DEPENDENCE OF PERCEPTUAL CHOICE ON NUMBER OF RESPONSE ALTERNATIVES AND FIDELITY OF EVIDENCE

by

Elisabeth Jeannette Ploran

B.A., Drew University, 2004

M.S., University of Pittsburgh, 2008

Submitted to the Graduate Faculty of
Arts & Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

School of Arts & Sciences

This dissertation was presented

by

Elisabeth Jeannette Ploran

It was defended on

June 22, 2010

Christian Schunn, Professor, Department of Psychology

Julie Fiez, Professor, Department of Psychology

Carl Olson, Professor, Department of Neuroscience

Dissertation Advisor: Mark E. Wheeler, Assistant Professor, Department of Psychology

DEPENDENCE OF PERCEPTUAL CHOICE ON NUMBER OF RESPONSE ALTERNATIVES AND FIDELITY OF EVIDENCE

Elisabeth Jeannette Ploran, Ph.D.

University of Pittsburgh, 2010

The study of perceptual decisions has been developed as a substitute for investigating more complex multiple attribute decisions. However, little attention has been paid to the similarity of results between the two literatures. Four separate behavioral experiments and a secondary trial-by-trial analysis investigated the sensitivity of perceptual decisions. Results were compared to both previous perceptual decision research and that of multiple attribute decisions in an effort to bridge the divide. The first experiment examined the effect of increasing the similarity of available response alternatives on accuracy and reaction time. The results suggest that high levels of similarity can begin to degrade the decision process by lowering accuracy and slowing reaction time; however these changes may be dependent on the extent to which the alternatives use overlapping neuronal pools. The second experiment examined the effect of increasing the number of response alternatives available for a single decision. The results suggest that increasing the number of alternatives may not affect performance until some critical point (in this case, eight alternatives). The third experiment examined how delay in the presentation of evidence compared to the start of the decision process affects the overall accuracy and reaction time once information is given. The results demonstrate that as the decision process extends in time, decisions are made faster and less accurately. Finally, the fourth experiment examined how the interrupting the incoming stream of information with either highly informative or highly misleading evidence would affect the decision. The results illustrated a complicated picture in which highly informative evidence accelerated decisions but misleading evidence failed to slow decisions. In addition to the individual aims, a secondary analysis investigated potential trial-by-trial variation in performance. There was some evidence that participants made ongoing adjustments to their strategy dependent on performance, but only when feedback was available; previous trial status (e.g., highly informative vs. highly misleading evidence) did not affect current trial performance. In sum the results demonstrate that perceptual decisions do show high levels of sensitivity to a variety of manipulations, but fail to replicate many of the results from more complicated multiple attribute decisions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First and foremost I have to thank Mark Wheeler for always knowing when I need a pep talk, a talking to, or just time to talk aloud. Your open door advising policy has been key to the dual sense of independence and support I have had throughout the past five years. The impact your advising has had on my work ethic and resulting product is immeasurable and will last a lifetime. I also have to thank Chris Schunn, Julie Fiez, and Carl Olson for their patience and enthusiasm throughout the dissertation process, with a special nod to Chris for numerous 10-minute chats that kept me on track. A large amount of appreciation goes out to Josh Tremel, Mandy Collier, and the many other undergraduates for whom I have had the privilege of supervising and guiding through various projects, some of which has morphed its way into the experiments presented here. I am a better teacher because of you. And a big thank you to all the other cognitive graduate students; I don't know how one could survive this process without the support, love, and inside jokes we have for each other.

I would also like to express my gratitude to Jay Anderson for putting up with a rollercoaster year of deadlines and documents. You have been a great support for all my needs from Excel formulas to al pastor tacos. And lastly I have to thank my family, who never doubted me when I decided to start this journey and still didn't lift an eyebrow the countless times I came home exhausted, deflated, and yet still tapping away at the computer despite the family festivities. Your love is truly irreplaceable.

1.0 INTRODUCTION

Imagine you are waiting for the bus at the end of a long workday. When it is sunny outside, you can clearly see the bus marquee as the vehicle comes closer, allowing you to easily determine the route number and consider the secondary decision of whether to hail the driver or step back from the curb. However, during a downpour or snowstorm, the bus marquee becomes difficult to see, with only bits of information reaching your eyes at any point in time. The decision about which route is designated on the marquee becomes difficult, even effortful, and you probably watch the bus for a longer amount of time in order to decipher the numbers. This example highlights how even simple perceptual acts like identifying the route number on a bus marquee can become a laborious decision process. Yet despite the increased difficulty and effort involved, the example fails to capture the additional complexity of the decisions that may follow the identification, such as determining whether or not that route will take you to your destination, how likely it is that a better bus will arrive soon, and remembering to take a particular route in order to stop by the grocery store on the way home. This dissertation will explore how closely perceptual decisions approximate the variability of the natural environment, such as the timing of incoming evidence, changes in the fidelity of that evidence, and similarity between potential alternatives. In addition, attention will be paid to whether or not the results from perceptual decision studies agree with the findings in multiple attribute decisions to determine whether perceptual decisions are truly a simplified version of the same process.

1.1 OVERVIEW

The decision process is a complicated series of stages that involves the reception of external input, the interpretation of that input, and the selection of a choice among output options (Schall, 1999). The perceptual decision process has been favored as a way to examine the decision process in a simplified environment, allowing for finer manipulation and study. However, there is some question as to whether the simplified perceptual decisions truly emulate the same behavioral characteristics as more traditional multiple attribute decisions. For example, past research has shown that the amount of noise in a stimulus can change both the accuracy and the reaction time of an impending decision (e.g., Ratcliff & Rouder, 1998). Although this lends insight into the decision process, previous studies have generally involved only two response options (e.g., bright/dark, red/green, left/right). This approach does not accurately reflect more common situations in the environment that involve the vetting of multiple possibilities at once. Research on multiple attribute decisions has shown that increasing the number of options can change decision dynamics, either slowing response time or causing adoption of strategies to speed decisions (Tversky & Sattath, 1979; Einhorn & Hogarth, 1981; Payne, Bettman, & Johnson, 1988). However, it remains unclear whether perceptual decisions will also show similar changes in behavior to the multiple attribute counterparts they are expected to simulate.

There are also other aspects of the perceptual decision literature that fail to capture some of the variability of the real world. Although different levels of stimulus fidelity may be assessed across trials (e.g., more or fewer white pixels compared to black pixels during brightness discrimination), little has been done to test changes in fidelity *within* a trial (e.g., a shifting amount of white versus black pixels during a single trial). When trying to recognize the bus marquee during the snowstorm as described above, the snow would be constantly shifting in the wind and causing moments of clearer views and more obstructed views. However, we must integrate information over the course of those more informative and less informative moments in order to reach our final identification. Therefore it is important to understand not only how perceptual decisions are reached with static amounts of evidence, but also how the process handles fluctuations in the fidelity of evidence over time.

In addition, the timing used in the majority of tasks has involved the presentation of a stimulus after a fixation period, creating an association between the start of a trial and the presentation of perceptual evidence. However, the natural environment rarely has clear signals for when evidence towards a decision may arrive. Before accepting that perceptual decisions adequately emulate the greater general decision process, particularly when discussing potential related brain mechanisms, this dissertation attempts to answer some of these open questions.

Examining perceptual decisions affords the researcher an opportunity to remove or control task and stimulus parameters, such as bias or expertise, which might be indefinable or uncontrollable in more complex situations. There is also the expectation that determining the mechanisms underlying basic perceptual decisions will serve as a theoretical base for more complex or abstract forms of decision-making, such as the determination of whether a presented item has been stored in memory or not. However, research on perceptual decisions spans several tasks, stimulus sets, and goals, making comparisons across studies difficult. Previous studies have demonstrated that even small changes in task structure can change the decision process (Ebbesen & Konecni, 1980), highlighting the importance of building a set of experiments using a single stimulus and task paradigm with only minor manipulations. Constraining the current series of experiments to the same stimulus and highly similar task structure should allow for cleaner examination of how each manipulation affects the decision process and limit the effects of extraneous variables.

The remainder of this chapter will be an overview of the extensive previous research on perceptual decisions to date, including behavioral studies, neurophysiological studies using macaques, and the limited human research thus far using electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). There will then be a brief description of the open questions this dissertation attempts to answer, with previews of the experiments and results presented throughout this document. There are two overarching purposes threaded throughout the experiments. First, the experiments are intended to serve as a bridge among isolated behavioral and neurophysiological experiments described throughout the introduction that involve a variety of stimuli and task structures. By presenting a series of experiments with the same basic stimulus and task paradigm but individual manipulations related to the literature, hopefully a more cohesive picture of what affects perceptual decisions will arise from the results. Second, references will be made throughout this work to findings from multiple attribute decision studies. By examining the parallels and divergences between the perceptual and multiple attribute decision literature, we can better understand how well studying the perceptual decision process emulates the general decision process and what kinds of generalizations are possible. However, before setting up the experimental crux of the problem, it is important to consider the theoretical roots. The following section will describe the theoretical underpinnings and assumptions used throughout this document, as well as some of the language from mathematical modeling that is used for descriptive and illustrative purposes.

1.2 THEORETICAL AND MATHEMATICAL ASSUMPTIONS ABOUT DECISIONS

Decisions can come about based on internal information, external stimulation, or some combination of the two, and may or may not carry an extrinsic reward. Finally, some decisions may be made instantaneously, while others can take several seconds, minutes, or longer. The process can be simplified into the basic model depicted in Figure 1; through a sequence of stages information is received, a decision process occurs in which that information is evaluated and evidence towards a particular alternative is extracted, and an output is selected. After some decisions, a reward may occur which reinforces the chosen output. For example, in previous work (Ploran et al., 2007), I used fMRI in humans to separate brain activity into patterns indicative of three stages of information processing: sensory input, evidence gathering, and decision (Figure 1, right panels). Fast-rising, sustained levels of activity in regions in the occipital lobes were indicative of sensory processing, activity in parietal and middle frontal areas for which the rate of increase in activity changed in accordance to the timing of the final response suggested an ongoing evidence gathering process, and areas in the frontal lobe responded transiently to the specific response output (Ploran et al., 2007; Ploran, Tremel, Nelson, & Wheeler, submitted). This separation of patterns lends support to the use of the model depicted in Figure 1, at least at a broad scale as required by the spatial limitations of fMRI. Due to the breadth of this phenomenon, this document will focus on the decision process after information has been received and prior to the selection of output, referencing data from healthy humans and macaque monkeys during perceptual decisions.

However, returning to potential theoretical assumptions without limitations, the evidence gathering stage may involve several subprocesses. Rachlin (1988) suggested separating the identification of alternatives, evaluation of alternatives, and a decision mechanism from one another (Figure 2). These subprocesses work well with the theme here that there are separate levels of complexity that can individually affect evidence gathering and the selection of output. In the bus marquee example used earlier, the decision process changed dependent on the clarity of the marquee through particular weather conditions (i.e., evaluation of alternatives). Yet, the decision process may also change depending on whether you know of four or eight routes that use the particular bus stop by which you are standing (i.e., identification of alternatives). Although each of these stages can individually affect the decision process, it is more likely that the quality and/or amount of detail contained in the input interacts with the number of possible alternatives to make the largest impact on the decision process. Due to the diversity of language used to describe the decision process, the following terms will be used throughout this document to represent the three processing stages: *information* will describe incoming sensory input, *evidence* will describe what the decision process derives from the information and integrates together, and *response* will be used to indicate the termination of the process and resulting output.

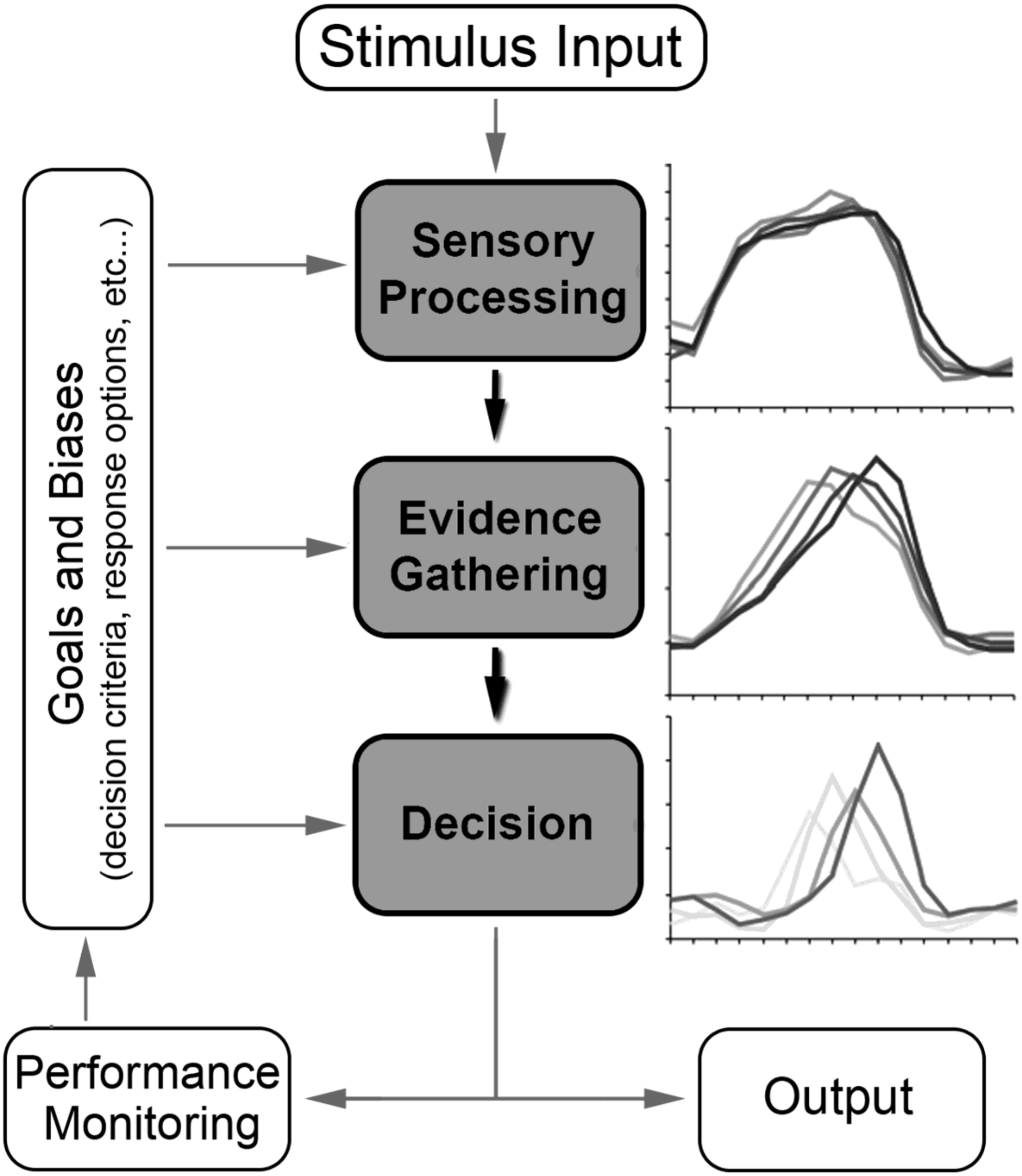


Figure 1. Simplified framework of the decision process with three major stages. To the right of the stages are fMRI results that lend to support to the dissociation of timing (Ploran et al., 2007).

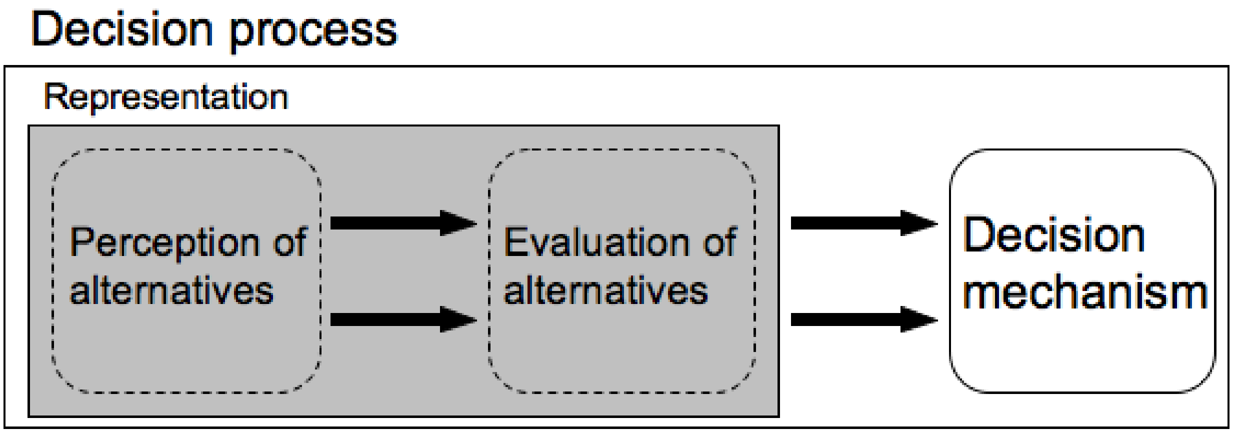


Figure 2. Separation of the evidence gathering stage into multiple subprocesses (modified from Rachlin, 1988).

The evidence gathering or evaluation stage of the decision process has been represented by variety of mathematical models; the two most popular are race models (Smith, 2000; Usher & McClelland, 2001; Ratcliff, 2006) and diffusion models (Ratcliff, 1978, 2002; Ratcliff & Rouder, 2000; Ratcliff & McKoon, 2008). Both types of models are sequential-sampling models that use representations of evidence that change with incoming information over time and have to pass a threshold in order to trigger a response. However, the models differ in the number of representations for response alternatives, the manner in which evidence is accumulated, and the placements of threshold. Due to the frequent use of these models to account for behavioral perceptual decision data (e.g., Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008) and neurophysiological data (e.g., Gold & Shadlen, 2001), several modeling terms will be used throughout this dissertation when discussing potential mechanisms and parameters of the neural underpinnings that could create the behavioral results seen here. The following will serve as a brief primer on the type of parameters these models use, the similarities between the two types, and how we might use these models to create predictions for the current studies.

While race models have a separate "counter" for information (or evidence) towards each possible response (Figure 3b, 3c), diffusion models use a single representation for evidence (Figure 3a); this leads to a major difference between race and diffusion models. Because race models have separate counters for each possible response, the decision is literally a race to see which counter reaches threshold first (Figure 3b, 3c). Each counter increases with supportive evidence, but decreases after opposing evidence. However, as more response options are added to the race model, increases and decreases of equal magnitude would make it increasingly impossible for any one counter to pass threshold due to low positive-to-negative evidence ratios. Therefore, negative evidence for each response must be scaled by the number of available responses (Usher & McClelland, 2001). It is possible, with the right number of responses and supporting evidence among those responses (i.e., even distribution of positive and negative evidence), to have a very close race in which the response that passes threshold is only slightly ahead of the next best option (for more detail on race / accumulator models, see Usher & McClelland, 2001).

In diffusion models, on the other hand, the evidence is gathered in only one parameter that represents the balance of information between the two possible responses. This leads to a tug-of-war in the direction of the parameter according to the available evidence; the evidence parameter can only pass threshold and finalize a decision after a continuous stream of evidence for the same response, rather than a noisy stream with evidence for both responses (Figure 3a). However, because diffusion models are only able to have two possible responses at the most (there are also single outcome diffusion models) and evidence is gathered in one parameter, the tug-of-war happens on an equal basis. Unlike the race model, there is no scaling of negative evidence in relation to the number of possible responses; evidence for and against a response in a diffusion model is weighted equally. This causes a situation in which as the signal-to-noise ratio decreases (and therefore some amount of evidence is available for each option), it becomes increasingly difficult for the parameter to gather enough evidence and pass threshold (Figure 3a, dashed line). On the other hand, if there is a high signal-to-noise ratio in favor of one response, the parameter will quickly drift towards the boundary for that response (Figure 3a, solid line; for more detail on diffusion models, see Shadlen et al., 2006; Busemeyer & Townsend, 1993).

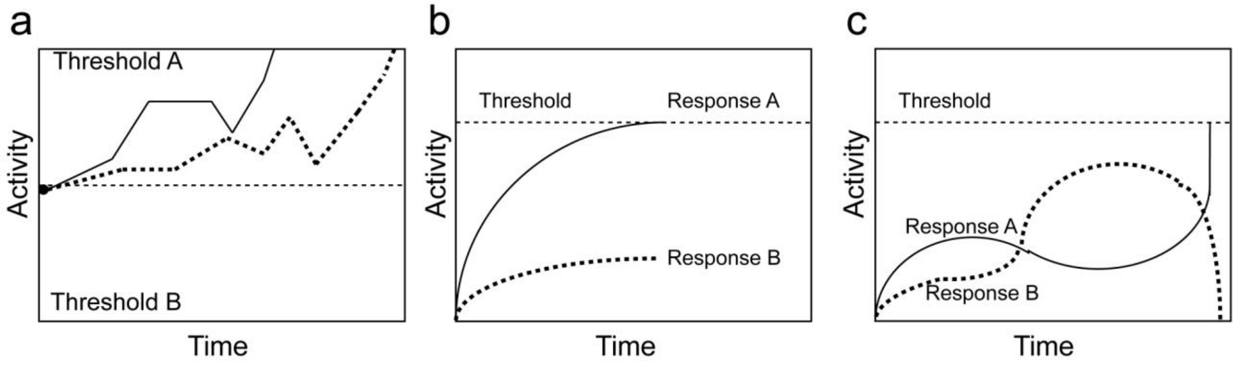


Figure 3. Examples of the mechanisms behind diffusion and race models. (a) Diffusion model with thresholds for each option and drift parameters representing two separate decisions: a fast decision (solid line) and a slow decision (dashed line); (b, c) Race model in which each line represents a counter and both accrue at the same time for either fast (b) or slow (c) decisions.

The main parameters that affect the trajectory within a diffusion model are starting point, drift rate, and threshold. The starting point is the point from which the decision parameter begins accumulating evidence. By allowing the starting point to vary trial-to-trial, diffusion models can adequately account for fast errors; if the starting point is much higher than usual, it only takes a few bits of information to cross the upper threshold even if the next bits of information would lead the model to cross the lower threshold instead. The drift rate is the average amount of increase towards a threshold per bit of information. Finally, the threshold is the level of activity needed in order to trigger a decision; the threshold can be manipulated to be more liberal or conservative depending on the requirements for the task (e.g., speed versus accuracy, Fitts, 1966; Ratcliff & McKoon, 2008). Aside from these parameters that vary according to task, there is usually also a constant (sometimes with its own trial-to-trial variability) that accounts for all the non-decision computation, such as the time it takes to execute a motor output indicating a decision. Although there has been a larger discussion of whether or not these are the only necessary parameters and the need for variation in them (Busemeyer & Townsend, 1993), these represent the parameters that overlap in most diffusion model accounts.

Despite differences, both types of models include a threshold and include some fluctuation in the decision parameter as a function of incoming stimulus noise. Although the race model would be more appropriate to use for decisions with more than two choices (an important constraint for diffusion models that has been recently explored using neurophysiological data, Churchland, Kiani, & Shadlen, 2008), it is outside the scope of this dissertation to test the models for efficacy. Instead, the concepts within the models will be used as references and good theoretical parameters to consider when examining behavioral data. The combination of the starting point, drift rate, and threshold concepts have allowed models to account for a variety of behavioral data, including both memory (Ratcliff, 1978) and perceptual decisions (Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). These parameters have also been used to fit neurophysiological data from perceptual decisions (e.g., Gold & Shadlen, 2001), demonstrating the flexibility of these parameters to account for both behavioral measures like accuracy and reaction time, as well as the underlying neural processes. Therefore it is these three concepts that will be used to create questions, predictions, manipulations, and ultimately possible explanations for the results.

1.3 EVOLUTION OF BEHAVIORAL STUDIES OF PERCEPTION

The study of perception harks back to the earliest psychophysical experiments examining anatomy, physiology, and behavior. However, the study of perceptual decisions, in which the subject must respond to the stimulus in some instructed manner, has a more research history. Beginning with stimulus-response paradigms in which subjects make quick pre-determined responses to the presentation of a particular associated stimulus (Hick, 1952; Hyman, 1953), the study of perceptual decisions has evolved into a closer investigation of what affects the interpretation of a stimulus that in turn creates the behavioral response. The following section will describe some of the evolution of the study of perceptual decisions from a behavioral standpoint. Succeeding sections will then extend this evolution to studies of neurophysiology in macaques and the use of neuroimaging techniques in humans.

An important consideration that should be maintained throughout reading this document is what exactly constitutes a perceptual decision, versus the pure act of perception without interpretation. Humans make decisions constantly, often without conscious awareness. For example, perceptual recognition is usually an automatic process; the ability to identify and respond with an object's name can occur within 1000 ms of presentation (Rossion & Pourtois, 2004). Similarly, word identification can occur within 750 ms of presentation, for words up to 6 letters in length and 2 syllables (Frederiksen & Kroll, 1976). However, if the object or word is occluded or incomplete, more calculations among the alternatives must be made before a final identification response is selected. Instead of the fast, easy recognition made when all the perceptual information is available, you might pause to gather more evidence towards the correct choice. When not given enough information, the act of perception becomes an effortful decision process that requires gathering of evidence from incoming stimulus information and interpretation of that evidence prior to the selection of a response.

Behavioral research on perceptual decisions has largely focused on two-choice paradigms using motion, color, or brightness discriminations (e.g., Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). These studies collectively demonstrate that the timing and accuracy of a perceptual decision are functions of uncertainty (i.e., noise), both in the stimulus itself and the neural processes underling the decision process. The uncertainty in the stimulus can be a result of actual sensory qualities (i.e., the number of white versus black pixels; Ratcliff & Rouder, 1998) or the duration of the stimulus prior to masking (Ratcliff & Rouder, 2000). Biological sources of uncertainty arise from the noise in neural pools encoding potential alternatives (Gold & Shadlen, 2001). In any case, higher uncertainty creates longer and less accurate responses.

However, there are limitations to the use of 2-choice paradigms when attempting to describe the perceptual decision process more generally. First, it is rare that one is presented with two discrete choices. Often we are presented with noisy sensory information without a clear delineation of the possible identity. If there are multiple response options to vet in relation to the presented stimulus, processing may take longer than when there are only two response options. In fact, stimulus-response paradigms have long demonstrated that as participants are required to respond with unique responses to an increasing number of stimuli, the reaction time after each presentation is longer than when there are fewer possible stimuli among the presentation set (Hyman, 1953). After initial training, this paradigm doesn't even require a decision so much as the automated response selection from an unambiguous stimulus, yet the number of possible stimuli still has a strong effect (accounting for up to 97% of the variance in reaction time).

It has also been shown in decisions involving multiple attributes that an increase in the number of response alternatives can spontaneously prompt the adoption of strategies and heuristics to aid in decision making, particularly when certain time pressures and other task demands are present (Payne, Bettmann, & Johnson, 1988). For example, the "elimination by aspects" strategy (Tversky, 1972) allows the decision maker to rule out one or more response options very early in the decision period, thereby rerouting information towards the remaining possibilities. Alternatively, the "majority of confirming dimensions" strategy (Russo & Dosher, 1983) considers pairs of alternatives out of the many available options and selects the best of each pair to remain available while discarding the worst of each pair. However, there are many more possible strategies than described here, and in addition, it is possible to combine strategies into a more comprehensive or efficient approach (e.g., eliminate unlikely possibilities and then compare pairs; Payne, Bettmann, & Johnson, 1988).

A recent study using 2-, 3-, and 4-choice letter discriminations found increasing reaction times and decreasing accuracy as the number of response alternatives increased (Leite & Ratcliff, 2010). This suggests that the decision process changed as there were more response alternatives available for matching the evidence. However, it is unclear whether four alternatives are enough to cause significant changes in strategy as observed in multiple attribute decisions. The impact of the number of alternatives on the perceptual decision process is investigated further in 3.0 , including a post-experiment questionnaire on possible strategy adoption to assess whether or not participants attempted to mitigate any detriment to performance as the number of response alternatives increased.

Another consideration for the validity of previous perceptual decision research is that 2-choice experiments often present opposing choices (e.g., rightward vs. leftward motion, light vs. dark pixels). However, the external world is created from a continuum of sensory stimuli, and often we must make very fine distinctions between two similar choices (e.g., “Does that bus marquee say 71B or 71D?”). One previous study (Ratcliff & Rouder, 2000) did assess easy vs. hard letter discriminations (e.g., E vs. C., or E vs. F) and found that similar letter pairs were associated with lower accuracy and longer response times. Unfortunately, the study was focused on validating a diffusion model of decision making and the results based on degrees of similarity were not fully explored.

While the binary choice studies have laid a strong foundation for the examination of perceptual decisions, related research has demonstrated that the internal decision process may not be strictly dichotomous. Experiments allowing subjects to indicate more than one guess on various temporal and spatial selective attention paradigms have found that while the individual responses were discretely categorical in nature and subjects did not report being “half aware” of the targets (Sergent & Dehaene, 2004), their responses arose from graded distributions (Vul, Hanus, & Kanwisher, 2009). These results suggest that there is some noise in the underlying cognitive process prior to the decision. If the distributions for two similar options overlap, responses may be more difficult to make compared with options with non-overlapping distributions. This possibility is examined in 2.0 , which assesses how performance changes as the two available alternatives become more similar to one another.

A last consideration is the effect of prior performance on current decisions. That is, aside from within-trial noise and variability in the stimulus, does across-trial noise and variability also affect decisions? That is, is each individual decision based solely on stimulus aspects from the current trial, or also from the characteristics demonstrated about a stimulus over previous trials as well? Both temporal and spatial selective attention responses appear to be affected by only uncertainty in the stimulus itself on a given trial (within trial variability), rather than potential across-trial variability (Vul, Hanus, & Kanwisher, 2009). That is, the uncertainty of a stimulus within the trial has a larger effect on the decision process than the accumulated known uncertainty of stimulus identification over the previous trials. However, previous theoretical accounts of effort and achievement behavior have considered a more dynamic selection of effort that is affected by task difficulty, relationship of the current task to previous tasks, previous outcomes, and possibly motivation (see Thomas, 1983 for review). Given the similarity between the models for selective attention and perceptual decisions, it is important to test how ongoing performance affects each new decision, as well as how bias in previous trials (in the form of highly coherent information for or against the correct alternative) can affect the treatment of information on new trials, as was done for several experiments in 6.0 .

Together, these behavioral studies demonstrate the breadth of manipulations that have been used to study perceptual decisions. In particular, consideration should be made regarding the variety of stimulus sets (e.g., bright/dark pixels, red/green pixels, letters, etc.). The physical characteristics involved in each of these stimuli result in a variety of brain areas necessary for processing the sensory input prior to interpretation during the decision process. Also, some stimuli involved quantifiable amounts (brightness patches) whereas others did not (letters). These issues make it difficult to compare across studies, a problem that the current set of studies attempts to address by using the same stimulus throughout. In addition, some of the literature involved results from multiple attribute decisions (e.g., Payne, Bettman, & Johnson, 1988), highlighting the question of whether perceptual decisions adequately emulate a more general decision process in simplified form. The discussion of results from the current experiments will include references to both the perceptual and multiple attribute decision research in an attempt to bridge between these two disparate fields.

1.4 NEUROPHYSIOLOGICAL SUPPORT FOR DECISION MODELS

There is a wealth of literature on the primary sensory, secondary sensory, evidence gathering, execution of response, and reward/feedback stages of perceptual decisions using neurophysiological recordings in macaques and rats. These studies span various visual paradigms (Shadlen & Newsome, 2001; Roitman & Shadlen, 2002; Horwitz, Batista, & Newsome, 2004; Lo & Wang, 2006), flutter/grating sensory discriminations (Romo & Salinas, 2003; Salinas, Hernandez, Zainos, & Romo, 2000), auditory discriminations (Lemus, Hernandez, & Romo, 2009), and even responses based on gustatory stimuli (MacDonald, Meck, Simon, & Nicolelis, 2009). From this broad testing of the various sensory circuits, there has been the emergence and collective adherence to a systems level theory of decision making that applies to all these situations. This research has supported the notion that decisions arise from a series of hierarchical brain areas, spanning sensory, evidence accumulation, and decision execution processes. An example of which brain areas may be involved and how each area might affect the decision process will be described presently.

