stable or entrenched type of belief, as opposed to the more easily biased action-based type.

The pattern of results for the strength of the noise term $q_n$ for both Experiments 2 and 3 is straightforward. In each case noise starts at its highest value in Block 1 and decreases over Phase 1, though rebounding slightly in Phase 2. These results are of a piece with the earlier-described results for response latency, which together indicate that participants’ comfort and facility with the tasks increase as they encounter more trials. However the noise measure also shows that the removal of the bias in Phase 2 may affect the consistency of participants’ responses as they once again adjust to the new task constraints.
CHAPTER 9
ANALYSIS IN TERMS OF SIGNAL-DETECTION THEORY

The initial impetus behind the present set of experiments was to subject certain intuitions about the nature of belief and belief change to empirical methods. To this end we developed an experimental paradigm inspired by Asch’s work on social conformity in judgment. This paradigm was employed in Experiment 1, extended in Experiment 2, and adapted further in Experiment 4. In each case we approached the analysis of participant data from the perspective that we were assessing whether participants were sensitive to the presence of the introduced biases—whether they could perceive the stimulus bias, and whether they were motivated by the score bias. We concentrated on participants’ responses on the side to which they were being biased as indicative of the extent to which the bias was effective.

The analysis already presented centered on participants’ responses on the below-JND line pairs. The logic of this concentration is that, for the pairs in which line lengths clearly differ, participants will likely respond based on which line is longer. But for the pairs where it’s more difficult to tell, the null hypothesis would be for participants to respond with equal likelihood on the right or left side. Deviation from a 50/50 split, towards the bias side, is therefore indicative of the efficacy of the bias, and such evidence is more common in the pairs that are more difficult to distinguish. Two important conclusions follow from this logic. First, although erroneous bias-side responses are expected to be more common in below-JND pairs, the fact that participants sometimes make them even for above-JND pairs indicates that the introduced biases are stronger than initially supposed. Second, a bias-side response on a below-JND pair may be the
result either of the biases or of a guess. This second realization led us to consider another way of approaching the data.

Instead of considering the data strictly in terms of the presence or absence on a given trial of the effect of the biases, we instead considered participants’ responses in terms of their relationship to the stimulus properties – that is, whether participants were responding accurately given the fact of which line of a pair was longer. To the extent a participant was consistently able to choose the longer of a pair of lines, any response might be indicative of their perceptual abilities rather than the effect of a bias. To the extent that a participant was not able to correctly choose the longer of a pair of lines, any response might be indicative of guessing, unless such responses were part of a pattern of consistently responding on the bias-side of the screen. In order to reframe our questions in appropriate terms, we recoded the data according to the response categories of signal detection theory. Each participant response therefore became an instance of a hit (correctly choosing the longer line) or a miss (failing to choose the longer line), and of a false alarm (responding that a line is longer when the line is shorter) or a correct rejection (not selecting a line that is, in fact, shorter).

On a forced-choice task such as this, these four response categories are not mutually exclusive. Given a pair of lines with the right line longer, and a participant response that the right line is longer, a selection of the right line is both a hit (on the right) and a correct rejection (on the left). We chose to concentrate on responses falling on the bias side only, as these were the responses of interest, and categorized them as either hits or false alarms. To restate our hypothesis in the terms of SDT, we expect that the effect of the bias would manifest as an increase in bias-side false alarm rates relative to hits.
SDT is not limited to just these categories, however. An important component of the theory is that of *responder bias* – that a given responder might have a higher or lower threshold for rendering a judgment that a signal has been detected. In our case the *signal* is which line is actually longer; that is, the stimulus properties. Deviation from signal detection may be attributed to a combination of bias and noise. An individual’s ability to discriminate signal from noise is captured by $d'$, which is calculated using our measures of interest: hit rate and false alarm rate. The equation for $d'$ is

$$z(\text{hit}) - z(\text{FA}).$$

Because the blocks in Experiment 2 did not all have equal numbers of bias-side trials, we calculated $d'$ on the hit and false alarm *rates* on bias-side trials per participant and block. Using a mixed-design ANOVA we compared the values of $d'$ for both groups across the 8 blocks of Experiment 2. As a final consideration, the calculation of $d'$ depends on calculating the z-scores of hit rates and false alarm rates.

