CONTEXT DEPENDENCE IN PERCEPTUAL AND PREFERENTIAL CHOICE

A Dissertation Presented

by

SEAN PATRICK CONWAY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2025

Psychological and Brain Sciences

CONTEXT DEPENDENCE IN PERCEPTUAL AND PREFERENTIAL CHOICE

A Dissertation Presented

by

SEAN PATRICK CONWAY

Approved as to style and content by:

Andrew L. Cohen, Chair

Youngbin Kwak, Member

Elizabeth Miller, Member

Jeffrey Starns, Member

Ilia Karatsoreos, Department Chair
Psychological and Brain Sciences

ACKNOWLEDGMENTS

Thank you to...

ABSTRACT

CONTEXT DEPENDENCE IN PERCEPTUAL AND PREFERENTIAL CHOICE

SEPTEMBER 2025

SEAN PATRICK CONWAY

, UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Andrew L. Cohen

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Overview

One known fact about human decision-making is that context can, in certain circumstances, systematically affect people's choices. That is, the relative likelihood of choosing one option over another can vary systematically with the menu of options, the *choice set*. These findings are known as *context effects*. One notable context effect, the attraction effect, occurs when the choice share of a *target* option is boosted upon the inclusion of a similar but inferior *decoy* option. Another finding, the repulsion effect, occurs when a decoy boosts the choice share of a *dissimilar* option, known as the *competitor*, rather than the target.

The attraction and repulsion effects, originally studied in preferential choice, have recently been shown to occur in simple perceptual decision-making (Evans et al., 2021; Liao et al., 2021; Spektor et al., 2018, 2022; Trueblood et al., 2013). This is theoretically interesting because it suggests that context effects are a theoretical primitive rather than simply a feature of high-level consumer or risky choice (Trueblood et al., 2013). The goal of this dissertation is to understand how and why these effects occur by employing well-studied statistical models from the psychology and economics literature. Additionally, this dissertation sets out to differentiate the perceptual from decision-making processes that may lead to context effects (specifically, the attraction and repulsion effects).

This dissertation is structured as follows. In Chapter 1, I introduce the relevant empirical and theoretical literature in context effects. In Chapter 2, I develop and test a statistical model of perceptual variability when applied to context effects. In Experiment 1, I first show that the types of stimuli used in perceptual choice context effects experiments are easily confusable and vary systematically with theoretically relevant properties of the stimuli. In Experiment 2, I use the results of a high-powered psychophysics experiment to show that the repulsion effect, but not the attraction effect, is naturally predicted by this statistical model. In Chapter 3, I further test the statistical model by applying it to best-worst choice (Flynn et al., 2007b). In Chapter 4, I generalize the model to preferential choice. Finally, in Chapter 5, I use a perceptual choice experiment to show presentation format affects choice when the decoy is equally similar to both focal options.

1.1.1 Introducing the Attraction Effect

In decision-making experiments, researchers present participants with a finite set of options on each trial and ask them to select a single option. Researchers universally assume that participants use the input they receive (i.e., the value of each option) to arrive at a choice. The study of choice spans multiple fields, including psychology, economics, marketing, and political science. In economics, for example, researchers have developed models based on the idea that, while preferences may vary from moment to moment, people generally make rational choices in any given choice setting (McFadden, 2001). In psychology and marketing, however, researchers have identified a set of phenomena that violate such assumptions, by showing that choices vary based on the *choice set*, or the menu of available options. This class of phenomena is known as *context effects* (Adler et al., 2024).

Context effects are interesting to decision-making researchers because they violate properties of large classes of choice models, such as Independence of Irrelevant Alternatives (IIA) (Ray, 1973) and regularity (McFadden, 2001). IIA states that the likelihood of selecting one option over another is invariant of other options available.

Regularity states that the probability an option is chosen cannot increase by adding more options to a choice set.

Cognitive psychologists have developed numerous process models to explain how context effects can arise (Bergner et al., 2019; Bhatia, 2013; Noguchi & Stewart, 2018; Roe et al., 2001; Trueblood et al., n.d.; Tversky & Simonson, 1993; Usher & McClelland, 2004; Wollschläger & Diederich, 2012). These models vary considerably in their mechanisms but align in the common assumption that decision makers take in the veridical values of each option on each attribute and use these values to arrive at a preference state, leading to a decision.

This dissertation will explore context effects in both perceptual and preferential decision-making. In particular, I will explore the attraction effect and its reversal, the repulsion effect. I summarize these effects and the relevant literature below.

To begin, I first demonstrate a notable context effect, the *attraction effect*. As a demonstration, see 1.1 (left panel), which shows a graphical configuration of various choice options. These options vary on two dimensions (or attributes), where higher values of an attribute are always preferred. I give these dimensions generic names to emphasize that they may be anything from the fuel efficiency and horsepower of cars in a consumer choice experiment to the height and width of rectangles in a perceptual choice experiment.

In Figure 1.1 (left panel), options A and B trade off on attributes. A is high on dimension 2 but low on dimension 1, while B is high on dimension 1 but low on dimension 2. A decision-maker who assigns equal importance to both dimensions should be indifferent between both options when presented with choice set $[A, B]$. Now, however, consider option D_A, which is inferior to A and B, but more similar to A than to B[1]. Similarly, D_B is inferior to both A and B but more similar to B. The

[1]Similarity here is negatively related to distance in attribute space (Shepard, 1987).

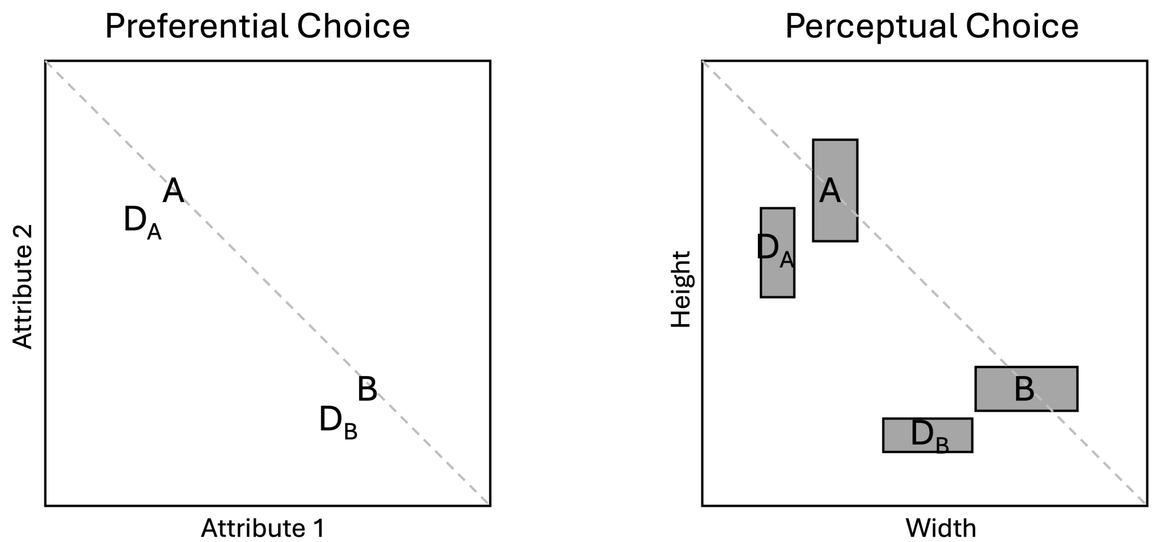

Figure 1.1. A graphical depiction of choice options in the attraction/repulsion effect. Left panel: preferential choice. Right panel: perceptual choice.

attraction effect is the finding that choice for A over B is greater given set A, B, D_A then given set A, B, D_B[2]. Choice models often, though not necessarily, assume the *Independence of Irrelevant Alternatives*(IIA) principle. IIA states that the relative likelihood of choosing a particular option over another is invariant of the choice set (Ray, 1973).

To demonstrat IIA, let A, B, D_A, and D_B be discrete choice options, [] denote the options in a choice set, and $P(A|[A, B])$ denote the probability of choosing option A from a set consisting of A and B. According to IIA:

$$\frac{P(A|[A, B, D_A])}{P(B|[A, B, D_A]} = \frac{P(A|[A, B, D_B])}{P(B|[A, B, D_B]} \tag{1.1}$$

However, in the attraction effect:

[2]This is the weak version of the attraction effect. A strong version requires the ordering of P(A) and P(B) to change with choice set. See Davis-Stober et al. (2023) for a discussion of similar issues.

$$\frac{P(A|[A, B, D_A])}{P(B|[A, B, D_A]} > \frac{P(A|[A, B, D_B])}{P(B|[A, B, D_B]} \tag{1.2}$$

Thus, IIA is violated.

In the context effects literature, it is common to refer to the similar, dominated option as *decoy*, the similar dominating option as a *target*, and the dissimilar dominating option as a *competitor*. I adopt this terminology through this dissertation. For example, in the choice set $[A, B, D_A]$, A is the target, B is the competitor, and D_A is the decoy. The decoy is dominated by the target, so no rational agent should intentionally select it over the target.

The attraction effect was first demonstrated by Huber et al. (1982) [3], who tested participants with duples and triples of choice options, using products such as cars, beers, and TV sets. The authors showed that the introduction of an asymmetrically dominated decoy tended to increase the choice share of a similar, target option. Such a result violates IIA but also a principle known as regularity, which states that the introduction of another option to a choice set cannot increase the probability of choosing any given option (CITE COLONIUS 1984, MACKAY AND ZINNES 1995, MARLEY 1989). That is to say:

$$P(A|[A, B]) \geq P(A|[A, B, D_A]) \tag{1.3}$$

Thus, Huber et al. (1982)'s finding that $P(A|[A, B]) \leq P(A|A, B, D_A)$ violates this assumption. Huber and Puto (1983) replicated these results and also showed that if the decoy has a relatively high value, it can actually take choice shares away from the target. This result suggests that the relative positioning of the decoy to the target can greatly affect patterns of choice, a finding explored by numerous other researchers which has strong theoretical consequences (see below).

---

[3] These authors referred to this finding as the asymmetric dominance effect. To stay consistent with contemporary research, I use the term attraction effect throughout this dissertation.

Numerous researchers have since demonstrated the attraction effect in preferential choice, including in real-world scenarios. Doyle et al. (1999) found an attraction effect in real world supermarket choices by adding a decoy option to an existing product set, where the decoy option was the same brand and price as the target, but of a lower volume. van den Enden and Geyskens (2021) showed that the attraction effect can be used to induce people to choose healthier food items. Slaughter et al. (1999) showed that the attraction effect can be found even without the explicit attribute descriptions commonly used in laboratory experiments, when participants must infer option attributes.

Context effects like the attraction effect have strong theoretical implications. Traditional models of choice, as used in economics and marketing research (McFadden, 2001), treat the *utility*, or value, of each option as a random variable whose parameters are to be estimated from choice data, and on each trial of a choice experiment the participant samples values from these distribution and deterministically chooses the option with the highest sampled value. These models are known as *Random Utility Models* (RUMs). When utilities are assumed to follow Type 1 Generalized Extreme Value distribution, the logit or softmax model is used (CITE). As will be the focus of much of this dissertation, the probit model assumes Gaussian distributed utilities (CITE). Often (though not necessarily) RUMs assume IIA, though this assumption can be relaxed by assuming set or alternative specific intercepts (CITE) or allowing correlations between options and/or attributes (CITE). Haaijer et al. (1998) showed that the probit model shows an improved fit to context effect data by allowing such correlations. I elaborate on this class of models below.

In cognitive psychology, researchers have developed cognitive process models that attempt to explain the cognitive processes that lead to context effects (Bergner et al., 2019; Bhatia, 2013; Noguchi & Stewart, 2018; Roe et al., 2001; Trueblood et al., n.d.; Tversky, 1972; Tversky & Simonson, 1993; Usher & McClelland, 2004;

Wollschläger & Diederich, 2012). These models differ, to varying degrees, in their explanations for the attraction effect. Many, however, rely on comparisons between the target and the similar, but inferior, decoy, which boost an internal preference state for the target. Roe et al. (2001)'s Multialternative Decision Field Theory (MDFT) model proposes that the similarity between target and decoy causes frequent target-decoy comparisons, and through a lateral inhibition, the negative valence for the decoy causes a boost to the preference state of the target at the expense of the competitor. Trueblood et al. (n.d.)'s Multiattribute Linear Ballistic Accumulator (MLBA) model proposes that pairwise attention weights, which are a function of the similarity between options, increase the importance of target-decoy comparisons and thus boost preference for the target.

This dissertation does not explore the predictive success of these models, nor does it incorporate model fitting to compare these models. Indeed, other researchers have done such analyses (Berkowitsch et al., 2014; Cataldo & Cohen, 2021; Evans et al., 2019; Molloy et al., 2019; Turner et al., 2018). Instead, I use behavioral experiments and statistical modeling to understand how context dependence arises in various domains.

The attraction effect is not limited to merely consumer choice. Researchers have found the attraction effect in risky choice (Mohr et al., 2017), policy choice (Herne, 1997), intertemporal choice (Marini et al., 2020), probability judgment (Cai & Pleskac, 2023), medical decision-making (Schwartz & Chapman, 1999), charitable donation (Pittarello et al., 2020), inference (Trueblood, 2012), job candidate selection (Highhouse, 1996), political choice (Pan et al., 1995), and, as will be the focus of much of this dissertation, perceptual choice (Evans et al., 2021; Liao et al., 2021; Spektor et al., 2018, 2022; Trueblood et al., 2013; Turner et al., 2018; Yearsley et al., 2022).

1.1.2 From preferential to perceptual choice

As mentioned above, researchers have begun studying context effects in perceptual choice. Trueblood et al. (2013) demonstrated the attraction effect in perceptual choice. In their experiments, participants saw three rectangles on each trial, arranged in a horizontal array, and they selected the option they believed to have the largest area. As a demonstration of this stimulus configuration, see 1.1 (right panel). Options A and B have equal area but trade off in height and width. [4] Notably, the title of their paper was "Not Just for Consumers: Context Effects Are Fundamental to Decision Making", and in their General Discussion, Trueblood et al. (2013) argued "our experiments suggest that these context effects are a general feature of human choice behavior because they are a fundamental part of decision-making processes. As such, our results challenge explanations of these effects exclusively in terms that are unique to high-level decision making and thus call for a common theoretical explanation that applies across paradigms." (p. 907). According to Trueblood et al. (2013), context effects are not idiosyncratic to high-level choices but are a fundamental part of the choice process. Trueblood et al. (2013) also used these perceptual results to argue against the context-dependent advantage (CDA) model developed by Tversky and Simonson (1993) as well as the Leaky Competing Accumulators model of Usher and McClelland (2004), as these models use "loss aversion" (i.e., the idea that an option's disadvantages on an attribute are weighted more strongly than its advantages on other attributes) to account for context effects.

Turner et al. (2018) replicated Trueblood et al. (2013)'s results and performed a large scale modeling study. The authors compared the ability of numerous mechanisms assumed by decision models to account for context effects. For example, Turner et al. (2018) concluded that pairwise comparisons on individual attributes greatly im-

[4] Trueblood et al. (2013) also demonstrated the similarity and compromise effects.