A large portion of the research on perceptual decisions has been conducted using neurophysiological measures of visual discrimination in macaques. Importantly, several researchers have adopted a motion discrimination paradigm as an easy way to manipulate the perceptual information available for the decision. The adoption of this stimulus presentation across many studies has created a cohesive narrative about the potential development of a decision across multiple brain areas. The knowledge of how various manipulations affect responses during this paradigm was the primary reason for its use throughout this dissertation. Therefore, it is important to consider both the behavioral and neural patterns found within the neurophysiological literature when constructing the experiments and predictions described later.

The motion discrimination paradigm has been used with both humans (Shulman et al., 1999; Shulman et al., 2003; Palmer, Huk, & Shadlen, 2005) and non-lesioned macaques (Celebrini & Newsome, 1994; Kim & Shadlen, 1999; Shadlen & Newsome, 2001), and lends itself to a variety of manipulations. The basic paradigm involves a screen with a large number of dots that move randomly, making the display look like a simple version of television static. Researchers can manipulate how many dots move together coherently in one direction, and subjects discriminate the trajectory of motion. By varying the coherence of the dots (i.e., how many dots are moving together), the quality of the motion signal is manipulated and the discrimination of the direction becomes more or less difficult. In an easy discrimination (>50% coherence), decisions are made quickly, because the signal clearly outweighs the noise. At lower levels of coherence (<20%), the signal of the motion is disrupted by the majority noise of the randomly moving dots and discrimination becomes difficult.

By examining the single-unit neurophysiology research in monkeys, it appears that perceptual decisions arise from activity spreading from the highly motion selective area middle temporal area (MT) (Celebrini & Newsome, 1994), to visually responsive areas involved in preparation of eye movements such as the lateral intraparietal area (LIP) (Shadlen & Newsome, 2001; Roitman & Shadlen, 2002; Hanks, Ditterich, & Shadlen, 2006) and superior colliculus (Horwitz, Batista, & Newsome, 2004; Lo & Wang, 2006), and finally into the dorsolateral prefrontal cortex (dlPFC), specifically the frontal eye fields (FEF). These final frontal areas are non-visually responsive areas that may be involved in motor planning, reward, or attention (Kim & Shadlen, 1999; Leon & Shadlen, 1999). Collectively, the main finding of these papers was that brain areas involved in the earlier stages of visual processing modulated activity based on the coherence of motion, while areas farther upstream reacted less to stimulus properties and more to the impending decision (i.e., the preparation of eye movement indicating the decision).

Early experiments testing motion discrimination in LIP demonstrated a difference in the increasing or decreasing of neural activity dependent on both the coherence of motion and the final decision (Shadlen & Newsome, 2001). Specifically, when highly coherent motion towards the response field caused neurons to increase in activity very quickly; however, when motion was weak (but still towards the response field), neurons increased in activity at a slower rate. Similar differences in the rate of change in neuronal activity were found during motion away from the response field, such that neurons decreased in activity more quickly for strong motion away from the response field than weak motion (Shadlen & Newsome, 2001). Interestingly, despite the differences in rate of increase or decrease, activity reached a common level by the end of the trial, prior to the saccade to indicate decision. In addition, a later study (Huk & Shadlen, 2005) incorporated short (100ms) bursts of highly coherent motion within a subset of trials and found the accumulation rate of neural firing changed in trajectory and reached the common level earlier than without the burst. That is, while initially accumulating normally, the burst caused the accumulation rate to increase mid-trial. However, in both these studies the monkeys were required to maintain fixation until a cue, so it is unknown whether or not this would have resulted in faster reaction times (e.g., faster rise in activity to the common plateau level would have triggered a threshold and caused a detection response).

The LIP studies demonstrated how neurophysiology can be an extremely useful methodology since it allows for close examination of neural activity on the millisecond level. This affords the opportunity to consider differences in activity between conditions in a very minute way. However, one limitation of the aforementioned LIP studies was the use of a delayed response task. As mentioned above, the monkeys were required to wait after seeing the motion display until a cue gave permission to make the saccade indicating the decision. This wait prior to response created a situation in which inferences were made about the decision process without knowing the precise moment at which the process ended, leading to possible confounding of decision processes and the maintenance of a motor plan. By using a reaction time experiment in which the monkeys made a saccade upon making a decision, Roitman & Shadlen (2002) were able to examine the activity leading directly up to the decision in a more natural timecourse. The comparison of activity for different levels of coherence illustrated the same modulation of activity throughout the trial, such that higher levels of coherence were associated with faster increases and lower levels of coherence with slower increases. However, the important comparison grouped trials by reaction time regardless of coherence, eliminating any modulation based on sensory efficacy of the stimulus. This comparison illustrated that faster increases were also associated with earlier response times, and slower increases were associated with longer response times. Combined with the modulation of activity based on coherence, the difference in relative timing of activity and its relation to response time is a good indicator that area LIP is involved in some stage of the decision-making process – potentially as an integrator between sensory information and the decision variable (Roitman & Shadlen, 2002; Mazurek et al, 2003).

Perhaps the most interesting results are from a very carefully controlled experiment attempting to eliminate the confounding of the decision choice with movement planning to indicate that choice (Horwitz, Batista, & Newsome, 2004). While the traditional motion discrimination paradigm presented both the central viewing window and the two dots marking the possible decisions in their relative locations, the new paradigm moved the dots around the screen as a pair so that the monkey had to select the dot based on its position relative to the other dot, rather than the viewing window. The change in the paradigm created a situation in which a rightward motion stimulus could result in a leftward eye movement, even though the eye movement would target the right dot out of the pair. Movement planning could now be in opposition to the perception of motion. However, the most crucial aspect of the new paradigm was that the decision targets were only shown after a delay during which no motion was present. This means that any potential evidence accumulation that occurs should only be related to the decision, and not the planning of the movement to indicate that decision. Indeed, the researchers found that a small subset (~10%) of neurons in the superior colliculus demonstrated modulation during the presence of motion that was indicative of the decision, even without available targets for which to plan an eye movement. This suggests that some neurons are able to hold abstract representations of a decision without the explicit targets necessary to execute action based on that decision (Horwitz, Batista, & Newsome, 2004).

Recent research has expanded the motion discrimination task from two to four alternatives (Churchland, Kiani, & Shadlen, 2008). The main finding was that increasing the number of alternatives resulted in a decreased initial firing rate towards the correct alternative, with a slower buildup to a stable threshold. Behaviorally, 4-choice decisions took longer and were less accurate than 2-choice decisions. This suggests that increasing the number of alternatives may change how neural areas process the task, altering accuracy and reaction time. This potential effect was examined more closely in 3.0 .

The control condition for the 2- vs. 4-choice experiment was to include a 2-choice trial in which the alternatives were perpendicular to each other rather than opposite (Churchland, Kiani, & Shadlen, 2008). These 90° trials resulted in an intermediate decrease in reaction time and firing rate buildup compared to the 180° trials. This suggests that there was a processing difference between the original opposing response alternatives and the slightly more similar perpendicular alternatives. The experiment in 2.0 examines how similarity of response alternatives may affect accuracy and reaction time, with the general prediction that as similarity increases the decision process becomes more difficult.

These neurophysiological experiments demonstrate several important characteristics described by the decision making models. First, the change in the rate of increase in firing dependent on the response may represent the evaluation of alternatives (Figure 2) that we would expect given different efficacies of evidence. Importantly, Roitman and Shadlen (2002) illustrated that these changes in rate are associated with changes in response time independent of stimulus efficacy. This ties into the second characteristic of decision making highlighted here: the need for a stable threshold. In these experiments, neurons throughout the brain appeared to reach a steady threshold level prior to a decision, in both the delayed and immediate versions.

Despite changes to the rate of increase in firing, the decision was not made until this threshold is reached, indicating a stable cut-off mechanism in the decision process. Therefore increases in activity in sensory areas like MT in response to increases in coherence of the stimulus in turn increase the rate of firing in evidence gathering areas like LIP and superior colliculus, causing the threshold to be reached sooner.

1.5 NEURAL BASIS OF PERCEPTUAL DECISIONS IN HUMANS

The neurophysiological literature highlights the important role that timing plays in determining whether or not an area of the brain is involved in sensory, decision, or movement planning processes. By using paradigms that attempted to separate the processes, the neurophysiology literature has made significant headway into determining the pathways for certain visual and somatosensory tasks. However, due to the invasive nature of the methodology, it limits these studies to the use of non-human primate subjects. Fortunately, advances have been made in the analysis of the blood oxygen level dependent (BOLD) signal within fMRI, high-density electroencephalogram (EEG) recording, and the concurrent use of fMRI and event-related potential (ERP) recordings. These advances now allow researchers to make some statements about the comparative timing within a region, and tentatively make comparisons across regions.

For example, McKeeff & Tong (2007) created a face/non-face paradigm using Mooney contorted images of faces; the Mooney transformation alters a grayscale image into pure black and white contrast, masking the usual contours and lines of the original image. Because facial features can often be very small in scale, altering the quality of the image by using this technique highly increases the difficulty of the discrimination. Interestingly, among the trials for which the stimulus was categorized as a face, there were timing differences in fusiform gyrus (a highly face-selective area) dependent on the response time. Faster face categorizations were associated with faster BOLD rise times, while slower discriminations had slower rise times. However, this difference was not present for trials in which the stimulus was categorized as a non-face. The authors suggest that this timing difference was indicative of top-down influences on fusiform face area as evidence of face contours was gathered from the otherwise indecipherable image (McKeeff & Tong, 2007). This suggests that since the participant was actively looking for a face, this top-down information drove the search for facial contours. However, the simpler answer may be that as lower visual areas parsed the contours of the incoming sensory information into recognizable facial features, more activity occurred in fusiform gyrus from feedforward connections. As soon as this activity reached some threshold, the brain could categorize the image as a face. If the activity in fusiform gyrus did not reach some predetermined threshold, the found contours were deemed noise, and the image was categorized as a non-face. Consequently, on trials when there is a faster increase in activity in fusiform gyrus, categorization occurs quicker than when activity rises more slowly because the threshold is met faster.

The quality of facial features was also used by Thielscher & Pessoa (2007) to create an emotion discrimination experiment using faces with varying levels of expression; however, in this case the images themselves were not transformed. Instead, the researchers maintained clear images, but manipulated the extent of the frown and openness of the eyes to create varying degrees of expression. By focusing the analysis on the brain areas active during the discrimination of neutral stimuli, the researchers were able to examine how activity evolved when there was little drive by the sensory stimulus. Instead, brain areas active during the categorization of neutral stimuli were assumed to represent the decision process. The main analysis again suggested a series of brain areas stretching from visual areas, through emotion-related areas, and finally to higher-level areas.

However, the more interesting analysis considered how the activation in a decision area should correlate with response time. The authors suggested that if a brain area is associated with the decision, the activation should mimic the response time curve. That is, trial-by-trial fluctuations in the BOLD response should correlate with trial-by-trial RT in a systematic way. Indeed, this is what Thielscher & Pessoa (2007) found; activity in the anterior cingulate, inferior frontal gyrus, and middle frontal gyrus correlated with response time such that longer RTs were associated with more activation. This runs counter to the argument presented by Heekeren, Marrett, Bandettini, & Ungerleider (2004), in which more activity in a decision area was thought to reflect a bigger difference in the two areas of comparison. If the brain area encoding one response (e.g., face) was significantly more active than the brain area encoding the competing response (e.g., house), the "decision" brain area would have an easier comparison and therefore an easier decision with most likely a faster response time. It is possible that both these theories are true, but simply represent different aspects of the decision process: signal detection via Heekeren et al. (2004) and some sort of continued processing via Thielscher & Pessoa (2007). Indeed, the region found by Heekeren et al. (2004) was in a separate area than the area found by Thielscher & Pessoa (2007), highlighting potential functional differences. This illustrates how elusive the process of decision-making can be, and further highlights the complexity of the process.

The observation of an increased BOLD activation that correlates with longer response times inevitably arouses the suggestion of an area related to task demand or task difficulty. While this does not preclude the area from involvement in the decision process, it does place certain boundaries on function. For example, when attempting to separate the temporal signatures of the decision process by using ERP during a face/object discrimination, Philiastides, Ratcliff, and Sajda (2006; see also Philiastides & Sajda, 2006) combined manipulation of difficulty within the task as well as difficulty between two tasks using the same phase-coherence shifted images. By using the scrambled images for both the difficult face/object and easy red/green discriminations, the researchers were able to separate the temporal signatures of visual areas responsive to the particular stimulus from general task difficulty areas. The results showed two points at which the ERP waveform separated dependent on condition. There was first an early component at 220ms that modulated based on task difficulty; there was increased negativity during the difficult face/object task with low coherence stimuli, but no increase in negativity during the easy face/object task with high coherence stimuli. Furthermore, when the low coherence stimuli were used in the easy red/green task (which did not require attention to details masked by the phase-shift), the negativity did not appear. This lack of negativity during the easier decisions suggests that this early component was only indicative of task difficulty, not the decision process. However, since it is difficult to isolate the source of ERP waveforms, it is not possible to know whether or not this component overlaps with the increased activation seen by Thielscher & Pessoa (2007).

Another interesting finding with the face/object paradigm was a late component (starting at 300ms) that appeared to correlate with response time. However, unlike the early component that modulated based on stimulus properties, the late component was present regardless of the coherence, suggesting it was not related to task difficulty. More importantly, the timing of onset for the late component was correlated with coherence level and the length of activity was correlated with response time. By comparing these parameters with a diffusion model based on the behavioral data, the authors found that the activity in the late component was highly correlated with the mean drift rate proposed by the model for each trial. While this appears to suggest that the late component was indicative of an evidence accumulator, the authors were careful to note that the activity was not found during the red/green color categorization task; this means that either the activity was only strong enough to detect during more difficult tasks, or the activity was explicitly involved in the face/object discrimination similar to the early visual peaks they found rather than a general mechanism involved in decisions (Philiastides, Ratcliff, & Sajda, 2006).

While ERP studies can lend some support for the relative timing of components, it is difficult and unreliable to identify the source of these components. This highlights the benefits of a multimodality approach across electrophysiological and neuroimaging methods. However, it is widely known that the underlying mechanics of fMRI do not easily fit into the discussion of timing within neural processes. Fortunately, recent studies have demonstrated that fMRI can be used to separate temporal signature given proper design and analysis techniques (Ploran et al., 2007; Wheeler et al., 2008: Ploran, Tremel, Nelson, & Wheeler, submitted). By obscuring images of common objects, these studies were able to draw out the object recognition process for up to 14 seconds. Comparisons of recognitions that occurred after different lengths of time revealed that the process extends through a series of brain regions starting in occipital lobe, through parietal lobe, and finishing in frontal lobe. In addition to the breadth of activity, these studies also demonstrated different patterns of activity suggestive of varying stages of the decision process. Occipital areas contained a sustained level of activity throughout the trial, indicating constant sensory stimulation. Parietal and some frontal areas demonstrated growing levels of activity that shifted in peak coinciding with the recognition response, suggesting the accumulation of evidence. Lastly, frontal and some subcortical areas were active transiently at the time of the decision, but not beforehand, indicating a tight coupling with response execution.

In addition, these patterns remained despite a manipulation in which the object identity was not consistent during the trial (i.e., the available information was constantly changing). Despite the increased difficulty, activity in parietal and frontal areas continued to demonstrate an increasing pattern dependent on the timing of the response as participants attempted to recognize the object. Importantly, activity during these inconsistent trials was not any greater than trials during which object identification was consistent, ruling out that the activity was due to effort. These diverse patterns not only illustrative of how and when different brain areas may be involved in the process, but also demonstrate that fMRI can be a useful technique for investigating timing differences if the task is adapted appropriately.

Collectively the neuroimaging studies on humans lend support to theory of perceptual decision described earlier, with information progressing through a series of stages (Figure 1). Changes in sensory areas related to the stimuli correlated with changes in anterior executive areas (Heekeren, Marrett, Bandettini, & Ungerleider, 2004). Additionally, there is evidence that sensory areas were active earlier than parietal and frontal areas that may be evaluating the incoming information for evidence towards the response (Philiastides, Ratcliff, & Sajda, 2006; Ploran et al., 2007; Ploran, Tremel, Nelson, & Wheeler, submitted). However, neuroimaging and electrophysiology highlighted activation in large areas of the brain. The lack of precision and general overlapping nature of neuroimaging results from a wide variety of studies can make interpretation of these results difficult. Before advancing further, it is important to understanding the types of changes certain manipulations to the stimulus and decision environment can create in order to better predict where and when changes may occur in the brain. By creating strong a priori hypotheses about the processes and potential brain areas involved based on the results of behavioral studies, neuroimaging results will gain strength and stability.

1.6 OPEN QUESTIONS AND OVERVIEW OF DOCUMENT

Although the research on perceptual decisions has been extensive, it is not without shortcomings. For example, the use of the 90° control condition by Churchland, Kiani, & Shadlen (2008) caused changes in neural activity that may have represented the unintended effects of similarity. As this was not the focus of the study it was not examined further, yet leaves an interesting question as to how the similarity of response options (in this case a smaller angle of disparity between two possible directions of motion) may affect the extraction of evidence from incoming sensory information. Also, the limited attempt to examine how the number of responses alternatives affects the decision process (Leite & Ratcliff, 2010) failed to consider how the use of letters as stimuli creates comparison problems to the more quantifiably manipulable stimuli like color or brightness patches from previous experiments (Ratcliff & Rouder, 1998).

This dissertation will address some of these shortcomings by systematically assessing aspects of perceptual decisions using a single quantifiable stimulus – random dot motion. The first experiment (2.0) will follow up on Churchland, Kiani, & Shadlen (2008) by testing how the similarity of two presented possible directions impacts the speed and accuracy of the decision process. The second experiment (3.0) will follow up on Leite & Ratcliff (2010) by testing whether increasing the number of choices to a greater extent continues to negatively impact performance. The remaining two experiments will start to explore perceptual decisions in less predictable environments. The first will assess how delaying information towards a decision after the decision process has started affects overall accuracy and speed once the information is presented (4.0). This represents a potentially more realistic assessment of the process, as the environment does not always give cues to when information to be evaluated is forthcoming. The second will assess how altering the fidelity of evidence during the decision process affects the ultimate decision (5.0). This manipulation mimics previous studies in macaques (Huk & Shadlen, 2005) in an effort to bridge between the human behavioral and macaque neurophysiological literature. Finally, a trial-by-trial analysis of the data will assess whether performance changes based on previous trial characteristics such as accuracy, timing of evidence, or stimulus fidelity. This analysis will address whether perceptual decisions follow similar behavioral patterns to multiple attribute decisions, which see variability, or visual and spatial attention studies, which do not. These questions are characterized more fully below.

1.6.1 Question #1

Research in multiple attribute decisions has shown that similarity of the alternatives can affect decision processes (Biggs, Bedard, Gabor, & Linsmeier, 1985; Einhorn & Hogarth, 1981). The study of perceptual decisions often uses a basic visual stimulus that varies along a single continuum (e.g., brightness). Sometimes these continuums are used to create stimuli that are difficult to discrimination (e.g., patches of almost equal brightness, Ratcliff & Rouder, 1998). More often, the continuum is ignored and choices reside at opposite ends of the scale (e.g., right vs. left motion in a variety of neurophysiological experiments such as Shadlen & Newsome, 2001). Although there has been some interest in how similarity affects perceptual discriminations, the limited studies have used letter tasks in which it is hard to quantify the similarity or manipulate it at fine grain levels (Ratcliff & Rouder, 2000). As mentioned above, one study which did start to capitalize on additional directions in the motion discrimination experiment found potentially interesting neurophysiological results that suggested a change in the process, but failed to fully explore the behavioral and neural results.

2.0 discusses an experiment that investigated the effect of similarity of options, in this case possible directions of motion, on accuracy and reaction time. It was predicted that as alternatives become more similar, the decision process must evaluate increasingly overlapping pools of neurons. The overlapping activity would then result in more time needed to make a decision with less accuracy. The results demonstrate that alternatives must be highly similar in order to affect decision processes. However, once there is a high enough similarity, performance is affected in both accuracy and reaction time measures.

1.6.2 Question #2

Until recently, the majority of perceptual decision experiments to date involve binary choices (c.f., Churchland, Kiani, & Shadlen, 2008; Leite & Ratcliff, 2010). While prior stimulus-response studies have demonstrated a linear relationship between number of possible stimulus and speed of response (Hyman 1953), this does not capture the process that occurs when stimuli are ambiguous and do not require a conditioned response. In order to understand the naturalistic dynamics of decision making, it is necessary to test how increased available response alternatives affect decision times. Assessing decision time as a function of response number will determine whether large numbers of options are all taken into consideration, or whether the response set is strategically reduced prior to the commitment to a decision. A linear increase in decision time with number of options would support the former, similar to the linear increase in memory search times with set-size increases found in the Sternberg task (Sternberg, 1966) and the stimulus-response paradigm (Hyman, 1953). On the other hand, a selective strategy might be used to eliminate clearly wrong answers, focusing assessment on the most likely answers (Payne, Bettmann, & Johnson, 1988). This may create non-linear asymptotic changes in speed as the number of response alternatives increases.

3.0 discusses an experiment in which 2-, 4-, and 8-choice versions of the motion discrimination paradigm were tested. Competing hypotheses predict that participants will either adopt strategies to mitigate the increasing evaluation load (by potentially eliminating obvious non-choices or other means), thus saving reaction time and performance from detriment as choices increase. On the other hand, participants may not adopt strategies and instead spread the decision process to vet each new response alternative, resulting in longer decision times and lower accuracy. The results suggest that increasing the number of possible response alternatives does affect both accuracy and reaction time, though not necessarily simultaneously. Increasing choices from 2 to 4 options decreased accuracy despite similar reaction times, whereas increasing from 4 to 8 choices both decreased accuracy and increased reaction time. A discussion of how increasing the number of choices and the accompanying increase in similarity of those choices is included.

1.6.3 Question #3

Previous studies of perceptual decisions appear to fall into two procedural categories: brief, isolated presentations of stimuli (e.g., Ratcliff, 2002; Shadlen & Newsome, 2001) or continuous presentations of stimuli until response (e.g., Ratcliff & Rouder, 1998; Roitman &

Shadlen, 2002). Importantly, past studies have used a predictable relationship between the start of a trial and the presentation of the stimulus with minimal variation within an experiment (c.f., Green, Smith, & von Gierke, 1983). However, the decision process should be able to handle a dynamic environment in which evidence occurs at any point past the initial knowledge of its impending arrival. In addition, it has been shown that time constraints can affect strategy selection in decision making (Payne, Bettmann, & Johnson, 1988). This leads to the open question of whether the timing of incoming evidence after search onset affects the speed with which a decision is made. If so, how does early information get incorporated into the longer decision process?

Two experiments were conducted to test this question. 4.0 describes an experiment in which the informative stimulus was presented at varying times in an otherwise random background to test isolated effects of timing. If the decision process involves a timeout function to limit the spending of mental resources on an unsuccessful evaluation, reaction times should max out across conditions in which information is presented very late in the trial. If there is no timeout function to preserve resources, there should be no change in performance levels once information is presented. The results demonstrate a decrease in accuracy and a decrease in reaction time as the decision extends in time. Two different theories that might account for these results are discussed.

In addition, 5.0 describes an experiment in which bursts of highly informative or highly misleading information are embedded early in the trial. It was predicted that altering the fidelity of information early in the trial should influence the overall process, but not cause an immediate termination. This should create faster, more accurate responses when the highly informative evidence agrees with the correct response, but slower, less accurate responses when it is contrary. The results present a complicated pattern whereby the ongoing mix of trials throughout each block appeared to inform strategy selection as the participant moved through the task.

1.6.4 Question #4

Prior experiments on effort and achievement behavior have suggested that completion of the current trial can be dependent on the difficulty, success, and other characteristics of the prior trial (Thomas, 1983). Given variability in perceptual stimuli, it is possible that the decision process adapts continuously according to recent trial performance. If this is true, reaction time and accuracy should fluctuate depending on characteristics of the previous trial, including successful completion and stimulus characteristics. Conversely, studies of temporal and spatial attention have not shown trial-by-trial variability in responses, and instead demonstrate that performance on a given trial is the direct result of variability of the stimulus within the trial (Vul, Hanus, & Kanwisher, 2009).

6.0 explores the possibility of trial-by-trial shifts in strategy by reanalyzing data from the 4 experiments contained in earlier chapters based on previous trial characteristics such as accuracy, speed, and trial status (e.g., informative burst of information vs. misleading burst of information). If participants were adjusting their strategy as suggested by work on multiple attribute decisions, there should be shifts in performance associated with the status of the previous trial. The results suggest that this may be partially true when participants are presented with feedback, as shown through changes in performance based on previous trial accuracy during the burst experiment (5.0). However, there were no trial-by-trial shifts in performance based on previous trial characteristics, for example whether or not the previous trial contained an informative or anti-informative burst of evidence. These paradoxical results highlight the need to further assess perceptual decisions behaviorally before attempting to understand the neural underpinnings. If the parameters of the decision are changing based on characteristics of the previous trial, hypotheses regarding brain activity should also change.

2.0 POINT OF PERCEPTUAL BREAK AND EFFECTS OF SIMILARITY

2.1 INTRODUCTION

The visual system is highly effective at identifying stimuli and does so through a series of specialized downstream regions for particular types of visual input (Ungerleider & Mishkin, 1982). For example, areas MT in the macaque and a homologue in humans (V5; Tootell & Taylor, 1995) have been shown to process motion stimuli (e.g., direction of a moving stimulus) (Newsome & Pare, 1988; Britten & Newsome, 1998). The physiology of this area will partially determine the downstream effects of the perceptual decision process regarding the random dot motion paradigm used throughout this document. Neurons in the middle temporal lobe fire preferentially to stimuli moving in a particular direction; however, there are variations in firing rate dependent on the level of coherence (Britten, Shadlen, Newsome, & Movshon, 1992). As the level of coherence decreases, and in turn the amount of noise present in the stimulus increases, preferential firing does not meet the same level of activity as when the stimulus has high coherence. Additionally, the ability to predict the impending behavioral choice regarding direction of motion from any one neuron begins to fail as coherence decreases. This suggests that decisions regarding motion at low levels of coherence must arrive from the integration of activity over a pool of neurons with similar direction sensitivity (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996). However, although these neuronal pools demonstrate a high rate of firing for the preferred stimulus, they also respond at reduced levels to directions adjacent to the preferred direction. Importantly, these tuning curves broaden as the coherence of the stimulus decreases (Britten & Newsome, 1998). This means that at low levels of coherence, the activity caused by a near-preferred stimulus will be closer to the activity caused by the preferred stimulus than if the two were presented at a higher coherence, potentially resulting in high levels of interference between pools during the evaluation stage of the decision process. The results from Britten & Newsome (1998) demonstrate that for a narrowly tuned neuron, activity falls by 50% of the peak amount for a stimulus 30° off the preferred direction when presented at 100% coherence. Yet, when coherence is decreased to 25.6% (the closest coherence level to the current studies), activity for the same 30° offset stimulus drops by approximately 10%. This variance in the sensitivity of the tuning curve has direct implications for the ability of the participant to accurately discriminate the direction of motion.

Before using the random dot motion paradigm for the remaining experiments, it is important to test the ability of participants to adequately (i.e., significantly above chance) discriminate the directions of motion within the display even at low coherences. The display used throughout the experiments in this dissertation is a stochastic random dot display that replots 1000 dots in a random fashion, with a specified subset moving in one direction together. Experiments hereafter will use 45° increments and greater as the minimum degree difference between options. The goal of this experiment is to identify the direction detection threshold in human participants using our random dot display. The hypothesis is that the threshold will be below 45 degrees of visual angle (the minimum angle used in the remaining experiments) based on the neurophysiological data described above and behavioral piloting results in the lab that have demonstrated above chance performance at 45° disparity. To assess this formally, the current experiment will test discrimination with two response alternatives at 22.5°, 45°, 67.5°, and 90° in disparity from one another.

A secondary goal of this experiment is to examine how changes in the similarity of the response options in this paradigm influences the accuracy and speed of discriminations. Previous studies have demonstrated that increasing the similarity of options can affect performance in both accuracy and reaction time. For example, similar letters (e.g., E and F) require more time to discriminate and result in lower accuracy compared to dissimilar letters (e.g., E and C; Ratcliff & Rouder, 2000). In addition, previous neurophysiological work using the stochastic random dot motion paradigm tested 90° discriminations compared to 180° discriminations. This resulted in slower neural firing buildup and slower reaction times, suggesting a more difficult decision (Churchland, Kiani, & Shadlen, 2008). The neurophysiological data suggests a potential mechanism by which similarity affects performance. As the options become more similar to each other, the potential for the comparison of overlapping neural pools increases. As described above, motion stimuli near the perceptual threshold for discrimination have the potential to activate more neurons due to the wide tuning curves (Britten & Newsome, 1998). In addition, studies of primary visual cortex have demonstrated that in certain perceptual discrimination tasks, perceptual ability relies on the steep slopes on either side of the tuning curve rather than the peak (Butts & Goldman, 2006). If activity occurs in a set of neurons that overlap from pools that code similar directions of motion, discrimination ability may suffer.

There are two possible ways this might affect performance. The first way is if perceptual decisions occur due to a simple passing of threshold mechanism (Mazurek, Roitman, Ditterich, & Shadlen, 2003); in this case, increasing similarity will likely result in decreases in accuracy but not reaction time. The activity from the neuronal pools will still occur at the same fidelity, with one reaching the threshold prior to the other. However, if the two options are so similar that the pools overlap, the wrong pool may reach threshold first due to variation in neural spiking (Butts & Goldman, 2006;). The second way is if discrimination occurs through a brain area (the “comparator”) that compares the relative activity in the pools that code the presented options (Heekeren, Marrett, Bandettini, & Ungerleider, 2004; Lo & Wang, 2006; Kiani, Hanks, & Shadlen, 2006). In this case, both accuracy and reaction time may be affected. During an easy discrimination (i.e., the compared pools do not overlap in preference), the comparator will increase in activity quickly, pass threshold, and designate the pool with greater activity as the winner (Figure 5a, comparing either the solid bars to each other or the striped bars). If the comparison is difficult (e.g., the activity in the two pools of neurons under comparison are very close in activity; Figure 5b), an executive function of the comparator may be to delay the decision in order to gather more information about the stimulus. Once the comparator passes threshold, the decision is executed in favor of the pool with more activity.

The results of this experiment should shed light on which of these two possibilities may be in action during perceptual decisions based on the resulting changes in performance as similarity increases. However, although decreases in accuracy may be expected by the increasing similarity of the options, performance in the 45° should remain significantly above chance levels in order to validate the use of that angle of disparity as the minimum between alternatives for the remaining experimental designs.