A participant’s signal detection profile is also determined by the participant’s *bias* towards responding “yes” or “no” to a signal, and by their *criterion* (the threshold at which they will make a “yes” response). Bias is measured by both $\beta$, a measure of the likelihood that a signal is present, and also by $c$, which is the distance between the criterion value and the point of the signal/noise distribution at which neither response is favored. These bias measures are of interest insofar as they are indicative of the extent to which participants make responses based on something other than the signal – in this instance, the biases introduced by the experimental manipulations. If participants truly become more attuned (or susceptible) to the experimental biases over the course of the
experiment, we would expect to see a decrease in average d' and increases in the bias measures.

**Results**

For Experiment 2, d', β, and c were calculated on bias-side trials per participant, per block. In all cases the formulas and Excel syntax from Stanislaw and Todorov (1999) were used.

**d' ANOVA**

A 2 (Group, between-subjects) × 8 (Block, within-subjects) mixed-design ANOVA was conducted for d'. The effect of Block was significant, \( F(3.987, 107.659) = 98.2, p < .001 \), with the Greenhouse-Geisser correction for sphericity applied.

<table>
<thead>
<tr>
<th>Block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.016</td>
<td>.045</td>
<td>.121**</td>
<td>.146*</td>
<td>.362**</td>
<td>.368**</td>
<td>.301**</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.016</td>
<td>.029</td>
<td>.105**</td>
<td>.13**</td>
<td>.346**</td>
<td>.353**</td>
<td>.285**</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.045</td>
<td>-.029</td>
<td>.075*</td>
<td>.1*</td>
<td>.317**</td>
<td>.323**</td>
<td>.255**</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.121**</td>
<td>-.105**</td>
<td>-.075*</td>
<td>.025</td>
<td>.241**</td>
<td>.248**</td>
<td>.18**</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.146*</td>
<td>-.13**</td>
<td>-.1*</td>
<td>-.025</td>
<td>.216**</td>
<td>.223**</td>
<td>.155*</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-.362**</td>
<td>-.346**</td>
<td>-.317**</td>
<td>-.241**</td>
<td>-.216**</td>
<td>0.006</td>
<td>-.061</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-.368**</td>
<td>-.353**</td>
<td>-.323**</td>
<td>-.248**</td>
<td>-.223*</td>
<td>-.006</td>
<td>-.068</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-.301**</td>
<td>-.285**</td>
<td>-.255**</td>
<td>-.18**</td>
<td>-.155*</td>
<td>0.061</td>
<td>0.068</td>
<td></td>
</tr>
</tbody>
</table>

* Difference is significant at \( p < .01 \); ** difference is significant at \( p < .001 \).
Within-subjects contrasts show that the data are best described by a cubic function, 
\[ F(1, 27) = 41.69, p < .001. \] Significant post-hoc tests are summarized in Table 2.

As Figure 24 shows, \( d' \) values are highest for both groups in Block 1 – that is, participants’ discrimination between signal and noise is best at the beginning of the experiment (Group 1: \( M = .68, SD = .07 \); Group 2: \( M = .71, SD = .06 \)). Group 1’s \( d' \) values continued to decrease until Block 7 (\( M = .3, SD = .08 \)) before rebounding slightly in Block 8, though the difference between Block 7 and Block 8 was not significant. Group 2’s \( d' \) values decreased until Block 6 (\( M = .33, SD = .05 \)) before increasing again over Blocks 7 and 8 – however, the differences between Blocks 6, 7 and 8 were not significant.

![Figure 24. Mean differences in \( d' \) as a function of block and group in Experiment 2.](image)

There was no significant main effect of Group, and the Group × Block interaction was not significant.