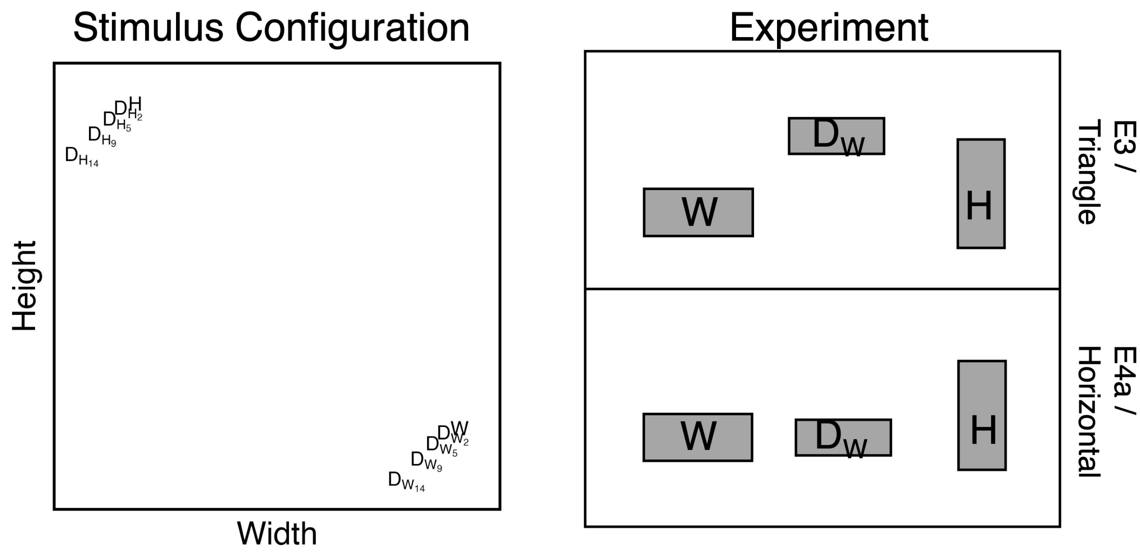

Figure 1.2. Stimulus configuration and example trials from Spektor et al. (2018), Experiments 3 and 4a.

proves models' ability to account for context effects. This may not be appropriate, however, given a perceptual domain where dimensions may not be separable (Ashby & Townsend, 1986).

Spektor et al. (2018) followed up on this work and demonstrated the *repulsion effect* in a rectangle choice experiment. In the repulsion effect, the competitor's choice share is higher than the target's choice share (Banerjee et al., 2024; Bhui & Xiang, 2021; Evans et al., 2021; Frederick et al., 2014; Liao et al., 2021; Simonson, 2014; Spektor et al., 2022). In Spektor et al. (2018)'s experiments, the target and competitor options varied in area, such that one option was always larger, but on average they were the same size. The researchers also varied the *target-decoy attribute distance* (TDD), the percentage difference between the target and decoy areas. For example, if TDD is 2%, the decoy is 2% smaller than the target.

Spektor et al. (2018) ran a total of five experiments, but I focus on their experiments 3 and 4a here. In Experiment 3, the authors varied TDD at four levels: 2%, 5%, 9%, and 14%. The rectangles were arranged in a trianglular display on the screen (see

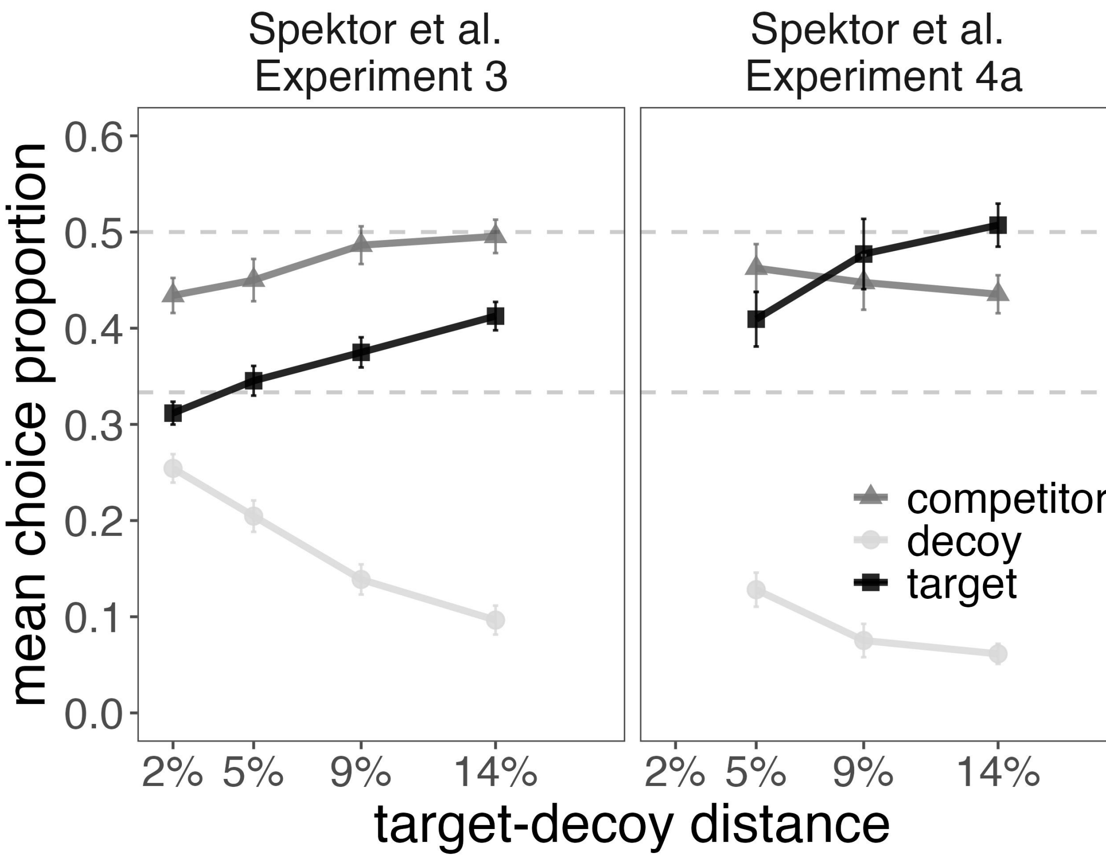

Figure 1.3. Data from Spektor et al. (2018), collapsed across choice set. Error bars are 95% CIs of the mean, computed using the within-subjects error correction from Morey et al. (2008). Dashed lines are drawn at .5 and .33.

Figure 1.2, Experiment 3), in contrast to Trueblood et al. (2013)'s horizontal display. Spektor et al. (2018) found an empirical repulsion effect such that the competitor was selected more than the target at all levels of TDD (see Figure 1.2).

Spektor et al. (2018) also ran a follow-up experiment using the horizontal diplay of Trueblood et al. (2013) (see Figure 1.2, Experiment 4a). Here, they varied TDD at 5%, 9%, and 14%. In Experiment 4a, the data show a slight repulsion effect at low TDD levels that eventually becomes an attraction effect at high TDD levels.

Liao et al. (2021) also replicated the general pattern of Spektor et al. (2018)'s results and also found a an inverse U-shaped relationship between TDD and RST[5]. Relatively low and extremely high TDD values created a repulsion effect, while intermediate TDD values created an attraction effect.

Spektor et al. (2022) demonstrated the repulsion effect in both preferential and perceptual choice, using similar stimuli and display configuration. In these experiments, each stimulus option contained a square with two partially filled bars. In perceptual choice, participants were to select the stimulus with the largest cumulative filled area. In the preferential choice scenario, participants were told that each filled bar represented one possible outcome of a 50-50 gamble, and they were to select the gamble with the highest expected value. Their results were similar to those of Spektor et al. (2018), however, where target and competitor choices increased with TDD, with target and decoy showing a particularly strong trade-off.

Researchers are clearly using perceptual experiments to demonstrate context effects and test theory. These results are clearly informative and theoretically interesting. I argue, however, that researchers should be cautious in assuming that decision-makers receive perceptual input with the same accuracy that they do in preferential choice experiments. Researchers should seek to separate the role of perceptual discriminability from decisional processes when understanding participants' responses. I elaborate on these ideas below, with a demonstration using the results of Spektor et al. (2018).

1.1.3 Understanding Perceptual Choice Experiments

I seek to understand the role of perception in perceptual choice context effect experiments. To do so, I take an extreme stance - that such experiments are solely perceptual experiments rather than decision-making experiments and that no high

[5] $RST = \frac{P(T|[T,C,D])}{P(T|[T,C,D) + P(C|[T,C,D])}$

level decision processes are occurring. This extreme assumption is incorrect, but I believe it is a good place to start in understanding existing data. I also demonstrate how and under what circumstances it is incorrect later in this dissertation. To begin, I return the results of Spektor et al. (2018).

As reported above, Spektor et al. (2018) demonstrate that a relatively small change in stimulus display (arranging stimuli in a triangle rather than a horizontal line) reverses the attraction effect. Why is this? To begin to answer this question, I highlight the differences between Spektor et al. (2018)'s data and previous context effect data. In preferential choice tasks, participants are given a set of options on each trial (e.g., laptops, apartments, washing machines), along with the attribute values associated with each option (e.g., 10 GB RAM, 1500 square feet, 2.7 cubic feet capacity). These attributes are typically represented numerically (Banerjee et al., 2024; Hayes et al., 2024) or with easily discriminable graphical representations (Cataldo & Cohen, 2019). The decoy option, therefore, is rarely selected (e.g., $< 5\%$ of all trials), and these selections are assumed to be the result of inattention or accidental responses. Researchers assume that participants are able to detect the dominance relationship between target and decoy. Perceptual choice tasks complicate participants' ability to detect this dominance relationship. In Spektor et al. (2018)'s experiments, the decoy is selected as often as 25% of all trials in some conditions. The decoy is selected less often in experiment 4a (triangle display) than in experiment 3 (horizontal display). Decoy selections also decrease as the difference between decoy area and target/competitor areas increases. Finally, though both target and competitor increase in choice share as TDD increases, the target choice share increases at a higher rate than does the competitor, suggesting a strong trade-off between target and decoy choices (stronger, indeed, than that of competitor-decoy choices). That is, the mean *Relative Share of the Target* (RST) (Berkowitsch et al., 2014) increases with TDD in both experiments

Clearly, perceptual discriminability plays a role in Spektor et al. (2018)'s results. Participants clearly are better able to discriminate the target and competitor from the decoy as the decoy decreases in size. Any reasonable account of these data should parse the out discriminability from genuine context effects.

There is a large body of psychological research, beginning with the work of Thurstone (1927), of treating the perception of a stimulus as a random variable. In his famous "Law of Comparative Judgment" paper, Thurstone (1927) first showed that researchers can use binary choice proportions to estimate the psychological distance between stimuli, by treating perceptual intensity as a Gaussian random variable. This work led to similar research using Signal Detection Theory (STD) (Hautus et al., 2021), which also treats psychological quantities (e.g., memory, perception), as random variables. Similarly, Ashby and colleagues' General Recognition Theory (GRT) models the perception of a stimulus as a multivariate normal random variable, where each dimension of the model is the perceived dimension of a stimulus (Ashby & Perrin, 1988; Ashby & Gott, 1988; Ashby & Townsend, 1986). In marketing and economics, researchers treat the utilities of choice options as random variables, which are often assumed to be Gaussian or Extreme Value Distributed and estimate choice models known as Random Utility Model (RUMs) (J. A. Hausman & Wise, 1978; McFadden, 2001; Train, 2009).Often, though not necessarily, these models share the common property that value (whether it is the utility of a consumer product, the perception of magnitude, or the memory signal in a recognition task) is stochastic while choice is deterministic (c.f.) (Benjamin et al., 2009) I consider this class of models throughout this dissertation.

CHAPTER 2

PARSING THE ROLE OF PERCEPTION AND DECISION IN CONTEXT-DEPENDENT CHOICE

2.1 Introduction

I now introduce the model I explore throughout the dissertation. This model is simplistic, as it treats the experiments of Trueblood et al. (2013) and Spektor et al. (2018) as perceptual, rather than decision tasks. This also completely eschews the possibility of higher level decision processes. I use this model to understand the extent to which. My model treats value (perceived area) as stochastic and choice as deterministic. In the current modeling framework, I do not treat height and width as independent attributes but rather consider perceived area to be unidimensional. In my model, I consider a perceptual choice experiment, where participants are presented with 3 perceptual stimuli on each trial. I assume that on each trial i with choice set K, The perception $\mathbf{X_i}$ of all 3 stimuli is sampled from a multivariate Gaussian distribution with a mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$:

$$\mathbf{X}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{2.1}$$

where $\boldsymbol{\mu}$ is a column vector consisting of:

$$\begin{pmatrix} \mu_T \\ \mu_C \\ \mu_D \end{pmatrix} \tag{2.2}$$

where the subscripts T, C, and D indicate target, decoy, and competitor, respectively, and $\mathbf{\Sigma}$ is a positive semi-definite 3 x 3 covariance matrix computed as:

$$\mathbf{\Sigma} = S\mathbf{\Omega}S \tag{2.3}$$

where S is a diagonal matrix consisting of:

$$\begin{pmatrix} \sigma_T & 0 & 0 \\ 0 & \sigma_C & 0 \\ 0 & 0 & \sigma_D \end{pmatrix} \tag{2.4}$$

and $\mathbf{\Omega}$ is a correlation matrix:

$$\begin{pmatrix} 1 & \rho_{TC} & \rho_{TD} \\ \rho_{TC} & 1 & \rho_{CD} \\ \rho_{TD} & \rho_{CD} & 1 \end{pmatrix} \tag{2.5}$$

with ρ_{TD}, for example, indicating the population-level correlation between target and decoy stimuli.

As mentioned above, the model assumes that value is stochastic while choice is deterministic[1]. The model always "selects" the option perceived the largest, regardless of the magnitude of the difference between the "winner" and "runners-up". That is, given a vector $\mathbf{X}_i$ of perceived areas on trial i with set K, the probability a participant selects stimulus i is:

[1]This also assumes ties are not possible, which is true if and only if perceived area is absolutely continuous.

$$P(i|K) = P(\mathbf{X}_i > \mathbf{X}_j, j \in K, i \neq j) \tag{2.6}$$

If all off-diagonal elements of $\mathbf{\Omega}$ are 0, the model collapses to the standard Thurstonian Case CHECK CASE XX (Thurstone, 1927) often used by cognitive psychology researchers. Models of this form have closed form solutions and their predictions are easy to compute.

On the other hand, if any elements of $\mathbf{\Omega}$ are non-zero, the closed form solution of this model does not exist, and to compute predictions and estimate model parameters, researchers must use simulation or numerical integration methods (Train, 2009). In all applications of these model through this dissertation, I use simulation to generate model predictions.