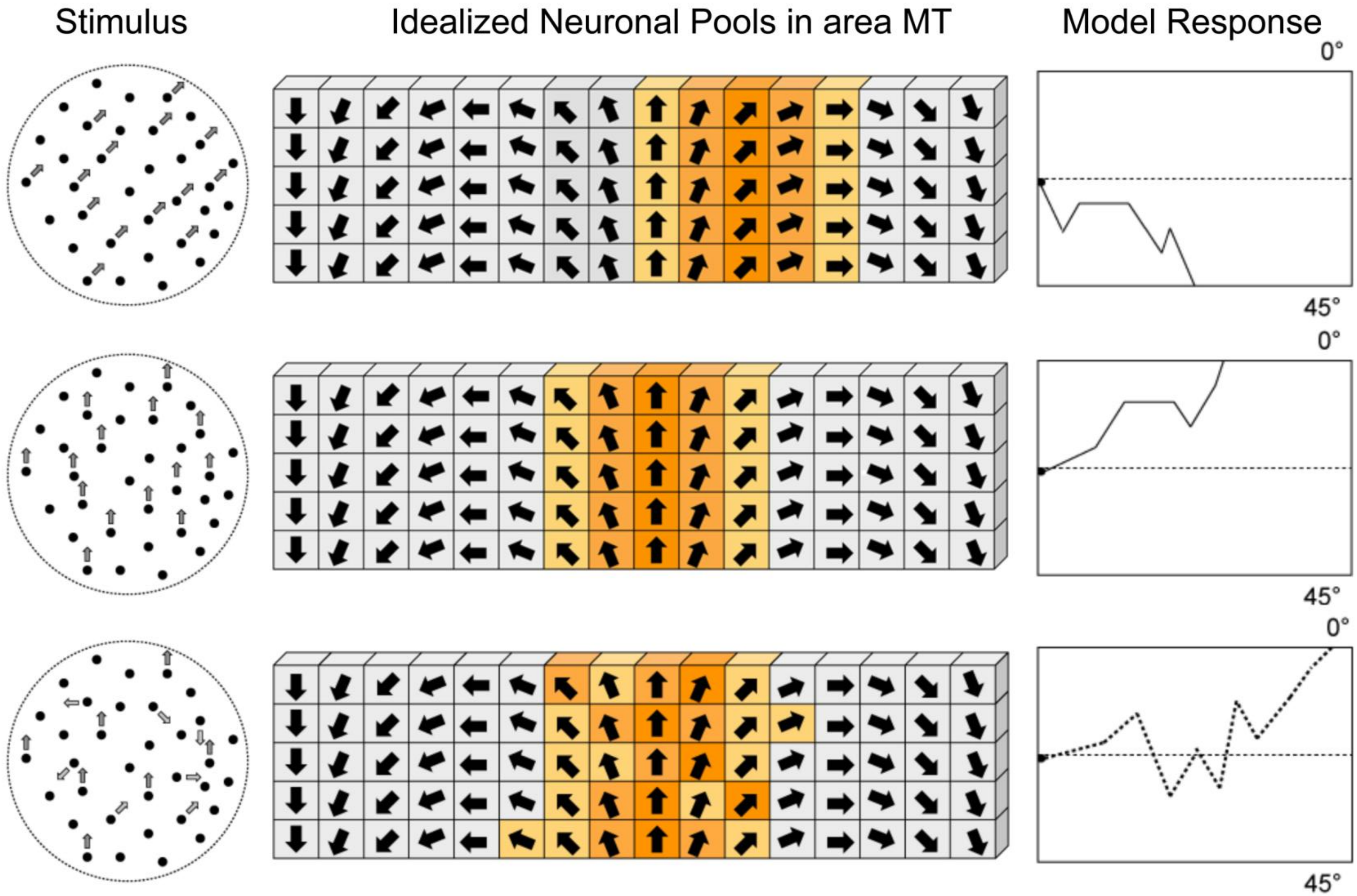


Figure 4. Examples of overlapping neural pools and the resulting decision process. *Left column*, the motion stimulus; *Center column*, graded responses due to tuning curve characteristics in an idealized neural pool; *Right column*, the resulting changes in neural firing as illustrated by a diffusion model. *Top row*, changes when a high coherence 45° stimulus is presented; *Middle row*, changes when a high coherence 0° stimulus is presented, *Bottom row*, changes when a low coherence 0° stimulus is presented. Note that the increased physiological noise and overlap in pools that may cause incorrect 45° decisions on some trials.

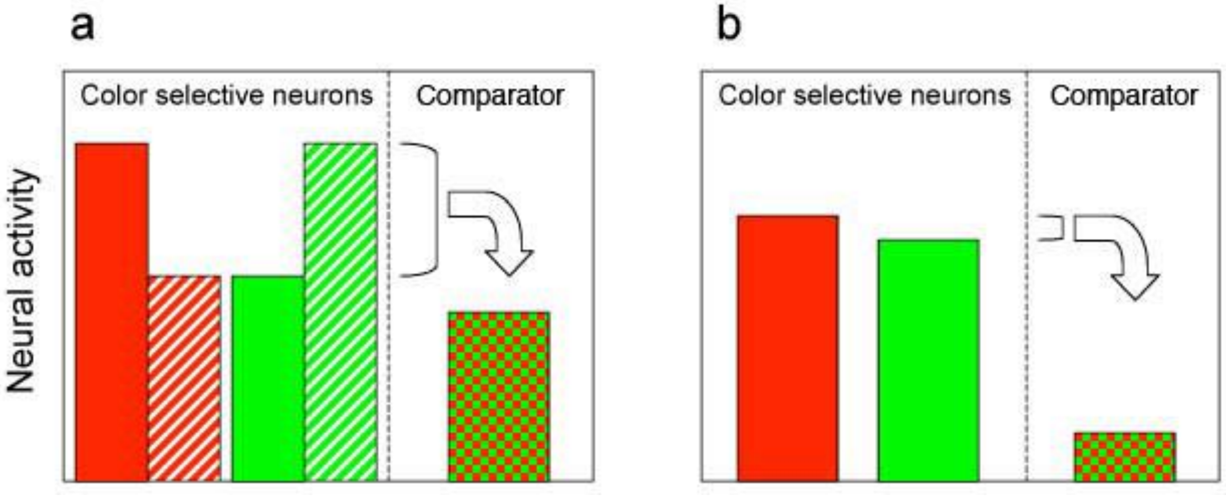


Figure 5. Example of resulting activity in a comparator brain area depending on similarity of alternatives. (a) Two examples of an easy decision: either the solid bars or the slashed bars are compared with each other, causing a high level of activity in the comparator region due to the large difference in activity between the red and green selective areas; (b) Example of a difficult decision that results in low activity in the comparator and would require more information before triggering a response. Note that both comparisons in (a) would result in the same amount of activity in the comparator, highlighting a problem with this theoretical approach.

2.2 METHODS

2.2.1 Participants

Forty-one healthy young adults (ages 18-25 years) were recruited from the Introduction to Psychology participant pool at the University of Pittsburgh. Participants were tested in groups of 7-22 in a computer lab, with each seated at their own computer. The experiment lasted one hour and participants received one credit towards their course requirement. Prior to beginning the experiment, all participants completed a consent form approved by the Internal Review Board at the University of Pittsburgh. Data from 4 participants was removed for chance performance that indicated a misunderstanding of the task.

2.2.2 Materials

The experiment was conducted using EPrime software (Psychology Software Tools, Pittsburgh, PA) on Dell computers. The display presented white dots on a black screen contained within a circle of space approximately 4” in diameter; a fixation cross was in the middle of the circle. The dots were replotted in a random direction at a rate 60Hz. The experimenter could control both the direction and coherence (i.e., how many dots were moving together in the same direction) of the display. To indicate the options on a particular trial, “1” and “0” were presented at locations adjacent to the circular motion frame in a location found by drawing a straight line from the center of the circle to the frame in the desired direction. Participants used a standard keyboard for response.

2.2.3 Design

Discriminations were made between 22.5°, 45°, 67.5°, and 90° pairs. The correct motion could be in any of 16 directions, starting at 0° (upward) through 337.5° in 22.5° increments. The other response option (the "comparator") could be in either the clockwise or counterclockwise direction. This resulted in a 16 (direction of motion) x 4 (degree of difference) design. The number indicating the correct response was randomly selected to reduce any bias effects. Due to the large number of trials this creates, and in order to get a sufficient number of trials in the 22.5° and 45° conditions to accurately assess the discrimination ability for the remaining experiments, trials were selected randomly from probability distributions with greater probabilities for selection of the higher similarity trials but equal probabilities for each direction of motion as the correct response.

2.2.4 Procedure

After completing the consent form, participants were given brief verbal instructions on the basic aspects of experiment including the task and response options. Participants were asked to place the index finger of their right hand on the "0" key of the number line at the top of the keyboard and the index finger of their left hand on the "1" key at the other end of the number line. These fingers were to remain in place throughout the experiment, and the experimenter reminded participants who removed their fingers from these locations throughout the session.

Participants then read a series of screens reiterating the verbal instructions about the task and were prompted to ask any questions they may have. Then they were given 10 practice trials designed to acclimate them to the task. The first trial was conducted at a high coherence level (80%), and succeeding trials decreased in coherence until the final three trials were given at the 25% coherence level that remained throughout the rest of the experiment (selection of 25% coherence was based on prior behavioral testing that demonstrated this level to elicit a variety of performance levels across participants that is neither at floor or ceiling). At the end of the practice session participants were again prompted to ask the experimenter any questions regarding the task they may have. If they did not have any questions, they proceeded with the experiment.

The experiment was divided into 8 blocks of 40 to 56 trials (an earlier version of the experiment (N = 15) had a shorter trial duration that allowed for more trials per block; this led to a high number of no-response trials so the trial length was increased for the remaining participants to increase responses). At the beginning of each block, participants read an instruction screen that reminded them of the response options; they pressed the space bar to advance to the block of trials. Blocks lasted approximately 6 minutes in duration, after which a screen instructed participants to briefly rest their eyes before moving on to the next block.

Each trial consisted of two parts. The first part was a 2000ms fixation period, during which the fixation cross was red in color and the dots were set to all move randomly. Participants had been previously instructed to not respond during these periods, as there was no direction possible when the fixation cross was red. The second part of the trial was a 6000ms motion period (4000ms in the earlier version) during which the fixation cross turned white and the dots were set to 25% coherence towards the selected direction. The “0” and “1” white markers for response options were displayed in the appropriate locations around the circular motion frame dependent on the condition at the start of the 6000ms motion period (Figure 6). Participants were instructed to respond during this period with the response option they believed corresponded with the overall motion of the dots. Trials lasted the entire 6000ms (or 4000ms) period despite earlier responses. No feedback on performance was given.

2.2.5 Data Scoring and Sorting

Each trial was marked based on the response. There were four potential responses: correct (response matched the marker for the correct choice), incorrect (response matched the marker for the incorrect choice), fixation (a response was made in the fixation period prior to the trial), or no response. Trials with fixation responses or no response were removed from the analysis. Trials with correct and incorrect responses also had an associated reaction time that was calculated by the amount of time between the start of the coherent stimulus and the button press in response.

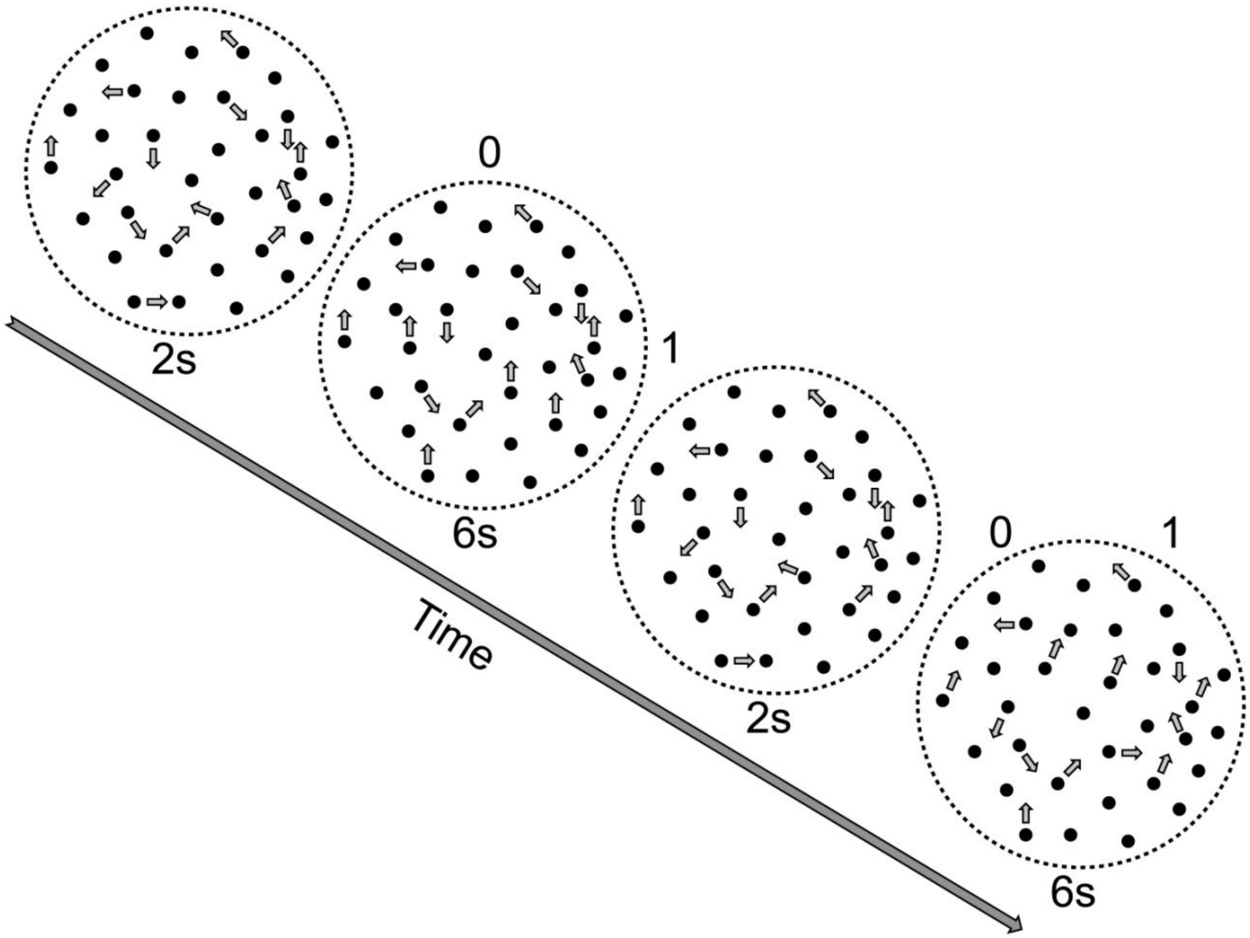


Figure 6. Timeline illustrating two trials of the perceptual break experiment.

2.3 RESULTS

2.3.1 Accuracy

Overall accuracy was high (M = 88.4%, SD = 6.6), indicating that participants were able to learn the task. Performance on the increasing similarity of options is shown in Figure 7. A repeated-measures ANOVA was used to analyze the accuracy rates for the degree of difference (4 conditions; 90°, 67.5°, 45°, and 22.5°), the correct direction of motion (16 conditions; 0° to 337.5° in 22.5° increments), and the interaction between the two variables. The main effect of degree of difference between response options was significant (F(3,99) = 61.881, p < .001). Accuracy for the 22.5° condition (M = 79.8%, SE = 2.0) was significantly lower than all other comparison conditions (P < .001 for all pairwise comparisons). In addition, the 45° condition (M = 91.2%, SE = 1.2) was significantly lower than the 90° condition (M = 93.9%, SE = 1.0; p = .01). However, the 67.5° condition (M = 93.6%, SE = 1.0) was not significantly different from either the 45° or 90° conditions. This demonstrates that at fine levels of similarity (22.5° difference in response options), accuracy begins to decline. However, larger degrees of difference have less or no effect on performance.

The main effect of direction for the correct option was also significant (F(15,495) = 4.521, p < .001; Figure 8). Post-hoc pairwise comparisons indicate that this effect is driven by the 135° (M = 84.5%, SE = 2.1) and 315° (M = 83.8%, SE = 2.2) directions, which are significantly different from 0° (M = 94.3%, SE = 1.3) and 180° (M = 93.4%, SE = 1.3)(p < .01

for all four pairwise comparisons). In addition, the 315° condition is also significantly different from trials in which 22.5° is correct, (M = 92.3%, SE = 1.3; p =.01) and 270° is correct (M = 91.9%, SE = 1.2; p = .05). All other pairwise comparisons were not significant, indicating that aside from the 135° and 315° conditions, there was not a bias in favor or against particular directions of motion.

The interaction between degree of difference and direction of correct motion was also significant (F(45,1485) = 2.02, p < .001). This again is driven by the 135° and 315° conditions in which the 90° comparisons for these directions decrease in accuracy compared to the 67.5° comparisons (86.5% vs. 89.5% and 83.6% vs. 90.2% respectively). Although this suggests there was an error in experimental coding, no error could be found.

2.3.2 Reaction Time

Responses were on average made about halfway through the trial (M = 2162.67ms, SD = 407.25). A repeated-measures ANOVA was used to analyze the reaction times for the degree of difference (4 conditions; 90°, 67.5°, 45°, and 22.5°), the correct direction of motion (16 conditions; 0° to 337.5° in 22.5° increments), and the interaction between the two variables. The main effect of degree of difference between response options was significant (F(3,99) = 32.17, p < .001). Post-hoc pairwise comparisons reveal the same effects as for accuracy. The 22.5° difference between response options (M = 2250.14ms, SE = 65.44) took significantly longer than all other differences (all comparisons p < .001). In addition, the 45° condition (M = 2051.57ms, SE = 64.52) took significantly longer than the 90° condition (M = 1962.25ms, SE = 57.87; p = .02). The 67.5° condition (M = 2025.30ms, SE = 59.09) was not different from either the 45° or

90° conditions (Figure 9). This increase in reaction time as options became more similar suggests that the decision process requires more time, and therefore potentially more information as it accumulates over time, than when the options are further apart.

The main effect of correct direction of motion was also significant ($F(15,495) = 4.78$, $p < .001$). This appears to be driven by shorter reaction times for stimuli on the vertical axis (up, $M = 1930.19$ms, $SE = 69.02$; down, $M = 1872.48$, $SE = 55.91$) compared to directions off the vertical axis. These effects are summarized in Figure 10. In addition, there was a significant difference between the 270° ($M = 2010.51$ms, $SE = 73.93$) and 292.5° ($M = 2205.56$ms, $SE = 75.79$; $p = .03$). The shorter reaction times for stimuli on the vertical axis suggests a processing preference, as has been previously demonstrated in object recognition (the "oblique effect"; see Appelle, 1972 for review).

The interaction between degree of difference and the motion of the correct direction was not significant ($F(45,1485) = 1.159$, $p = .22$). All directions of motion had increasing reaction times as similarity increased.

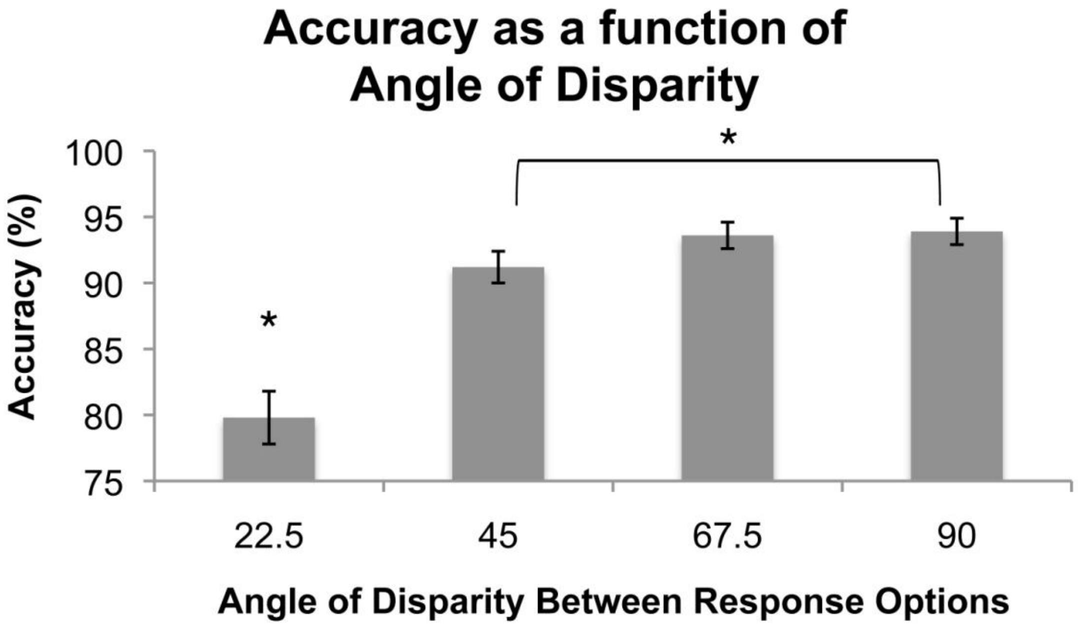


Figure 7. Accuracy as a function of angle of disparity.

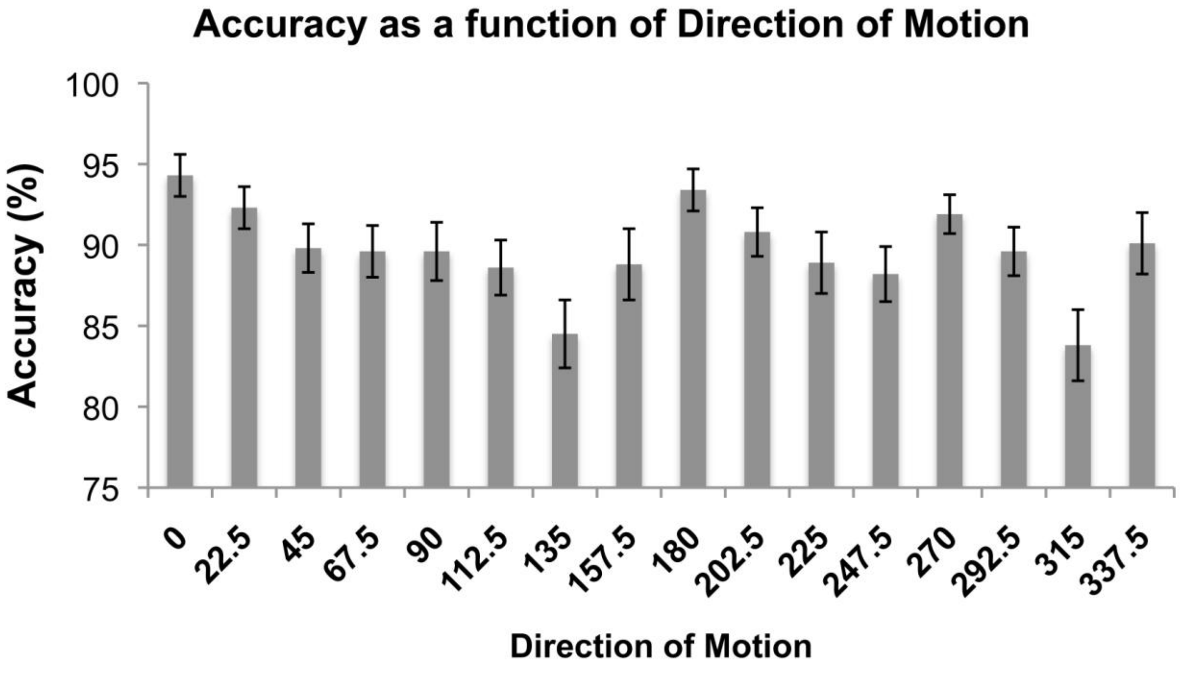


Figure 8. Accuracy as a function of direction of motion (0° is up, 90° is right).

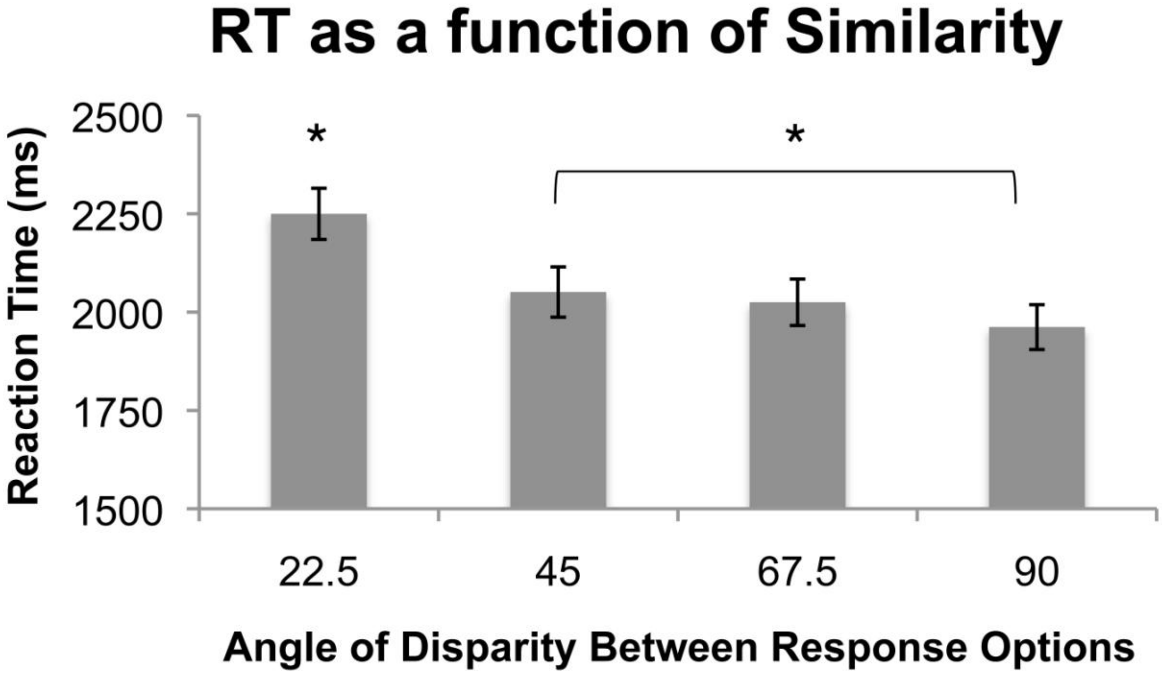


Figure 9. Reaction time as a function of similarity.

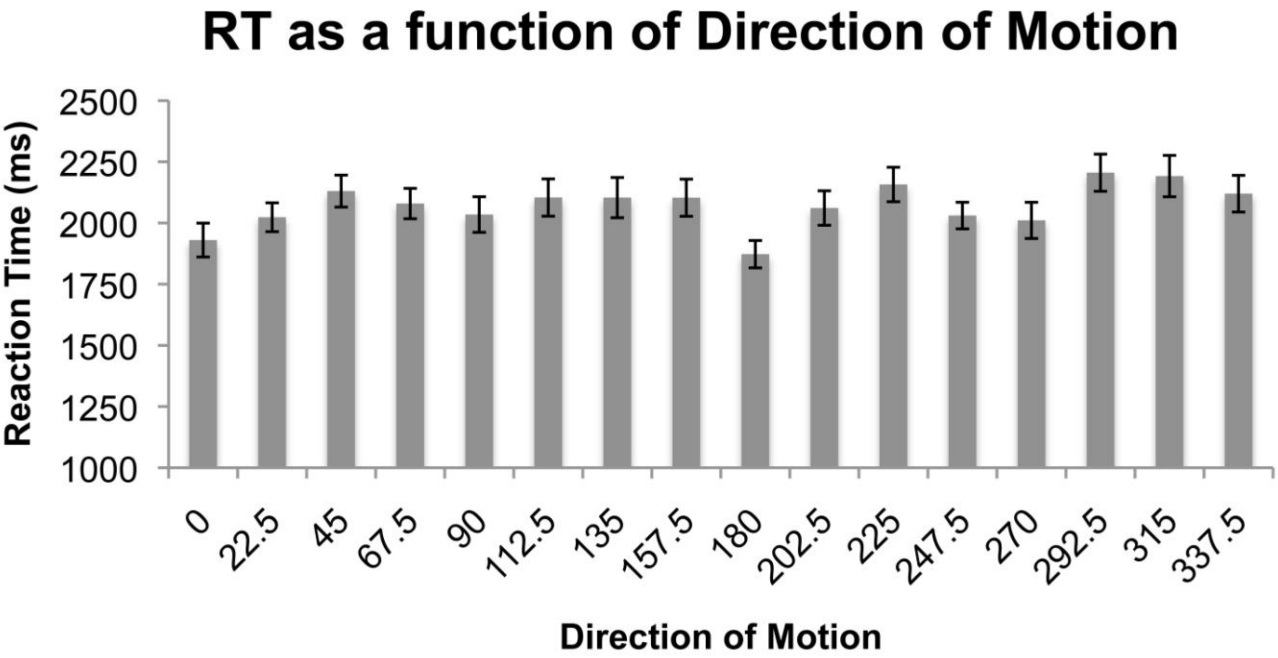


Figure 10. Reaction time as a function of direction of motion (0° is up, 90° is right).

2.4 DISCUSSION

The results demonstrate only minor changes in performance at angles of disparity of 45° or greater, with the effect of similarity only significantly decreasing accuracy and increasing reaction time at very high levels of similarity (22.5° angle of disparity between the two possible response options). However, these changes to both measures of performance agree with the effects of response option similarity found in research on both perceptual decisions (Ratcliff & Rouder, 2000) and higher levels of decision making (e.g., medical diagnosis based on a list of symptoms; Biggs, Bedard, Gaber, & Linsmeier, 1985). In addition, the changes in response time directly agree with previous work with the random dot motion task in behavioral (Ball & Sekuler, 1980) and neurophysiological (Churchland, Kiani, & Shadlen, 2008) studies. Although increasing the similarity of choices did not appear to affect accuracy and reaction time in a linear fashion, it is likely due to the large grain size of the comparisons. As the degree of difference decreased down to 45° and 22.5°, performance began to decrease in accuracy and increase in reaction time. Had the degree of difference between possible options decreased even further (e.g., 11.25° between response options), the pools of neurons involved in the decision (Mazurek, Roitman, Ditterich, & Shadlen, 2003) would involve higher overlap in tuning. In that case, it is likely that the effects would have grown larger.

It is possible that the motion stimulus does not have enough variation in order to induce these similarity differences. Since coherence was kept constant in order to equate neural activity levels across comparisons, the only attribute altered was the degree of difference between the options. Previous studies of similarity in decision making have often used multiple attribute contexts (like medical diagnoses based on a variety of symptoms; Biggs, Bedard, Gaber, &

Linsmeier, 1985), but the motion stimulus does not have as many degrees of freedom. This lack of variety may constrain the extent to which similarity can have an affect in the current context.

While the effect of similarity was an important secondary goal of this experiment, the primary goal was to test whether or not the remaining experiments would be compromised by the minimum 45° angle of disparity between response options when 8 options are presented. The high performance rate in the 45° condition confirmed that the difference is distinguishable and therefore able to be used in the remaining experiments. However, there was a significant change in performance between the 90° and 45° conditions, so caution will be taken when discussing changes in performance when there are also changes in the angle of disparity between response options (e.g., between the 4-choice and 8-choice conditions in 3.0). In addition, there appears to be a bias towards motion along the vertical axis, particularly in the speed of decisions. While all the experiments use counterbalancing to avoid these effects overall, any indication that direction of motion is affecting the results above the intended manipulations in the remaining studies will be examined in a separate analysis.

3.0 INCREASING CHOICES

3.1 INTRODUCTION

The previous chapter investigated how increasing the similarity of options might affect reaction time and accuracy. The results indicated that similarity might not affect performance until the identified alternatives are highly similar (e.g., 22.5° in disparity or less for the present paradigm). While the effect was not as strong as hoped, it still demonstrated that aspects of the response options themselves might affect the decision process. In addition to similarity between possible response options, it is also likely that the number of available options could make a difference in performance. In simple stimulus-response paradigms, increasing the number of possible items to be displayed increased reaction time linearly (Hyman, 1953). This suggests that processing slowed as there are more associations to consider upon stimulus presentation. However, stimulus-response paradigms do not require interpretation of the stimulus, necessarily. If the stimulus is presented unambiguously (i.e., without added noise or disruption), it is a matter of conditioned execution of the correct response. The current chapter examines how increasing the number of possible response options affects the speed of interpretation and execution of a decision for an ambiguous perceptual stimulus.

The majority of previous studies on perceptual decisions require a choice between two dichotomous options (e.g., face/house, Heekeren, Marrett, Bandettini, & Ungerleider, 2004;

face/car, Philiastides & Sajda, 2007; Philiastides, Ratcliff, & Sajda, 2006). However, we do not always have the benefit of determining binary decisions. The primary goal of this experiment is to test how the increase from 2 response alternatives to more (in the present experiment, 4 or 8 alternatives) affects both accuracy and reaction time. While this does not assess completely blind decisions in which the number of alternatives is unknown or infinite (e.g., object identification), it should begin to establish how performance changes as possibilities increase under controlled conditions. As mentioned above, increasing the number of items in the stimulus-response paradigm slowed responses but did not change accuracy (Hyman, 1953). However, the lack of a change in accuracy appears to be largely due to the task instructions, which required a low error rate and specific timing of responses after the stimulus was presented. In fact, the design of Hyman's experiment was an attempt to adjust the task based on changes in accuracy in an earlier paper (Hick, 1952) that he believed failed to capture the true effect because it allowed for strategies that increased speed and errors at the same time. Therefore, without the strict response instructions (i.e., accuracy is unconstrained in the current experiment), it will be interesting to see if increasing the number of response alternatives in an ambiguous stimulus situation affects both reaction time and accuracy, as has been found in 4-choice decision tasks (Churchland, Kiani, & Shadlen, 2008; Leite & Ratcliff, 2010).