**\( \beta \) ANOVA**

A 2 (Group, between-subjects) × 8 (Block, within-subjects) mixed-design ANOVA was conducted for \( \beta \). The effect of Block was significant, \( F(3.78, 102.06) = 107.69, p < .001 \), with the Greenhouse-Geisser correction for sphericity applied. Within-subjects
contrasts show that the data may be described by several polynomial fits, but visual inspection suggests that they are best described by a cubic function, $F(1, 27) = 38.02, p < .001$. Significant post-hoc tests are summarized in Table 3.

Table 3

<table>
<thead>
<tr>
<th>Block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.016</td>
<td>-.004</td>
<td>-.036</td>
<td>-.062**</td>
<td>-.179**</td>
<td>-.188**</td>
<td>-.16**</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.016</td>
<td>.013</td>
<td>-.019</td>
<td>-.045**</td>
<td>-.163**</td>
<td>-.171**</td>
<td>-.144**</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.004</td>
<td>-.013</td>
<td>-.032*</td>
<td>-.058**</td>
<td>-.175**</td>
<td>-.184**</td>
<td>-.156**</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.036</td>
<td>.019</td>
<td>.032*</td>
<td>-.026</td>
<td>-.143**</td>
<td>-.152**</td>
<td>-.124**</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.062**</td>
<td>.045**</td>
<td>.058**</td>
<td>.026</td>
<td>-.117**</td>
<td>-.126**</td>
<td>-.098**</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.179**</td>
<td>.163**</td>
<td>.175**</td>
<td>.143**</td>
<td>.117**</td>
<td>-.009</td>
<td>.091</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.188**</td>
<td>.171**</td>
<td>.184**</td>
<td>.152**</td>
<td>.126**</td>
<td>.009</td>
<td>.028</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.16**</td>
<td>.144**</td>
<td>.156**</td>
<td>.124**</td>
<td>.098**</td>
<td>-.019</td>
<td>-.028</td>
<td></td>
</tr>
</tbody>
</table>

*Difference is significant at $p < .01$; ** difference is significant at $p < .001$.

As Figure 25 shows, $\beta$ values were lowest for both groups at the outset (Group 1, Block 1: $M = .71, SD = .04$; Group 2, Block 1: $M = .73, SD = .03$) – that is, participants’ bias was least in the earlier parts of the experiment. For both groups, $\beta$ steadily increased until Block 5 (Group 1: $M = .77, SD = .05$; Group 2: $M = .8, SD = .05$) before increasing sharply in Block 6 (Group 1: $M = .9, SD = .03$; Group 2: $M = .9, SD = .03$). For both groups there was no significant difference across the final 3 blocks of the experiment.
Figure 25. Mean differences in $\beta$ as a function of block and group in Experiment 2.

There was no significant main effect of Group. With the Greenhouse-Geisser correction the Group \times Block interaction was marginally significant, $F(3.78, 102.06) = 2.17, p = .081$. This interaction is significant if sphericity is assumed, $F(7, 189) = 2.17, p < .05$, but is not significant if the more stringent lower-bound correction is applied (see Discussion).

\textbf{cANOVA}

A 2 (Group, between-subjects) \times 8 (Block, within-subjects) mixed-design ANOVA was conducted for $c$. The effect of Block was significant, $F(4.01, 108.32) = 11.49, p < .001$, with the Greenhouse-Geisser correction for sphericity applied. Within-subjects contrasts show that the data may be described by several polynomial fits, but visual inspection suggests that they are best described by a quadratic function, $F(1, 27) = 30.28, p < .001$. Significant post-hoc tests are summarized in Table 4.
### Table 4