This model is capable of generating a(n) attraction (repulsion) effect by assuming $\mu_T > \mu_C$ ($\mu_C > \mu_T$), i.e., that on average target and competitor stimuli differ inherently in their perceived areas. This, however, is an ad hoc assumption that may describe the data well but will generate limited theoretical insight. Moreover, I wil later present empirical data in this dissertation that, generally speaking, $\mu_T = \mu_C$.

I now consider the role of perceptual correlations between all pairs of stimuli, i.e., ρ_{TC}, ρ_{TD}, and ρ_{CD}. I vary both ρ_{TD} and ρ_{TC} from -1 to 1; in other words, all rectangles oriented the same way share one correlation and those oriented differently share another correlation. I show model predictions that result from varying these correlations in Figure 2.1. Here, I assume that $\mu_T = \mu_C > \mu_D$ and that $\sigma_T = \sigma_C = \sigma_D$[2].

[2]In Experiment 2 I present evidence supporting the former assumption. I also relax the latter assumption when modeling Experiment 2

2.1 shows model predictions in the form of RST (Relative Share of the Target), where RST values above .5 indicate an attraction effect and values below .5 indicate a repulsion effect. My model can, depending on the relationship between ρ_{TD} and ρ_{TC}, predict a repulsion, attraction, or a null context effect. If $\rho_{TD} > \rho_{TC}$, the model predicts a repulsion effect. If the target and decoy are correlated more strongly than competitor and decoy, it is more likely that if on a particular trial the target perception is large, that the decoy is even *larger*, causing the decoy to "steal" choice shares from the target more than the competitor, i.e., a repulsion effect.

If $\rho_{TD} < \rho_{TC}$, the model predicts an attraction effect. This is because $\rho_{TC} = \rho_{CD} > \rho_{TD}$, so the decoy "steals" choice shares from the competitor more than the target.

Finally, if $\rho_{TD} = \rho_{TC} = \rho_{CD}$, the model predicts a null effect. In this case, no pair of stimuli are more correlated than any other pair, so the predictions are identical to a model where $\rho_{TD} = \rho_{TC} = \rho_{CD} = 0$ model. This model collapses to an Independent Normal Model.

2.1.1 Accounting for perception

I propose that these perceptual correlations may be driving the repulsion effect in Spektor et al. (2018)'s data. The decoy option is smaller than the the target and competitor options and is thus not always discriminated. The triangle configuration makes discriminability particularly difficult for participants (as I show below in my Experiment 1). Simultaneously, however, the fact that target and decoy share an orientation (i.e., both wide or both tall) makes the comparison of these two options easier. In statistical terms, the perception of the decoy is more strongly correlated with the perception of the target than with perception of the competitor. Given these correlations, the perceived area of the decoy is more likely to exceed the target than the competitor. Conceptually, an empirical repulsion effect may be driven by

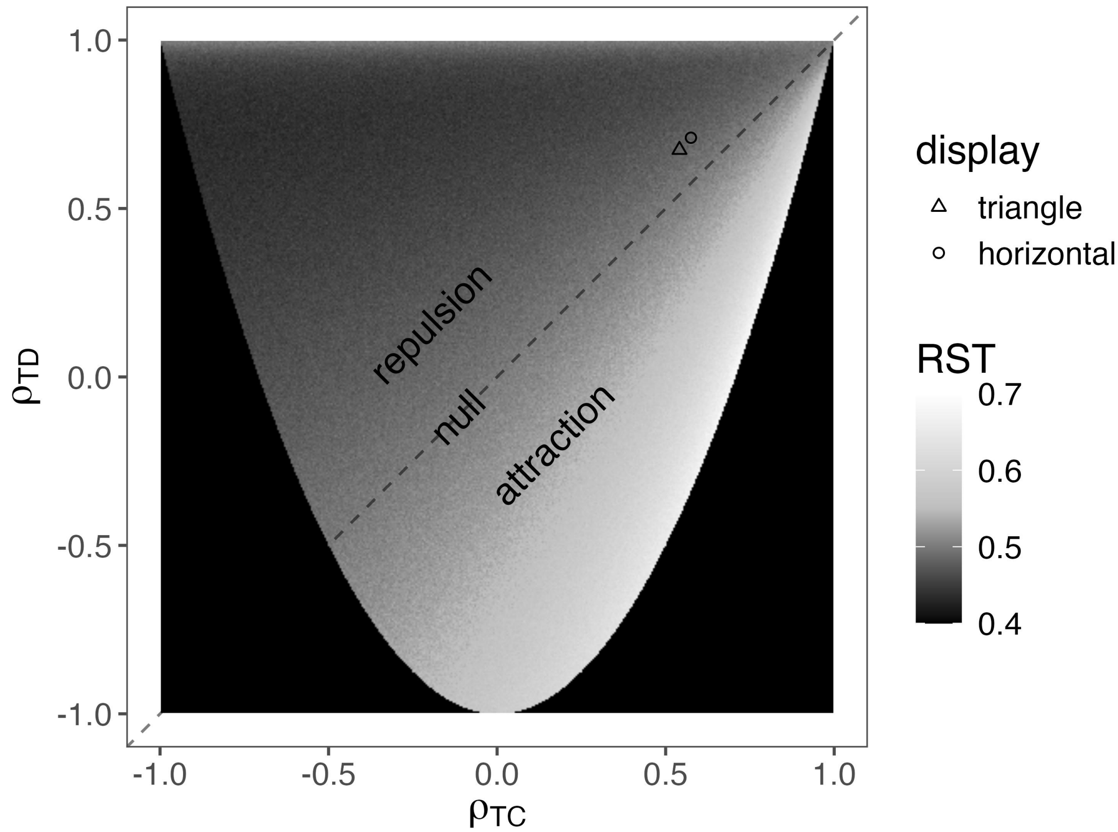

Figure 2.1. Model simulations for the attraction and repulsion effects based on the variation of ρ_{TD} and ρ_{TC}. "Regions" of the plot are labeled based on their qualitative predictions for attraction, null, and repulsion effects. The black region is the area where, due to extreme correlations, a positive semi-definite variance-covariance matrix could not be formed and predictions are unavailable. The triangle and circle mark the observed mean correlations from the Experiment 2 triangle and horizontal conditions respectively.

perception rather than the "high-level" effects typically considered in the decision-making literature.

Modern psychological models of context effects often assume an attribute-wise comparison process (Bhatia, 2013; Roe et al., 2001; Trueblood et al., 2013; Usher & McClelland, 2004). Under this class of models, participants arrive at a decision by comparing pairs of options on a single attribute, where the modeller assumes attribute values are veridical. This assumption is quite reasonable when modeling choices where each attribute is presented separately and discriminability issues are minimal or non-existent. In this article, I present evidence that this assumption may not be appropriate. ELABORATE HERE

In Experiment 1, I first present results from a two-alternative forced-choice experiment to show that these stimuli are easily confusable and that the triangle display of Spektor et al. (2018) decreases discriminability relative to the horizontal display. I also show that, consistent with the predictions of a perceptual model where $\rho_{TD} > \rho_{TC} > \rho_{CD}$, target-decoy discriminability is in fact *easier* than competitor-decoy discriminability and that target-decoy discriminability increases with TDD.

In Experiment 2, I combined a psychophysics task with a choice task to estimate the parameters of my perceptual model. In the first phase of the experiment, participants estimate the size of target, decoy, and competitor rectangles on each trial. In the second phase of the experiment, I conducted a standard choice experiment, replicating Spektor et al. (2018)'s results. I use the data from the first phase of Experiment 2 to obtain stable estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in my perceptual model. Finally, I show that the model, conditioned on the observed parameter estimates, naturally predicts a repulsion effect but not an attraction effect.

2.2 Experiment 1

The goal of Experiment 1 was to test participants' ability to discriminate between rectangles in the perceptual choice tasks of Trueblood et al. (2013) and Spektor et al. (2018). I do, however, acknowledge the possibility that the presentation of three options (rather than just two) may impact perception. On each trial, I presented participants with three options (target, competitor, and decoy). After a short delay, I highlighted two of the three rectangles and asked participants to indicate which rectangle was larger. Additionally, I used a within-subjects manipulation to compare discriminability in both the triangle display of Spektor et al. (2018), Experiment 3, and the horizontal display of Spektor et al. (2018), Experiment 4a[3]. Otherwise, with a few exceptions discussed below, I follow the stimulus construction and experimental design of Spektor et al. (2018), Experiment 3.

2.2.1 Methods

2.2.1.1 Participants.

Data collection took place at the University of Massachusetts Amherst. 86 undergraduate students participated in exchange for course credit. 1 participant who achieved less than 80% accuracy on catch trials (see below) was excluded from all analyses. Trials with response times (RTs) $<$ 100ms or $>$ 10000ms were also excluded from all analyses.

2.2.1.2 Stimuli.

The experiment had two types of trials: critical trials and catch trials. On each critical trial, the target and competitor had the same area[4] but differed on orientation, with one stimulus being wide and the other tall. The decoy always had the

[3]see also Trueblood et al. (2013), Experiment 1.

[4]Here I simplify Spektor et al. (2018)'s design by ensuring both focal stimuli had the same area.

same orientation as the target. The height and width of the decoy were reduced proportionally so that the decoy area was always 0%, 2%, 5%, 9%, or 14% of the target areas. Because the target and competitor always had the same area, this means that the decoy was also 0%, 2%, 5%, 9%, or 14% of the competitor area. These are the TDD values from Spektor et al. (2018), plus a 0% level which acted as a baseline[5].

2.2.1.3 Design.

There were 5 blocks of trials. In each block there were 60 critical trials, 12 at each TDD level, and 30 catch trials. Of the 12 critical trials at each TDD level, 6 were presented in a triangle and 6 were presented horizontally. Finally, 3 of the 6 targets in each display condition at each TDD level were wide and 3 were tall. Of each of these 3, one was a target-decoy comparison, one was a target-competitor comparison, and one was a target-competitor comparison. Trial order and rectangle order within each trial were randomized.

On each catch trial, there was one large rectangle and two much smaller rectangles. The large rectangle was $260 \pm U(-40, 40)$x$200 \pm U(-40, 40)$ pixels, with a random orientation. The smaller rectangles were $180 \pm U(-40, 40)$x$120 \pm U(-40, 40)$ pixels, one wide and one tall.

On every trial, the rectangles were displayed in either a triangle or horizontal display (see 1.2). The horizontal distance between all rectangles was constant, but 25 pixels of jitter was added to each rectangle's vertical location.

Stimuli were presented on computer monitors with a resolution of 1920 x 1080 pixels. The experiment was programmed with jsPsych (De Leeuw, 2015).

---

[5]When TDD=0%, the target and decoy are identical, so labeling is arbitrary.

2.2.1.4 Procedure.

On each trial, participants saw three rectangles, labeled 1, 2, and 3 (from left to right). The rectangles appeared for 1825ms total, but after 500ms, two of the rectangles changed to a darker shade. After all three rectangles disappeared from the screen, participants were asked to select which of the two darker rectangles had the larger area.

2.2.2 Results

2.2.2.1 Catch Trials.

Participants performed well on the catch trials. The mean percentage correct in the triangle display was 92.6%($SD = 3.77$), and the mean percentage correct in the triangle display was 93.2%($SD = 3.52$).

2.2.2.2 Critical Trials.

I first checked the baseline TDD level data (TDD=0%) across to make sure that participants were indifferent between pairs of options when they had identical area. The mean percentage of target choices in target-competitor trials was 48.99% (SD=10.18). The mean percentage of competitor choices in competitor-decoy trials was 49.80% (SD=11.30). The mean percentage of target choices in target-decoy trials was 49.47% (SD=12.06). Participants were indifferent between all pairs of options in the TDD = 0% trials, so I do not consider these trials further.

The primary analysis was performed on the target-decoy and competitor-decoy trials, excluding the TDD=0% trials. In these trials participants' task is simply to not select the decoy option on a given trial. I present mean choice proportions across conditions in Figure 2.2. Participants' performance indeed improves with TDD. Furthermore, their performance is better when stimuli are displayed in the horizontal configuration than in the triangle configuration, and it is also better in target-decoy trials compared to competitor-decoy trials. Finally, there is an interaction, such that

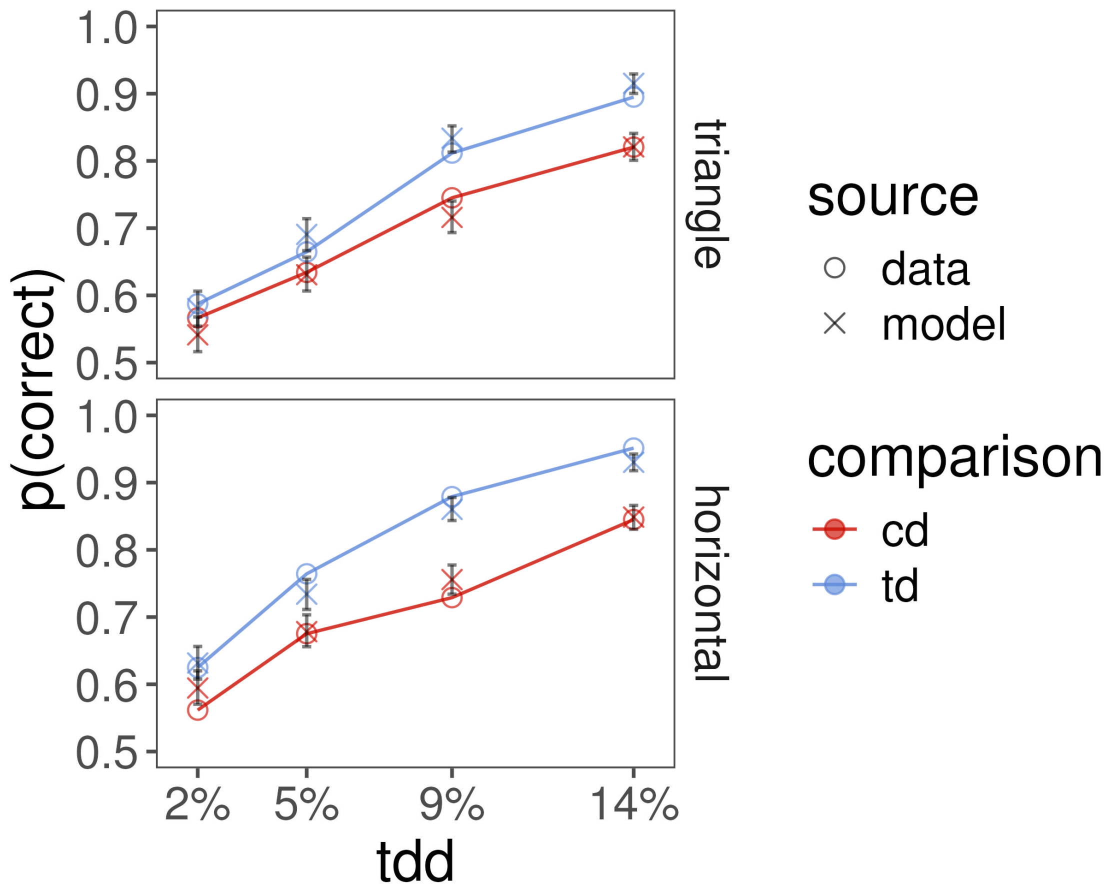

Figure 2.2. Experiment 1, mean choice proportions by stimulus display, TDD, and comparison. td=target-decoy trials, dc=competitor-decoy trials. Model predictions come from the Bayesian hierarchical logistic regression presented in Appendix. Error bars are 95% HDIs on the mean.

as TDD increases, the target-decoy performance is even better than the competitor-decoy performance. See Appendix XXX for inferential statistics which support these conclusions.