Although increasing the number of response alternatives in this type of motion discrimination experiment inevitably increases the similarity of the alternatives, the previous experiment demonstrated that there is only a slight decrease in performance from a 90° to a 45° angle of disparity in two-choice decisions. Therefore it seems safe to assume that as long as the minimum degree of difference between alternatives is 45°, there should not be adverse affects of the increase in similarity in addition to the increasing number of alternatives. Just as the previous experiment began to find decreases in accuracy and increases in reaction time as similarity increases, it is expected that adding response alternatives will create similar results. However, because the 45° difference condition did not large changes in performance, changes in the current experiment above those levels can be attributed to the addition of response alternatives instead of similarity.

It is expected that response time will increase as there are more alternatives to vet (Leite & Ratcliff, 2010; Churchland, Kiani, & Shadlen, 2008). There are two potential increasing patterns. First, response time may increase in a linear fashion with the number of alternatives, demonstrating that each added alternative adds a constant load to the process, similar to the results of the stimulus-response paradigm (Hyman, 1953). This appears to be true for when moving across 2-, 3-, and 4-choice decisions (Leite & Ratcliff, 2010), but it is unclear how this will extend to 8-choice trials. Therefore, the second possibility is that response time may reach a plateau as alternatives increase as the subjects adopt a decision strategy (e.g., use of early information to eliminate clear non-options on a particular trial to reduce number of alternatives; Payne, Bettman, & Johnson, 1988). Previous decision research has shown the use of preference trees as a fast way to eliminate improbable options when considering decisions with multiple relevant aspects (Tversky & Sattath, 1979); it is possible that a similar heuristic may be used during simple perceptual decisions. The pattern of reaction times based on the number of possible options should dissociate these two theories. A consistent linear increase in reaction time as the number of alternatives increases will support the theory that each alternative adds a consistent amount of processing to the cognitive load (Figure 11, solid line). On the other hand, if the increase in reaction time slows or reaches a plateau, a strategy account will be supported (Figure 11, dashed line).

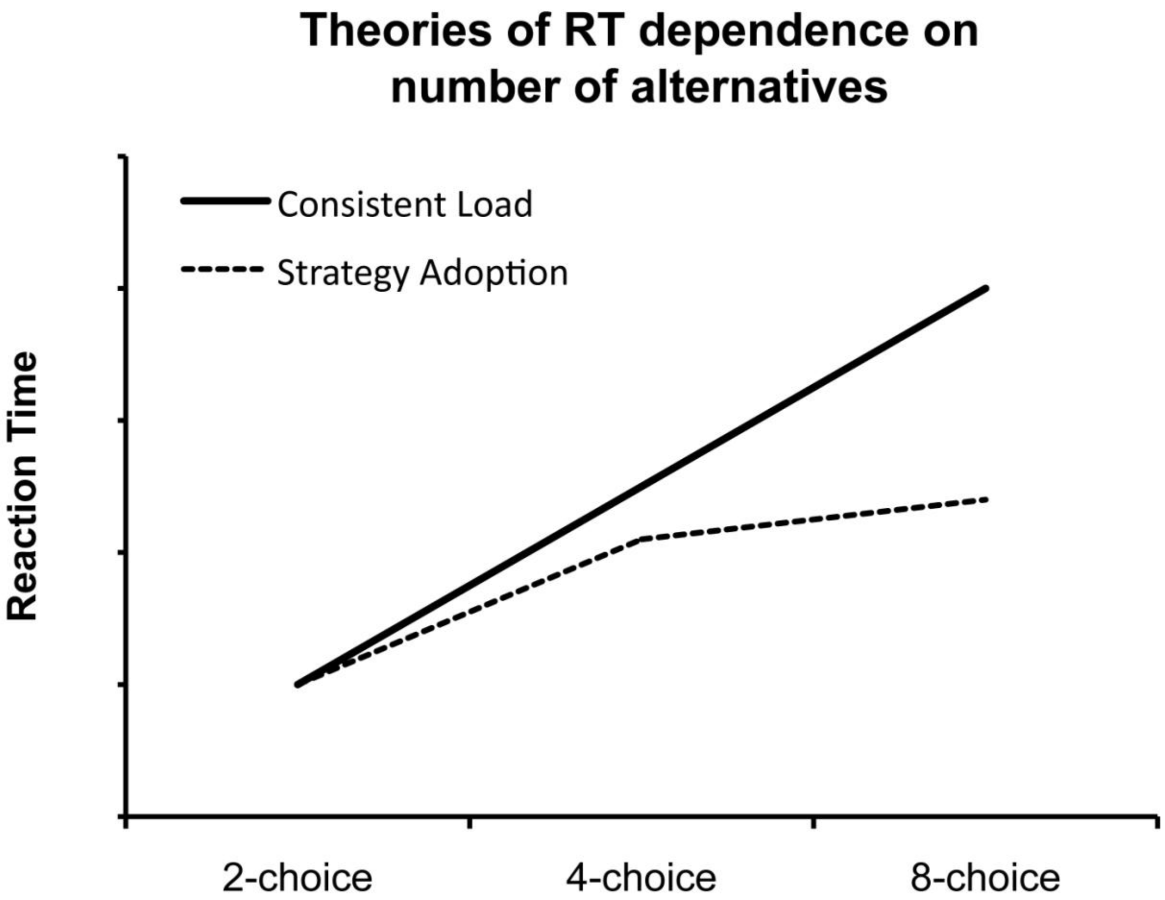


Figure 11. Predicted differences in reaction time depending on either a consistent load with each new alternative or strategy adoption.

A secondary goal of this experiment is to gauge whether or not strategy towards the task changes as the stimulus space is crowded with more possible options. Previous research in decision making suggests that people use heuristics or strategies to solve difficult decisions quickly (Tversky & Sattath, 1979; Einhorn & Hogarth, 1981; Payne, Bettman, & Johnson, 1988). As the number of possible alternative increases, it is likely that participants are adopting one or more strategies to simplify the problem space. These strategies may include ruling out obvious non-options, dividing the problem space a priori to bin evidence quickly, or starting with certain biases towards the probability of one result over the others. This will be assessed by subjective feedback on strategy approach, with questions directed at these possibilities.

A final goal of this experiment is to test how the requirements for response output might affect speed and accuracy. Previous experiments on perceptual decisions have used a single button to indicate the discrimination, which may leave doubt as to the veracity of the decision (Ploran et al., 2007; Ploran et al., in preparation). While the previous experiments on object recognition did garner enough error trials to conduct a separate analysis (Wheeler et al., 2008), there remains concern that without a definitive option-related response output the decision is made under slightly different conditions (e.g., a lower threshold for making the initial response because there is less concern regarding performance). In theory, while there might be marginal differences in reaction time based on different motor output (one button responses might be faster than selection of the appropriate button out of up to 8 options), accuracy should be similar. However, this assumes that participants are not motivated to boost their reported accuracy in the one-button version of the task (see Methods). The following experiment will examine how the reported accuracy in the one-button version of the task compares to actual accuracy computed through definitive response options. It will also assess the potential timing differences between using one motor action to respond versus several motor actions, each associated with a unique response option.

3.2 METHODS

3.2.1 Participants

Seventy-one healthy young adults (ages 18-25 years) were recruited from the Introduction to Psychology participant pool at the University of Pittsburgh. Participants were tested in groups of 7-22 in a computer lab, with each seated at their own computer. The experiment lasted one hour and participants received one credit towards their course requirement. Prior to beginning the experiment, all participants completed a consent form approved by the Internal Review Board at the University of Pittsburgh.

One set of participants (N = 20) was tested in the one-button version; one participant was removed from the analysis for performance outside 2 standard deviations below the mean. Another set of participants (N = 27) was tested in the definitive response version of the experiment. Data for two participants was removed for performance outside 2 standard deviations below the mean that indicated a misunderstanding of the task or response options. A final set of participants (N = 24) was also tested on the one-button version of the experiment; in addition, these participants were given a strategy assessment at the end of the session. Unfortunately, due to experimental error, the accuracy and reaction time data for these participants is unusable, but the strategy assessments remain intact.

3.2.2 Materials

The experiment was conducted using EPrime software (Psychology Software Tools, Pittsburgh, PA) on Dell computers. Responses were made via the keyboard attached to the computer. The display presented white dots on a black screen contained within a circle of space approximately 4” in diameter; a fixation cross was in the middle of the circle. The dots were replotted in a random direction at a rate of 60 Hz. The experimenter could control both the direction and coherence (i.e., how many dots were moving together in the same direction) of the display.

3.2.3 Design

Trials were broken into blocks with instructions indicating whether the block would include 2, 4, or 8 choices. In the 2- and 4-choice conditions, the instructions also indicated the possible directions of the choices (e.g., up/down or left/right). Only the vertical and horizontal axes were used in the 2-choice condition; the 4-choice condition used the cardinal directions in one case (i.e., up, down, left, and right) and the diagonal directions in another case. In order to equate and counterbalance the number of occurrences per each response choice (e.g., left) in the 4- and 8-choice blocks, the number of trials and number of blocks were different for each number of choices. There were two 2-choice blocks with 20 trials each, three 4-choice blocks with 28 trials each, and four 8-choice blocks with 32 trials each. This resulted in blocks approximately 3 to 6 minutes in length.

3.2.4 Procedure

After reading and signing consent forms according to the rules of the Internal Review Board at the University of Pittsburgh, participants were verbally instructed about the basic aspects of the experiment including a description of the experiment, task, and response options. Participants then read a series of instruction screens reiterating the verbal instructions and were prompted to ask questions if necessary. They then started a three-minute practice session containing 10 trials in the 8-choice version of the experiment. The coherence level started at 80% for the first trial and shifted down in increments to 25% for the last three trials in order to acclimate the participants to the task. After completing the practice session participants were again prompted to ask any questions they may have regarding the task. They then proceeded into the experiment and responded according to their particular instructions and trial structure, elaborated below. After completion of the experiment, participants were thanked for their time and allowed to leave the computer lab.

3.2.5 One-button Version

In this version of the experiment, participants responded using the Space bar and the task structure was slightly different. The trial started with 2000ms of random motion and a red fixation cross; the period of motion designated by the white fixation cross was broken into two parts. The first part was 6000 ms of motion at 25% coherence during which participants were instructed to press the Space bar when they were pretty confident as to the direction of motion. After the 6000 ms allowed for the initial decision, 2000 ms of 90% coherent motion was presented for the participants to verify their response. If their initial guess was correct, participants pressed the Space bar again. If their initial guess was incorrect, participants did not make a second response. Verbal instructions prior to the experiment emphasized the likelihood of incorrect responses and stressed the need to be honest over the need to report high accuracy.

3.2.6 Definitive Response Version

In this version of the experiment, participants responded using the number pad on the right side of the keyboard. Instructions designated the “8” key as up, the “2” key as down, the “6” key as right, the “4” key as left, and the remaining “1”, “3”, “7”, and “9” keys as the respective diagonal directions. Participants were to rest their right index finger on the “5” key (in the middle of the pad) and press the appropriate button to indicate the direction of motion on each trial.

Each trial began with 2000ms of 0% coherence motion with a red fixation cross. Participants had been instructed that periods with a red fixation cross were at 0% coherence and therefore did not require a response as there was no motion present. The fixation cross then changed to white for 8000 ms and the coherence level was raised to 25%. Depending on the block, the direction of motion was randomly selected from among the possible choices, with equal numbers of trials per direction chosen throughout the block. Participants responded by pressing the appropriate button on the number pad as described above.

3.2.7 Strategy Assessment

A portion of the participants was asked to report their general task strategy at the end of the experiment. The debriefing questionnaire began with general open-ended questions regarding confidence level at the decision point for each block type. It then presented several options for how participants might have divided the visual space in order to complete the task effectively in the 2- and 4-choice blocks (the 8-choice condition was left out due to time constraints; Section 1.01(a)(i)Appendix B). Participants were asked to give a percentage for each option as to how close it was to their actual strategy. The questionnaire presented several complementary pairs of strategies for the 2- and 4-choice versions to investigate whether participants maintained a similar strategy across blocks or changed strategies according to the number of options.

3.3 RESULTS

3.3.1 One-Button Version – Accuracy

Overall performance was extremely high (M = 91.47%, SD = 8.15), indicating there may be some motivation to misreport accuracy. A repeated-measures ANOVA found that the main effect of number of choices was not significant (F(2,38) = 2.23, p = .12). Accuracy was similar in the 2-choice (M = 91.31%, SD = 12.72), 4-choice (M = 88.57%, SD = 12.54), and 8-choice (M = 90.47%, SD = 10.29). As already mentioned, this suggests some motivation (or lack of motivation) to misreport accuracy and complete the task dishonestly.

3.3.2 One-Button Version – Reaction Time

Overall reaction times were well within the 6000ms of motion during the trial prior to the 2000ms verification stage (M = 2154.44ms, SD = 553.46). However, a repeated-measures ANOVA found that the main effect of number of choices was not significant (F(2,36) = .373, p =

.69). The 2-choice (M = 2194.10, SD = 679.37), 4-choice (M = 2123.31, SD = 642.31), and 8-choice (M = 2186.56, SD = 495.25) were all similar in speed. It is unclear why reaction times did not change. Even if participants were misreporting accuracy, there could still be some changes in the initial response phase. It is likely that participants, realizing that accuracy was self-reported, weren't bothering to attend to the task at all, causing faster responses in which the participants' were not as confident as required for the definitive response version.

3.3.3 Definitive Response Version - Accuracy

Overall performance was high (M = 82.06%, SD = 14.92), indicating that participants were able to complete the task effectively. Unlike the one-button version, a repeated-measures ANOVA found that the main effect of number of choices was significant (F(2,48) = 29.60, p < .001); the 2-choice condition was the most accurate (M = 93.11%, SD = 8.4), followed by the 4-choice condition (M = 89.75%, SD = 11.81), and the 8-choice was least accurate (M = 74.62%, SD = 19.60; Figure 12). While the 2- and 4-choice conditions were not significantly different from each other, both were significantly more accurate than the 8-choice condition (both p < .001). This suggests that when participants are required to make a specific response for each option, adding choices to task causes a decrease in the accuracy of the decision process.

3.3.4 Definitive Response Version – Reaction Time

Overall reaction times were well within the 8000ms of motion during the trial (M = 3074.68ms, SD = 568.52). A repeated-measures ANOVA found that the main effect of number of choices was significant (F(2,48) = 27.26, p < .001; Figure 13). The 2-choice (M = 2671.49, SD = 695.56) and 4-choice (M = 2674.29, SD = 663.77) conditions were equal in speed, while the 8-choice condition was significantly slower (M = 3449.65, SD = 617.59; both post-hoc comparisons were p < .001). This suggests that the decision process is affected by the number of choices available, but perhaps not in a linear fashion as expected from the stimulus-response literature.

3.3.5 Definitive Response version – Angular disparity of errors

As more response alternatives are available, evidence is divided among finer grain distinctions. Therefore, we would expect that errors should be made through selection of a response alternative similar to the correct alternative, not a random other possibility. To assess this, errors in each block were sorted by the angle of disparity from the correct response alternative (regardless of clockwise or counterclockwise direction) and then entered into a one-way repeated measures ANOVA per block (Figure 14). The data was analyzed per block to avoid data loss due to high performance in the 2-choice condition, which would eliminate data points from other conditions if a 3x4 ANOVA was performed instead (i.e., participants with 100% correct performance in the 2-choice condition would be removed from all conditions). The numbers of retained participants per analysis are reported with the statistics below.

There are two predictions of interest. First, if participants are performing the task correctly, errors should be isolated to 180° in disparity in the 2-choice block and 90° and 180° in the 4-choice block. Errors that are 45° in disparity from the correct response alternative are likely the result of motor error rather than selection error. Second, in the 8-choice block, in which all buttons are associated with a response, there should be a higher proportion of incorrect responses made with a 45° angle of disparity off the correct alternative than other angles of disparity. Importantly, errors of 45° in the 2- and 4-choice can be used to calculate a more accurate depiction of selection error in the 8-choice condition by subtracting the 2- and 4-choice 45° error baseline rate from the 45° error rate in the 8-choice block.

The one-way repeated-measures ANOVA (N = 16) with 4 levels of angular disparity (45°, 90°, 135°, and 180°) for the 2-choice block resulted in a significant main effect (F(1.23, 18.45) = 116.54, p < .001). Post-hoc pairwise comparisons illustrate that the majority of errors were made at 180° in disparity from the correct response alternative, with no difference among the rates for 45°, 90°, and 135° disparities (all comparisons p > .5). Only 2.5% of errors were made at 45° in disparity from the correct response, suggesting that motor error was low.

The one-way repeated-measures ANOVA (N = 22) for the 4-choce block also resulted in a significant main effect (F(1.65,34.72) = 12.71, p < .001). Post-hoc pairwise comparisons highlighted a higher proportion in the 90° and 180° disparities versus the 45° and 135°disparities (p < .05), with no difference within the pairs. This result combined with the finding of high errors in the 180° disparity in the 2-choice block confirms that participants were responding according to the directions for available response alternatives per block. In addition, there again were a low proportion of errors at 45° in disparity (5.6%).

The one-way repeated-measures ANOVA (N = 25) for the 8-choice block again resulted in a significant main effect (F(1.70,40.87) = 31.99, p < .001). Post-hoc pairwise comparisons highlighted no difference in the proportion of errors at 90°, 135°, and 180° in disparity (all p > .2), but significant differences between each of these angles and errors at 45° in disparity (all p < .001). These results remain significant in the same pattern even when the proportion of errors in the 45° condition in the 8-choice blocks were adjusted by subtracting the average proportion of errors at 45° in disparity in the 2- and 4-choice blocks on a subject-wise basis. In addition, the proportion of 45° errors appears to double in magnitude from the 2- to 4-choice blocks, so a second adjustment was made to the 8-choice block by subtracting double the proportion of 45° errors from the 4-choice block to mimic this increasing trend. Even with this liberal adjustment, the pattern remains the same (though the main effect did decrease to $p = .04$). The continued significance despite adjustments for baseline motor error confirms that errors in the 8-choice block were likely due to the decision process rather than an increase in motor errors.

3.3.6 Strategy Assessment

Participants responded an average confidence level at the initial button press of 53.75% (SD = 24.85). While this is lower than expected (past object recognition experiments with a similar structure have reported confidence of approximately 85%; Ploran et al., 2007), this appears to be due to a handful of participants reporting confidence levels in the 10-30% range. Instructions emphasized maintaining a high level of confidence, so these low reported values suggest either a low level of motivation or a lack of understanding of either the instructions or question.

Excerpts from the open-ended section of the questionnaire are included at the end of the chapter. A large proportion of the participants (11 out of 24) explicitly mentioned focusing on the fixation cross, which suggests the instructions to maintain fixation throughout the experiment were heeded appropriately. In addition, several subjects (6 out of 24) mentioned trying to relax their vision or not to focus too hard in order to see the whole circle of dots. In general, the answers to the open-ended question did not vary substantially, and several participants spontaneously mentioned strategies similar to those assessed in the close-ended questions discussed below.

In the close-ended section of the questionnaire, participants rated a set of pictures depicting possible ways to divide the circle of dots, with marks to indicate areas of attention and inattention (1.0). Of the three options participants rated for the 2-choice decision (Figure 15), the third option was the closest to their chosen strategy (M = 54.77%, SD = 39.95), followed by the second option (M = 35.91%, SD = 34.35), and then the first (M = 25.45%, SD = 35.15). A repeated-measures ANOVA found that the main effect of strategy choice was significant (F(2,42) = 4.02, p = .02); participants attended to a narrow band of the space while ignoring the outside in order to choose between the two options. For the 4-choice strategies, the two options were equal in similarity to the chosen strategy (Figure 15; M = 52.05%, SD = 39.48; M = 50.23%, SD = 35.74, respectively) and a repeated-measures ANOVA found that the main effect was not significant. There seems to be some preference for dividing the space, but there was less agreement on whether attempts were made to ignore some parts of the space in order to simplify the problem.

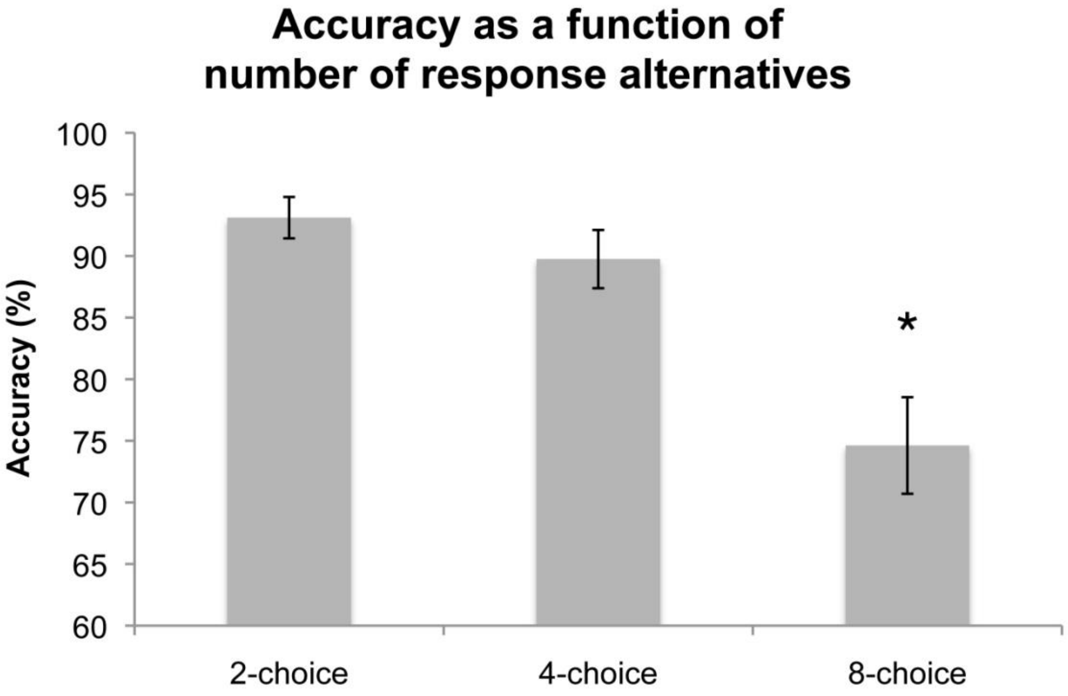


Figure 12. Accuracy as a function of number of choices in the definitive response version.

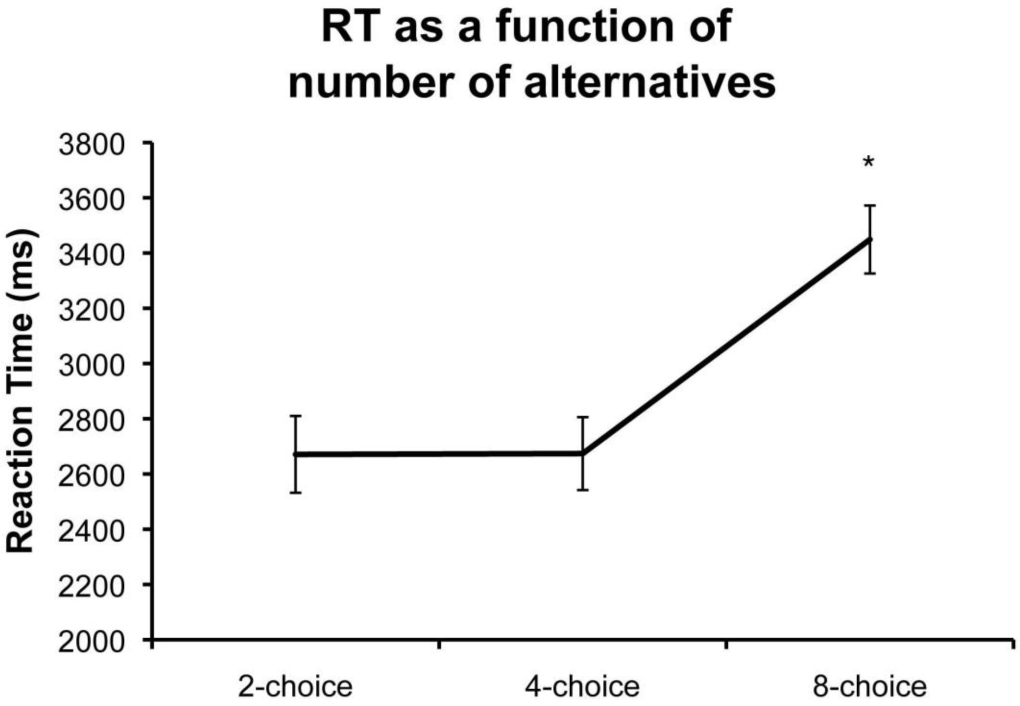


Figure 13. Reaction time as a function of choices in the definitive response version.

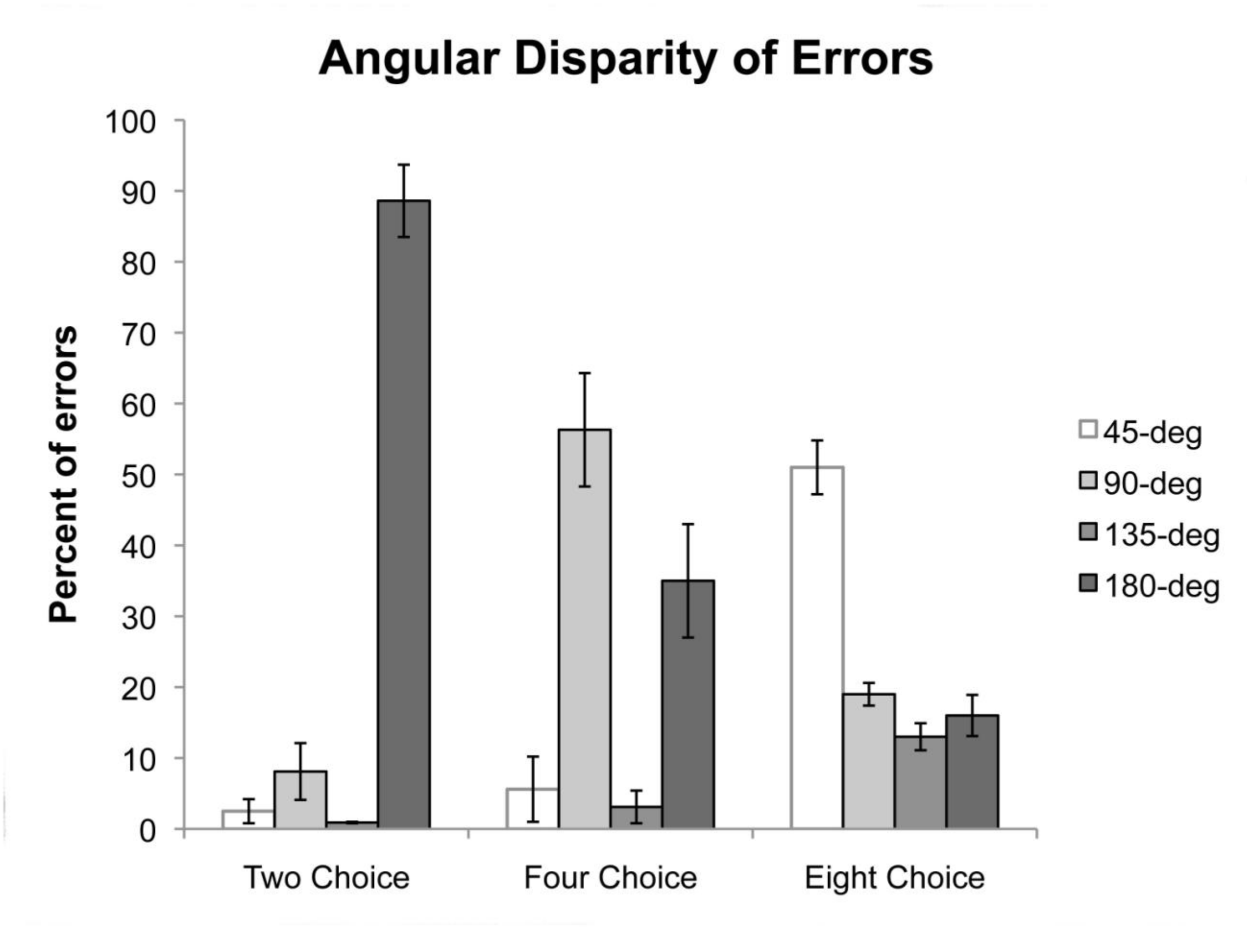


Figure 14. Angular disparity of errors from correct response alternative.

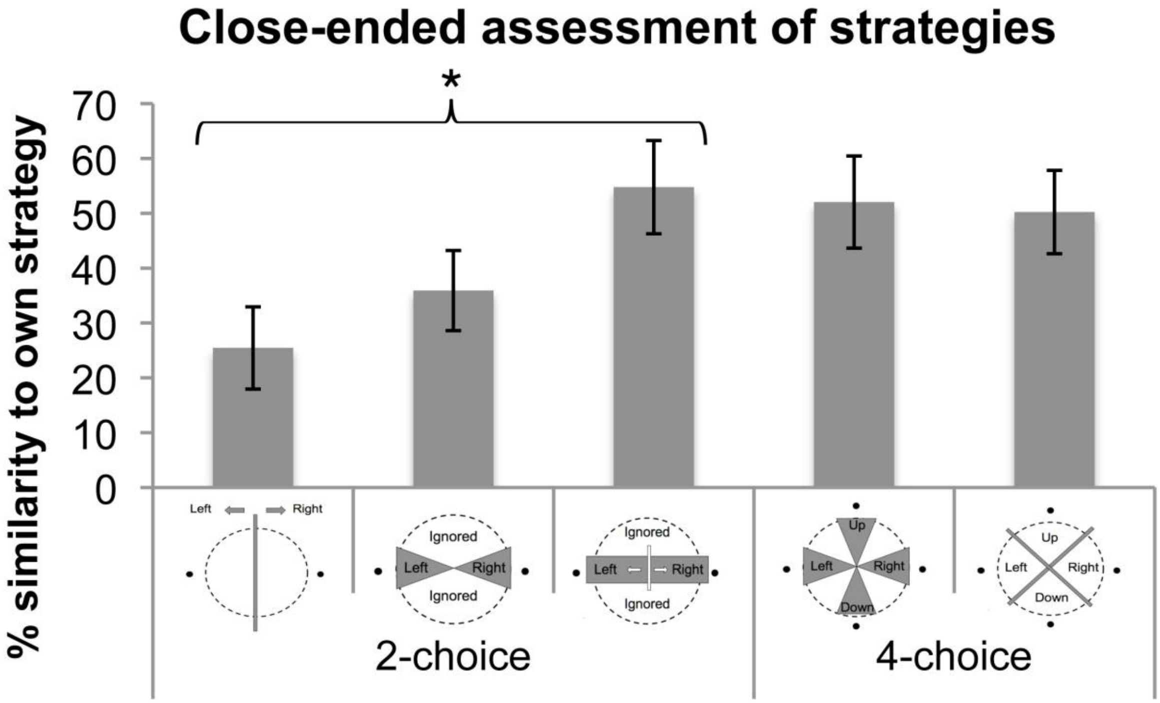


Figure 15. Subjective assessment of possible strategies to divide the screen as the number of alternatives increased.

3.4 DISCUSSION

The comparison of the one-button versus the definitive response versions of the experiment presents a stark contrast. While overall performance still remains high when participants press a particular button associated with each possible response, it does not match the self-report of those in the one-button condition. In addition, accuracy decreases and reaction time increases in the definitive response condition as more options are added to the problem space, yet both measures remain the same in the one-button version. Importantly, the analysis on the disparity of error responses from the correct response suggest that faulty motor execution is not likely able to account for the total decrease in performance. This suggests that participants are misreporting accuracy in the one-button task across all conditions, despite strong instruction to the contrary. The stable reaction time data also suggests that participants were not being as careful with signaling when they were confident of the direction of motion as those in the definitive response condition. This weighs heavily in favor of either always designing tasks with more definitive response options when possible, or at least training participants until they are reporting accuracy correctly based on verbal response at button press or comparison to a definitive response group. This highlights similar concerns to the design of stimulus-response paradigms as examined in Hick (1952) and Hyman (1953). When participants were forced to respond under conditions that ensured their button press was made at the moment of decision (by requiring a fast release of one button and selection of another with low error rates; Hyman, 1953), response timing demonstrated a different pattern than a less controlled version of the experiment (Hick, 1952). Unfortunately, the requirements used by Hyman inherently negated the ability to observe changes in accuracy.