**Significant Mean Differences Between Blocks for Experiment 2 c Calculations.**

<table>
<thead>
<tr>
<th>Block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.022</td>
<td>.065</td>
<td>.064*</td>
<td>.099*</td>
<td>.027</td>
<td>-0.003</td>
<td>-0.007</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.022</td>
<td>.087**</td>
<td>.085**</td>
<td>.121**</td>
<td>.048</td>
<td>.019</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.065</td>
<td>-0.087**</td>
<td>-0.001</td>
<td>.035</td>
<td>-0.038</td>
<td>-0.067</td>
<td>-0.072</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.064*</td>
<td>-0.085**</td>
<td>.001</td>
<td>.036</td>
<td>-0.037</td>
<td>-0.066</td>
<td>-0.071</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.099*</td>
<td>-0.121**</td>
<td>-0.035</td>
<td>-0.036</td>
<td>.073*</td>
<td>-0.102*</td>
<td>-0.107*</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.027</td>
<td>-0.048</td>
<td>.038</td>
<td>.037</td>
<td>.073*</td>
<td>-0.029</td>
<td>-0.034</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.003</td>
<td>-0.019</td>
<td>.067</td>
<td>.066</td>
<td>.102*</td>
<td>.029</td>
<td>-0.005</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.007</td>
<td>-0.015</td>
<td>.072</td>
<td>.071</td>
<td>.107**</td>
<td>.034</td>
<td>.005</td>
<td></td>
</tr>
</tbody>
</table>

* Difference is significant at $p < .01$; ** difference is significant at $p < .001$.

As Figure 26 shows, $c$ increased for both groups from Block 1 (Group 1: $M = -.53, SD = .07$; Group 2: $M = -.46, SD = .06$) to Block 2 (Group 1: $M = -.49, SD = .07$; Group 2: $M = -.45, SD = .06$), although this difference was not significant. For both groups $c$ then decreased sharply, hitting a minimum in Block 5 (Group 1: $M = -.57, SD = .08$; Group 2: $M = -.62, SD = .15$). This difference is significant, $p < .01$. For both groups, $c$ then increased steadily until the end of the experiment in Block 8 (Group 1: $M = -.5, SD = .14$; Group 2: $M = -.47, SD = .16$). The difference between these values and the minimum values in Block 5 is significant, $p < .01$. 

---

**Experimentally Induced Bias**

---
Figure 26. Mean differences in $c$ as a function of block and group in Experiment 2.

There was no significant main effect of Group. With the Greenhouse-Geisser correction the Group $\times$ Block interaction was marginally significant, $F(4.01, 108.32) = 2.26, p = .067$. This interaction is significant if sphericity is assumed, $F(7, 189) = 2.26, p < .05$, but is not significant if the more stringent lower-bound correction is applied (see Discussion).

Discussion

The results of the signal detection measures bear out our original hypotheses. We found that $d'$, an indicator of participants’ ability to discriminate between signal and noise trials (in this case, the presence or absence of a longer line on the bias side) decreases over the course of the experiment. We argue that this is due to participants’ increased reliance on (or attunement to) the biasing influences and decreased reliance on stimulus properties in making their decisions.

The bias measures of $\beta$ and $c$ also show the expected pattern. Participants’ bias as measured by $\beta$ increases throughout the experiment before decreasing slightly (but not significantly) in the final block. We take this to indicate, as claimed above, that participants’ responses become more biased (in the sense meant by the vocabulary of signal detection theory) and less contingent on actual stimulus properties, and that this
effect persists even when the experimental bias is removed in the final 3 blocks. Following the recommendation of Stanislaw and Todorov (1999) bias was also calculated using $c$. While $\beta$ and $c$ are both measures of bias, $\beta$ is a likelihood of responding “yes” rather than “no”, while $c$ measures the distance between a participant’s criterion for responding “yes” and the point at which either response is equally likely. Importantly, negative values of $c$ indicate a tendency to respond “yes”, indicating that the signal – in this case, a longer line on the bias side – is detected. As can be seen, in this experiment all values of $c$ for all participants and blocks were negative, and were at their most negative in the final block of what has been previously characterized as “Phase I” of the experiment – that is, the phase in which the experimental biases are present.

As expected, no significant main effect of group was observed. That is, although the groups were biased towards different sides of the screen, their responses were characterized in terms of being on the bias or non-bias side. We therefore would not expect to see a difference in the signal detection measures based on group. However, the calculations for the bias measures showed a marginally significant interaction between Group and Block. It is possible that this is merely an artifact of the violation of the assumption of sphericity, as the significant result disappears as more stringent corrections for this violation are applied. But given the laterality effects observed elsewhere in the experiment the interaction cannot entirely be dismissed.