2.2.2.3 Discussion

In Experiment 1, I showed that participants are not always able to discriminate between target-decoy and competitor-decoy stimuli. I also show that this discriminability increases with TDD and that overall discriminability is better in the horizontal compared to the triangle display. Finally, through the interaction of comparison-pair and TDD, I show that target-decoy discriminability increases with TDD at a higher rate than competitor-decoy discriminability.

These results, in fact, are natually predicted by a model where $\rho_{TD} > \rho_{CD}$ (see Appendix XXX for simulations - NEED TO DO THIS).

2.3 Experiment 2

I continue this line of research in Experiment 2, where I used a psychophysics task to estimate the mean perceived area and correlations between perceived area to the target, competitor, and decoy rectangles. Experiment 2 used the *method of cross-modal matching* (Stevens & Marks, 1965), where participants adjusted the size of a circle to match the perceived area for each rectangle. On each trial, participants saw three rectangles and three circles, each labeled 1, 2, and 3. Participants adjusted the size of the circle corresponding to each rectangle, until they believed the two to have equal area. I also replicate Spektor et al. (2018)'s choice data in a second phase of the experiment. Finally, I used a between-subjects manipulation to display the rectangle stimuli in either the horizontal or triangle displays of Spektor et al. (2018).

2.3.1 Methods

2.3.1.1 Participants.

Data collection took place at the University of Massachusetts Amherst. 521 undergraduate students participated in exchange for course credit. 68 participants did not complete the full experiment within the 1 hour time limit and were removed from all analyses.

2.3.1.2 Stimuli.

In the circle adjustment phase there were three types of trials: critical trials, filler trials, and catch trials. On each critical trial, the target and competitor had the same area but differed on orientation, with one stimulus being wide and the other tall.

The decoy always had the same orientation as the target. I varied TDD from 2%, 5%, 9%, and 14%. I also varied the target, competitor, and decoy stimuli to fall on three diagonals. In pixels, the small and larger focal stimulus dimension values on the lower, middle, and upper diagonals were $[60, 135]$, $[90, 165]$, and $[120, 195]$. I reduced the absolute size of the target/competitor stimuli from Experiment 1 to Experiment 2 to accomodate the circle adjustment phase (see procedure below).

On filler trials, I randomly sampled three rectangles by sampling a height and width from the distribution $U(56, 195)$, encompassing the full range of stimuli from the critical trials.

On the catch trials, I randomly sampled one rectangle from the lower diagonal and two from the upper diagonal, such that one stimulus was clearly larger than the other two stimuli.

The choice phase had identical trial types with the exception that there were no catch trials, only critical and filler trials.

2.3.1.3 Design.

Across both phases, I varied display condition between-subjects and TDD, diagonal, target-decoy orientation within-subjects. After removing subjects, there were 218 participants in the horizontal display condition and 225 participants in the triangle display condition.

In the circle adjustment phase, there were 4 blocks, each 40 trials. Each block consisted of 24 critical trials, 14 filler trials, and 2 catch trials. Within the critical trials, there were 6 trials at each level of TDD. In 3 of these 6 trials the target and decoy were oriented wide (choice set $[w, h, d_w]$), and in the other 3 target and decoy were oriented tall (choice set $[w, h, d_h]$).

In the choice phase, there were 4 blocks, each with 34 trials. 24 of these trials were critical trials and 10 were filler trials. Of these 24 critical trials, there were 6 trials at each level of TDD. Within each 6, there were 3 trials where target and decoy were oriented wide and 3 were target and decoy were oriented tall.

2.3.1.4 Procedure.

The experiment took place in two phases:

On each circle adjustment trial, three gray rectangles appeared in the lower left corner of the screen, either in a triangle or horizontal display. In the upper right, three gray circles appeared in the upper right of the screen, in the same display as the rectangles (see Figure 2.3). A small amount of jitter ($U(-15, 15)$px) was added to the position of each rectangle and the corresponding circle. Each circle started with an area of 78 square pixels (i.e., with a radius of 5), the minimum size I allowed in the experiment. Participants used the mouse to adjust the circle. Within a single trial, they were free to adjust the circles in any order they liked or to go re-adjust a circle as much as they liked. There was no time limit to each adjustment trial. The maximum circle area allowed was 65144 square pixels[6]. When a participant finished adjusting the circles on a trial, they clicked the "Submit" button located on the lower right hand corner of the screen.

The circle adjustment phase began with three practice trials, followed by the 4 blocks of experimental trials. At the beginning of each experimental block, participants completed 3 calibration trials. Calibration trials were identical to filler trials, with the caveat that I provided feedback after participants' responses. After participants submitted their responses on a trial, a red circle appeared around each adjusted circle, showing the true area of the corresponding rectangle.

Throughout the circle phase, I kept track of the deviations between the true rectangle areas and the participants' adjusted circle areas. At the end of each block, I

[6] I arrive at this number based on the maximum area the circles could be while only appearing on the right half of the screen and maintaining the same horizontal distance from each other as the corresponding rectangles.

computed the current mean deviation, and, depending on this value, told the participant that they were either over or under-adjusting, on average.

The choice phase began with 3 practice trials. Participants were not provided feedback during these practice trials.

On each choice trial, three rectangles appeared in the center of the screen in a horizontal or triangle display. There was no vertical jitter added here. Participants were told to select the rectangle with the largest area by clicking on it.

At the end of the choice phase, I let each participant know their percentage correct from the choice phase. Note that in a critical trial, a correct response is simply one where they did not select the decoy, as the target and competitor rectangles always had the same area.

Stimuli were presented on computer monitors with a resolution of 1920 x 1080 pixels. The experiment was programmed with GNU Octave (Team, 2019) and PsychoPhysics Toolbox (Brainard, 1997).

2.3.2 Results

2.3.2.1 Data Processing

Given the difficulty of the circle adjustment task, the data required processing to ensure that outlier trials and participants did not influence our estimates of **Omega**.

First, I removed all participants who did not give correct responses on 75% (6/8) of the catch trials. A correct response entails estimating the larger rectangle to be the largest on the trial. 10 participants were removed from all analyses after they failed to achieve at least 75% correct on catch trials (see below for details). This left a total of 443 participants.

Next, from the remaining participants, I first natural log-transformed all responses. I then dropped all trials where at least one circle was not adjusted (i.e., at least one circle was left at the starting size).

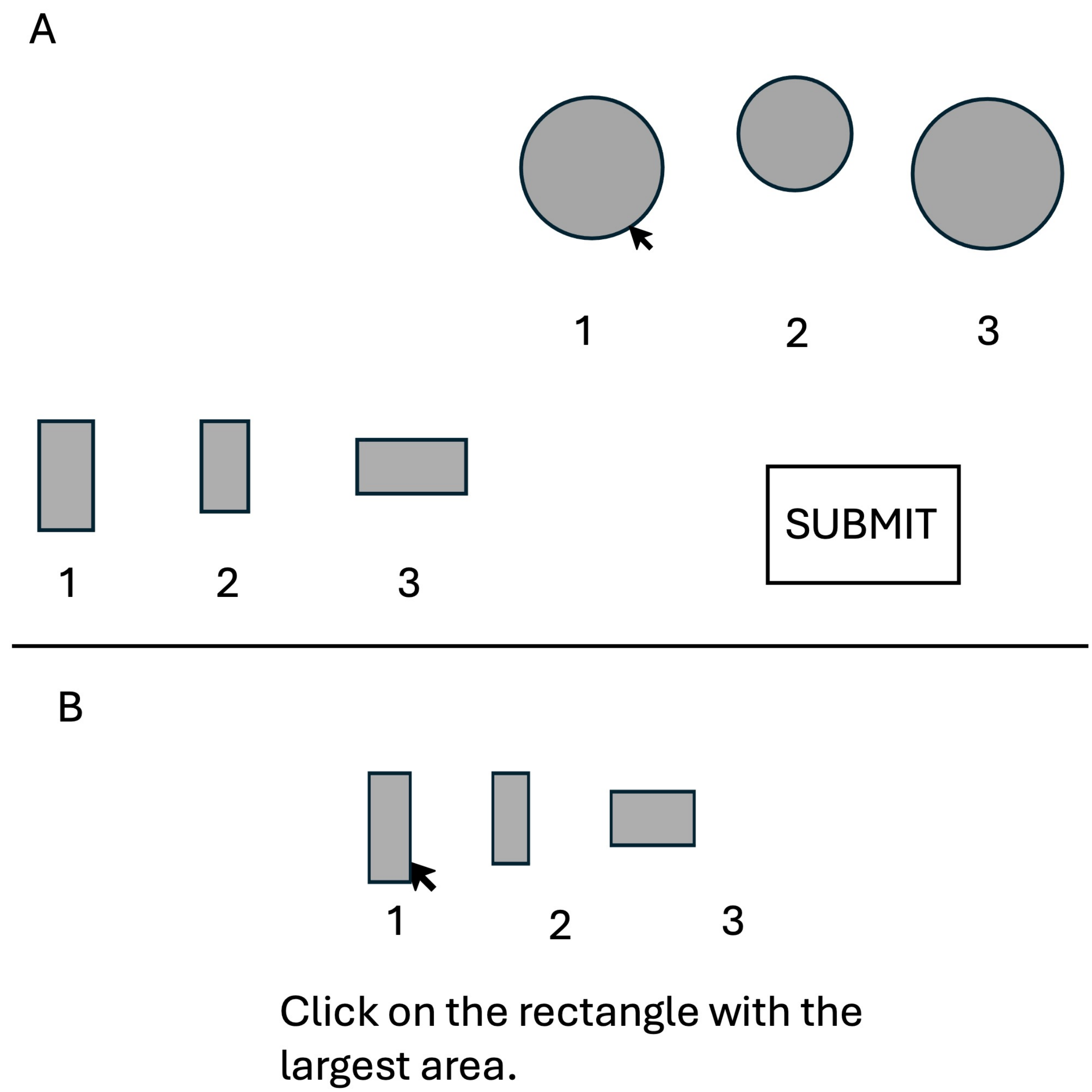

Figure 2.3. Example trials from Experiment 2. A: Circle adjustment phase. B: Choice phase. This an example of trials in the horizontal display condition.

I then removed outlier participants using the following procedure:

I fit a linear regression to each individual participants' data, regressing each log circle area on each corresponding log rectangle area. I then computed an R^2 for each participant. I then removed all participants whose R^2 fell below the 5% quantile for all R^2s, which in this case was .3975. This removed 23 subjects, leaving us with a total of 420 participants, 213 in the triangle display condition and 207 in the horizontal display condition. Of the remaining participants, R^2 values were high ($M = .67, SD = .12$), indicating they could generally perform the task.

From the 420 partipants whose data I analyzed, I removed outlier trials from the critical trial data. I did so to ensure that any outliers do not influence our estimates of ρ_{TD}, ρ_{TC}, or ρ_{CD}. I z-transformed all log circle areas within each participant and diagonal. I remove all critical trials where at least one z-score had an absolute value above 3.5. This led to 102 trials being dropped. I dropped 0, 1, 2, and 4 critical trials from 339, 62, 18, and 1 participants, respectively.

After all circle phase data processing, I were left with 20371 trials in the triangle display condition and 19809 trials in the horizontal display condition.

For the choice phase, I only retained participants whose data I retained in the circle phase.

2.3.2.2 Circle Phase Results - Critical Trials

Before modeling our data, I wanted to ensure that participants could successfully perform the task. While I do not expect perfect performance in an absolute sense, I do require adequate relative performance.

To assess performance on the critical trials, I computed the mean difference between actual log area and estimated log area for each subject, stimulus pair (i.e., target-competitor, target-decoy, competitor-decoy), and actual difference. I plot these via a set of boxplots in Figure 2.4. Although participants vary considerably in their

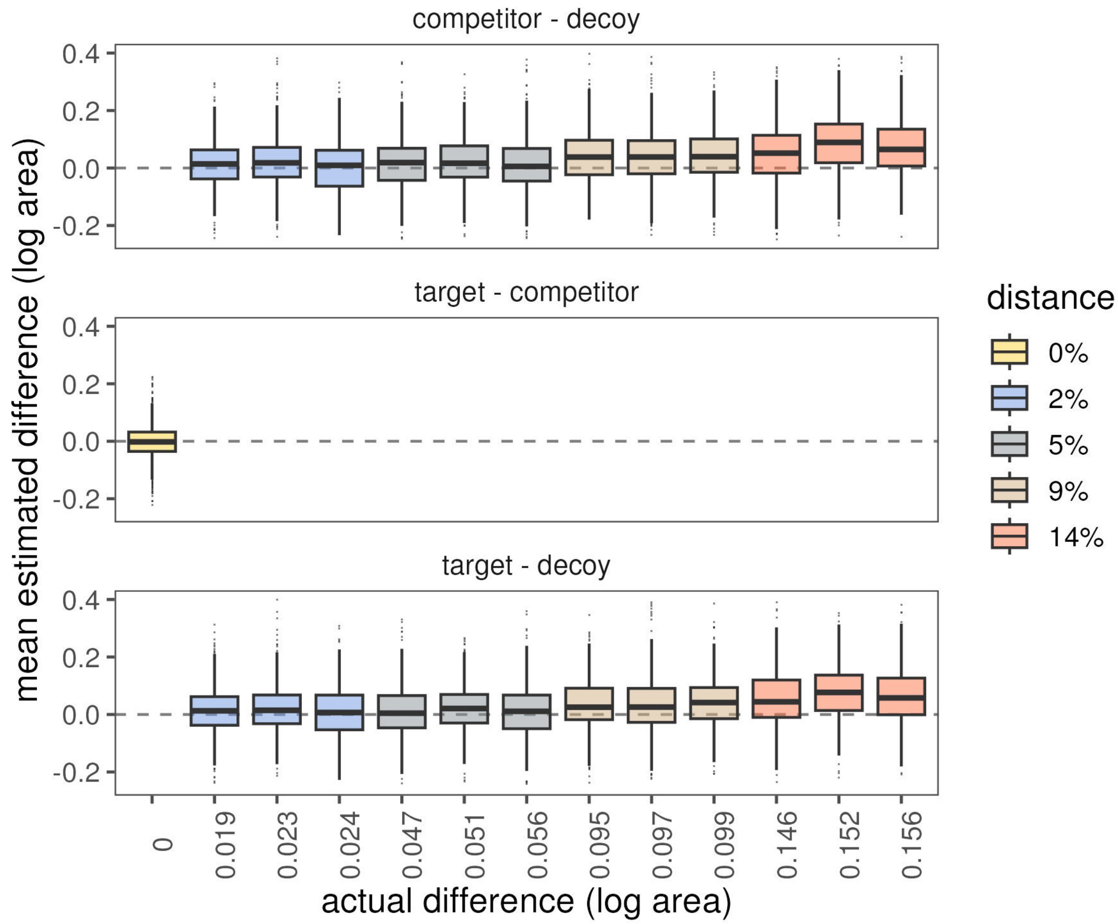

Figure 2.4. Boxplot of subject-level mean error in area estimations, split by stimulus pair, TDD, and absolute discrepancy in rectangle area. Note that because the target and competitor rectangles always had equal areas, the true difference is always 0.

judgments, I find that on average, participants' adjusted circle areas increase with the absolute size of rectangles.