However, despite these differences in performance between the two conditions of the current experiment, the results from the definitive response condition still describe how increasing the number of choice alternatives affects performance. As mentioned above, accuracy decreased and reaction time increased as more choices were added to the problem space. This mimics the pattern found in previous perceptual decision experiments (Churchland, Kiani, & Shadlen, 2008; Leite & Ratcliff, 2010). It is important to note, though, that despite the decreases in accuracy, performance in the 8-choice condition was still well above chance levels (12.5%). In addition, the 2- and 4-choice decisions did not differ from each other in reaction time, suggesting an initial accuracy cost prior to a time cost associated with increasing available options. Given the high performance in the 90° condition in the similarity experiment (2.0) and the 4-choice experiment here (in which the options are 90° separated from each other), it is possible that this initial drop in accuracy was due to the increasing number of choices and not the increasing similarity of those options. The large decrease in accuracy from 4-choice to 8-choice (compared to a smaller decrease from 2- to 4-choice) and the significant increase in reaction time may then reflect both the increasing number of choices and the increasing similarity of options (similar to the changes in accuracy and reaction time between the 90° and 45° conditions in the similarity experiment).

It is interesting to see that performance in the 4-choice condition reported here was approximately the same as the 90° condition in the similarity experiment reported in 2.0 . However, performance in the 8-choice condition was far worse than the complementary 45° (2-choice) decision in the similarity experiment. Conservatively, accuracy is 15% lower and reaction times are 1400ms slower, on average. These changes over and above the possible changes due to increasing similarity may be attributable to the increasing number of response alternatives. Considering the structure of the two popular mathematical models of the decision process may aid in explaining how adding options can cause a change in reaction time. In a race model of decision making, adding more response options would require more counters (one for each option), which may require more intense monitoring of ongoing activity if there is not an automatic trigger when a counter reaches threshold. It also means that the incoming sensory information and extracted evidence will have to be divided among those counters in smaller increments. In a diffusion model of decision making (assuming the possibility of a multidimensional diffusion model rather than the standard binary instantiation), increasing the number of response alternatives increases the space in which activity can drift. This would likely lead to longer paths to a boundary. In either case, decisions are likely to take longer, even if accuracy remains high (i.e., the threshold of activity needed to make a decision remains similar across trials despite increasing options).

It is when considering this change in accuracy that the assessment of strategy may become relevant. The popular choices in the 2- and 4-choice conditions both involved the participant attending to certain parts of the field and ignoring others (Figure 14); this deviates from previously reported decision strategies involving the elimination of unlikely responses in order to concentrate evidence towards more likely alternatives (Tversky & Sattath, 1979; Payne, Bettman, & Johnson, 1988). However, if the common strategy involved ignoring parts of the stimulus field, this may account for the unexpected decreases in accuracy as the number of options increased. As the space is divided into smaller and smaller areas according to the number of response alternatives, participants attend to motion towards each option in a narrower space.

This results in less information and evidence per alternative despite the consistent coherence level across blocks and presence of motion on any trial being displayed throughout the circle. While a few participants reported focusing on the whole circle as their main strategy, the majority either spontaneously reported a division strategy or rated the division strategies highly. Because participants were instructed to do their best to respond on every trial, it is possible that under circumstances of low evidence (due to both the coherence level and the division strategy), participants were forced to make decisions prior to reaching threshold. This would result in the decreasing accuracy as number of options (and therefore division of the evidence) increased, and may require a timeout mechanism (or some other parameter) to execute the best decision possible prior to passing threshold. This potential for a timeout function in decision making will be assessed in the next experiment.

3.5 EXCERPTS FROM STRATEGY ASSESSMENT

- I tried to see if the dots would make smaller patterns of lines or streams in a consistent direction.
- Pick one dot and follow it across
- General feeling of movement, where did I perceive I'd be moving compared to the dots
- Divided the screen into sides or quarters (for the 2- and 4-choice, respectively)
- I focused on the dots nearest to the cross to see if I could figure out the direction.
- I just stared at the cross and looked at the dots in my periphery

4.0 TIMING OF EVIDENCE

4.1 INTRODUCTION

Previous studies of perceptual decisions appear to fall into two procedural categories: brief, isolated presentations of stimuli (e.g., Ratcliff, 2002; Shadlen & Newsome, 2001) or continuous presentations of stimuli until response (Ratcliff & Rouder, 1998; Roitman & Shadlen, 2002). Importantly, past studies have used a predictable relationship between the start of a trial and the presentation of the stimulus with minimal variation within an experiment (c.f., Green, Smith, & von Gierke, 1983). This is also true of the experiments in 2.0 3.0 here. However, there are often instances in which we are aware a stimulus is coming but do not know the specific timing of its presentation. We are therefore waiting for information without the ability to prepare for response execution. This experiment will assess how the timing of information relative to the expectation of its arrival affects speed of response.

In 3.0 we saw a decrease in accuracy in the 8-choice condition, despite an increase in reaction time that should have allowed for enough evidence to accrue. It is possible that due to the division of evidence among many bins, participants may have been unable to gather enough evidence for any particular alternative among the 8 options. In order to execute a response prior to the end of each trial, participants may have used a timeout function to end the search before the threshold was passed. This in turn would cause the decrease in accuracy in addition to the increase in reaction time. Several models of memory describe the possibility of a timeout function that serves to truncate search in favor of selecting the most likely option when it is clear no new information is forthcoming (Gillund & Shiffrin, 1984; Anderson, 2007). It is possible that perceptual decisions behave similarly. When confronted with ambiguous evidence towards an eventual decision, a timeout function would ensure that a decision is made even though the incoming sensory information does not provide enough evidence to pass a pre-defined threshold for response. This could be true both for cases in which the available evidence is ambiguous (like the experiment in 3.0) or, as in the present experiment, the availability of evidence itself is uncertain due to shifts in the timing.

Before testing how the timing of evidence affects perceptual decisions, it is useful to consider how the process might change. In diffusion models, decisions arise from the accrual of evidence in a drift parameter that eventually meets a threshold. Once activity meets the threshold, the decision is executed in favor of the response alternative that threshold represents (Figure 16a). Using this framework, a timeout function could operate under two different guises. First, the neural mechanism that processes the decision could designate a relatively constant amount of time allowed for a decision; if this time is reached prior to activity passing a decision threshold, the response matches the closest threshold the activity would pass if given more time/evidence (Figure 16b). This is similar to the truncation of a memory search when no memory trace appears to be available (Gillund & Shiffrin, 1983; Anderson, 2007). On the other hand, the threshold needed to execute a decision could decrease over time, similar to how the start point and drift rate can vary in traditional diffusion models (Ratcliff & McKoon, 2008). Although many decision models include thresholds, there is often more discussion about the shift in start point due to the need for simplicity and the inclusion of as few integrals into the mathematical instantiation of the model as possible (Ratcliff & McKoon, 2008). However, constraining the model for the purposes of simplicity may also be constraining interpretation of the results. For instance, changes in the start point are typically thought to occur as an indication of bias (Ratcliff, 1985; Voss, Rothermund, & Voss, 2004), much like low levels of neural firing for a particular option that may be forthcoming (Schultz, 2000). While this is likely the case when only two alternatives are present and the participant can keep track of the general ongoing ratio, it is unclear whether this continues to be true when more response alternatives (in the present case, eight) are available. Instead of changing the start point, it is possible that certain tasks may require a change in threshold. Previous accounts of diffusion models have allowed for changes in threshold, but only under particular instructions such as the speed/accuracy trade-off paradigm (Fitts, 1966; Ratcliff & McKoon, 2008). A potential mechanism for the timeout function, however, could be that as the decision process extends in time, the threshold the accumulating evidence needs to pass decreases, thereby increasing the chances of passing the boundaries and making a response (Figure 16c). This would be similar to the changes in threshold needed to increase speed under speed-emphasized instructions versus the threshold for accuracy in the complementary instructions (Fitts, 1966; Ratcliff & McKoon, 2008), but would occur without explicit instruction.

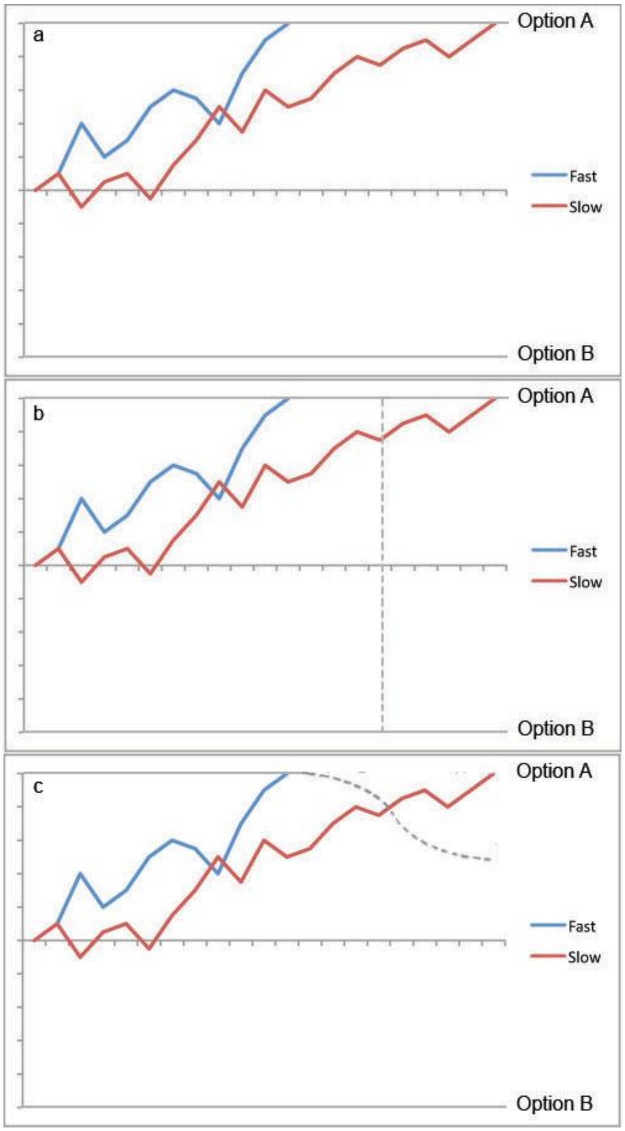


Figure 16. Examples of how diffusion models may include a timeout function. A fast decision (e.g., low ambiguity evidence) is represented in blue and a slow decision (e.g., high ambiguity evidence) is represented in red; (a) Traditional diffusion model without a timeout function in which evidence accrues in the drift parameter and the decision is executed upon reaching the upper and lower bound; (b) Model with a predesignated time (vertical dashed line) at which a decision will be made in favor of the closest bound if the threshold has not yet been passed; (c) Model with a decreasing threshold as time extends (note: sigmoidal shape is for illustration purposes only).

While both of these possibilities would result in lower accuracy due to insufficient perceptual evidence to pass the initially high decision threshold, there are slightly different predictions with regard to reaction time. For the timeout option, with a definitive pre-designated time at which a decision is made, we would expect a consistent mean reaction time across responses when evidence is not forthcoming. By this view, in the current line of investigation in which a motion signal can appear transiently at any point in a long time window, we would expect participants to consistently false alarm earlier in the trial for trials in which motion is presented very late. The time at which participants false alarm should be similar across the later conditions (e.g., when motion is presented 6 seconds or 8 seconds into the trial) if a pre-designated time is being used. For the variable threshold option, in which the threshold continually lowers as time goes on, response times should be related to the meeting of increasing evidence and the decreasing threshold. Under the same assumptions as random walk models, responses should vary in time based on physiological noise in interpreting incoming information and extracting evidence (Ratcliff, 1978; Smith, 1986; Ratcliff & McKoon, 2008), and may allow for longer waiting periods than a pre-designated timeout. This variation dependent on physiological noise would create a wider distribution of response times across conditions, without clustering around one specific time point late in the trial. The following experiment will test how the timing of incoming evidence is incorporated into the final decision, and will examine the resulting changes in accuracy and reaction time. Reaction times that tightly cluster together late in the trial with a large number of false alarms will support the timeout model, while spread out reaction times more associated with the timing of the stimulus presentation will support the variable threshold model.

4.2 METHODS

4.2.1 Participants

Sixty-one healthy young adults (ages 18-25 years) were recruited from the Introduction to Psychology participant pool at the University of Pittsburgh. Participants were tested in groups of 10-22 in a computer lab, with each seated at their own computer. The experiment lasted one hour and participants received one credit towards their course requirement. Prior to beginning the experiment, all participants completed a consent form approved by the Internal Review Board at the University of Pittsburgh.

4.2.2 Materials

The experiment was conducted using EPrime software (Psychology Software Tools, Pittsburgh, PA) on Dell computers. Responses were made via the regular keyboard attached to the computer. The display presented white dots on a black screen contained within a circle of space approximately 4” in diameter; a fixation cross was in the middle of the circle. The dots were replotted in a random direction at a rate of 60Hz. The experimenter could control both the direction and coherence (i.e., how many dots were moving together in the same direction) of the display.

4.2.3 Design

A 2000ms section of 25% coherent motion was embedded within each 10000ms trial. This section of motion was implemented at 5 different windows at equal intervals. In order to allow for enough time to respond, only the first 4 windows will be examined (see explanation of analysis parameters in Results). No other manipulations were included. There were 42 trials of each condition, for 210 total trials combined across conditions. Trials were chosen randomly and were broken into seven blocks of 30 trials each (approximately 6 minutes in length).

4.2.4 Procedure

After completing the consent forms, participants were given verbal instructions regarding the basic task structure, response options, and length of the experiment. Next the participants read a series of instruction screens reiterating this information with more detail; at the end they were asked to alert the experimenter of any questions they may have. Then they started a practice session consisting of 5 trials; afterwards they were again prompted to ask any remaining questions they may have. When the participants felt comfortable they proceeded on to the experiment, which was divided as described above. After each block of trials participants were asked to rest their eyes; at the end of the experiment they were thanked for their time and asked to leave the computer lab.

Each trial began with 2000ms of random motion and a red fixation cross. Participants were instructed that periods with the red fixation cross contained no motion and no responses should be made at that time. The fixation cross would then change to white for 10000ms, and participants were to make a motion discrimination response once per trial (Figure 17). The motion was set to 0% coherence except the 2000ms window of 25% motion that moved dependent on timing condition. When the participant felt confident about the direction of motion, he/she would press the corresponding number on the number pad at the right side of the keyboard. For example, “2” was down and “6” was right. Motion could be in any of 8 directions (up, down, left, right, and the diagonals). When not making responses, participants keep their finger on the “5” key in the middle of the number pad.

4.2.5 Data Scoring and Sorting

Trials were sorted based on when the 2-second bin of motion occurred (either the 1st 2 seconds, the 2nd 2 seconds, and so on up to the 5th 2 seconds of the trial). Each trial was then scored based the response; there were five possible responses: correct (number pressed on number pad matched direction of motion), incorrect (number response did not match direction of motion), fixation (response occurred during the fixation period prior to the start of trial), double response (two responses were made during the trial), and no response. Fixation, double response, and no response trials were removed for further analyses. The remaining correct and incorrect trials were sorted based on when the responses was made during the trial. Responses prior to the presentation of motion were labeled as false alarms. Responses during the 2 seconds of motion presentation and the 2 seconds immediately succeeding the motion stimulus (a 4 second window) were labeled as hits or misses according to accuracy. Responses more than 4 seconds after the presentation of motion were also discarded as it is difficult to assess whether these are late hits or misses, or false alarms after having failed to detect the motion.

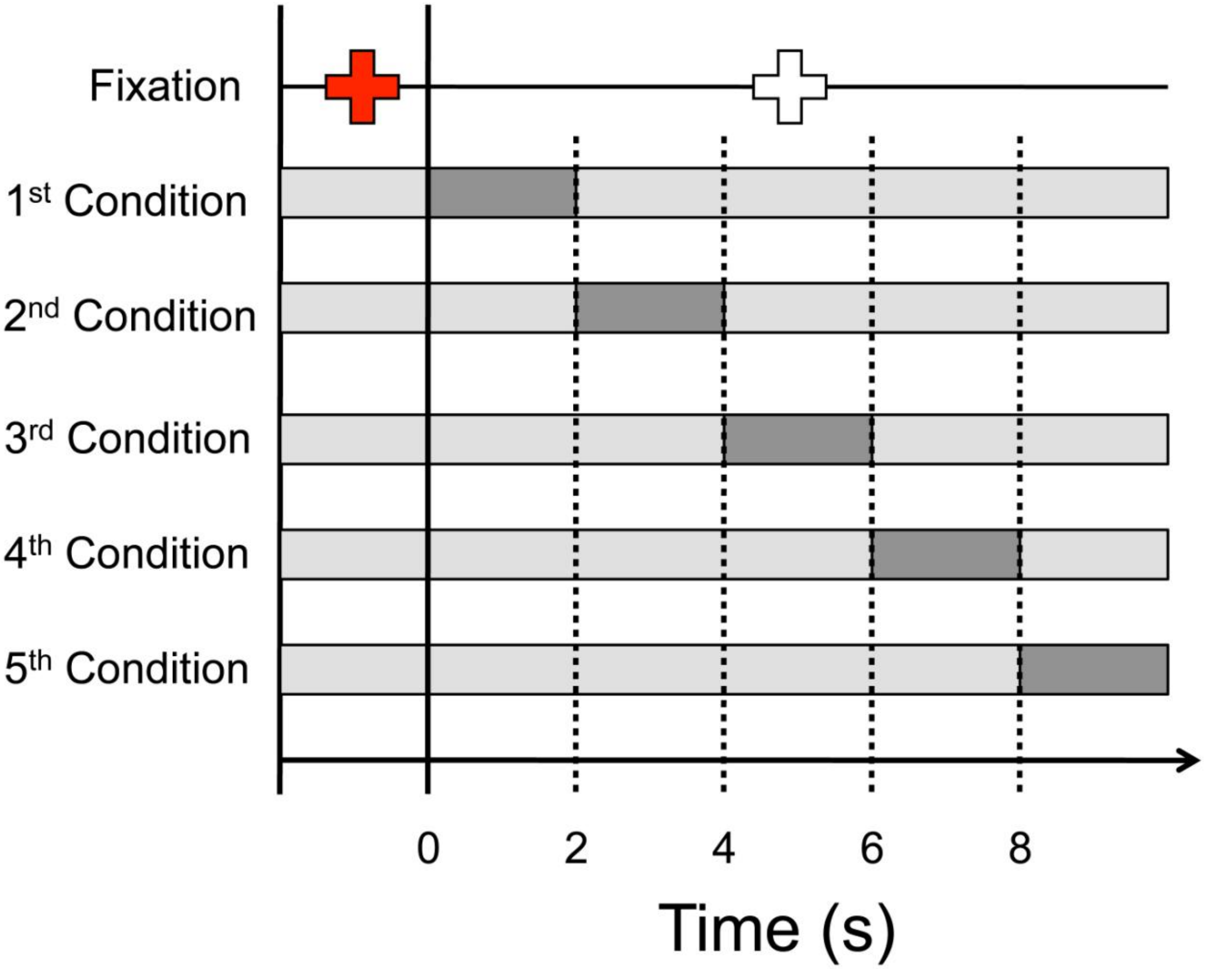


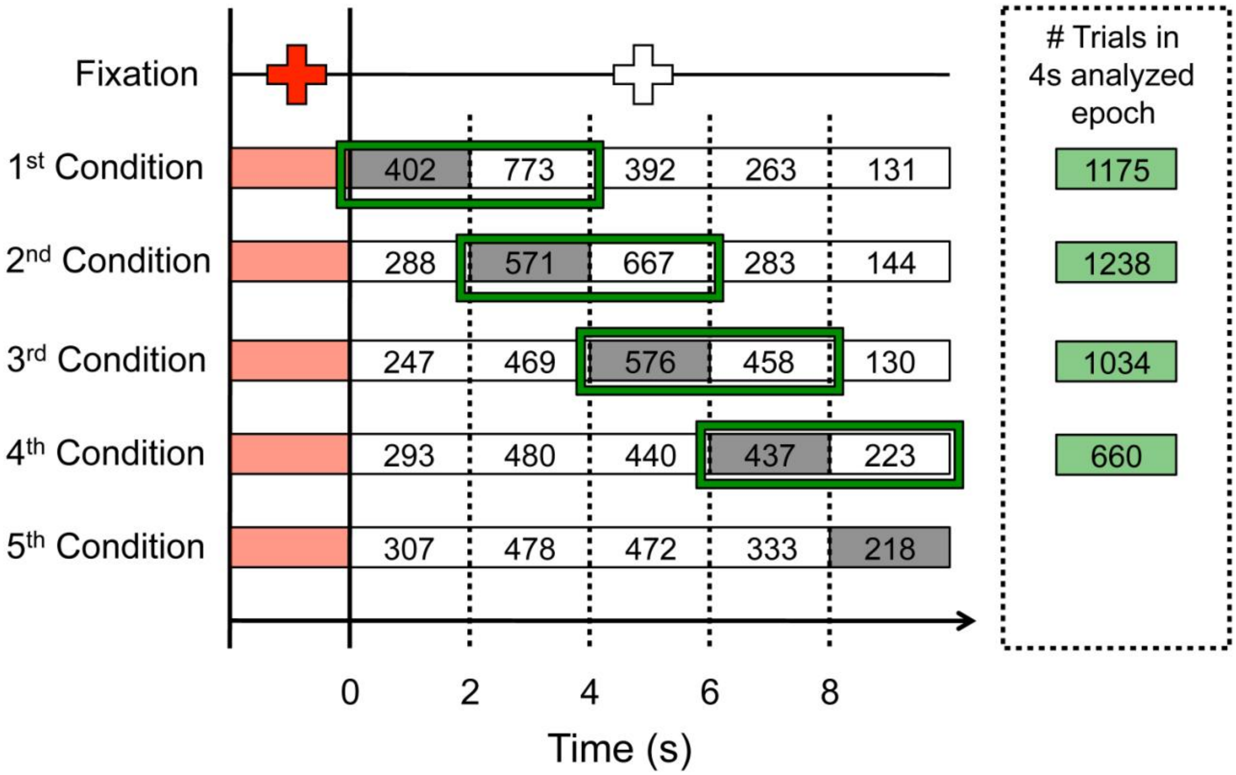
Figure 17. Representation of the task. Light grey boxes indicate 0% (random) motion, dark grey boxes indicate 25% motion in any of 8 directions. Fixation crosses are the same for all trials.

4.3 RESULTS

4.3.1 Accuracy and Reaction Time

Five participants were removed from the analysis due to missing data in one or more conditions; this indicates that they did not learn to wait for the coherent motion and instead always responded within the first few seconds of each trial. Three more participants were removed from the analysis because their average accuracy across conditions was not above chance (note: in this particular experiment chance is 12.5% because there are 8 possible options at all times). In order to ensure accurate measurements of accuracy and reaction time, participants were required to have at least 3 trials per bin of data in order to be included. Using this guideline, 23 participants with fewer than three trials per bin per accuracy status were removed from the analysis. The following analyses were completed on data from the remaining 30 participants.

To equate the amount of time available for response across conditions, only the 2000ms of motion (grey boxes, Table 1) and the 2000ms immediately afterwards will be analyzed for each condition (the maximum time possible in the 4th bin condition; see Data Scoring and Sorting). For example, when motion was presented in the 2nd condition, there were 571 trials during which participants attempted to make a response. In addition, the following 2000ms window (white box immediately to the right of the grey box) included another 667 responses to the motion. This division of the data results in roughly equivalent numbers of trials, though the 4th bin condition has slightly fewer (right column, Table 1). The diminished trial count in the 4th bin is due to an increase in false alarms during the 6 seconds prior to presentation of coherent motion. These false alarms will be analyzed in an ANOVA and incorporated into receiver-operating characteristic analyses at the end of this section.



						# Trials in 4s analyzed epoch
1st Condition	402	773	392	263	131	1175
2nd Condition	288	571	667	283	144	1238
3rd Condition	247	469	576	458	130	1034
4th Condition	293	480	440	437	223	660
5th Condition	307	478	472	333	218	

Table 1. Total number of responses across participants in each time window depending on trial type. Numbers in dark grey boxes indicate responses during actual motion, numbers within the green boundaries indicate analyzed responses (summed up in the column of green boxes to the right.

Accuracy was calculated by dividing the number of hits by the total number of trials in which the participants responded during the 4-second window during and after motion presentation as specified above. Accuracy decreased as the wait time prior to coherent motion increased (F(3,63.31) = 22.29, p < .001). Motion within the first 2000ms (1st bin) resulted in high accuracy; as motion moved later in the trial, accuracy started to decrease (Table 2). Post-hoc pairwise comparisons reveal that the 1st bin is more accurate than the others (all p < .001), the 2nd bin is more accurate than the 3rd and 4th bins (both p < .01), but the 3rd and 4th bins have similarly low accuracy (p > .10). This suggests that the decision process was operating with a higher threshold earlier in the trial that demanded more information in order to execute a response, leading to more accurate responses.

Reaction time was calculated from the start of coherent motion in each trial to equate time across the conditions. A repeated-measures ANOVA indicated that overall reaction time, regardless of accuracy, decreased as the timing of motion occurred later in the trial (F(3,68.06) = 44.54, p < .001); all post-hoc pairwise comparisons were significant (p < .03). This main effect of timing of the motion on reaction time is consistent for both Correct (F(3,87) = 15.30, p < .001) and Incorrect (F(3,87) = 31.073, p < .001) trials, indicating that the change in reaction time is an effect of the decision process as a whole, and does not change depending on how well the participant is able to interpret the incoming information. There was also a main effect of accuracy on reaction time (F(1,29) = 41.642, p < .001), such that Incorrect trials are faster than Correct trials (Table 2). However, an interaction was present (F(3,87) = 11.461, p < .001) and post-hoc pairwise comparisons revealed that the difference in reaction time depending on accuracy was only present for the 3rd and 4th bins (both p < .01), suggesting that it is after waiting

Timing of Motion	**Accuracy**	**Correct Reaction Time**	**Incorrect Reaction Time**
1st bin	69.2% (3.2)	2611.60ms (65.15)	2580.32ms (145.93)
2nd bin	54.5% (2.9)	2349.87ms (92.09)	2216.02ms (113.93)
3rd bin	46.5% (3.1)	2272.18ms (67.09)	1573.57ms (90.71)
4th bin	44.1% (2.7)	2059.20ms (64.42)	1299.15ms (75.20)

Table 2. Accuracy rates and reaction times depending on timing of motion (numbers in parentheses are standard error). Note: chance in this experiment is 12.5%.

for long periods of time prior to coherent information is presented that the decision process begins to lose fidelity. Overall, the data indicate that as decisions extend in time, the process through which those decisions are made changes; these changes result in faster, but less accurate responses once information is presented.

4.3.2 False Alarm Rates

One indication of a timeout function would be increased false alarm rates as the wait time before motion presentation increases. The longer the wait, the more likely a false alarm will occur. An analysis of the 5th condition in which motion did not occur until the last 2 seconds of the trial, thus requiring 8 seconds of waiting, demonstrated an increasing number of false alarms as wait time increased. Trials were separated in 2-second bins to complement the timing of motion in the other conditions, and then the false alarm rate was calculated by dividing the number of false alarms by the number of trials that had not yet received a response (i.e., opportunities to false alarm). A repeated-measures ANOVA demonstrated that there was a higher percentage of false alarms at later times in the trial ($F(3,87) = 65.18$, $p < .001$). All pairwise comparisons were significant; descriptive statistics are presented in Table 3. Again this suggests that the decision process loses fidelity as the process extends in time, allowing for more premature executions of the decision (i.e., false alarms) prior to the presentation of information.

Time in Trial	Mean False Alarm Rate (SE)
0-2 sec	5.50% (1.2)
2-4 secs	14.60% (1.2)
4-6 secs	22.00% (1.5)
6-8 secs	31.80% (1.8)

Table 3. False alarm rates calculated from the 5th condition in which motion is not presented until 8 seconds into the trial.

4.3.3 Receiver-Operating Characteristic Analysis

Sensory discrimination tasks often use a receiver-operating characteristic (ROC) analysis to assess participants' sensitivity (d-prime, or d') and selectivity (criterion, or beta). Due to the nature of the task, there is not an obvious way to calculate d'. Unlike tasks in which each trial presents a single stimulus (e.g., motion present or absent) and requires a single response (e.g., "present" or "absent"), the extended trial structure means that on any given trial participants have varying durations in which they may false alarm at the beginning of the trial and only one chance to make a hit or miss later in the trial. In addition, an early false alarm on a trial removes the opportunity for a hit once motion is presented due to instructions to only respond once per trial. Therefore the participant must continuously correctly reject (i.e., withhold response) for several seconds in order to have the opportunity to make a hit or miss once motion is presented. This overlapping dependence of response options may create several confounds for the ROC analysis. In order to capture as much as possible with this analysis, three different d' and beta calculations were made.

The first analysis assessed false alarm and hit rates for each trial type and calculated d' and beta accordingly. Because it was not possible to make a false alarm when motion was presented in the first 2 seconds, this analysis only includes trials when the motion was in the 2nd, 3rd, or 4th bin. This analysis found that both d' ($F(2,58) = 98.58$, $p < .001$) and beta ($F(2,58) = 151.05$, $p < .001$) decrease as time waiting for motion increases. All post-hoc pairwise comparisons were significant (all $p < .001$; Table 4).

The second d' analysis focused on time within the trial, collapsed across conditions. This analysis calculated the probability of hits from conditions with coherent motion in the 1st, 2nd, 3rd, or 4th bins. Hit rates were calculated for each time epoch from trials in which motion was presented in that epoch, with responses in both that epoch and the immediately succeeding epoch collapsed together to account for RT differences (e.g., if motion was presented in the first bin, responses in the 1st and 2nd bins were counted for the first epoch). In order to separate the opportunity to false alarm at any given point with the opportunity to make a hit later in the trial, the false alarm rate was calculated from the 5th bin. This analysis found a main effect of time within the trial on d' (F(1.82, 52.70) = 69.59, $p < .001$). D-prime decreased in each succeeding time epoch within the trial (Table 5), with all post-hoc pairwise comparisons resulting in significance (all $p < .01$). Beta values also decreased as a function of epoch (F(3,87) = 82.39, $p < .001$); all post-hoc pairwise comparisons were again significant (all $p < .01$; Table 5).

The third analysis redefined hits as any response made during the 4-second analysis window during and after motion presentation for each condition in relation to the total number of trials at which point motion had already been presented. In this case, accuracy was disregarded and any response during the 4-second window was considered a "hit". False alarm rates were calculated using responses prior to motion onset in the other conditions during the 2-second time window for the calculated condition (e.g., for the 1st condition hits were responses in the 1st and 2nd bins when motion was presented in the 1st bin, false alarms were responses during the 1st bin during the 2nd-5th timing conditions at which point motion had not yet been presented) in relation to the total number of trials for which a response had not yet been made (Figure 18). This analysis yielded some interesting results. Although there was a main effect of timing of motion on d' (F(1.39,40.46) = 9.92, $p = .001$), this was driven by lower d' values in the 2nd condition than the 1st and 3rd conditions (both comparisons $p < .05$) while there was no difference between the 1st and 3rd conditions ($p = 1.0$). In addition, there was also a main effect of timing on beta values (F(1.41,41.01) = 118.47, p < .001), with all post-hoc pairwise comparisons significant (Table 6).

While the results from these ROC analyses suggest that both sensitivity and selectivity decrease as time spent on the decision increases, it is important to note the issues described earlier. In the first analysis, there is a greater opportunity to false alarm in the 4th condition than the 2nd condition. The participant must withhold response (i.e., correctly reject) for the equivalent of 3 trial windows (6 seconds) to reach the information in the 4th bin, while they only need to withhold response for 1 trial window (2 seconds) prior to motion in the 2nd bin. While on the one hand this describes exactly the issue at hand (i.e., whether or not waiting longer causes more false alarms), the analysis may be biasing the results by including more opportunity. There was an attempt to control for this in the second analysis by using only the associated bin of the 5th condition for the false alarm rate, while still maintaining the hit rate from the appropriate 1st, 2nd, 3rd, or 4th condition. The d' and beta values using this technique again demonstrated decreases in both measures as the decision process extended in time, suggesting that participants are losing sensitivity and selectivity over time.