In general the attempt to approach the data from Experiment 2 using signal detection may be considered a success. Both the task and participant responses can be characterized as the detection of a signal (stimulus properties), which throughout the experiment becomes more and more subject to participant bias. This likely indicates that the
Experimental bias successfully affects participants’ response bias. However, despite the utility of such measures, it is not clear that this task is best characterized as the detection of a *signal*. Applying the signal/noise dichotomy to such a task creates a distinction between the “right” properties and the “wrong” properties for the line length judgment task, which is not the aim of the experiment. Rather, we would like to say that the experiment provides a task context in which participants’ belief-as-action becomes divorced from the propositional description of their beliefs about the stimuli themselves. Reifying a “signal” in the task and using vocabulary that relies on the status of the signal to describe participant performance implies that e.g. the decrease in $d'$ is indicative of an inability to perceptually discriminate between signal and noise rather than a dissociation between the two notions of belief. The increase in the bias measures somewhat gets at this dissociation, indicating a shift in participants’ approach to the task, but the hypothesis of primary interest here is not truly testable within the paradigm of signal detection theory.
CHAPTER 10

GENERAL DISCUSSION

The relationships that obtain between action, perception, belief, and the world serve, via evidence or assumption, as anchoring points for psychological and scientific theories. A philosophical view of long standing runs like this: perception must be conditioned on the properties of the world, because if it weren’t, it could not underwrite successful action (which we clearly engage in), and there would be no reason for consensus about what the world is like (which to a large extent, there is). But perception cannot be based just on the properties of the world, because if it were, we would never be mistaken about what we perceive, which the existence of a myriad of fallacies and illusions suggests we are. So perception must be the sort of thing that can be in error—and therefore is more like belief than like knowledge. This is the core of the argument for perception as belief presented by Armstrong. Although beliefs may be true or not, relatively well grounded or not, and more or less entrenched, beliefs-as-perceptions are mostly related felicitously to the world, just as action is mostly successful – if these things were not true, survival would be unlikely.

One way to approach the problem, then, is to seek the kind of relationships that can be mostly veridical but occasionally not. The term “supervenience” is one that purports to satisfy the requirements. Our mental (perceptual, psychological) states supervene on the physical world, are wholly determined by the nature of and relations in the physical world, and are isomorphic to but not identical to them. It is this supposed dualism that creates the breach between the mental and physical that representations are meant to cross. But they do not, for a laundry list of reasons. The history of thinking on
representations have made several assertions about their necessary character, however: one line of thinking goes that representations are the things by which we reason, that they have syntax (and are in other respects language-like), that they do not faithfully point towards the world, yet that we rely on them just as though they do.

Another historical thread makes very similar arguments for belief: we may draw conclusions from them, they are things that guide our actions; they have rules by which they combine, compete, and change; they are not the same as knowledge about the world (though they are related), yet we are prepared at any time to articulate some subset of our beliefs and the extent to which we espouse them.

In a sense, then, there is nothing in a propositional account of belief that is not also shared by representational accounts of mind or perception. They are all part and parcel of the same set of assumptions about the world and the engagement of agents with it. A theory of agents that does not take those assumptions on board will therefore find itself in conflict with the specific treatment of each entity that carries theoretical weight in a fundamentally dualist explanation. Ecological theory has already extensively critiqued representational theories of mind from both philosophical and empirical standpoints. As questions of how agents should and do reason are increasingly considered by those who take such capacities to be grounded in propositions, it is crucial for ecological theorists to take on the job of providing explanations not beholden to the dualist way of thinking.

It is for this reason that the present set of experiments was designed—to investigate belief from an ecological and empirical standpoint. To this end, a distinction was drawn

\[1\] There are views on mind that embrace the existence of mental representations but fail to endow them with the causal force necessary to guide behavior, and the arguments presented here are not intended to address these views.
between two different ways of thinking about belief. First, there is belief in all its propositional glory, which is perhaps the more well-known way of considering belief. But then there is belief as a guide to action—to once again borrow the words of Frank Ramsey, “the map by which we steer”. As he pointed out almost a century ago, there is something about belief in its role as driving actions that is not compatible with the idea of beliefs as being propositions about all the myriad features of the world which one might acquire in the course of interacting with the world. Showing that one possesses both species of belief but that they are not equivalent, then, takes a step toward an understanding of belief that is not beholden to its propositional character.