I next present scatterplots of all pairwise circle areas from each trial, see Figure 2.5. I present these to be transparent about the raw data and to illustrate the necessity of a statistical model to understand these correlations.

Computing raw correlations, without accounting for subject and trial-level differences, will grossly inflate the size of these correlations. Moreover, any differences between, say, ρ_{TD} and ρ_{CD} are bound to be small. I used Bayesian hierarchical mod-

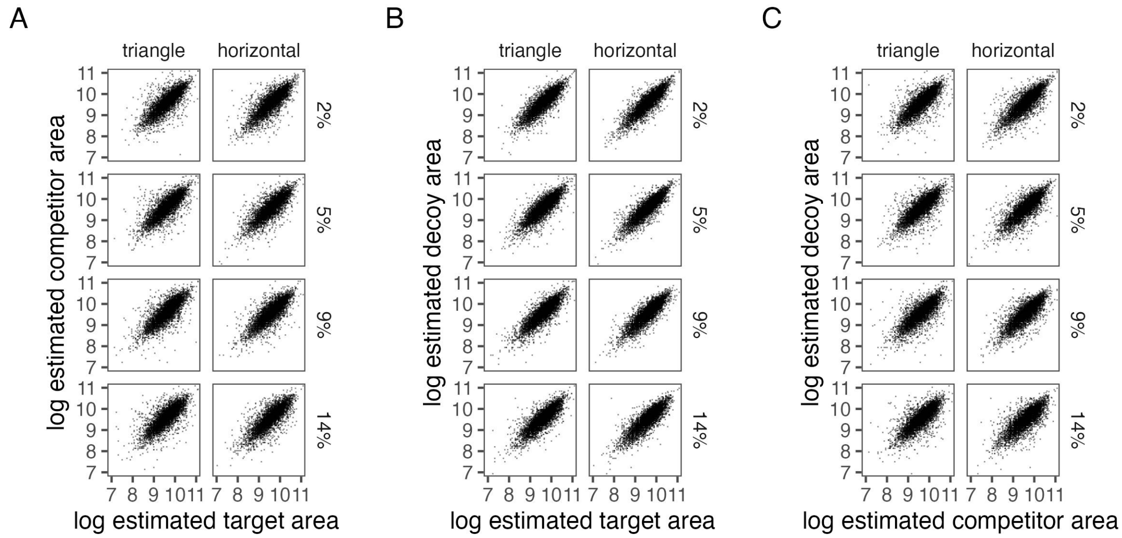

Figure 2.5. Scatterplots of target-competitor (A), target-decoy (B), and competitor-decoy (C) correlations, split by display condition and TDD.

eling to estimate the parameters of a multivariate normal distribution (as outlined in the introduction), with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. I present the details of the model in Appendix XXX but present the main results below.

I assume that, for participant i, on each critical trial j, the perceived target, competitor, and decoy areas $\mathbf{X}_i$ are sampled from a multivariate normal distribution with mean vector $\boldsymbol{\mu_{ij}}$ and variance-covariance matrix $\boldsymbol{\Sigma}$.

As discussed in the Introduction, I decompose sigma into the $S\boldsymbol{\Omega}S$, where the diagonal elements of S are population standard deviations and the off diagonal elements are 0. $\boldsymbol{\Omega}$ is 3 x 3 correlation matrix.

I focus on the estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ in the main text and discuss the details of the modeling, along with § estimates, in Appendix XXX.

I show mean estimates of $\boldsymbol{\mu}$ in Figure 2.6 and show estimates of $\boldsymbol{\Omega}$ in Figure 2.7. In both conditions, ρ_{TD} is larger than both ρ_{CD} and ρ_{TC}, while ρ_{CD} and ρ_{TC} do not differ from each other.

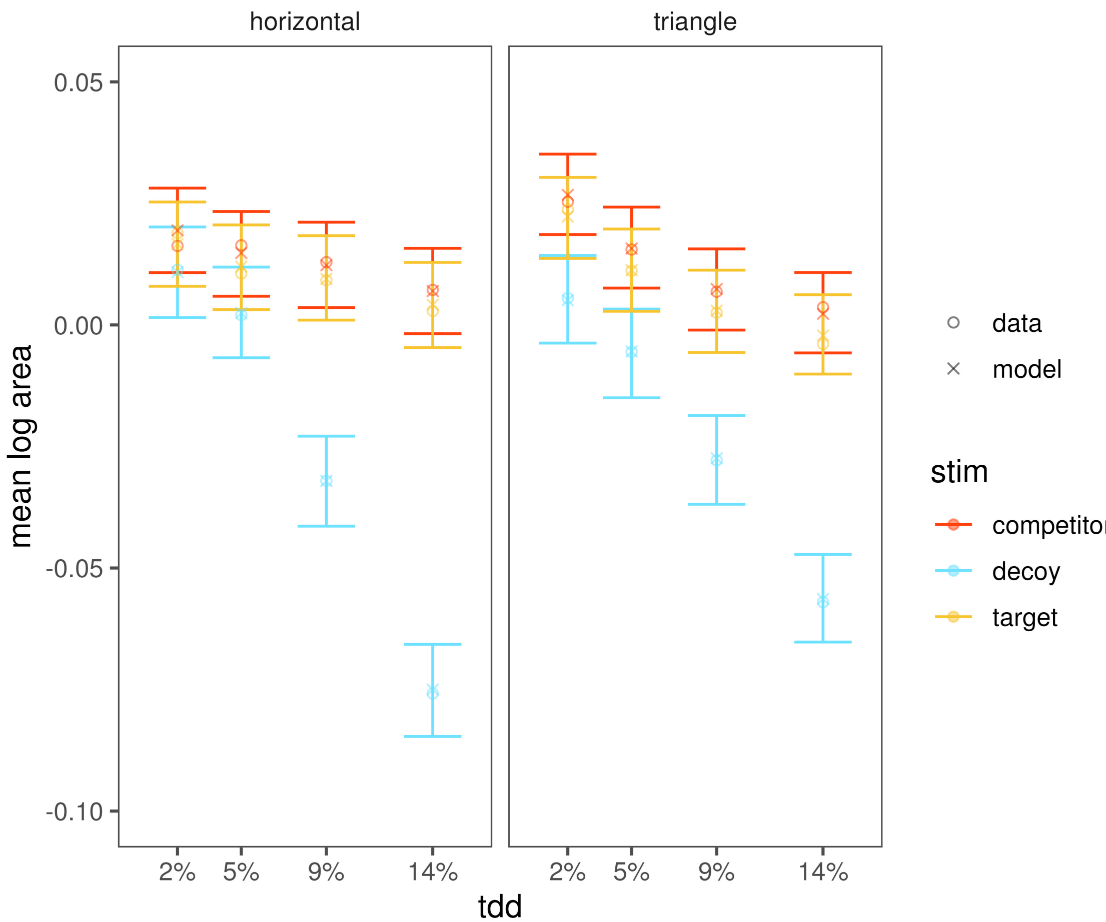

Figure 2.6. μ estimates from Experiment 2.

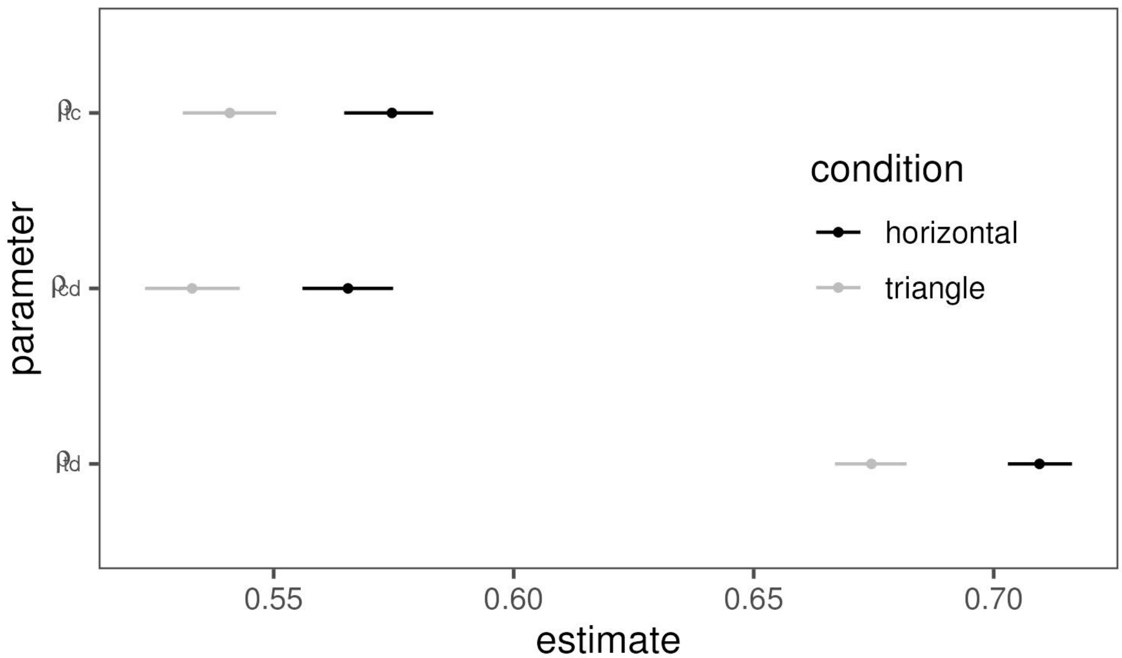

Figure 2.7. Posterior estimate of $\mathbf{\Omega}$ (off-diagonal parameters) across display conditions. Lines show 95% HDIs. Dots indicate means.

2.3.2.3 Choice Results

I present mean choice proportions across display conditions and TDD in Figure 2.8. I replicate the qualitative results of Spektor et al. (2018). At low levels of TDD, I find a repulsion effect in both display conditions. At higher levels of TDD, I either find a null effect (triangle condition) or an attraction effect (horizontal condition).

To ensure that this result is not an artifact of averaging across choice sets, I present mean changes in choice proportion for the options w and h across the two choice sets $[w, h, d_w]$ and $[w, h, d_h]$ (see **??**). These results also show that low levels of TDD create a repulsion effect, while higher levels create a null or attraction effect. See Appendix XXX for inferential statistics that support these conclusions.

2.3.3 Model Simulations

After estimating the parameters of the perceptual model and analyzing the choice data, I sought to test whether the model can predict the choice data. Though I

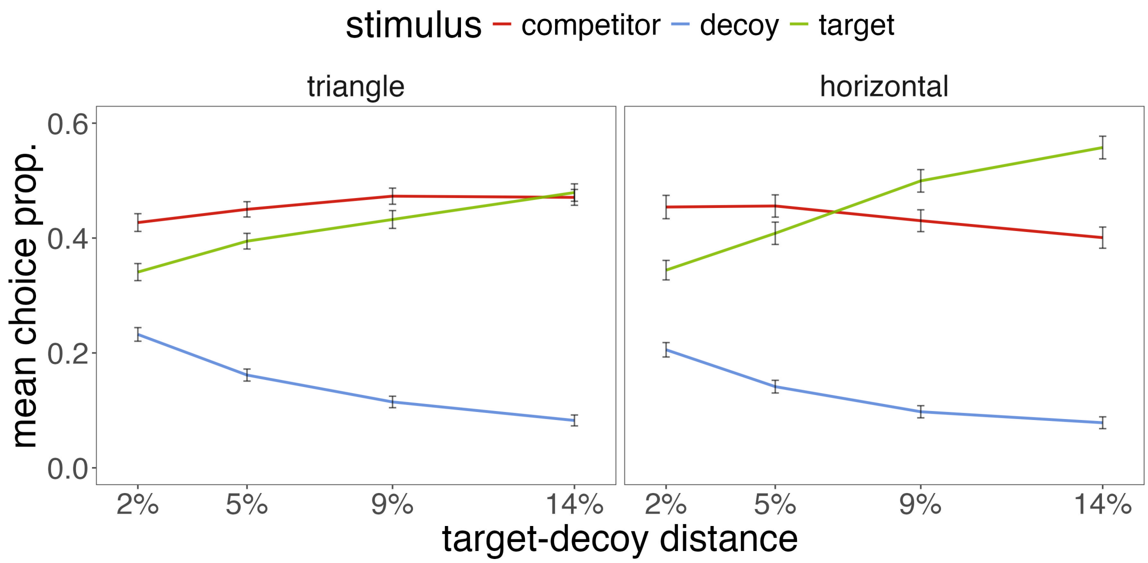

Figure 2.8. Mean choice proportions for target, competitor, and decoy options, by TDD and display condition. THIS FIGURE IS A PLACEHOLDER. WILL CHANGE WITH HDIs FROM STAT. MODEL.

showed in the introduction that this is possible, it was an open question whether the parameters estimated from actual data would produce qualitatively interesting predictions.

I used mean estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (see Appendix XXX) to generate predictions at each level of TDD in both display conditions. I present the predictions in Figure **??**.

Given the estimated parameters, the model is able to produce a repulsion effect, but not an attraction effect. This aligns with our predictions from the introduction; the repulsion effect, at least in some forms, can be generated by a higher correlation between target and decoy stimuli compared to target-competitor and competitor-decoy pairs.