The third analysis presents a potentially more complicated story. Again, there are constraints on the validity of this analysis due to different numbers of trials per calculation of hit rate and false alarm rate. However, there is some possibility that d' was remaining stable while beta values decreased. This would complement the hypothesis that the threshold for response is decreasing over time (represented by beta), while overall approach to the task is remaining constant (represented by d').

Timing of Motion	d' (SE)	beta (SE)
2nd bin	1.65 (.16)	1.52 (.10)
3rd bin	.69 (.14)	.78 (.10)
4th bin	-.17 (.11)	-.01 (.08)

Table 4. D-prime and beta values calculated per condition.

Epoch in Trial	d' (SE)	Beta (SE)
0-2 secs	2.40 (.17)	1.82 (.09)
2-4 secs	1.24 (.12)	1.11 (.06)
4-6 secs	.71 (.12)	.80 (.05)
6-8 secs	.35 (.10)	.50 (.06)

Table 5. D-prime and beta values calculated using hit rates from the appropriate timing condition and false alarm rate from the 5th condition.

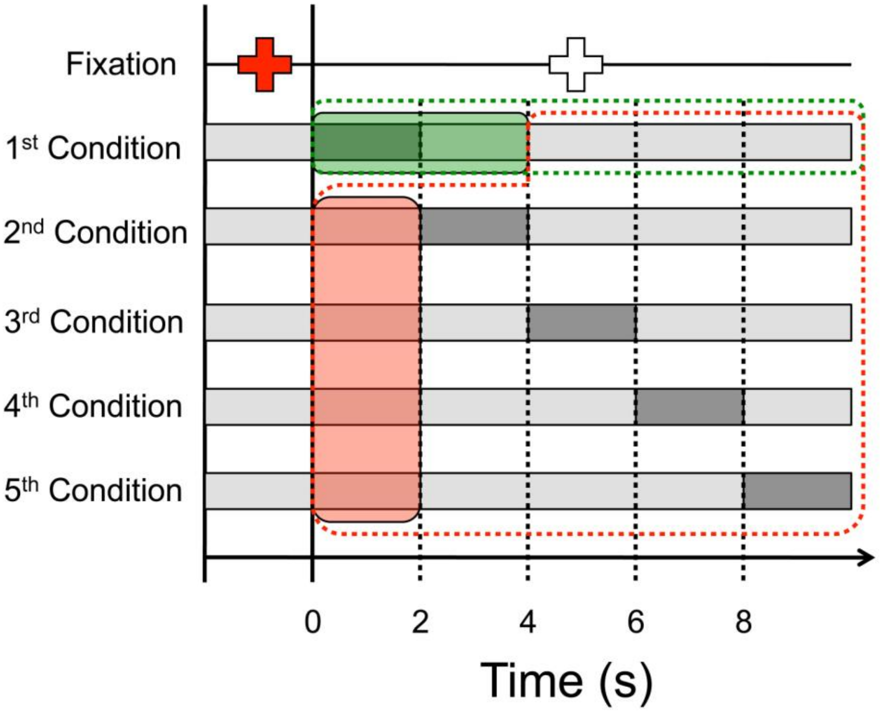


Figure 18. Depiction of analysis windows for third d' and beta analysis using the 1st condition as an example. Green box indicates 4-second window for hits in the 1st condition; these responses were divided by the total number of responses in the dashed green window. The red box indicates the responses that were considered false alarms for the 1st bin, and were divided by the total number of trials in the red dashed window.

Timing of Motion	**d' (SE)**	**Beta (SE)**
2nd bin	1.85 (.11)	1.88 (.14)
3rd bin	1.16 (.07)	.67 (.07)
4th bin	1.68 (.15)	.13 (.05)

Table 6. Average values for the third d’ analysis that disregarded accuracy.

4.3.4 Mean Reaction Times Across Response Types

One of the potential instantiations of a timeout function is a pre-designated time at which a decision must be executed regardless of what evidence has been presented. This would result in a relatively stable mean response time among the later conditions coinciding with the same timing as highest percentage of false alarm rates as indicated above. Alternatively, a variable threshold without a definitive time associated with it would result is a larger distribution of responses, with each later condition resulting in a later mean reaction time. To assess reaction times among the conditions, the mean of each participant's RT distribution for each motion condition was calculated regardless of response type (e.g., false alarm, hit, or miss) and the pattern of RTs was assessed across bins 1-5. A repeated-measures ANOVA found that the mean response time decreased as the wait before motion was presented increased ($F(2.42,70.06) = 32.12$, $p < .001$). This is not surprising because the informative portion of the trial moved later across conditions, thereby moving the response times for hits and misses. However, the post-hoc pairwise comparisons highlight a more complicated story. While the 1st condition is faster than all others (all comparisons $p < .001$) and the 2nd condition is faster than the 3rd, 4th, and 5th conditions (all comparisons $p < .05$), the latter conditions are not significantly different from each other (Figure 19). Importantly, the average reaction times in the 4th and 5th conditions were *prior* to the presentation of coherent motion in those conditions, highlighting that decisions were consistently made prior to the presence of coherent information. In addition, the similarity of the mean reaction times for both the 4th and 5th conditions suggests that participants were unwilling to wait longer than a certain amount of time (approximately 5500ms per trial in the current experiment).

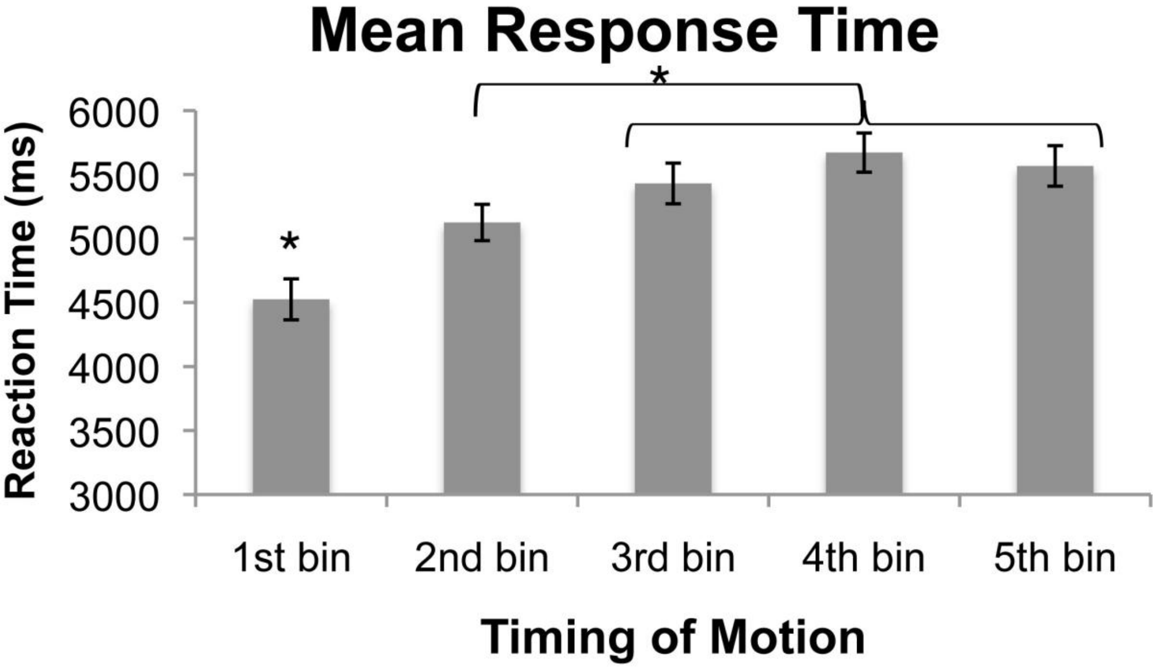


Figure 19. Mean reaction time depending on timing of motion, regardless of accuracy or false alarm status.

4.4 DISCUSSION

These results highlight some interesting aspects of how the timing of incoming sensory evidence in relation to the start of the decision-making period can affect the end response out of that decision. The significant decreases in both accuracy and reaction time as the timing of evidence came later in the trial suggest that the decision process changed as it extended in time. Two competing hypotheses might account for these results. One theory would require that the timeout occur at a consistent, pre-designated time in the decision process that would function to maximize efficiency when it appears no new evidence will be presented, similar to discussions of some memory models (Gillund & Shiffrin, 1983; Anderson, 2007). At the pre-designated time, selection would be made of whichever response alternative was the most likely solution based on evidence that had accrued to that point. The other theory suggests that the threshold used to execute the decision process lowers over time, making it easier to execute the decision on little evidence but not requiring a response at a particular time. This is similar to how the threshold changes under speed instructions compared to accuracy instructions (Fitts, 1966; Ratcliff & McKoon, 2008), but would occur without explicit instruction on the part of the experimenter. While both hypotheses would predict decreases in accuracy due to response execution prior to the initial high threshold is passed, the pre-designated timeout would predict a consistency in mean reaction times among the later conditions while the lowering threshold would predict a larger distribution of reaction times across conditions.

Not surprisingly, there was a decrease in accuracy as the timing of motion was later in the trial. In addition, two ROC analyses demonstrated decreases in both d’ and beta values, although a third analysis presented the possibility of stable d’ values while beta decreased. Overall, the decrease in beta values suggests a lowering criterion to allow more responses as the trial continues. However, as mentioned in the Results section, it is unclear whether these values have any meaning given the task structure. Each trial had multiple opportunities to correctly reject or false alarm prior to the opportunity to make a hit or miss. The participant had to be able to continuously correctly reject the random motion over the course of the trial (up to 8 seconds in the 5th condition) in order to give a single response during the coherent motion section. This overlapping nature involved in each trial made calculating the necessary rates difficult to do in isolation. In addition, the theory behind d’ assumes a 50% possibility for a hit given a stimulus (e.g., “old” vs. “new” responses to a studied item, or “present” vs. “absent” to a visual stimulus). The task in this experiment has only a 12.5% possibility of a hit given a stimulus (i.e., the participant identifies not only that the stimulus is present, but also the particular direction out of eight possibilities). Although it is possible to calculate signal detection measures for cases with more than two response alternatives, in which each pair of alternatives receives its own d’ measure (Macmillan, 1991), it continues to be unclear whether or not the task structure itself negates the use of a signal detection calculation.

The results also demonstrated that when participants did wait long enough to see the coherent motion, reaction times to motion later in the trial were shorter than to motion early in the trial when measured from motion onset. That is, responses to motion in the 4th bin were faster on average than responses to motion in the 1st bin. This suggests that participants were using a less stringent decision rule as the trial went on in order to maximize response efficiently later in the trial. As a result, and as described above, accuracy was lower later in the trial. Aside from faster reaction times once motion was presented, overall reaction times (collapsed across false alarms, hits, and misses) illustrated an interesting pattern. While average reaction times for the 1st and 2nd conditions were after the presentation of motion, average reaction times for the 3rd condition were in the middle of the motion block. This is not surprising given the increase in speed of responses just described. However, average reaction times for the 4th and 5th conditions were *prior to* the presentation of motion, meaning that on average participants were responding before they saw any meaningful evidence (i.e., the majority of their responses in these conditions were false alarms). The tendency to respond prior to motion onset is confirmed by the increase in false alarm rate as the random motion at the beginning of the trial extended in time. What is particularly interesting is that the trial did not end upon response; participants still had to wait the full 10 seconds. This has two implications of note. First, there was no incentive for participants to make early responses as it would not shorten participation time. And second, participants would see the motion presentation late in the trial during the 4th and 5th conditions, which could in theory cause participants to learn to wait longer. Yet despite seeing motion at the end of the trial, it is clear given the average reaction time results that participants did not learn to wait longer.

All of these findings lend support to both the lowering threshold and pre-designated timeout hypotheses. However, the overall reaction time data may highlight the stronger theory. There was no difference in mean reaction time in the 3rd, 4th, and 5th conditions. Although the mean reaction time was within the motion period for the 3rd condition, it was before motion was presented in both the 4th and 5th conditions (Figure 20). In addition, this mean reaction time is consistent with the high rate of false alarms during the 3rd and 4th epochs. A reanalysis of the data in 1 second epochs (half the epoch size as the original analysis) demonstrated that the highest rates of false alarms occurred between 5000 and 8000ms, concurrent with the 5600ms average response time across conditions (Figure 21). A repeated-measures ANOVA indicated that there was a main effect of timing on the false alarm rate ($F(3.39,98.57) = 25.33$, $p < .001$); in addition, the linear trend was also significant, indicating that the false alarm rate rose at each successive step ($F(1,29) = 109.45$, $p < .001$).

The overall reaction time results suggest that the decision process in this experiment maintains a pre-designated time by which a decision must be made, regardless of the presence or absence of informative evidence. The consistency of average reaction times in the 3rd, 4th, and 5th conditions, in addition to the fact that the average times in the 4th and 5th conditions is prior to the availability of evidence, suggests that there is some mechanism pressuring for the execution of a response after a certain about of time (in this case, approximately 5.5 seconds). It should be noted that it is unlikely that this result is due to time pressure from the impending end of the trial (Payne, Bettman, & Johnson, 1988). The average response time is only a little more than half the total trial length. All trials were 10 s in length, regardless of condition or response, so there should not have been an inherent pressure to respond prior to, on average, 6 seconds as the average reaction time suggest. In some ways, the decrease in reaction time dependent on wait length is a bit surprising. Previous studies that used a variable wait prior to the presentation of the stimulus found very little shift in reaction time after an initial speed-up (Green, Smith, & von Gierke, 1983); however, the longest wait period in that study was slightly over 3 seconds. It is possible that the extension of the wait period to more than twice that length on some trials is triggering a mechanism not previously found.

This experiment examined how perceptual decisions behave when confronted with a dearth of evidence for a required decision. In the trials presented here, participants often began with a lack of information for an extended period of time, yet were required to make a response.

When the timing of incoming evidence is ambiguous, some mechanism must be in place to efficiently execute a response as soon as evidence in available. If there were insufficient evidence to pass threshold (even a lowered one), a timeout function as described here would serve as an efficient way to limit decision time. Although there may be instances in which long decision times do not result in adverse consequences, it is better for cognition to err on the side of limiting processing so as to execute responses in a fast amount of time and free up resources for other matters (Kiani, Hanks, & Shadlen, 2008). In the results shown here, this may be the case even when the inevitable end of the trial is far from the time at which this mechanism acts. The next experiment will examine how flexible the decision process is during the earlier portion of the trial, which should be unaffected by any regulatory mechanisms (as illustrated by the low level of false alarms here).

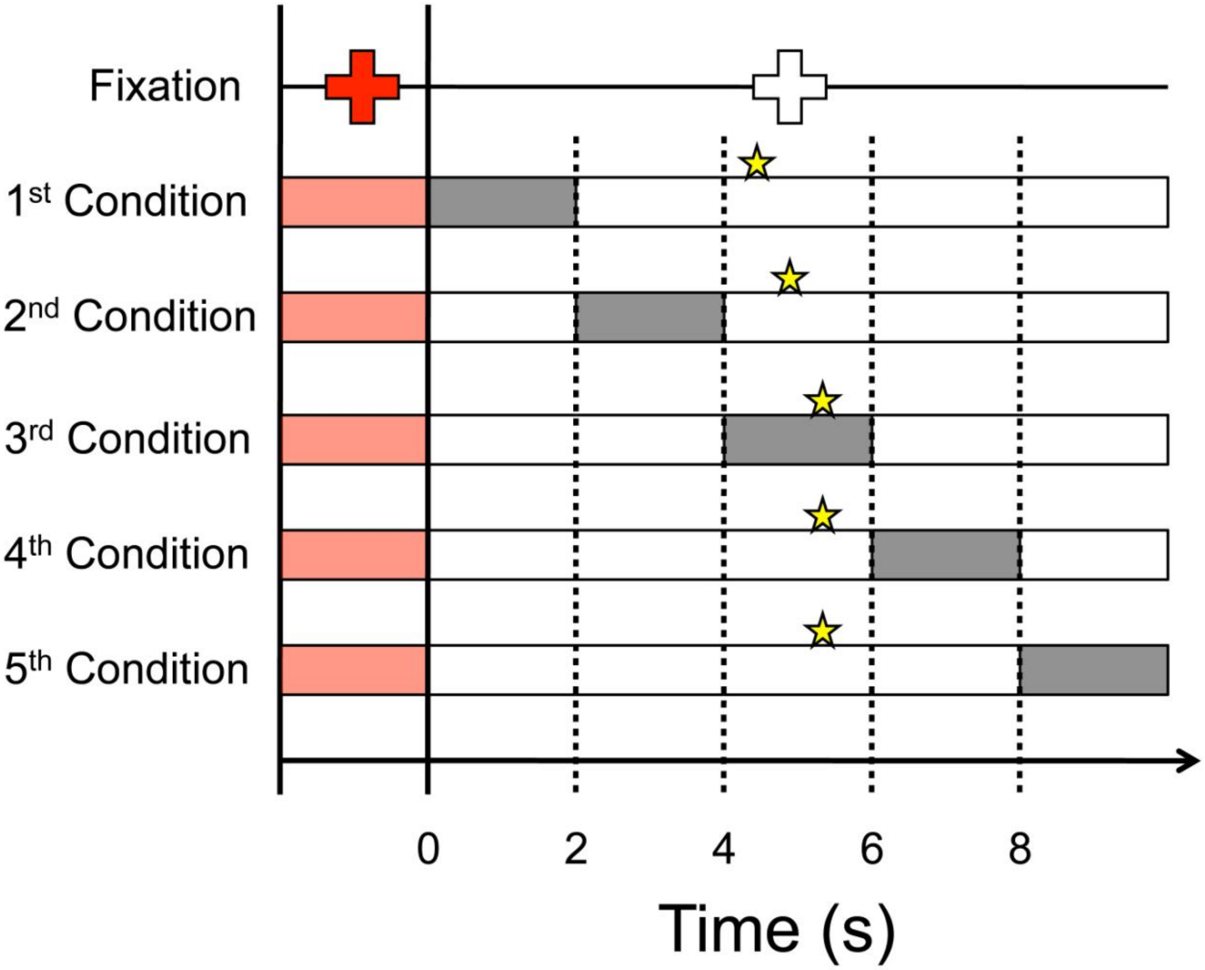


Figure 20. Mean reaction time in relation to the task structure. Yellow stars indicate the mean reaction time for each condition; note that stars for the 4th and 5th conditions are located prior to the grey box indicating the 25% coherent motion stimulus.

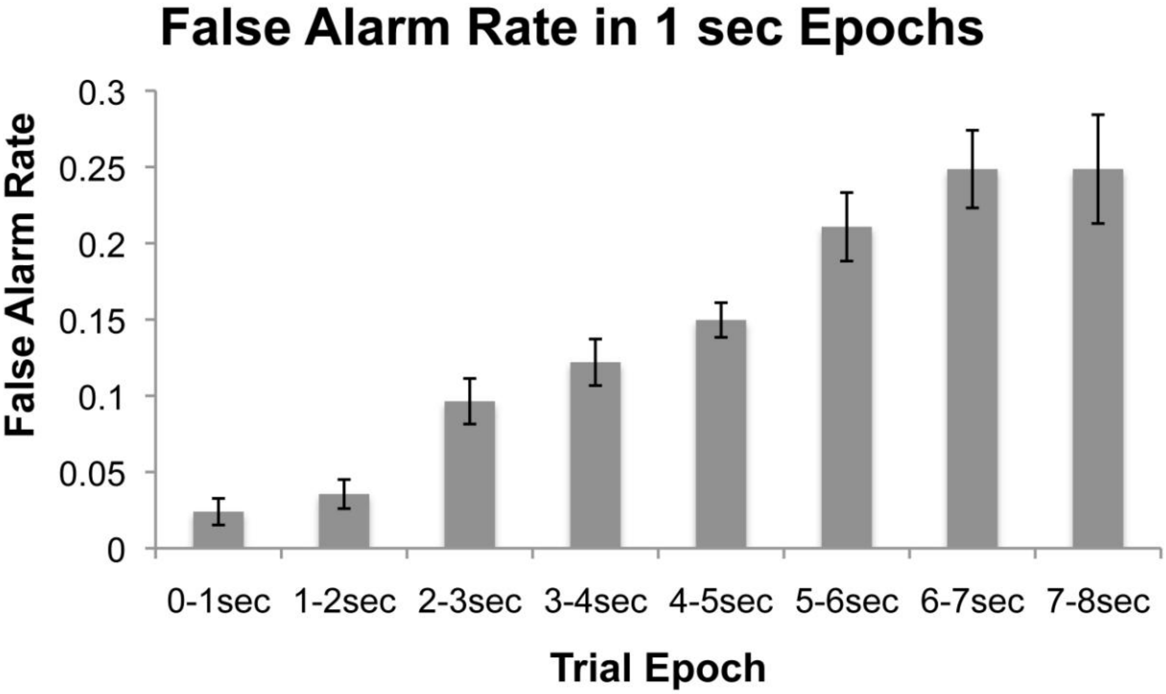


Figure 21. False alarm rate collapsed across conditions, divided into 1-second epochs.

5.0 FLUCTUATIONS OF EVIDENCE

5.1 INTRODUCTION

When faced with imperfect evidence, the perceptual decision process attempts to gather evidence over time to maximize the amount of useful information available towards the decision. However, as seen in the previous chapter, the parameters used to make that decision may change as this process is drawn out. When coherent motion was presented later in the trial decisions were faster and less accurate, suggesting a change in the threshold used to assess the data and execute the decision. On the other hand, coherent motion presented early in the trial was processed through a slower, more accurate decision process. This suggests that the timing of incoming information may have different effects on the decision process. This leaves open the possibility that other slight changes in the stimulus may influence the process, such as changes in the fidelity of incoming information as is examined here.

Attempts have been made to alter the decision process through various means. Research is several areas of perception have tested the ability to influence decisions by electrical microstimulation of relevant brain areas. For example, microstimulation of the lateral intraparietal sulcus can change the speed of responses dependent on whether the stimulation is affecting neurons that are excited or inhibited by the stimulus (Hanks, Ditterich, & Shadlen, 2006). Also, previous research has shown that presenting short, presumably imperceptible, bursts of strongly coherent evidence during an otherwise difficult decision can influence the final decision in motion discrimination paradigms in macaques without the aid of microstimulation (Huk & Shadlen, 2005; Kiani, Hanks & Shadlen, 2008). Huk and Shadlen (2005) found that the bursts created changes in the increasing trajectory of neural firing that caused it to reach the common level earlier than without the burst. That is, while initially accumulating normally, the burst caused the accumulation rate to increase. However, the monkeys were required to maintain fixation until a cue, so it is unknown whether or not this would have resulted in faster reaction times (e.g., faster rise in activity to the common plateau level would have triggered a threshold and caused a detection response). In addition, Kiani, Hanks, & Shadlen (2008) found that the effectiveness of the burst depended on the efficacy of the stimulus it interrupted. If the burst was inserted into a stimulus that was already above threshold, it did not affect the result. On the other hand, inserting a burst into a low coherence stimulus resulted in a longer effect, and was able to cause that effect later in the trial (i.e., presenting the burst later in time was still effective).

The current experiment assesses whether or not changes in the stimulus (e.g., embedded bursts of motion like Huk & Shadlen, 2005) can create changes in reaction time, like those found from direct microstimulation in Hanks, Ditterich, & Shadlen (2006). Without artificial stimulation of the neuronal pools involved in the decision, the current experiment relies on natural increases in firing due to the stimulus itself. If the increase in firing due to increased coherence of the stimulus is strong enough to pass the decision threshold, we should see similar changes in reaction time compared to the microstimulation experiment (i.e., faster when highly informative evidence is present than not). This will be examined by inserting bursts of strong evidence early in the trial, at a time when threshold levels still appear to be high based on 4.0 . The goal is for the burst to be short enough to not cause the immediate execution of the decision, but rather to influence the overall process and alter reaction time compared to trials without a burst of strong evidence. If the burst is long enough to immediately execute the decision, we will be unable to see how intermediate pieces of evidence influence the overall decision process, particularly the comparison of evidence in favor of and against the correct decision, described presently.

The influence of the highly informative, but shortly presented evidence should affect both decisions in accordance with the extra evidence and in the opposite direction as well. If early highly coherent information has the ability to influence a later decision we should see earlier responses in pro-burst blocks and later responses in anti-burst blocks compared to blocks with only regular non-burst trials, representing shifts in the speed of integration based on the highly coherent information. However, as long as the burst itself does not cause the termination of the decision, accuracy should remain approximately the same as the threshold for the decision should remain the same throughout the experiment. (If the burst does terminate the decision, anti-burst trials will not be accurate.) That is, the veracity of the early information should not affect the overall decision, just the speed with which it is made; this would be similar to a change in the drift rate in a diffusion model (Ratcliff, 1978). Another possibility is that while pro-bursts blocks will demonstrate faster responses, anti-burst blocks will be no different than regular blocks. This would support the theory that subjects ignore information that is contradictory to their current hypothesis (Mynatt, Doherty, & Tweney, 1978).

5.2 METHODS

5.2.1 Participants

Fifty-six healthy young adults (ages 18-25 years) were recruited from the Introduction to Psychology participant pool at the University of Pittsburgh. Participants were tested in groups of 10-22 in a computer lab, with each seated at their own computer. The experiment lasted one hour and participants received one credit towards their course requirement. Prior to beginning the experiment, all participants completed a consent form approved by the Internal Review Board at the University of Pittsburgh.

5.2.2 Materials

The experiment was conducted using EPrime software (Psychology Software Tools, Pittsburgh, PA) on Dell computers. Responses were made via the keyboard attached to the computer. The display presented white dots on a black screen contained within a circle of space approximately 4” in diameter; a fixation cross was in the middle of the circle. The dots were replotted in a random direction at a rate of 60 Hz. The experimenter could control both the direction and coherence (i.e., how many dots were moving together in the same direction) of the display.

5.2.3 Design

Three different trial types were used in this experiment; all had eight possible directions of motion (the four cardinal directions and the four diagonals, separated by 45°). The basic (“regular”) trial lasted 16 seconds and included eight 2000ms segments. Over the course of these segments, coherence of motion began at 5% and increased in 5% increments to 35% and direction of motion was always the same across segments. The final 2000ms segment was 85% coherent motion, during which participants verified their response (see Procedure). On some trials (“pro-burst”), the third segment (beginning 4000ms into the trial) was altered to include a 50ms burst of 85% coherent motion. The burst occurred 200ms into the segment to reduce visibility of the manipulation, and was in the same direction as the rest of the segments. On other trials (“anti-burst”), the trial structure was the same as for “pro-burst” but the direction of motion in the burst was 180° in rotation in comparison to the rest of the trial.

Trials were separated into four types of blocks. “Regular” blocks contained only “regular” trials to assess baseline performance on the task. “Pro” blocks contained a 25:75 mix of “pro-burst” and “regular” trials. “Anti” blocks contained a 25:75 mix of “anti-burst” and “regular” trials. “Mixed” blocks contained a 25:25:50 mix of “pro-burst”, “anti-burst”, and “regular” trials. The ratios were selected to limit how noticeable the burst manipulation was to participants, thus avoiding changes in strategy as much as possible. Participants completed 8 blocks total, two for each type. The Regular, Pro, and Anti blocks had 20 trials each; the Mixed blocks had 24 trials each.

5.2.4 Procedure

After completing the consent forms, participants were given verbal instructions regarding the basic task structure, response options, and length of the experiment. Next the participants read a series of instruction screens reiterating this information with more detail; at the end they were asked to alert the experimenter of any questions they may have. Then they started a practice session consisting of 13 trials; afterwards they were again prompted to ask any remaining questions they may have. When the participants felt comfortable they proceeded on to the experiment, which was divided as described above. After each block of trials participants were asked to rest their eyes; at the end of the experiment they were thanked for their time and asked to leave the computer lab.

Each trial began with 2000ms of random motion and a red fixation cross. Participants were instructed that the red cross indicated random motion and they should not respond during these periods. The motion discrimination period started once the fixation cross turned to white; each trial lasted 16 seconds total, divided into eight 2000ms segments. During the first segment, motion was presented at 5% coherence. Over the course of the succeeding segments motion increased in coherence in 5% increments through 35% at the seventh segment. The eighth (and last) segment consisted of motion at 85% coherence (Figure 22, “Regular” row). Depending on the trial status, a 50ms burst of motion may have occurred in the third segment as previously described in the Design section (Figure 22, “Pro Burst” and “Anti Burst” rows). The direction of motion was the same throughout all eight segments, except for “anti-burst” trials in which the burst of motion was in the opposite direction as the rest of the trial.

Participants were instructed to press the space bar when they could indicate the direction of motion. At the end of the trial during the 85% coherent 2000ms segment, participants were to press the space bar a second time if the direction of motion matched their earlier guess. If the direction did not match, participants withheld a response.

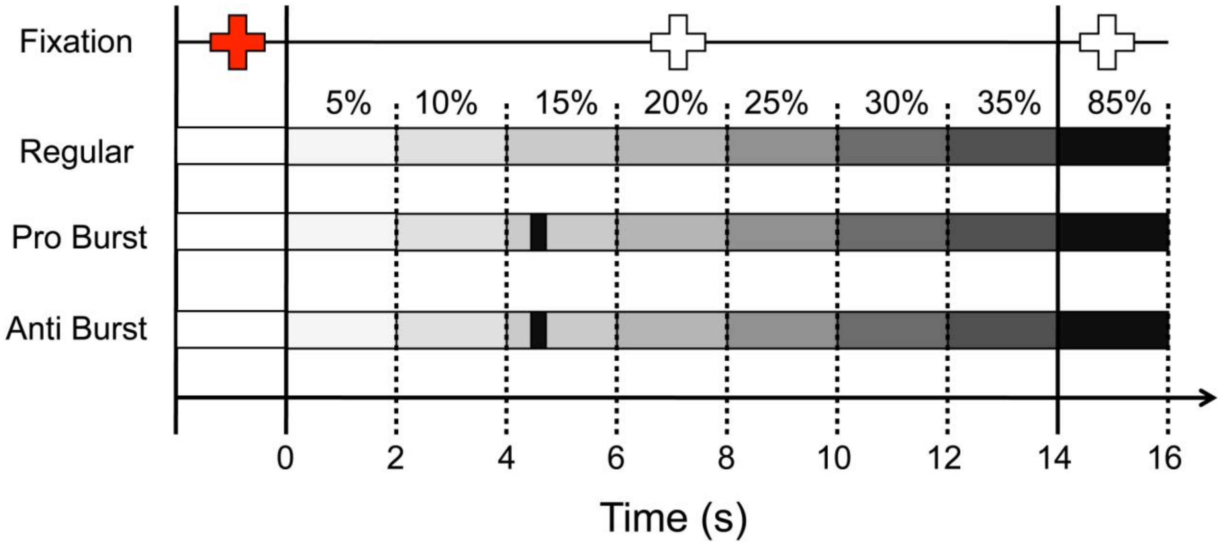


Figure 22. Illustration of task structure for the burst experiment. *Fixation row*, color of the fixation cross depending on break versus task periods; *Regular row*, coherence levels as they increased in 2000ms blocks, indicated by increasing saturation of grey; *Pro Burst* and *Anti Burst rows*, burst trials included an 85% coherence segment, indicated by the black bar, during the third portion of the trial (other 15% coherence). Note that verification responses were made during the final 2000ms of the trial during 85% coherent motion, illustrated to the right of the 14s marker.

5.3 RESULTS

Data from two participants was removed due to data recording problems. Data from two more participants was removed due to 0% accuracy scores, indicating a lack of learning the task properly. All data from the remaining participants (N = 52) was included in the following analyses; Greenhouse-Geisser corrections for sphericity violations were used when appropriate.

Overall accuracy was high but not perfect (M = 77.12%, SD = 14.52), indicating that participants were not inflating their self-reported accuracy as much as those in 3.0 . The self-reported accuracy rate is similar to an object recognition experiment with the same response structure (Ploran, Tremel, Nelson, & Wheeler, submitted), suggesting that with proper instruction and oversight, this response method can elicit good estimations of performance.