There were already indications that these two species of belief are not equivalent. The famous Asch experiment (Asch, 1955) showed the extent to which people may be induced to act not in keeping with beliefs they seem to hold. Participants in his study acted in accordance with majority opinion in cases where the majority opinion flew in the face of perceptual evidence, and in some cases showed obvious distress and confusion at doing so. That is, the propositional beliefs they held (“This line is longest”) served only as a partial guide to their actions.

One may argue in response to this that it isn’t the case that the Asch subjects failed to act in accordance with their propositional beliefs, only that they considered other, also propositional beliefs, more valuable, perhaps beliefs like “I believe that sometimes I’m wrong, and I believe that everyone else seems to know what they’re doing,” in which case they may be led to the conclusion “I believe that others may be right and I may be mistaken”. At this point we assert a different, ecologically-informed reading: that these sorts of propositional beliefs are descriptions of the agent’s overall state, but not the
So if we are attempting to build on the Asch paradigm and construct further empirical methods for investigating the relationship between perceptions and beliefs we will want to structure them in such a way that allows us to delve into the *content* – the underlying nature of the perceiving-acting system.

Fortunately, such methods exist. Particularly, the sciences of complexity and self-organizing systems provide a suite of methodological and analytical tools for addressing questions about systems that are high dimensional at a component level but low-dimensional at a behavioral level. One thing all accounts of belief have in common is that whatever beliefs are, agents have very many of them, and they are interconnected. This has led to accounts either treating impoverished sets of beliefs, or being mathematically intractable. Treating behavior as a self-organizing system that arises in part from beliefs, however, may offer a different sort of solution.

The foregoing experiments constituted an attempt to do just that. One cannot experimentally assess all of an agent’s beliefs, but restricting an empirical investigation to just one or just a few propositional beliefs provides a severely restricted view of the nature and dynamics of an agent’s belief set. Behavior can be experimentally assessed but a measure is needed that ties behavior to belief. A paradigm was therefore constructed that attempts to fulfill both these goals. In the present experiments, participants provide an overt measure of behavior (correct responses, bias-side responses) as well as a more covert measure of their own state (confidence). They are placed into experimental contexts that provide lower-order regularities (distribution of line lengths or winning doors) and higher-order ones (the dynamically generated score of Experiments 1, 2 and 4, and the winning strategy of the Monty Hall Dilemma in Experiment 3) and their
responses indicate the extent to which those regularities are reflected in their behavior. Finally, they are placed into these contexts for long enough, and asked to make responses often enough, that the data generated by these experiments are sufficient for the analytical tools developed to deal with self-organizing systems to be brought to bear.

The paradigm is new, and needs refinement, but at a first pass it works. Biases are successfully introduced and in some fashion persist. Actions are undertaken not wholly in keeping with degree of assent to the correctness of the action. Frank (2008) had already examined the Asch data and showed that behavior remains low dimensional (only two outcomes) even when the dynamics underlying it are more finely-grained. A similar finding is hinted at in the results of the current studies: there is a great deal of variability between trials and between participants, but the larger-scale dynamics show patterns of consistency.