2.3.4 Discussion

The results of Experiments 1 and 2 show that participants are not always able to discriminate the decoy from the target and the competitor, and, that target-decoy

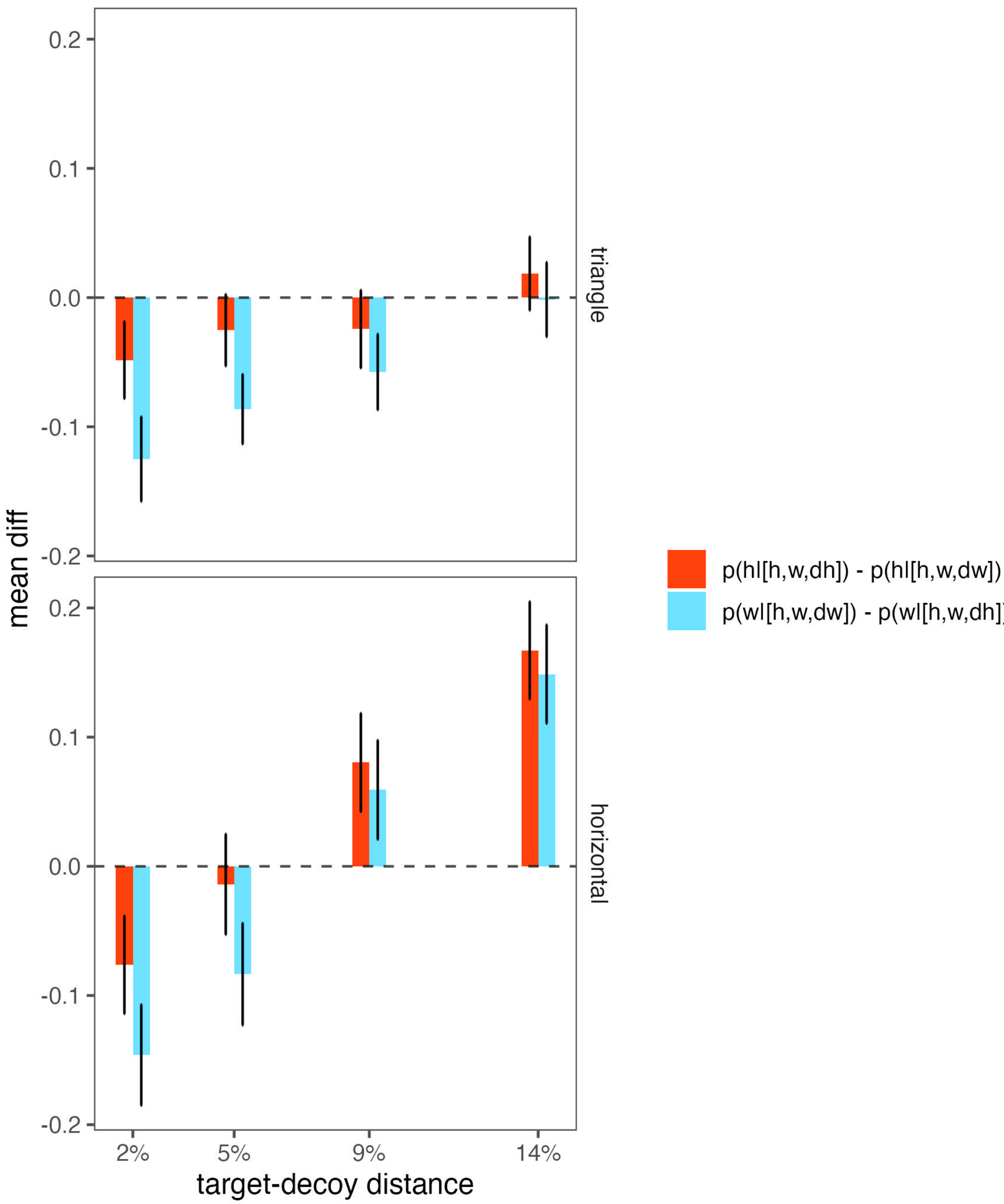

Figure 2.9. Mean changes in choice proportions across choice sets. FIGURE IS ALSO A PLACEHOLDER

perceptions appeared to be correlated. The observed correlations can, in turn, naturally produce the repulsion effect but not the attraction effect.

CHAPTER 3

EXTENDING A PERCEPTUAL MODEL TO BEST-WORST CHOICE

3.1 Introduction

In Chapters 1 and 2, I presented a model of perceptual choice and showed it can systematically predict the repulsion effect, but not the attraction effect. In this chapter, I test another prediction of the model while demonstrating an important empirical result in another domain: best-worst choice.

3.1.1 Introducing Best-Worst Choice

Best-worst choice (also known as best-worst scaling) is an experimental paradigm where participants select their most and least preferred options from a choice set. Originally proposed by Finn and Louviere (1992), best-worst choice is widely used in a number of applied fields, such as transportation (Beck & Rose, 2016) and healthcare economics (Cheung et al., 2016; Flynn et al., 2007a). One key advantage here, when compared to traditional discrete choice research, is that researchers can use best-worst choices to gain information about participants' ranking of options while never requiring them to complete a full ranking task, which may be quite difficult (Marley & Louviere, 2005a).

In addition to the empirical applications, researchers have developed theoretical results on best-worst choice, with many models relating best-worst choices to an underlying utility function. Marley and Louviere (2005a) developed a class of models known as ”maxdiff” (maximum difference) models of best-worst choice[1]. According to the maxdiff model, given choice set K, the probability of selecting option x as best and option y as worst (where $x \neq y$) is defined(Hawkins, Marley, et al., 2014):

$$BW_K(x, y) = \frac{e^{u_x - u_y}}{\sum_{\substack{p,q \in K \\ p \neq q}} e^{u_p - u_q}} \tag{3.1}$$

where u_i is the utility of option i. This model proposes a single utility function that determines best and worst choices. Specifically, it proposes that best-choice probabilities are an increasing function of u, while worst-choice probabilities are a decreasing function of u. The use of the exponential function means that the maxdiff model is another form of the widely used multinomial logit (MNL) choice model (J. Hausman & McFadden, 1984). Furthermore, so long as u does not vary based on choice set, the maxdiff model predicts a monotonic relationship between best-choice probabilities and worst-choice probabilities.

There are many variations on this model (Flynn & Marley, 2014; Flynn et al., 2007b; Marley & Louviere, 2005b; Marley & Pihlens, 2012; Marley et al., 2008), though the maxdiff model remains the dominant model for analyzing best-worst choice data.

Researchers have explored whether this monotonicity holds empirically. Hawkins, Marley, et al. (2014) examined both preferential and perceptual best-worst choice data using response time modeling. They used the linear ballistic accumulator model (LBA) Brown and Heathcote (2008), which casts the decision process as a race between ”accumulators” towards a threshold, where the average accumulation across trials is captured by the drift rate parameter. Modeling datasets containing both preferential and perceptual best-worst choice data, they were able to successfully account

[1]Note that the term maxdiff is sometimes erroneously used to refer to best-worst experiments in the generic sense. Following Marley and Louviere (2005a), I use maxdiff to refer to a specific class and parameterization of choice model.

for choice data by assuming a parallel race between "best" and "worst" accumulators for each option. Furthermore, they showed that the utility values estimated for each option using a MNL model were positively linearly related to the log drift rate values from the LBA, suggesting an underlying utility representation that captures choices.

In a follow-up Hawkins, Marley, et al. (2014) found that, collapsing across choice sets, best-choice probabilities are (negatively) monotonically related to worst-choice probabilities. They also showed that, using the parallel best-worst LBA as a model, the drift rate parameter for worst choice can be parameterized as the reciprocal of the best choice drift rate. Formally, if $d_b(i)$ is the drift rate for selecting option i as best, then $d_w(i) = 1/d_b(i)$, where $d_w(i)$ is option i's drift rate for best choices.

We can think of the parallel best-worst LBA as a process implementation of the maxdiff model (Hawkins, Marley, et al., 2014), which proposes set independence. While researchers have proposed models that allow set dependence (Marley et al., 2008), these models still predict a monotonic relationship between best and worst choices.

It is not always the case, however, that a single latent variable (i.e., utility) underlies choices. Indeed, as I show below, the model from Chapters 1 and 2 predicts, under certain conditions, a dissociation between best and worst choices.

3.1.2 Model Predicted Dissociations Between Best Choices and Worst Choices

Let K be a choice set consistings of options T, C, and D (i.e., target, competitor, and decoy). As in Experiments 1 and 2, these are rectangles in a perceptual choice experiment. As above, I asssume that on each trial i with choice set K, The perception $\mathbf{X_i}$ of all 3 stimuli is sampled from a multivariate Gaussian distribution with a mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ (see 2.1).

$\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are parameterized the same as in Chapters 1 and 2.

Figure 3.1. Best-worst choice simulations. Each row is a different TDD value from Experiment 2.

In this chapter, I apply the model to best worst choice by assuming that, given a vector $\mathbf{X_i}$ of perceived areas on trial i with set K, the probability a participant selects stimulus i is as best is:

$$P(i|K) = P(X_i > X_j, j \in K, i \neq j) \tag{3.2}$$

while the probability of selecting stimulus k as worst (where $i \neq k$) is:

$$P(k|K) = P(X_k < X_j, j \in K, k \neq j) \tag{3.3}$$

As it happens, the correlations (i.e., Ω) estimated from Experiment 2 predict that, in a best-worst choice paradigm, best and worst-choice probabilities are non-monotonically related. I demonstrate this using simulations.

To simulate best-worst choice, I simply used the mean parameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) estimated from Experiment 2[2] and simulated a large ($N = 10000$) number of choice trials. I collapse over choice set (as in previous reported simulations) and focus on target, competitor, and decoy choice proportions at each level of TDD. I show these results in figure 3.1, in a state-trace plot (Newell & Dunn, 2008). State-trace analyses plot the values of two dependent variables against each other for a particular experimental condition. State-trace analysis can be controversial (Ashby, 2019; Ashby & Bamber, 2022; Stephens et al., 2020), and statistical inference on state-trace data is not straightforward (Davis-Stober et al., 2016; Sadil et al., 2018). In principle, however, if the analyst can reliably conclude that the data points do not fall on a single curve, they conclude that the data vary on at least 2 dimensions.

[2] I used only those estimated from the triangle condition (See Experiment 2).

The model, conditioned on the estimated parameters, predicts an interesting result. Although the competitor is most frequently chosen as best, due to the repulsion effect from Experiment 2 and from Spektor et al. (2018), it is not, however, least frequently chosen as worst. Specifically, $B(C) > B(T) > B(D)$, while $W(T) < W(C) < W(D)$, where $B(i)$ and $W(i)$ are the probabilities that option i is chosen as best and worst, respectively.

The reason for this prediction is that because $\rho_{TD} > \rho_{CD} \approx \rho_{TC}$, on the (relatively few) trials where X_D is largest, it is more likely that $X_D > X_T > X_C$ than $X_D > X_C > X_T$. In other words, the high ρ_{TD} value "pulls up" the target more than the competitor.

This dissociation is subtle, and the predicted effect size is small. Indeed, all predicted $W(C) - W(T)$ probabilities were $< .05$. However, In Experiment 3, I show the empirical and modeling results from a best-worst choice experiment designed to test this prediction. I show that the dissociation between best and worst choices does indeed occur, and the maxdiff model cannot account for these results.

3.2 Experiment 3

The goal of Experiment 3 was to test the predictions of the perceptual choice model. Specifically, the perceptual model predicts that $B(C) > B(T) > B(D)$, while $W(T) < W(C) < W(D)$. I conducted a best-worst perceptual choice experiment. I used stimuli identical to those of Experiment 2 and presented stimuli in the triangle display of Experiments 1 and 2. I show that 1) this prediction (generally) holds empirically and 2)the maxdiff model cannot account for these results, even its parameters are free to vary.

3.2.1 Methods

3.2.1.1 Participants.

Data collection took place at the University of Massachusetts Amherst. 392 undergraduate students participated in exchange for course credit. 23 participants who achieved less than 80% accuracy on catch trials (see below) were excluded from all analyses. Trials with response times (RTs) < 100ms or > 10000ms were also excluded from all analyses.

3.2.1.2 Stimuli.

The experiment had three types of trials: critical trials, filler trials, and catch trials. On each critical trial, the target and competitor had the same area but differed on orientation, with one stimulus being wide and the other tall. The decoy always had the same orientation as the target. I varied TDD at 2%, 5%, 9%, and 14%. I also varied the target, competitor, and decoy stimuli to fall on three diagonals. Note that these stimuli are identical to those of Experiment 2.

On each filler trial, three stimuli were uniformly sampled the space between the largest and smallest diagonals.

On each catch trial, one stimulus was sampled from the largest diagonal, while two stimuli were sampled from the smallest diagonal.

3.2.1.3 Design.

There were 8 blocks of trials. In each block there were 24 critical trials, 6 at each TDD level. There were 8 trials per diagonal. There were 10 filler trials and 3 catch trials per block.

Participants were randomly assigned into one of two conditions: best-worst or worst-best. On each trial, participants in the best-worst condition initially chose the largest rectangle and then chose the smallest rectangle. Participants in the worst-best condition chose in the opposite order. The condition factor was included to account for the possibility that best-worst choice order impacts choice.

After removing poor performing participants, there were 185 participants in the best-worst condition and 184 participants in the worst-best condition.

Stimuli were presented on computer monitors with a resolution of 1920 x 1080 pixels. The experiment was programmed with GNU Octave and Psychtoolbox (Brainard, 1997; Team, 2019).

3.2.1.4 Procedure.

The experiment began with three practice trials, which were identical to the filler trials.

On each trial, participants saw three rectangles, labeled 1, 2, and 3 (from left to right), arranged in the triangle display. Participants in the best-worst/worst-best condition saw a prompt asking them to select the largest/smallest rectangle on screen. Participants used the mouse to click on their chosen rectangle. After they made their choice, this rectangle changed color to indicate that it was no longer available as an option. Next, participants in the best-worst/worst-best condition selected the smallest/largest rectangle, at which point the trial ended. See 3.2 for an example trial graphic.

Stimulus order was randomized on each trial.

Participants were told their percentage correct of best choices, worst choices, and overall choices at the end of the experiment.

3.2.2 Results

Participants did not meaningfully differ in their choices by condition, so I collapse over condition for all reported analyses.

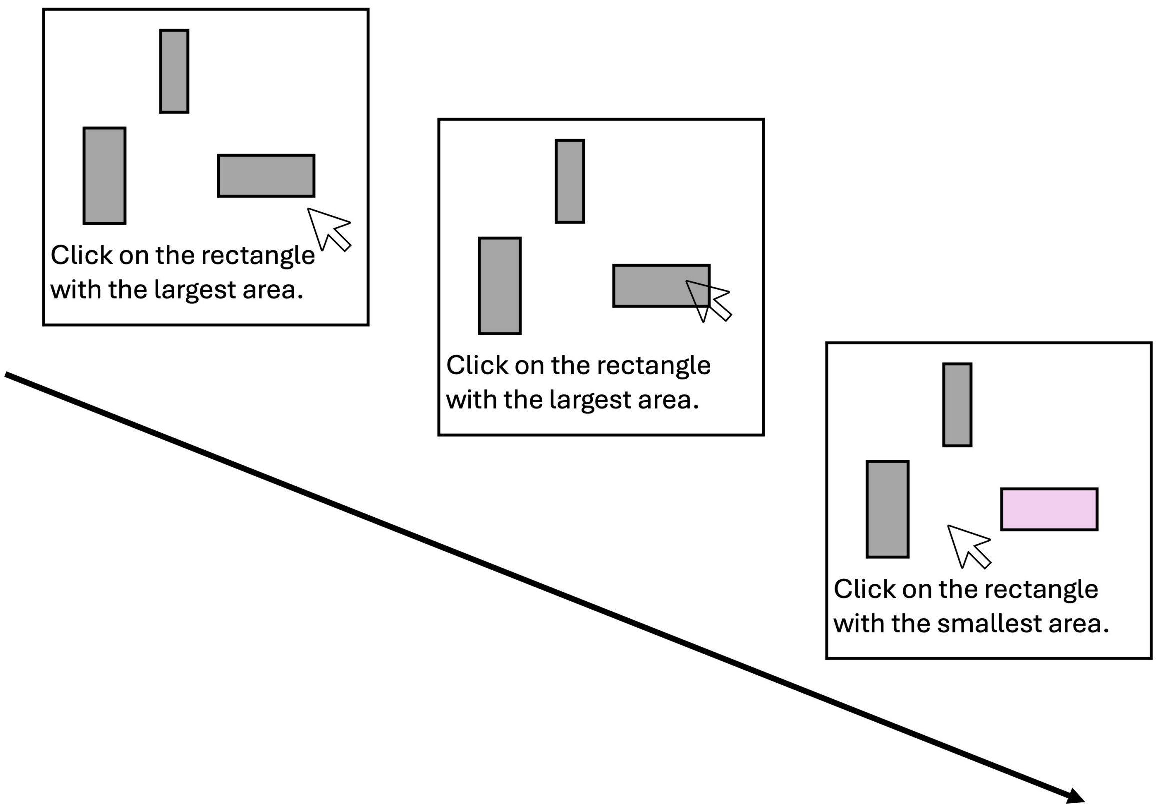

Figure 3.2. An example experimental trial for Experiment 3. Note that this is a trial in the best-worst condition.