5.3.1 Changes in performance by block

Descriptive statistics for accuracy and reaction time broken down by trial type and block are presented in Figure 23 and Figure 24. It was expected that changes in the stimulus should only affect reaction time and not accuracy; therefore it is important to assess accuracy across the different blocks prior to analyzing specific trial types. Accuracy was only marginally different dependent on block (F(2.55, 130.30) = 2.54, p = .069). Post-hoc comparisons indicate that this effect is driven by lower performance in the Pro Block (M = 74.63%, SD = 14.43) compared with the Mixed Block (M = 78.31%, SD = 14.47; p = .03) and the Regular Block (M = 78.60%, SD = 13.91; p = .04). There were no other significant pairwise comparisons.

Despite the only marginal differences in accuracy dependent on block, there were significant differences in reaction time (F(2.66,135.97) = 17.38, p < .001). The Anti (M = 8005.64ms, SD = 1241.02) and Regular (M = 8151.16ms, SD = 1151.06) blocks, while not different from each other, are significantly slower than the Mixed (M = 7313.07ms, SD = 1389.61) and Pro (M = 7492.58ms, SD = 1446.95) blocks. This suggests that there were at least some differences in the decision process, even if performance remains at the same high level.

5.3.2 Comparison within blocks

Within the Mixed block, there was a main effect of trial type on accuracy (F(2,102) = 6.6, p = .002). Anti trials were the least accurate (M = 73.98%, SD = 19.44), then Regular trials (M = 79.35%, SD = 18.07), and Pro trials were the most accurate (M = 81.69%, SD = 13.97). While the difference between Regular and Pro trials was not significant (p = .86), both were significantly better than Anti trials (p = .06 and p = .002, respectively). In addition, a main effect of trial type was also found when examining reaction time (F(2,102) = 30.42, p < .001). All pairwise comparsions are significant (p < .05); Regular trials were the slowest (M= 7904.01ms, SD = 1311.42), then Anti trials (M = 7443.22ms, SD = 1705.59), and Pro trials were the fastest (M = 6587.19ms, SD = 1621.70). This indicates that the burst was affecting the timing of the decision (i.e., causing faster decisions for both Pro and Anti trials), but was also affecting the fidelity of the decision (i.e., lower accuracy for Anti trials).

There were no differences in accuracy between Anti trials and Regular trials in the Anti block (t(51) = -1.06, p = .29). However, there was a difference in reaction time (t(51) = -2.82, p = .007) such that Anti trials (M = 7650.58ms, SD = 1697.25) were faster than Regular trials (M = 8174.13ms, SD = 1136.36). A similar pattern of results was found for the Pro block. There was no significant difference in accuracy between Pro and Regular trials (t(51) = .009, p = .99), but there was a significant difference in reaction time (t(51) = -5.80, p < .001). Pro trials (M = 6611.56ms, SD = 2003.63) were faster than Regular trials (M = 7908.15ms, SD = 1360.48).

5.3.3 Changes across blocks in performance by trial type

A repeated-measures ANOVA failed to find a significant difference among accuracies for Regular trials across blocks (F(2.30,117.34) = 2.27, p = .10). There was also no significant difference in accuracy for Anti trials across blocks (t(51) = 0.41, p = .68). However, there was a significant difference in accuracy for Pro trials across blocks (t(51) = 2.94, p = .005); Pro trials in Pro blocks were less accurate (M = 74.52%, SD = 19.52) than Pro trials in Mixed blocks (M = 81.69%, SD = 13.97). For reaction time, again there was no significance difference for Regular trials dependent on block (F(2.74,139.66) = 2.18, p = .10). Comparisons for Anti trials (t(51) = .896, p = .37) and Pro trials (t(51) = .12, p = .91) also lacked significant differences in reaction time dependent on block.

For the next analyses, performance Regular and Anti trials were collapsed across blocks, respectively. Pro trials, due to the significant difference in accuracy depending on block, remained separated by block (Pro vs. Mixed). The comparison of accuracy depending on trial type revealed a significant difference (F(2.14,109.15) = 5.36, p = .005). Post-hoc comparisons found that Pro trials in Mixed blocks were more accurate than all other trial types (all p < .05; Table 7). The comparison of reaction time depending on trial type also revealed a significant difference (F(2.03,103.60) = 28.99, p < .001). Post-hoc comparisons found that both Pro trial types (in Pro blocks and Mixed blocks; p =1.00) were faster than Anti trials, which in turn were faster than Regular trials (all p < .01; Table 7).

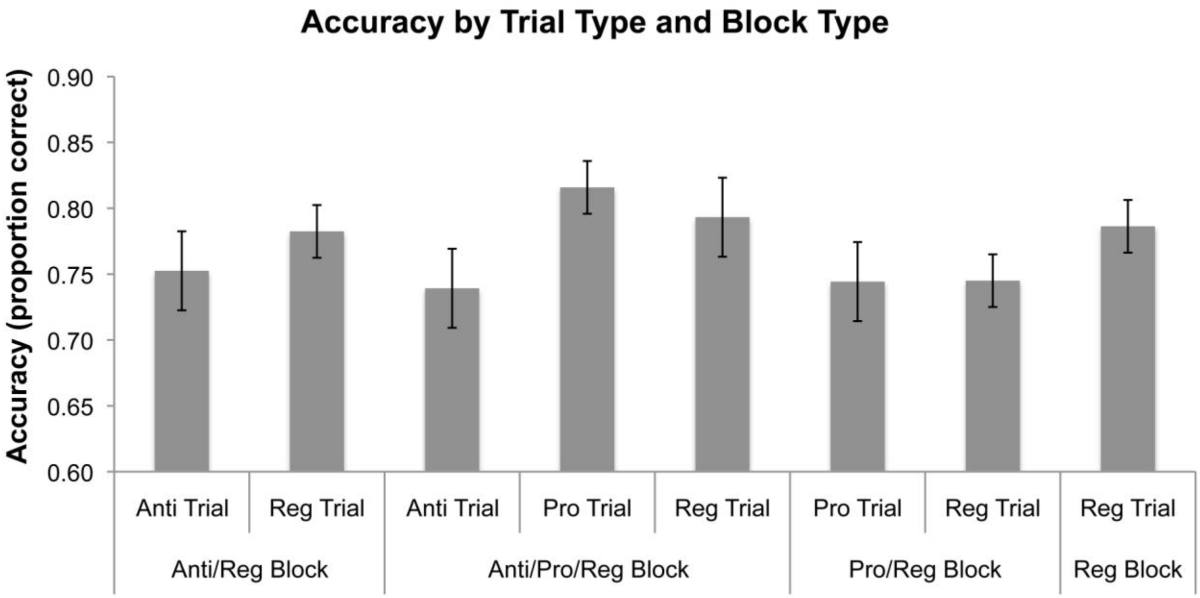


Figure 23. Accuracy by trial type and block type.

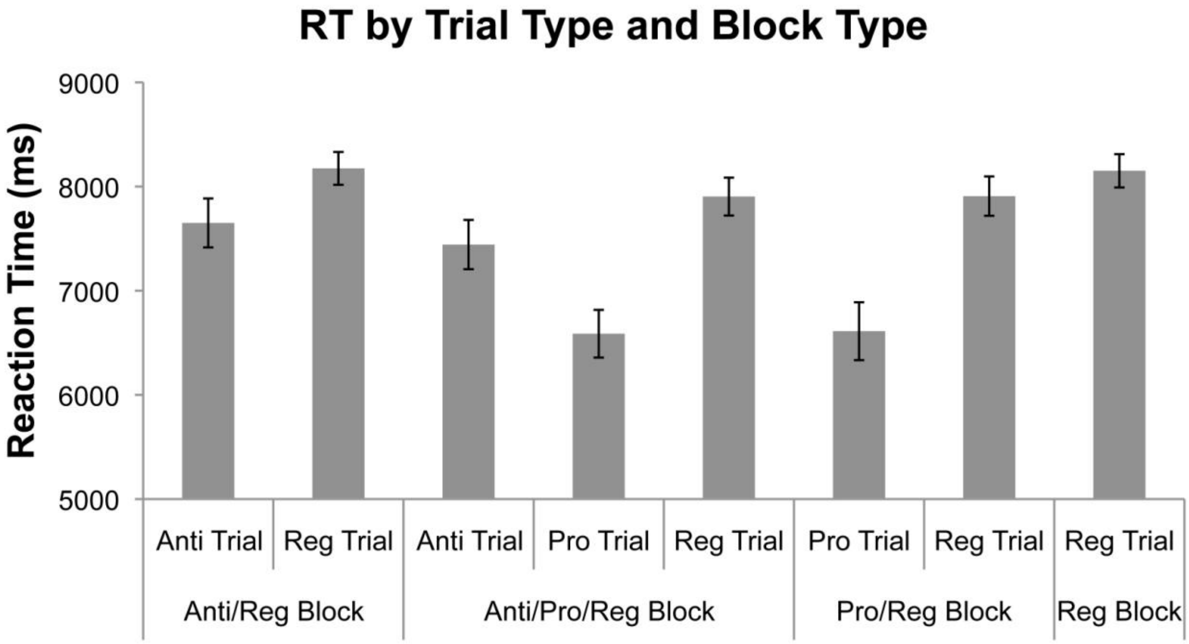


Figure 24. Reaction time by trial type and block type.

	All Anti Trials	All Regular Trials	Pro Trials, Mixed Block	Pro Trials, Pro Block
Accuracy (SE)	74.6% (2.5)	77.7% (1.7)	81.7% (1.9)	74.5% (2.7)
Reaction Time (SE)	7546.90 (205.62)	8034.36 (148.92)	6587.19 (229.05)	6611.56 (277.75)

Table 7. Average accuracy and reaction times for Anti trials collapsed across blocks, Regular trials collapsed across blocks, and the two Pro trial types separated by block.

5.3.4 Changes reaction time by accuracy and trial type

For each of the following analyses, trials for which there was an invalid response (e.g., no response, response during fixation) were discarded. The remaining trials were sorted within each block by trial type and accuracy and entered into an ANOVA (or t-test in the case of the Regular block) to test effects of accuracy on reaction time dependent on trial type (Figure 25). Post-hoc pairwise comparisons were made when appropriate.

5.3.4.1 Regular Block

A paired samples t-test comparing reaction times for correct versus incorrect trials was significant ($t(50) = 2.40$, $p = .02$). When no bursts were present at any time during the block, incorrect trials were faster than correct trials. This suggests that the errors may arise through a faulty decision process that triggers the decision before enough evidence is gathered.

5.3.4.2 Anti Block

A 2x2 repeated-measures ANOVA tested the effects of accuracy and trial type (anti-burst versus regular) on reaction time. While there was no difference in reaction time dependent on trial type ($F(1,40) = 1.22$, $p = .28$), there was a main effect of accuracy ($F(1,40) = 8.18$, $p < .01$). There was not a significant interaction ($F(1,40) = .10$, $p = .75$), demonstrating that incorrect trials were faster than correct trials without a dependence on trial type. This matches the results of regular block.

5.3.4.3 Mixed Block

A 2x3 repeated-measures ANOVA tested the effects of accuracy and trial type (anti-burst, regular, and pro-burst) on reaction time. As previously reported, the main effect of trial type was significant ($F(2,74) = 11.33$, $p < .01$) and post-hoc pairwise comparisons demonstrate that pro-burst trials were faster than both regular and anti-burst trials, but there was no difference between anti-burst and regular trials. In addition, there was a main effect of accuracy on reaction time ($F(1,37) = 9.67$, $p < .01$), such that incorrect trials were faster than correct trials. There was not a significant interaction ($F(2,74) = .265$, $p = .77$), illustrating that changes in reaction time dependence on performance were independent of changes in reaction time due to burst status.

5.3.4.4 Pro Block

A 2x2 repeated-measures ANOVA tested the effects of accuracy and trial type (pro-burst and regular) on reaction time. As previously reported, the main effect of trial type was significant ($F(1,45) = 15.05$, $p < .001$), with pro-burst trials resulting in faster times than regular trials. In addition, the main effect of accuracy was also significant ($F(1,45) = 21.99$, $p < .001$), continuing to demonstrate faster reaction times for incorrect trials than correct trials. Again, there was no interaction between the two factors ($F(1,45) = .58$, $p = .45$).

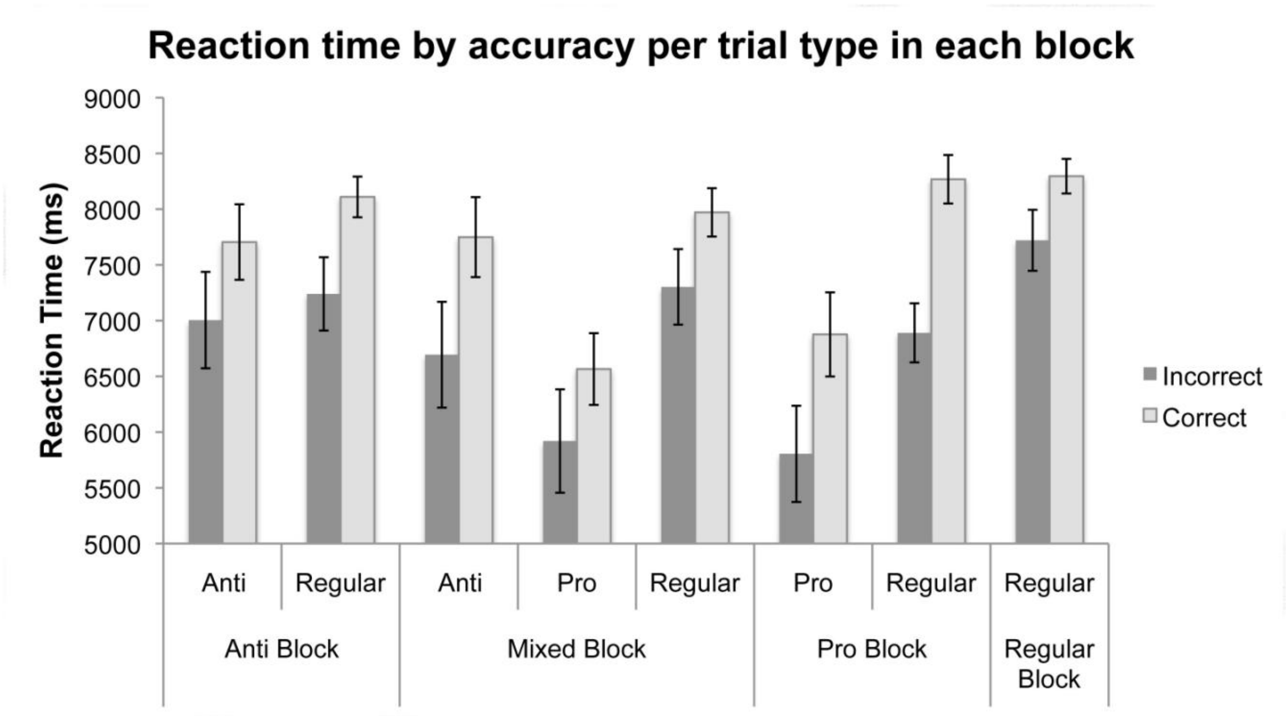


Figure 25. Reaction time by accuracy per trial type in each block.

5.4 DISCUSSION

Overall accuracy was the same regardless of trial type or block status aside from the Pro trials in the Mixed block that demonstrated an aberrant increase in performance. This suggests that the participants maintained the same decision threshold throughout the experiment regardless of trial type or the mix of trials per block. However, despite the similar threshold leading to similar accuracy, there were several differences in reaction time depending on both trial type and block status. Although reaction times were consistently faster for incorrect trials than correct trials, an interaction between this effect and trial status was not found, indicating that effects to the decision process did not differentially affect accuracy depending on the presence or absence of a burst. Taken together, the change in reaction time despite the consistent performance level indicates that the journey through the decision process was altered depending on the trial, not the ultimate threshold executing a decision. This agrees with the fluctuations in drift rate but steady threshold used in diffusion models (Ratcliff, 1978).

For example, in Mixed blocks that included Regular trials as well as trials with bursts both in the same direction of motion (Pro) and against the direction of motion (Anti), reaction time changed according to the validity of evidence in the trial. Pro trials, with a burst of highly coherent motion in the same direction as the rest of the trial, resulted in faster decisions than Regular trials. In contrast, Anti trials presented evidence against the rest of the trial and resulted in slower decisions than Regular trials. These results uphold the hypothesis that bursts of motion (i.e., momentary increases in the fidelity of the stimulus) would influence the speed of the perceptual decision without affecting the accuracy. Although the data has not been modeled, it is useful to consider what changes to the three basic diffusion model parameters (start point, drift rate, and threshold) would create this pattern of data. It is unlikely that the start point was shifting, as the direction of motion could be in one of eight directions on any given trial. Shifting the start point would make it harder to reach threshold on most trials (i.e., 7/8 of trials on average) regardless of burst status. It is also unlikely that the threshold changed between trials, as the accuracy remained that same regardless of burst status. This suggests that participants required the same amount of evidence for each decision, whether or not there was a burst adding to that overall information. This leaves open the possibility of changes to the drift rate. Strong Pro evidence would increase the rate; strong Anti evidence would decrease the rate. This altered rate would then take less or more time to reach the common threshold causing the reaction time patterns seen here, and would also mimic the neural firing patterns previously reported (Huk & Shadlen, 2005). It also refutes the notion that participants ignore evidence counter to their inclination (Mynatt, Doherty, & Tweney, 1978), in which case we would have expected reaction times for Anti trials to be similar to Regular trials.

Although the results in the Mixed block appear to fully support the hypothesis about how highly coherent bits of information influence the overall decision process, there were also results that did not necessarily agree with this suggestion. During the Anti block, Anti trials were expected to be slower than Regular trials due to the influence the burst should have on the drift rate; however, the results indicate that Anti trials were actually faster than Regular trials. Despite the best efforts to make the burst of coherent motion consciously imperceptible to participants, it is possible that there were some who noticed the change in trial structure. For those who were able to both recognize the existence of the burst and identify the direction of motion during the burst (50ms in duration), it is possible they adopted different strategies depending on the mix of trials included in the block. For blocks in which the burst always agreed with the final decision (e.g., the Pro block), participants could rely on the burst and make a quick answer. For blocks in which the burst always disagreed with the final decision (e.g., the Anti block), participants could again rely on the burst to indicate the answer – only in the opposite direction of the burst motion. However, in Mixed blocks the burst does not reliably indicate the direction of motion, so participants can only partially rely on the burst evidence. Because the participants were naïve to which trials would be included in any particular block, this change in strategy would have to develop of the course of a the first few trials in each block. Therefore it will be important to test the trial-by-trial fluctuations in performance to see if there are markers for strategy change. This will be addressed in the next chapter, which examines how strategy may develop over the beginning of each block and how the status of a previous trial (e.g., burst present or absent, performance, speed of decision) may influence performance on the next trial.

6.0 ONGOING ADJUSTMENTS TO THE DECISION PROCESS

6.1 INTRODUCTION

The previous four chapters have discussed results from experiments that are averaged over whole blocks or experiments. In fact, with notable exceptions in which single trials of neural firing are assessed and responses are predicted after a model has been built (Kiani, Hanks, & Shadlen, 2008) or single trial simulated data using a model based on averaged real data (Mazurek, Roitman, Ditterich, & Shadlen, 2003), research on perceptual decisions appears to always average over blocks of trials. This assumes that the decision process in any given experiment is static from beginning to end. In addition, the model created by Mazurek and colleagues (2003) used a stationary threshold based on the best fit to the averaged behavioral data. This approach, while useful for the model, likely fails to capture some of the ongoing fluctuations present in the process; in fact, the previous chapter on the burst experiment found differences in performance depending on the mix of trials in a block, indicating some potential shifts in strategy throughout the experiment. This is not surprising given the shifts in strategy and use of heuristics previously found during multiple attribute decisions (Payne, Bettmann, & Johnson, 1988). When confronted with changing task demands and uncertain stimuli, it is likely more efficient to adopt a variable strategy depending on growing knowledge of the overall task structure through recent trial history. In addition, effort and motivation, as affected by previous trial history, may also affect each new decision (Thomas, 1983).

As discussed in 2.0 , the perceptual stimulus used here does not afford a large amount of variability. While multiple attribute decisions may require strategy adoption for efficient and efficacious selection of a response alternative, it is not clear that this is true for simple perceptual decisions as studied here. However, given the number of parameters often used to model the decision process (e.g., start point, drift rate, threshold; Ratcliff, 1978), it seems naïve to assume these parameters remain static. Indeed, some early models involved presampling of the problem space in such a way that there were trial-to-trial correlations of response time (Laming, 1979). Yet it is important to remember that parameters in a mathematical model are designed to account for the data, and that does not necessarily mean those parameters are present in the neuronal pools underlying the responses. In fact, as mentioned in the introduction, studies of temporal and spatial attention (potentially more similar to the tasks used in the current set of studies than the multiple attribute studies of real estate, etc.) have not shown trial-by-trial variability in responses (Vul, Hanus, & Kanwisher, 2009), nor have some studies of choice reaction time for auditory stimuli (Green, Smith, & von Gierke, 1983).

The previous literature sets up an interesting contrast between instances of strategy adoption and studies without across-trial variability. Due to the experimental designs here, participants were usually aware of the manipulation and experienced more than one condition in each block (with the exception of the increasing choices experiment in 3.0). To assess how these theories may apply to perceptual decisions, this chapter will assess potential trial-by-trial fluctuations in performance depending on previous trial status (e.g., burst or non-burst, difficulty of pairing) and previous trial performance (e.g., correct vs. incorrect).

6.2 ADJUSTMENTS BASED ON MANIPULATIONS

As mentioned above, three out of the four main experiments presented here included mixes of trial types within each block (easy vs. difficult pairs of alternatives in 2.0 , early vs. later presentations of motion information in 4.0 , and pro- or anti-informative bursts in 5.0). Although the participants are not explicitly told the composition of each block, through experience they could develop a sense of the structure (with potentially the exception of the burst experiment which was designed to avoid detection by the majority of participants). If the participants understood the mixture of trials in each block, they may have used this knowledge to adjust their strategy on a trial-by-trial basis. In a sense, participants may have gambled on what the next trial would be, and thereby adjust their threshold for response higher (for predicted difficult trials) or lower (for predicted easy trials). Changes of the threshold in this way would result in changes in reaction time. The following section includes analyses of reaction time for the similarity, timing of evidence, and burst experiments based on the status of the previous trial based on the appropriate manipulation. If participants were adjusting their threshold throughout the experiment, we would expect to see shifts in reaction time as a result.

6.2.1 Similarity Experiment

One interesting aspect of the experiment in 2.0 is that participants experienced a mix of easier (e.g., 67.5° and 90° comparisons) and more difficult (e.g., 45° and 22.5°) comparisons. This embedded nature of the manipulation may cause some of the trial-by-trial changes in strategies based on effort (Thomas, 1983). After participants spend time on a difficult decision, they may be more patient with the decision process in the following decision. Conversely, an easy decision may lead to more brash responses due to overconfidence.

To assess this, trials were organized based on current difficulty and the difficulty of the previous trial, resulting in 16 categories (4 levels of difficulty for both current and previous trial status). The mean reaction time was then calculated for each participant in each of the 16 categories. These were then analyzed in a repeated-measures ANOVA using Greenhouse-Geisser corrections when appropriate. As described in 2.0 , the main effect of current trial status was significant ($F(2.17,78.24) = 28.23$, $p < .001$), confirming that reaction time did change dependent on trial difficulty. Difficult trials (22.5° in difference between the two response options) had longer reaction times than easier trials (90° in difference between options). However, the main effect of previous trial difficulty was not significant ($F(2.32,83.68) = 2.36$, $p = .09$), suggesting that the difficulty of the immediately preceding trial did not affect the length of the decision process on the current trial. In addition, the interaction between current trial and previous trial difficulty was also not significant ($F(6.47,232.81) = .74$, $p = .63$). This suggests that participants did not adjust their strategy towards the current trial based on the previous trial differently for harder or easier current trials. For example, even when faced with a difficult trial, participants did not alter their decision process based on previous trial status.

6.2.2 Timing of Evidence Experiment

The manipulation in 4.0 was also embedded. Participants responded to a mix of trials in which the motion was presented at different times such that they may see motion in the first 2000ms (the 1st bin) on trial 1, but in the last 2000ms (5th bin) of trial 2. As was suggested in 4.0 , participants did not learn to wait longer in order to respond during the motion on trials with motion in the 4th and 5th bins. However, this failure to learn is surprising considering they were forced to watch each whole trial, regardless of when they responded. Over the course of the experiment they would have seen many instances (1/5 of the total number of trials) in which motion was not presented until the last 2000ms of the trial. While not reflected in the average reaction time, it is possible that after watching a 5th bin trial, participants would be more patient on the following trial (i.e., have a longer reaction time).

To assess changes in strategy, each trial was coded with its timing of motion and the status of motion in the previous trial (e.g., motion in the 1st bin, 2nd bin, etc.). This created a 5 (current status) x 5 (previous status) table for each participant with the mean reaction time calculated for each cell. These values were entered into a repeated-measures ANOVA to test for changes in strategy depending on the previous trial status. If participants demonstrate some change in strategy depending on the ongoing manipulations, we would expect trials presented after one in which motion was late (e.g., the 4th and 5th conditions) would have longer reaction times than trials presented after one in which motion was early. Unfortunately, this does not appear to be the case. While average reaction time for the current trial was significant as demonstrated in 4.0 ($F(2.82,81.83) = 27.79$, $p < .001$), reaction did not change dependent on previous trial status ($F(3.22,93.31) = .414$, $p = .79$). Although we may expect changes in reaction time only after certain types of trials (e.g., the 4th and 5th conditions as specified above), the interaction term was also not significant ($F(8.78,254.54) = 1.13$, $p = .32$), rejecting this hypothesis as well. Participants were not shifting their decision process to allow for more data processing even after watching examples in which motion was presented late and in the absence of a benefit for early responding. This continues to suggest that participants may cease to allocate resources to an ongoing decision to use those resources elsewhere (Kiani, Hanks, & Shadlen, 2008), even when there is knowledge that more information is forthcoming.

6.2.3 Burst Experiment

Similar to the two previous manipulations, the burst manipulation was also presented in a mix of trials. The specific mixture of trials depended on block; Pro blocks had Pro and Regular trials, Anti blocks had Anti and Regular trials, and Mixed blocks had all three trial types. 5.0 describes how participants appeared to be most affected by the burst manipulation during the Mixed block, resulting in changes in accuracy and reaction time depending on trial type (both $p < .01$). This is not surprising, as the Mixed block included all three trial types, thus making the burst unreliable as a source of information either for or against the decision. To assess whether or not the status of the previous trial had an effect on performance on the current trial, data from the Mixed block were divided according to current and previous trial status (3 x 3 table). Due to this increased division of the data, some participants had to be excluded for low trial counts in particular category based on the combination of random sampling of the trials and invalid responses (e.g., responses during fixation, double responses, no response). The following analyses were conducted on a subset of participants who had sufficient trials (>3) in each cell (N = 38).

If participants are changing their strategy during the Mixed block depending on which trials they have seen we would expect to see increases in speed after Pro trials and decreases in speed after Anti trials. These changes would reflect shifts in confidence relating to performance and the fidelity of the burst towards the correct solution. However, a repeated-measures ANOVA failed to demonstrate the main effect of previous trial type ($F(1.88,69.76) = .504$, $p = .61$; Greenhouse-Geisser corrected). In addition, the interaction with current trial status was also not significant (F(3.39,125.38) = .713, p =.584). This suggests that although accuracy and reaction time changed dependent on current trial status, it was not due to the status of the previous trial. Participants were not relying on the presence and fidelity of the burst during the previous trial to adjust their strategy towards the current trial.

6.3 ADJUSTMENTS BASED ON PREVIOUS TRIAL ACCURACY

Although status of the previous trial did not affect any of the three experiments analyzed above, this does not preclude adjustments to threshold on the part of the participant. Thomas (1983) suggested that effort and motivation play a role in the adoption of a strategy. As seen in 5.0 , the Mixed blocks that included Pro, Anti, and Regular trials led to the most changes in reaction time, possibly indicating true changes drift rate. In addition, Anti trials caused significantly lower accuracy than Regular or Pro trials, indicating that the burst was effective in disrupting the decision process. Due to the nature of the response structure for the burst experiment, participants saw a highly coherent motion stimulus at the end of the trial and were required to respond if it matched their guess earlier in the trial. This end of trial verification could act as feedback to the participants; it forced them to identify their accuracy, which could in turn lead to changes in strategy on the next trial due to motivation to perform well. The following analysis assesses whether or not reaction time changes based on the accuracy of the previous trial.

6.3.1 Burst Experiment

As participants moved through the experiment, they verified each response at the end of the trial. This consideration of performance may have led to changes in threshold to maximize performance on the following trial. If the previous trial was correct, participants may have experienced a boost in confidence in their motion discrimination skills. This, in turn, could have led to a lowering of threshold for faster responses. On the other hand, incorrect performance would have lowered confidence and led to raising of the threshold to increase accuracy on the succeeding trial.

To assess this, each trial was labeled with the accuracy of the previous trial. After this labeling, trials that had invalid responses (at fixation, double responses, and no response) were removed from the data due to a lack of a valid response time. Data was also separated based on block type (Anti, Mixed, Pro, Regular), as the results in 5.0 suggest there were differences in performance dependent on block. The average reaction time data was then calculated for each participant based on block and previous trial performance (4 x 2 design). Nineteen participants had to be excluded due to a high number of invalid responses that caused insufficient trials in one or more cells. Data from the remaining 33 participants was entered into a repeated-measures ANOVA. The main effect of block on reaction time was again significant (F(3,96) = 5.92, p = .001). However, the main effect of previous trial accuracy on reaction time was not significant (F(1,32) = .21, p = .65). Interestingly, though, the interaction between block and previous trial accuracy was significant (F(3,96) = 3.53, p = .02). Post-hoc t-tests indicated marginal, though opposite, effects in the Anti (t(32) = 1.81, p = .08) and Mixed (t(32) = -1.75, p = .09) blocks. There was no effect for the Pro (p = .13) and Regular (p = .85) blocks. In the Anti blocks reaction time decreased after correct trials compared to incorrect trials. Conversely, in the Mixed blocks reaction time increased after correct trials compared to incorrect trials (Figure 26). These results suggest that there are subtle changes in threshold when participants are forced to verify their accuracy. This verification stage is likely acting as feedback, causing adjustments to performance under the general motivation to perform well. However, the reverse effects between the Anti and Mixed blocks indicate a more intricate story potentially based on previous trial status (e.g., Anti or Regular trial in Anti blocks; Anti, Pro, or Regular trial in Mixed blocks). Unfortunately, further dividing the data by previous trial type to tease apart these contrary effects results in too few participants that survive the minimum trial count to get reliable results.