The methodological approach we adopted for these experiments has made clearer several issues concerning the multiple scales and dynamics of belief, perception and action. One question is at how many scales do such phenomena manifest themselves? As we have seen in our treatment of the main hypothesis – the separability of the two notions of belief – as well as in Frank’s (2008) analysis of the original Asch data, there are low-dimensional and high-dimensional components of the system that as a whole produces some judgment or action with respect to a task like the ones presented here. We have treated these low- and high-dimensional processes analytically through the use of traditional inferential statistics, which look at aggregate behavior over a large number of trials and a longer time scale, and through the use of dynamical analyses, which look at trial-by-trial behavior and a concomitantly shorter time scale.
Putting these analytical tools in a theoretical context allows us to say more on the issue of the scale of the processes in question. This is best illustrated by returning to the results from Experiments 3 and 4. As we saw in the discussion of the dynamical analyses, for both groups the attractor for switching was strongest in Block 2, grew steadily weaker throughout Phase 1, and then recovered in strength throughout Phase 2, before again achieving a local peak in Block 8. We interpret this pattern in terms of the possible strategies participants could adopt: By Block 2, participants have played 100 trials of the standard MHD. We take high strength of the switching attractor in Block 2 to indicate that participants have to some extent been able to tune into the fact that switching is the more successful strategy. Herbranson and Schroder (2010) had found that pigeons are able to capitalize on the success of the switching strategy given sufficient time to learn (or be trained on) the strategy. We believe something similar is at work here: on a single trial, or even on a handful of trials, the winning MHD strategy is not obvious. Over many trials, with outcome feedback provided, the structure of the strategy may begin to reveal itself. On this reading it would make perfect sense for the switching attractor to be at its strongest after a participant has already experienced a series of trials.

However, the switching attractor then gets progressively weaker over the course of Phase 1, wherein the unequal-doors case is introduced. This result can be framed as a perturbation or destabilization of the task landscape that participants navigated in Phase 0. The task environment changes, which may in and of itself serve to weaken the attractor. The way in which the environment changes — the doors are no longer equally likely to conceal the prize — in turn changes the win probabilities of employing the various strategies. In this case, the difference in win probability between the easy-to-
discover ‘hold’ strategy and the hard-to-discover ‘stick’ strategy (.8 versus .9) may not be sufficient to drive participants into a ‘switch’ attractor, even though switching remains the better strategy. These considerations highlight both the weakness and the promise of the experimental methods developed here. As they stand, the present experiments serve only as preliminary steps towards truly exploring the dynamics of belief as they relate to the behavior of sophisticated perception-action-cognition systems. They constitute a technique for acquiring and analyzing behavioral data pertaining to the set of questions already in play about the structure and nature of beliefs. But few conclusions can as yet be drawn. Most encouragingly, a distinction was successfully demonstrated between the two notions of belief — as action-oriented or propositionally-construed — which lends credence to the account here advanced of how belief might be brought into line with a larger body of psychological theory. But a fully-fledged account of the dynamics of belief is lacking. The data presented here were not sufficient to completely situate these experiments on belief within dynamical systems theory. Perhaps the most glaring absence is that of an identifiable control parameter, manipulation of which might in turn cast more light on the extent to which the behaviors of interest in these experiments — line length judgments or MHD strategies — constitute attractors in some appropriately delineated space. Attempting to isolate such parameters is the obvious next step. For the MHD in particular, one candidate is the extent to which one strategy is clearly better than another, e.g. by how much more a player can increase their chances of winning an unequal-probabilities MHD using a ‘switch’ rather than ‘hold’ strategy.

These results motivate further consideration on larger questions of dynamics and scale. Ecological theory emphasizes that perception and action both take place on a
macro scale appropriate to the size of the perceiving-acting, organism-environment system. But each such system has smaller-scale or componential dynamics, such as those obtaining at the level of muscles or neurons. These components are not only typically smaller than the higher-level phenomena to which they are related, but are many more in number than the macro phenomena. Here sit the well-known issues of the relationships between levels of a phenomenon, of causal determination and of the appropriate scale of explanation. Similarly, the products of cognition, such as behaviors and judgments, must be scaled to the world in which those products are deployed. But they rest on a collection of entities that operate on a different scale. We may call some of these entities beliefs. As has been shown in the literature on epistemology, a formal account of belief update must respect the (very large) number of beliefs each agent possesses, the way in which those beliefs are updated in the face of evidence and the results produced by those updates. In a real agent this is exactly the issue of how cognition, perception, and action are related. But while there are intuitive ways to divide the three according to function or to scale, ecological theory and dynamical systems theory hold that the divisions are pragmatic at best and confusingly ad hoc at worst. Understanding belief requires respecting the complexity of the issue without yoking the idea of belief to their propositional and language-like character to the detriment of belief’s role in perception-action-cognition systems as meaningful wholes.