3.2.2.1 Catch Trials.

Participants performed well on the catch trials. The mean percentage correct for best choices was $97.97\%(SD = 14.09)$, and the mean percentage correct for worst choices was $98.26\%(SD = 13.09)$. The mean percentage correct for both best and worst choices (i.e., the mean percentage of the trials on which participants were able to correctly identify the largest and smallest rectangles) was $96.98\%(SD = 17.12)$.

3.2.2.2 Filler Trials.

Participants performed worse on the filler trials compared to the catch trials, but still well above chance. The mean percentage correct for best choices was $89.83\%(SD = 30.23)$, and the mean percentage correct for worst choices was $88.95\%(SD = 13.09)$. The mean percentage correct for both best and worst choices was $96.98\%(SD = 17.12)$.

3.2.2.3 Critical Trials.

I first consider participants choice proportions, conditioned on TDD and choice set. Mean choice proportions for these data are plotted in 3.3.

The results show a consistent bias to choose the w (i.e. the option wider than tall) as largest, a finding also shown in Experiments 1 and 2. Participants also (on average) regularly choose the decoy rectangle as smallest, with the exception of the choice set h, w, d_w and $TDD = 2\%$, where they select the h rectangle as smallest, on average. This can be attributed to the difficulty of the $TDD = 2\%$ condition and the overall wide rectangle bias. However, consistent with the predictions of the model, the target is still less likely to be chosen as worst than the competitor, $W(h|h, w, d_h) < W(h|h, w, d_w)$ and $W(w|h, w, d_w) < W(w|h, w, d_h)$, while the competitor option is more likely to be chosen as best, $B(h|h, w, d_w) > B(h|h, w, d_h)$ and $B(w|h, w, d_h) > B(w|h, w, d_w)$. See Appendix XXX for inferential statistics which support these conclusions.

These results are more easily understood by plotting mean target, competitor, and decoy choice proportions across TDD levels, collapsed over choice set. See 3.4 for these data. First, the best-choice proportions replicate the repulsion effect initially found by Spektor et al. (2018) and replicated in the current Experiment 2, where the competitor is more likely to be chosen as best at low TDD levels, while the target and competitor are chosen equally often at high TDD levels. Decoy best-choice proportions also decrease systematically with TDD. See Appendix XXX for inferential statistics which support these conclusions.

Furthermore, the target is always more likely to be chosen as worst, compared to the competitor and decoy, at all TDD levels, $W(T) < W(C) < W(D)$, as predicted by the perceptual model outlined in Chapter 2. This model still cannot predict the null repulsion effect when $TDD = 14\%$, as discussed in Chapter 2, which suggests that effect may be due to higher level decision processes.

3.2.2.4 Modeling.

To analyze the data, I fitted both the maxdiff model and a Dircihlet-multinomial model XXX CITE. The goal of this modeling analysis is to test the ability of a standard best-worst choice model to account for the observed dissociations in the data. The Dirichlet-multinomial is a flexible model that estimates probability distributions around the choice proportions for each participant in each condition. I first consider the maxdiff model.

Maxdiff Modeling

I first turn to the maxdiff model (Marley & Louviere, 2005a), which was outlined in the introduction to this chapter. This equation predicts that the probability of choosing options x and y, $x \neq y$ increases monotonically with the difference in their estimated utilities (see **??**). This model is the most commonly used (and arguably the simplest) analysis technique for best-worst choice data. I applied this model to the

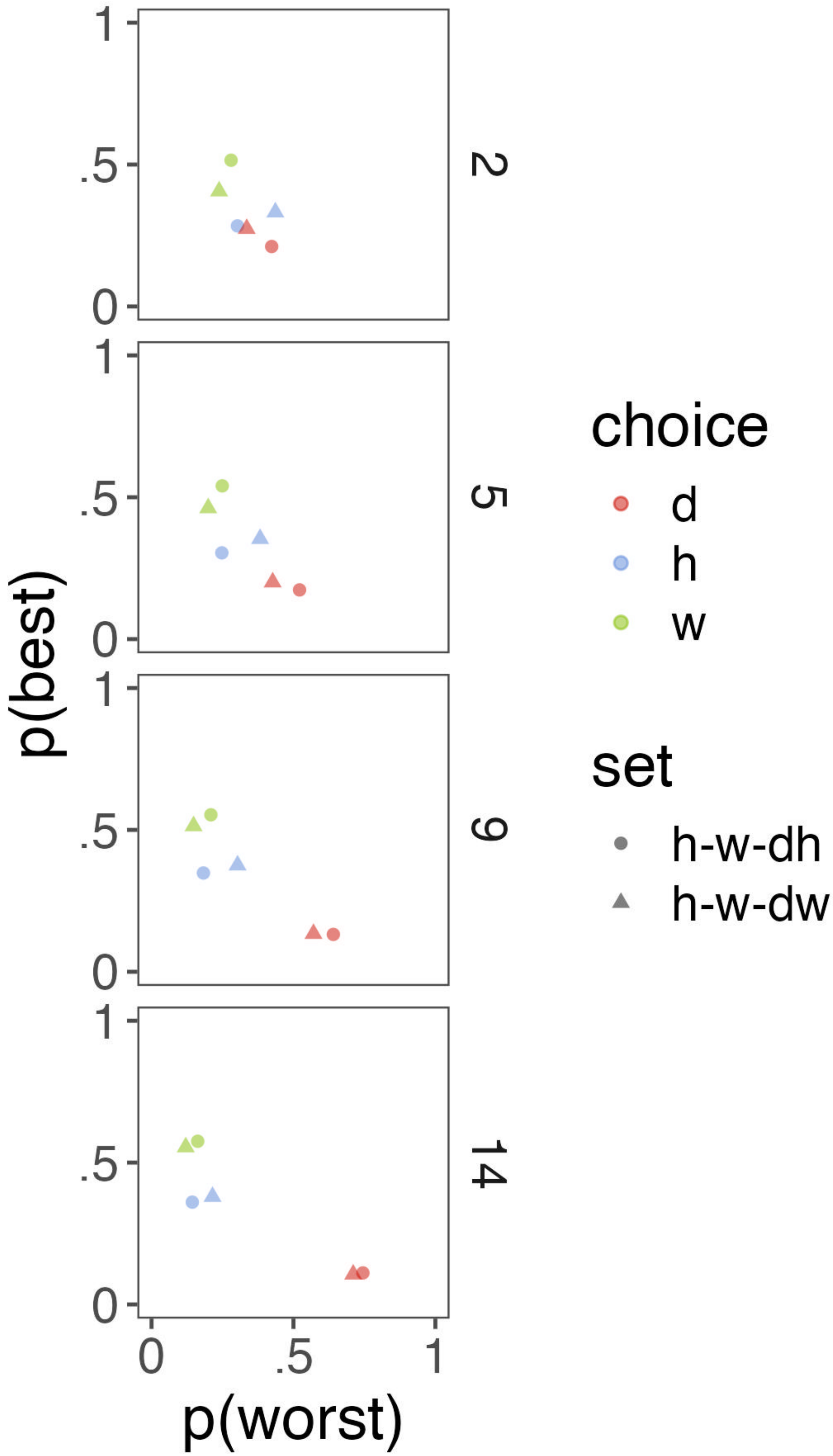

Figure 3.3. Mean best and worst-choice proportions for the h, w, and d rectangles, conditioned on TDD (rows) and choice set (shapes).

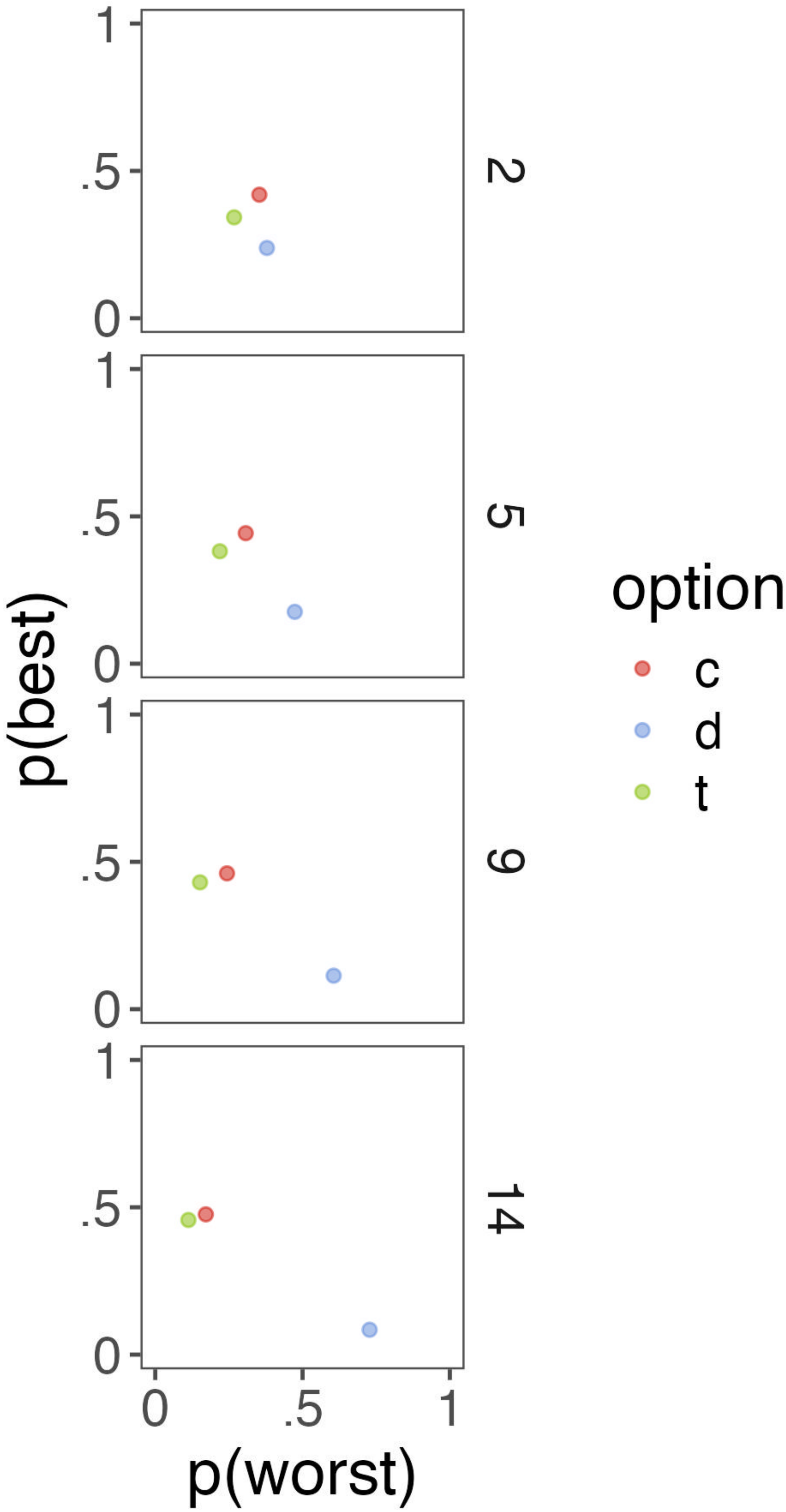

Figure 3.4. Mean best and worst-choice proportions for the target, competitor and decoy rectangles, conditioned on TDD (rows).

current experiment and show that it is unable to predict the observed dissocations in best-worst choices, even with its best fitting parameters.

I implemented this model as a Bayesian hierarchical model. I show the details of the model fitting procedure, including parameterization, parameter estimates, and all priors in Appendix XXX and focus on the model predictions in the main text. The model predictions for the mean best and worst choices are shown in 3.5.

The model clearly mispredicts the data. It predicts that target and competitor are chosen at the nearly the same rate for both best and worst choices. It fares better at predicting decoy choices but is still quantitatively off.

The target-competitor misprediction stems from the fact that the model is predicting the utility of each option through a linear combination of experimental factors, including target/competitor/decoy status. The model could, if the data suggest it, predict that the target has greater utility than the competitor or vice versa. However, because best-choice proportions are positively related to utility and worst-choice proportions are negatively related to utility, the model cannot simultaneously predict $B(C) > B(T)$ and $W(T) < W(C)$.