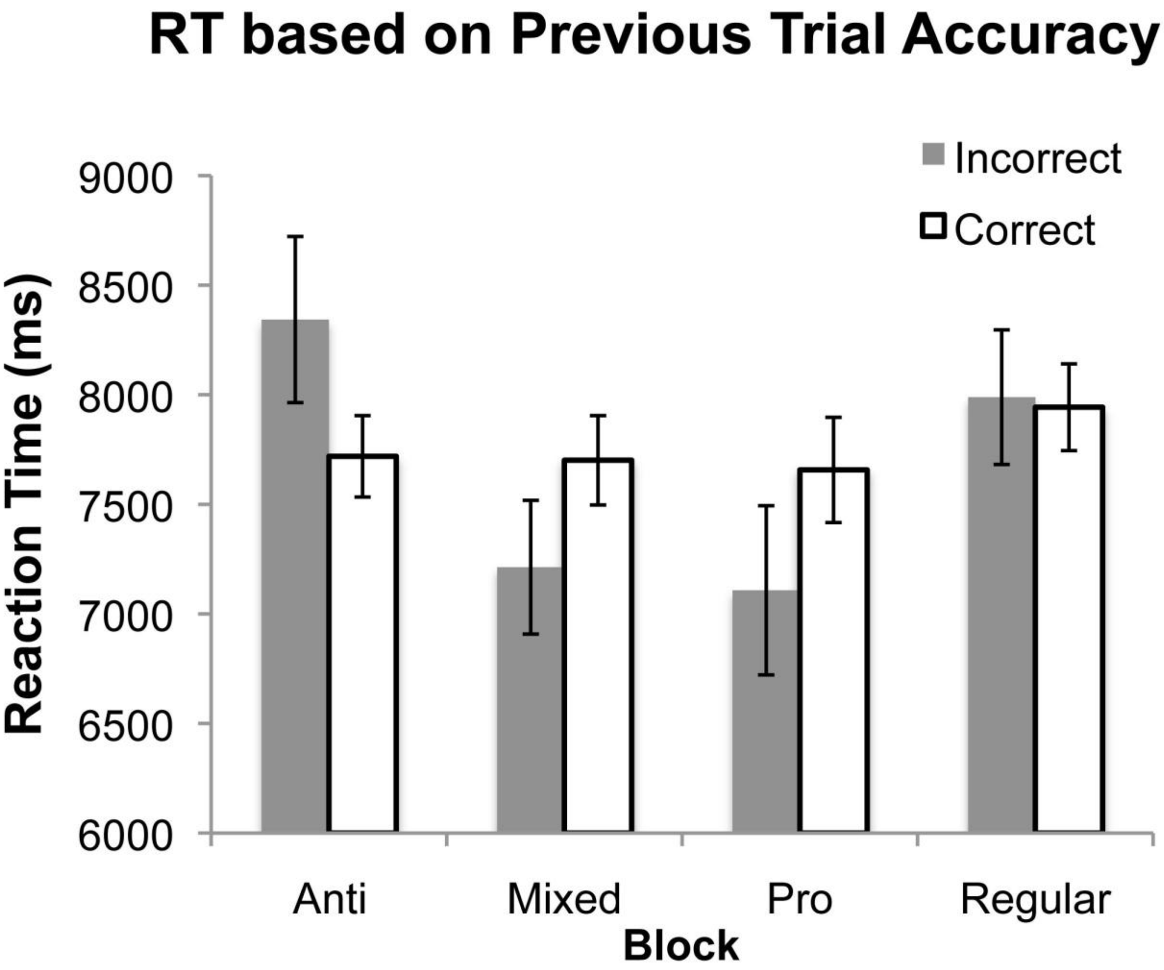


Figure 26. Average reaction time based on previous trial accuracy and block status.

6.4 DISCUSSION

This chapter contains two new analyses of the data from the various motion discrimination experiments described throughout this document. First, data from the similarity, timing of evidence, and burst experiments were reanalyzed to check for changes in reaction time based on the manipulation status of the previous trial (e.g., whether or not a burst was present). Previous research has suggested that participants may change their decision process depending on task structure, effort, and motivation (Payne, Bettman, & Johnson, 1988; Thomas, 1983). Each of the manipulations in these experiments may cause changes in effort (e.g., very similar vs. far apart options in the similarity experiment or early presentation of motion vs. late presentation of motion in the timing of evidence experiment). In addition, the burst manipulation may cause changes in motivation if participants noticed the manipulation (which may have caused some of the reaction time results as discussed in 5.0). However, the results from the analysis conducted here do not demonstrate the changes in reaction time that would be expected due to shifts in threshold based on strategy adoption. Instead, the results agree with those of Vul, Hanus, & Kanwisher (2009) and Green, Smith, & von Gierke (1983) which also failed to show trial-to-trial variability.

The second analysis in this chapter assessed whether or not feedback from the preceding trial regarding accuracy would cause shifts in threshold, and thus reaction time, on the next trial. In theory, correct performance would cause increases in confidence and associated decreases in threshold. Conversely, incorrect performance would cause decreases in confidence and increases in threshold. A strategy along these lines would optimize performance for both speed (in high confidence cases) and accuracy (in low confidence cases). Although there were no changes in reaction time based on previous trial accuracy with the Pro and Regular blocks, there were marginal effects in the Anti and Mixed blocks. Further inspection of these effects found an interesting contrast: the predicted effect (faster responses after correct trials, slower responses after incorrect trials) did occur in Anti blocks, but the opposite occurred in Mixed blocks. These results suggest that there was some trial-to-trial variability in the decision process, potentially in the threshold of activity necessary to execute a decision. Due to the paradoxical nature of the results, more experimentation is necessary to tease apart the underlying effects. It would be particularly interesting to conduct the burst experiment with only the Mixed block composition in order to have enough trials to test if and how previous trial status and accuracy may be interacting with one another.

7.0 GENERAL DISCUSSION

This dissertation sought to explore the perceptual decision process using a systematic approach in an effort to bridge and extend previous behavioral, neurophysiological, and neuroimaging work. By isolating manipulations but maintaining the same stimulus and general task structure, the results of each individual experiment can be discussed as a whole, an approach that has been lacking in the prior behavioral literature. In addition, this dissertation gives a broad account of how perceptual decisions are affected by minor changes in the decision environment. Building upon previous approaches using the random dot motion task, new manipulations of the similarity between response alternatives, number of response alternatives, timing of incoming evidence, and fidelity of that evidence add to the knowledge about how perceptual decisions behave. Importantly, these manipulations take steps towards a better understanding of perceptual decisions under more naturalistic conditions. Finally, consideration of results from prior work in both the perceptual decision and multiple attribute decision research was intended to push towards defining how the perceptual decision process emulates the general decision process used for more complicated situations like multiple attribute decisions. In sum, these four experiments and the trial-by-trial analysis determined that while perceptual decisions do often behave like multiple attribute decisions, this is not always the case, and more appreciation of the differences between these two processes is necessary.

7.1 SUMMARY OF FINDINGS

The first experiment (2.0) involved a perceptual decision between two specified directions of motion. The manipulation of interest was how similar the options were on each trial: 22.5°, 45°, 67.5°, or 90° apart. Although the main purpose of this experiment was to confirm that the 45° angle of disparity between options in the remaining three experiments was not going to retard performance in addition to the specific manipulations used in each, it also highlighted how similarity between options may affect performance. As the angle of disparity between options lessened from 45° to 22.5°, accuracy decreased and reaction time increased. This is likely due to the activation of neurons in motion-selective area MT (medial temporal lobe, or the human homologue in this case) that are sensitive to both directions of motion (Britten & Newsome, 1998; Butts & Goldman, 2006). A previous behavioral example of this type of interference was the effects on accuracy when comparing similar (e.g., E vs. F) vs. dissimilar (e.g., E vs. C) letter pairs (Ratcliff & Rouder, 2000). The first pair consist of almost identical lines with the exception of the bottom horizontal line for “F” compared to “E”; on the other hand, with only minor overlap in the upper and lower horizontal zones and the left vertical zone, “E” and “C” do not share as many features. This would alter how many neurons are sensitive to both letters in the pair. If the decision process in the brain makes a comparison between pools of neurons sensitive to the two choices and selects the more active pool, the activity levels in the “E” and “F” comparison will be closer than those in the “E” and “C” comparison, thus making it harder to identify the highest level of activity. In terms of motion selection, the superior colliculus has been identified as a potential monitor of the lateral intraparietal area (LIP) and MT (Lo & Wang, 2006) and might be involved in the comparison required in the current experiment.

The second experiment (3.0) explored how the addition of alternatives would affect the decision process. Initial work into perceptual detection (e.g., stimulus-response paradigms) found that as the number of potential stimuli increases, so does response time (Hyman, 1953). Limited work on perceptual decisions has shown some detriment to performance when alternatives increase from the traditional 2-choice paradigm up to 4 alternatives (Churchland, Kiani, & Shadlen, 2008; Leite & Ratcliff, 2010). The current experiment sought to extend this further by testing up to 8-choices; it was predicted that participants may adopt strategies to limit the problem space when given more than 4 choices to vet at once, as has been seen in more complicated multiple attribute decision experiments (Tversky & Sattath, 1979; Einhorn & Hogarth, 1981; Payne, Bettman, & Johnson, 1988). The results demonstrated an interesting two-stage change in performance. As the number of alternatives increased from 2 to 4 choices there was an associated decrease in accuracy. However, as the alternatives increased further to 8 choices there was both a greater drop in accuracy and an increase in reaction time. An analysis of the angular disparity of the errors found that these changes were present over and above any potential problems with motor execution as the response buttons were closer together. This suggests that it is unlikely participants were adopting a strategy to compensate for the increased number of alternatives available. In addition, the change from 4 to 8 response alternatives may have also been affected by the unavoidable increase of similarity between the alternatives, mimicking the changes in accuracy and reaction time seen in Chapter 2.

The third experiment (4.0) examined how the timing of informative evidence can affect the decision process. We are not usually in control of when the environment presents us with the perceptual information we must discriminate, so it seems unusual to use tasks in which the start of trial is also indicative of when information will be forthcoming. When the coherent motion was moved around in time, the results illuminated some interesting changes in performance. First, presenting motion soon after the start of the trial resulted in higher accuracy, but also slower decisions. Participants were more willing to wait and consider the evidence when it was delivered promptly. Conversely, evidence presented later in the trial resulted in less accurate, faster decisions. In addition, as they waited for informative evidence to be presented, the longer the wait before the presentation of coherent motion, the more likely participants would false alarm despite the lack of evidence available. Interestingly, there appeared to be a focus of when these false alarms were likely to occur that was not quite 2/3 of the way through the trial length even though participants were required to watch the whole trial regardless of response. Previous research has found similar results during neurophysiological recording of perceptual decisions in macaques and suggested that there may be motivation to reserve mental resources and thus truncate ongoing unsuccessful decisions (Kiani, Hanks, & Shadlen, 2008).

The final experiment (5.0) explored how insertions of high quality evidence early in the decision could influence the impending response. Previous neurophysiological studies have demonstrated that bursts of information can change the rate at which activity in certain brain areas (e.g., lateral intraparietal cortex) accrues, potentially altering the timing of the decision (Huk & Shadlen, 2005). The implementation of the burst early in the trial was of particular interest given the results of the third experiment (4.0), which suggested that the decision process maintains integrity and is harder to terminate early in the trial compared to later. If the process maintains a high level of fidelity early in the trial, extra information should only influence the process and not cause the execution of the response. However, the relevance of the information (either towards or against the correct response) should change the speed of the response. This hypothesis was only partially upheld, as the results told a more complicated story. Although trials with high coherent bursts in the direction of the correct response did result in faster decisions than trials with no burst, there was not a slowing of response when the burst was against the correct direction. In addition, the effects were most strongly demonstrated when all three trial types (Pro, Anti, and Regular) were mixed into the same block, rather than only two types. This suggests that despite best efforts and piloting of the experiment, some participants were still able to see the burst and use it to their advantage even when it was against the correct direction (though not perfectly: Anti trials resulted in lower accuracy). The shifts in results based on the composition of each block may highlight potential strategies on the part of the participants to limit the resources and time necessary to complete the decision (Tversky & Sattath, 1979; Einhorn & Hogarth, 1981; Payne, Bettman, & Johnson, 1988; Kiani, Hanks, & Shadlen, 2008).

Finally, an additional analysis (6.0) assessed whether or not trial-to-trial variability in reaction time occurred depending on previous trial status due to manipulation changes (similarity, timing of evidence, and burst experiments) or previous trial accuracy when feedback was available (burst experiment). While there has been some evidence of shifts in performance during multiple attribute decisions (Thomas, 1983; Payne, Bettman, & Johnson, 1988), trial-to-trial shifts have not been found in previous spatial attention (Vul, Hanus, & Kanwisher, 2009) and auditory choice (Green, Smith, & von Gierke, 1983) experiments. The results of the analysis failed to indicate trial-to-trial variability in reaction time based on previous trial status in the similarity experiment (close options vs. far apart options), timing of evidence (early vs. late presentation of motion), or burst (no burst vs. pro burst vs. anti burst) experiments. However, the second analysis regarding previous trial accuracy did highlight some changes in reaction time during the burst experiment in which participants were required to verify the accuracy of their response at the end of the trial. This verification stage forced participants to create their own feedback on their progress, which in turn appeared to change performance in the Anti and Mixed blocks. However, as noted in 6.0 , the resulting changes in reaction time were paradoxical. Performance was faster following correct trials than incorrect trials during Anti blocks, suggesting that correct trials increased confidence and lead to lower thresholds on succeeding trials. Interestingly, though, the reverse was true during Mixed blocks; performance was slower following correct trials than incorrect trials.

7.2 RELATIONSHIP TO PREVIOUS LITERATURE

It is often stated that the goal of perceptual decision research is to understand the complicated decision process by studying it in a simplified state (Gold & Shadlen, 2007). However, the collective literature on perceptual decisions to this point lacks a strong connection to the findings and theories about multiple attribute decisions, nor does it use consistent methods that allow for cross-comparison and consensus building among experiments. The primary goal of this dissertation was to systematically test perceptual decisions with an eye towards both the previous perceptual decision research as well as research on more complicated decision problems. The result was an interesting concert of findings that both agrees with and disconfirms some of the theories behind multiple attribute decisions. For example, although the perceptual stimulus does not have as many degrees of freedom for manipulation as multiple attribute problems, there were still some effects on performance when response alternatives became more similar (2.0). This agrees with prior work on multiple attribute decisions and similarity of options (Biggs, Bedard, Gaber, & Linsmeier, 1985). On the other hand, participants failed to adopt strategies that increased efficiency and performance despite the increasing number of response alternatives entered into the decision process (3.0). This goes against previously demonstrated changes in performance dependent on the number of alternatives to be considered in any one decision (Tversky & Sattah, 1979; Einhorn & Hogarth, 1981). Lastly, the trial-by-trial analysis in 6.0 only partially demonstrated variability in performance; the observed changes were not due to effort or motivation (Thomas, 1983) or the experience of certain conditions (Payne, Bettman, & Johnson, 1988), but rather prior trial performance. Collectively these results indicate that the assertion that perceptual decisions are a simplified version of the general decision process may not be fully valid. Therefore future studies of perceptual decisions should be careful about broad claims regarding the overall decision process.

This dissertation also presents significant extensions of previous work on perceptual decisions. While there has been some recent interest in how the number of response alternatives affect performance (Churchland, Kiani, & Shadlen, 2008; Leite & Ratcliff, 2010), the majority of studies thus far have included binary alternatives, often from stimuli without a quantifiable spectrum of similarity. 3.0 demonstrated that performance in perceptual decisions decreases as the number of response alternatives increases, suggesting that future studies should consider how stimulus selection and task structure affect the process in addition to any particular manipulation of interest. Additionally, 4.0 investigated how the timing of evidence affects the decision process, another rarely considered task parameter (c.f., Kiani, Hanks, & Shadlen, 2008). Because we are not usually in charge of when the environment presents us with information to discriminate, it is important to consider how the timing of evidence compared to the start of the decision process may affect the overall decision. Interestingly, it appears that there may be a timeout function that terminates the decision process after a certain amount of time, potentially in order to save mental resources for future processing (Kiani, Hanks, & Shadlen, 2008). Although the idea of a timeout function is relatively new to discussion of the perceptual decision process, it has been previously discussed in regards to memory function and unsuccessful searches (Gillund & Shiffrin, 1984; Anderson, 2007).

7.3 FUTURE DIRECTIONS

The results presented here highlight two potentially important avenues of future study. First, the paradoxical results of the trial-by-trial analysis in 6.0 suggest an interesting story about the role of feedback during ongoing decision processes and how it might alter future decisions. Investigating this effect could lead to previously unconsidered implications of how feedback and/or reward (e.g., trial-by-trial juice rewards in neurophysiological studies) might affect the overall results. At the very least, it would be interesting to reanalyze the neurophysiological data to see if prior trial performance affected neural firing rates in a way that is complementary to the changes in reaction time demonstrated here. These changes may come in either shifts in baseline firing activity at the start of the trial or the level of activity necessary to execute a response shift depending on prior trial performance. Examining which of these two changes occur can help separate exactly what is causing the changes in performance. Changes in baseline firing rate at the start of the trial would indicate some sort of preparation prior to the consideration of evidence. On the other hand, a shift in the level of activity necessary would indicate the shifting of threshold as theorized here, potentially an indication of confidence in ability.

Second, the evidence for a potential timeout function described behaviorally here and neurally elsewhere (Kiani, Hanks, & Shadlen, 2008) highlights a little examined factor in the decision process. It will be important to further investigate how time and expectations of forthcoming evidence impact the overall process. For example, the perceptual decisions used thus far in the current projects and previous literature were generally low-cost tasks in which there may be no reward (as presented here) or very little reward (a few drops of juice for each correct trial in the case of the macaque neurophysiological literature). If a decision has a higher cost or gain associated with it, it is possible the decision process will maintain fidelity later in time in order to maximize performance. Consider the difference between seeing someone familiar in a crowd while walking in your hometown versus seeing that someone while lost in a foreign city. In the former case it may not make a difference whether or not you accurately recognize the person, as it is unlikely to alter your interaction with the environment. However, in the latter case you may strain to determine whether or not the person is an acquaintance in order to call out and gain some help you may not otherwise be able to secure.

7.4 CONCLUSIONS

The study of perceptual decisions as a simplified example of the general decision process continues to be a useful exertion. However, caution must be taken as to how generalizable the results truly are to the more complicated multiple attribute process. In addition, the use of rewards during the task may be changing the decision process in an ongoing fashion that has otherwise not been accounted for in past analyses. This has potentially important implications towards the validity of previous findings. Future work should continue to taken these points into consideration, as well as further explore the relationship between the timing of evidence and the decision process to more closely approximate how the process operates in vivo.

APPENDIX A

EXPERIMENT INSTRUCTIONS

This section contains the text for each experiment as presented in the instruction screens. Participants were first verbally instructed by the experimenter and then read through the screens at their own pace. The screens were intended to reiterate and reinforce the verbal instructions.

A.1.1 Opening welcome screen

The following was presented as the opening welcome screen for all experiments:

> Hello and welcome to the experiment.
>
> We are going to begin with some practice trials.
>
> You are going to see a series of images made up of moving dots.
>
> There is one basic task in this study described over the next few screens.
>
> Press the SPACE BAR for more instructions.

After this screen was presented, each experiment then had two screens of experiment-specific instructions, described in separate sections below.

A.1.2 Instructions for the perceptual break experiment (2.0)

First screen:

When the cross at the center of the screen is red for two seconds, the dots do not move together and only random motion is present. DO NOT press anything when the cross is red.

Focus on the period in which the cross turns white for 6 seconds. During this time some of the dots will move in the same direction together. It is your job to figure out which direction.

You should make a response during every trial (i.e., each time the cross turns white). Each trial lasts 6 seconds, so you do not have to respond immediately.

Press SPACE to continue.

Second screen:

When the cross turns white you will also see two possible directions labeled “1” and “0”. You must decide which of these options is the correct one and press the corresponding key. The practice trials will start off relatively easy and work down to the difficult level of the actual study trials. Also, you will notice that the “1” and “0” will move closer and farther apart across trials.

Please only respond once during each trial; note, nothing will change when you press the button. As long as you think you pressed it hard enough for it to register, you should be fine.

Keep your eyes focused on the cross in the center of the screen.

Press the SPACE BAR to try some practice trials

A.1.3 Instructions for the number of alternatives experiment (3.0)

First screen:

When the cross at the center of the screen is red for two seconds, the dots do not move together and only random motion is present. DO NOT press anything when the cross is red.

Focus on the period in which the cross turns white for 10 seconds. During this time some of the dots will move in the same direction together. It is your job to figure out which direction.

You should make a response during every trial (i.e., each time the cross turns white). Each trial lasts 10 seconds, so you do not have to respond immediately.

Press SPACE to continue.

Second screen (one button version):

When you are reasonably confident you know the direction of motion, press the SPACE bar. At the end of the trial, the dots will start to gather around the center. If they are going in the same direction as your guess, press the SPACE bar again. If they are going in a different direction, don’t do anything and wait for the cross to turn red.

Note, nothing will change when you press the button. As long as you think you pressed it hard enough for it to register, you should be fine.

Keep your eyes focused on the cross in the center of the screen.

Press the SPACE BAR to try some practice trials.

Second screen (definitive response version):

When you are reasonably confident you know the direction of motion, press the corresponding key on the number pad (e.g., 8 = up, 2 = down, 7 = diagonal up/left).

Please only respond once during each trial; note, nothing will change when you press the button. As long as you think you pressed it hard enough for it to register, you should be fine.

Keep your eyes focused on the cross in the center of the screen.

Press the SPACE BAR to try some practice trials.

A.1.4 Instructions for the timing of evidence experiment (4.0)

First screen:

When the cross at the center of the screen is red for two seconds, the dots do not move together and only random motion is present. DO NOT press anything when the cross is red.

Focus on the period in which the cross turns white for 10 seconds. During this time some of the dots will move in the same direction together for about 2 seconds. It is your job to figure out which direction.

You should make a response during every trial (i.e., each time the cross turns white). Each trial lasts 10 seconds, so you do not have to respond immediately.

Press SPACE to continue.

Second screen:

When you are reasonably confident you know the direction of motion, press the corresponding key on the number pad (e.g., 8 = up, 2 = down, 7 = diagonal up/left).

Please only respond once during each trial; note, nothing will change when you press the button. As long as you think you pressed it hard enough for it to register, you should be fine.

Keep your eyes focused on the cross in the center of the screen.

Press the SPACE BAR to try some practice trials.

A.1.5 Instructions for the burst experiment (5.0)

First screen:

When the cross at the center of the screen is red for two seconds, the dots do not move together and only random motion is present. DO NOT press anything when the cross is red.

Focus on the period in which the cross turns white for 16 seconds. The motion will be faint and hard to see at first, but will slowly grow into coherent motion in one of eight directions (the 4 cardinal directions and the diagonals).

You should make a response during every trial (i.e., each time the cross turns white). Each trial lasts 16 seconds, so you do not have to respond immediately.

Press SPACE to continue.

Second screen:

When you are reasonably confident you know the direction of motion, press the SPACE bar. The trial will continue and the motion will still get clearer.

When the motion is extremely obvious at the end of the trial, press the SPACE bar again if your original guess was CORRECT. Do not press anything if you were incorrect.

Please respond once during each trial and once at the end of the trial when the motion is obvious if you were CORRECT; note, nothing will change when you press the button. As long as you think you pressed it hard enough for it to register, you should be fine.

Keep your eyes focused on the cross in the center of the screen.

Press the SPACE BAR to try some practice trials.

APPENDIX B

STRATEGY ASSESSMENT

Increasing Choice Experiment

Subject # _________

Date _____________

1. Describe any strategy you had for the 2-choice blocks (either right vs. left OR up vs. down).

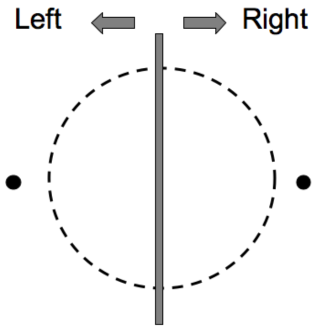
2. Describe any strategy you had for the 4-choice blocks (up/down/left/right OR the diagonals).

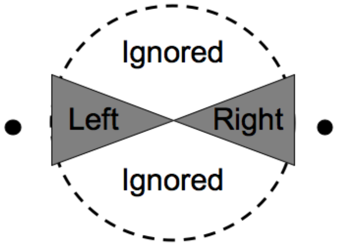
3. Describe any strategy you had for the 8-choice blocks.

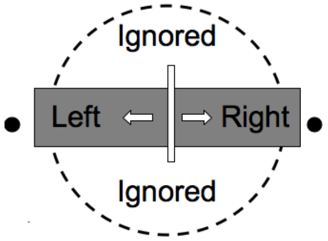
4. On average, how confident were you when you made your response (out of 100%)?

We suspect that people might be dividing the screen as part of their strategy. For each of the pictures below, give a percentage out of 100% for how similar it is to the way you were dividing up the screen.

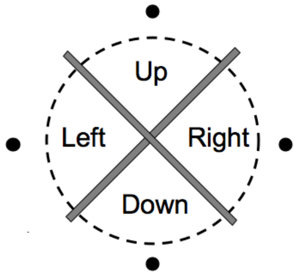
2-choice blocks

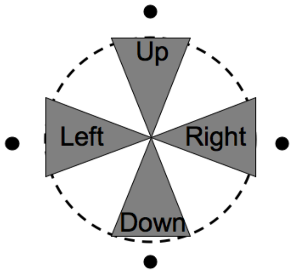






4-choice blocks





REFERENCES

Anderson, J.R. (2007). How can the human mind occur in the physical universe? New York: Oxford University Press.

Ball, K. & Sekuler, R. (1980). Models of stimulus uncertainty in motion perception. *Psychological Review, 87*(5), 435-469.

Biggs, S.F., Bedard, J.C., Gaber, B.G., & Linsmeier, T.J. (1985). The effects of task size and similarity on the decision behavior of bank loan officers. *Management Science, 31*(8), 970-987.

Britten, K.H. & Newsome, W.T. (1998). Tuning bandwidths for near-threshold stimuli in area MT. *Journal of Neurophysiology, 80*(2), 762-770.

Britten, K.H., Shadlen, M.N., Newsome, W.T., & Movshon, J.A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*(12), 4745-4765.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., & Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13,* 87-100.

Busemeyer, J.R. & Townsend, J.T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Butts, D.A. & Goldman, M.S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS Biology, 4*(4), e92.

Celebrini, S. & Newsome, W.T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *Journal of Neuroscience, 14,* 4109-4124.

Churchland, A.K., Kiani, R., & Shadlen, M.N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11*(6), 693-702.

Ebbesen, E.B. & Konecni, V.J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T.S. Wallsten (Ed.), Cognitive processes in choice and decision behavior. Hillsdale, N.J.: Lawrence Erlbaum Associates, 21-45.

Einhorn, H.J. & Hogarth, R.M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Reviews Psychology, 32,* 53-88.

Fitts, P.M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology, 71*(6), 849-857.

Fredericksen, J.R. & Kroll, J.F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 361-379.

Gillund, G. & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*(1), 1-67.

Gold, J.I. & Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*(1), 10-16.

Gold, J.I. & Shadlen, M.N. (2007). The neural basis of decision-making. *Annual Reviews Neuroscience, 30*, 535-574.

Green, D.M., Smith, A.F., & von Gierke, S.M. (1983). Choice reaction time with a random foreperiod. *Perception & Psychophysics, 34*, 195-208.

Hanks, T.D., Ditterich, J., Shadlen, M.N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neuroscience, 9*(5), 682-689.

Heekeren, H.R., Marrett, S., Bandettini, P.A., & Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431,* 859-862.

Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11-26.

Horwitz, G.D., Batista, A.P., & Newsome, W.T. (2004). Representation of an abstract perceptual decision in macaque superior colliculus. *Journal of Neurophysiology, 91*, 2281-2296.

Huk, A.C. & Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience, 25*, 10420-10436.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*, 188-196.

Kiani, R., Hanks, T.D., & Shadlen, M.N. (2006). When is enough enough? *Nature Neuroscience, 9*, 861-863.

Kiani, R., Hanks, T.D., & Shadlen, M.N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience, 28*(12), 3017-3029.

Kim, J.N. & Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience, 2*, 176-185.

Laming, D. (1979). A critical comparison of two random-walk models for two-choice decisions. *Acta Psychologica, 43,* 431-453.

Leite, F.P. & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics, 72,* 246-273.

Leon, M.I. & Shadlen, M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron, 24,* 415-425.

Lemus, L., Hernandez, A., & Romo, R. (2009). Neural encoding of auditory discrimination in ventral premotor cortex. *PNAS, 106*(34), 14640-14645.

Lo, C. & Wang, X. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience, 9*(7), 956-963.

MacDonald, C.J., Meck, W.H., Simon, S.A., & Nicolelis, M.A. (2009). Taste-guided decisions differentially engage neuronal ensembles across gustatory cortices. *Journal of Neuroscience, 29*(36), 11271-11282.

Macmillan, N.A. & Creelman, C.D. (1991). Detection theory: A user's guide (2nd edition). Mahwah, N.J.: Lawrence Erlbaum Associates.

Mazurek, M.E., Roitman, J.D., Ditterich, J., & Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex, 13*(11), 1257-1269.

McKeeff, T.J. & Tong, F. (2007). The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cerebral Cortex, 17,* 660-679.

Mynatt, C.R., Doherty, M.E., & Tweney, R.D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology, 30,* 395-406.

Newsome, W.T. & Pare, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8,* 2201-2211.

Palmer, J., Huk, A.C., & Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5,* 376-404.

Payne, J.W., Bettman, J.R., & Johnson, E.J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534-52.

Philiastides, M.G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *Journal of Neuroscience, 26*(35), 8965-8975.

Philiastides, M.G. & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Ceberal Cortex, 16,* 509-518.

Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., & Wheeler, M.E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual decision processes using fMRI. *Journal of Neuroscience, 27*(44), 11912-11924.

Ploran, E.J., Tremel, J.T., Nelson, S.M., & Wheeler, M.E. (submitted). High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex.

Rachlin, H. (1988). Judgment, decision, and choice: a cognitive/behavioral synthesis. New York: W.H. Freeman.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R. (1985). Theoretical interpretations of speed and accuracy of positive and negative responses. *Psychological Review, 92*, 212-225.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review, 9*(2), 278-291.

Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology, 53*, 195-237.

Ratcliff, R. & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 83-922.

Ratcliff, R. & Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347-356.

Ratcliff, R. & Rouder, J.N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 127-140.

Rossion, B. & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217-236.

Roitman, J.D. & Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22*(21), 9475-9489.

Romo, R. & Salinas, E. (2001). Touch and go: decision-making mechanisms in somatosensation. *Annual Reviews Neuroscience, 24*, 107-137.

Russo, J.E. & Dosher, B.A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 676-696.

Salinas, E., Hernandez, A., Zainos, A., & Romo, R. (2000). Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *Journal of Neuroscience, 20*(4), 5503-5515.

Schall, J.D. (1999). Weighing the evidence: How the brain makes a decision. *Nature Neuroscience, 2*(2), 108-109.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience, 1*, 199-207.

Sergent, C. & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science, 15*, 720-728.

Shadlen, M.N. & Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neuroscience, 86*, 1916-1936.

Shadlen, M.N., Hanks, T.D., Churchland, A.K., Kiani, R., & Yang, T. (2006). The speed and accuracy of a simple perceptual decision: a mathematical primer. In Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya, K., Ishii, S., Rao, R., & Pouget, A. (Eds.). Cambridge: MIT Press.

Shulman, G.L., Ollinger, J.M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Petersen, S.E., & Corbetta, M. (1999). Areas involved in encoding and applying directional expectations to moving objects. *Journal of Neuroscience, 19*(21), 9480-9496.

Shulman, G.L., McAvoy, M.P., Cowan, M.C., Astafiev, S.V., Tansy, A.P., d'Avossa, G., & Corbetta, M. (2003). Quantitative analysis of attention and detection signals during visual search. *Journal of Neurophysiology, 90,* 3384-3397.

Smith, G.A. (1986). Inspection time and response strategies in a choice response task. *Personality and Individual Differences, 7*(5), 701-707.

Smith, P.L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44*(3), 408-463.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652-654.

Thielscher, A. & Pessoa, L. (2007). Neural correlates of perceptual choice and decision making during fear-disgust discrimination. *Journal of Neuroscience, 27,* 2908-2917.

Thomas, E.A.C. (1983). Notes on effort and achievement-oriented behavior. *Psychological Review, 90*(1), 1-20.

Tootell, R.B.H. & Taylor, J.B. (1995). Anatomical evidence for MT and additional cortical visual areas in humans. *Cerebral Cortex, 5*, 39-55.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*(4), 281-299.

Tversky, A. & Sattath, S. (1979). Preference trees. *Psychological Review, 86*(6), 542-573.

Ungerleider, L.G. & Mishkin, M. (1982). Two cortical visual systems. In Analysis of Visual Behavior, Ingle, D.J., Goodale, M.A., Mansfield, R.J.W. (Eds.). Cambridge, MA: MIT Press.

Usher, M. & McClelland, J.L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*(3), 550-592.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*(7), 1206-1220.

Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: Selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General, 138*(4), 546-560.

Wheeler, M.E., Petersen, S.E., Nelson, S.M., Ploran, E.J., & Velanova, K. (2008). Dissociating early and late error signals in perceptual recognition. *Journal of Cognitive Neuroscience, 20*, 2211-2225.