I also show participant-level predictions in 3.6. The generally does a poor job at accounting for participant worst-choice proportions, though it performs fair in the best choice

Dirichlet-Multinomial Modeling

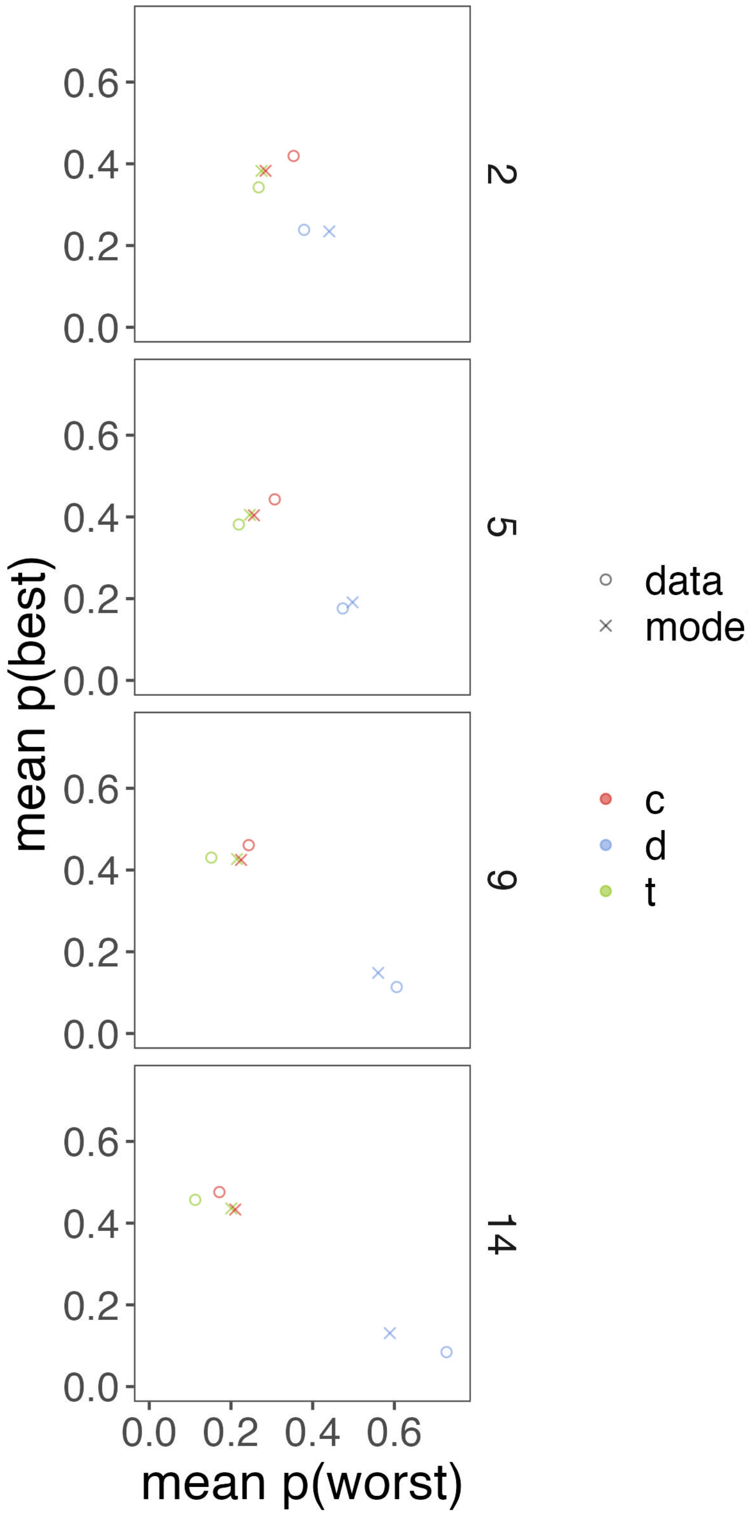

Figure 3.5. Maxdiff model predictions for the mean target, competitor, and decoy best-worst choice proportions.

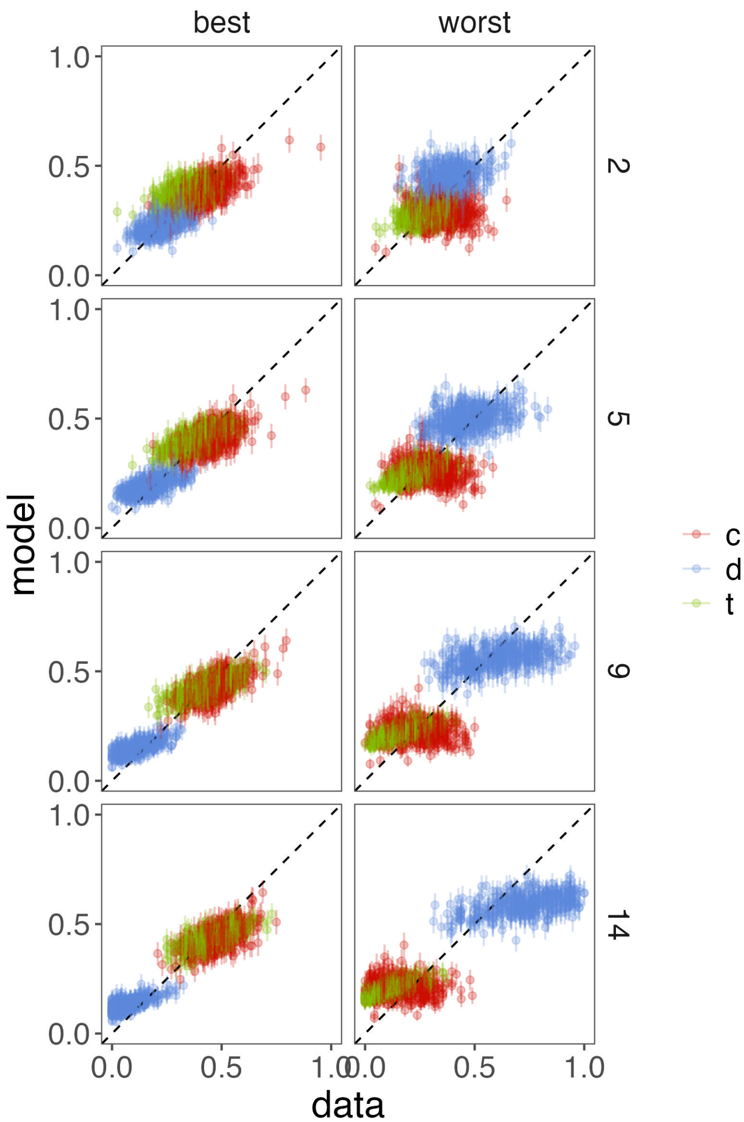

Figure 3.6. Maxdiff model predictions for the mean target, competitor, and decoy best-worst participant-level choice proportions, conditioned on TDD (rows) and choice type, i.e. best v. worst (columns). Vertical error bars are 95% HDIs.

BIBLIOGRAPHY

Adler, S. J., Schöniger, M. K., Lichters, M., & Sarstedt, M. (2024). Forty years of context effect research in marketing: A bibliometric analysis. *Journal of Business Economics*, *94*(3), 437–466.

Ashby, F. G., & Perrin, N. A. (1988). Toward a Unified Theory of Similarity and Recognition.

Ashby, F. G. (2019). State-trace analysis misinterpreted and misapplied: Reply to stephens, matzke, and hayes (2019). *Journal of Mathematical Psychology*, *91*, 195–200.

Ashby, F. G., & Bamber, D. (2022). State trace analysis: What it can and cannot do. *Journal of Mathematical Psychology*, *108*, 102655.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 33.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*(2), 154–179. https://doi.org/10.1037/0033-295X.93.2.154

Banerjee, P., Hayes, W. M., Chatterjee, P., Masters, T., Mishra, S., & Wedell, D. H. (2024). Factors that promote the repulsion effect in preferential choice. *Judgment and Decision Making*, *19*, e11. https://doi.org/10.1017/jdm.2023.46

Beck, M. J., & Rose, J. M. (2016). The best of times and the worst of times: A new best–worst measure of attitudes toward public transport experiences. *Transportation Research Part A: Policy and Practice*, *86*, 108–123.

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological review*, *116*(1), 84.

Bergner, A. S., Oppenheimer, D. M., & Detre, G. (2019). VAMP (Voting Agent Model of Preferences): A computational model of individual multi-attribute choice. *Cognition*, *192*, 103971. https://doi.org/10.1016/j.cognition.2019.05.008

Berkowitsch, N. A. J., Scheibehenne, B., & Rieskamp, J. (2014). Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, *143*(3), 1331–1348. https://doi.org/10.1037/a0035159

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, *120*(3), 522–543. https://doi.org/10.1037/a0032457

Bhui, R., & Xiang, Y. (2021). A rational account of the repulsion effect.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. https://doi.org/10.1016/j.cogpsych.2007.12.002

Cai, X., & Pleskac, T. J. (2023). When alternative hypotheses shape your beliefs: Context effects in probability judgments. *Cognition*, *231*, 105306. https://doi.org/10.1016/j.cognition.2022.105306

Cataldo, A. M., & Cohen, A. L. (2019). The comparison process as an account of variation in the attraction, compromise, and similarity effects. *Psychonomic Bulletin & Review*, *26*(3), 934–942. https://doi.org/10.3758/s13423-018-1531-9

Cataldo, A. M., & Cohen, A. L. (2021). Modeling Preference Reversals in Context Effects over Time. *Computational Brain & Behavior*, *4*(1), 101–123. https://doi.org/10.1007/s42113-020-00078-8

Cheung, K. L., Wijnen, B. F., Hollin, I. L., Janssen, E. M., Bridges, J. F., Evers, S. M., & Hiligsmann, M. (2016). Using best–worst scaling to investigate preferences in health care. *Pharmacoeconomics*, *34*, 1195–1209.

Davis-Stober, C. P., Marley, A., McCausland, W. J., & Turner, B. M. (2023). An illustrated guide to context effects. *Journal of Mathematical Psychology*, *115*, 102790.

Davis-Stober, C. P., Morey, R. D., Gretton, M., & Heathcote, A. (2016). Bayes factors for state-trace analysis. *Journal of mathematical psychology*, *72*, 116–129.

De Leeuw, J. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*, 1–12.

Doyle, J. R., O'Connor, D. J., Reynolds, G. M., & Bottomley, P. A. (1999). The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases. *Psychology and Marketing*, *16*(3), 225–243. https://doi.org/10.1002/(SICI)1520-6793(199905)16:3⟨225::AID-MAR3⟩3.0.CO;2-X

Evans, N. J., Holmes, W. R., Dasari, A., & Trueblood, J. S. (2021). The impact of presentation order on attraction and repulsion effects in decision-making. *Decision*, *8*(1), 36–54. https://doi.org/10.1037/dec0000144

Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin & Review*, *26*(3), 901–933. https://doi.org/10.3758/s13423-018-1557-z

Finn, A., & Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy & Marketing*, *11*(2), 12–25.

Flynn, T. N., Louviere, J. J., Peters, T. J., & Coast, J. (2007a). Best–worst scaling: What it can do for health care research and how to do it. *Journal of health economics*, *26*(1), 171–189.

Flynn, T. N., Louviere, J. J., Peters, T. J., & Coast, J. (2007b). Best–worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics*, *26*(1), 171–189. https://doi.org/10.1016/j.jhealeco.2006.04.002

Flynn, T. N., & Marley, A. A. (2014). Best-worst scaling: Theory and methods. In *Handbook of choice modelling* (pp. 178–201). Edward Elgar Publishing.

Frederick, S., Lee, L., & Baskin, E. (2014). The Limits of Attraction. *Journal of Marketing Research*.

Haaijer, R., Wedel, M., Vriens, M., & Wansbeek, T. (1998). Utility covariances and context effects in conjoint mnp models. *Marketing Science*, *17*(3), 236–252.

Hausman, J., & McFadden, D. (1984). Specification tests for the multinomial logit model. *Econometrica: Journal of the econometric society*, 1219–1240.

Hausman, J. A., & Wise, D. A. (1978). A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica: Journal of the econometric society*, 403–426.

Hautus, M. J., Macmillan, N. A., & Creelman, C. D. (2021). *Detection theory: A user's guide*. Routledge.

Hawkins, G. E., Marley, A. A. J., Heathcote, A., Flynn, T. N., Louviere, J. J., & Brown, S. D. (2014). The best of times and the worst of times are interchangeable. *Decision*, *1*(3), 192–214. https://doi.org/10.1037/dec0000012

Hawkins, G. E., Marley, A., Heathcote, A., Flynn, T. N., Louviere, J. J., & Brown, S. D. (2014). Integrating Cognitive Process and Descriptive Models of Attitudes and Preferences. *Cognitive Science*, *38*(4), 701–735. https://doi.org/10.1111/cogs.12094

Hayes, W. M., Holmes, W. R., & Trueblood, J. S. (2024). Attribute commensurability and context effects in preferential choice. *Psychonomic Bulletin & Review*, 1–12.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, *13*(3), 575–589. https://doi.org/10.1016/S0176-2680(97)00020-7

Highhouse, S. (1996). Context-Dependent Selection: The Effects of Decoy and Phantom Job Candidates. *Organizational Behavior and Human Decision Processes*, *65*(1), 68–76. https://doi.org/10.1006/obhd.1996.0006

Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research*, *9*(1), 90. https://doi.org/10.1086/208899

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of consumer research*, *10*(1), 31–44.

Liao, J., Chen, Y., Lin, W., & Mo, L. (2021). The influence of distance between decoy and target on context effect: Attraction or repulsion? *Journal of Behavioral Decision Making*, *34*(3), 432–447. https://doi.org/10.1002/bdm.2220

Marini, M., Ansani, A., & Paglieri, F. (2020). Attraction comes from many sources: Attentional and comparative processes in decoy effects. *Judgment and Decision Making*, *15*(5), 704–726. https://doi.org/10.1017/S1930297500007889

Marley, A., Flynn, T. N., & Louviere, J. (2008). Probabilistic models of set-dependent and attribute-level best–worst choice. *Journal of Mathematical Psychology*, *52*(5), 281–296. https://doi.org/10.1016/j.jmp.2008.02.002

Marley, A., & Louviere, J. (2005a). Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology*, *49*(6), 464–480. https://doi.org/10.1016/j.jmp.2005.05.003

Marley, A., & Louviere, J. (2005b). Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology*, *49*(6), 464–480. https://doi.org/10.1016/j.jmp.2005.05.003

Marley, A., & Pihlens, D. (2012). Models of best–worst choice and ranking among multiattribute options (profiles). *Journal of Mathematical Psychology*, *56*(1), 24–34. https://doi.org/10.1016/j.jmp.2011.09.001

McFadden, D. (2001). Economic choices. *American economic review*, *91*(3), 351–378.

Mohr, P. N., Heekeren, H. R., & Rieskamp, J. (2017). Attraction effect in risky choice can be explained by subjective distance between choice alternatives. *Scientific reports*, *7*(1), 8942.

Molloy, M. F., Galdo, M., Bahg, G., Liu, Q., & Turner, B. M. (2019). What's in a response time?: On the importance of response time measures in constraining models of context effects. *Decision*, *6*(2), 171–200. https://doi.org/10.1037/dec0000097

Morey, R. D., et al. (2008). Confidence intervals from normalized data: A correction to cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

Newell, B. R., & Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in cognitive sciences*, *12*(8), 285–290.

Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, *125*(4), 512–544. https://doi.org/10.1037/rev0000102

Pan, Y., O'Curry, S., & Pitts, R. (1995). The attraction effect and political choice in two elections. *Journal of Consumer Psychology*, *4*(1), 85–101.

Pittarello, A., Caserotti, M., & Rubaltelli, E. (2020). 'three is better than two': Increasing donations with the attraction effect. *British Journal of Psychology*, *111*(4), 805–822.

Ray, P. (1973). Independence of irrelevant alternatives. *Econometrica: Journal of the Econometric Society*, 987–991.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, *108*(2), 370–392. https://doi.org/10.1037/0033-295X.108.2.370

Sadil, P., Cowell, R., & Huber, D. E. (2018). A hierarchical bayesian state trace analysis for assessing monotonicity while factoring out subject, item, and trial level dependencies.

Schwartz, J. A., & Chapman, G. B. (1999). Are more options always better? the attraction effect in physicians' decisions about medications. *Medical Decision Making*, *19*(3), 315–323.

Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. *Science*, *237*(4820), 1317–1323. https://doi.org/10.1126/science.3629243

Simonson, I. (2014). Vices and Virtues of Misguided Replications: The Case of Asymmetric Dominance. *JOURNAL OF MARKETING RESEARCH*.

Slaughter, J. E., Sinar, E. F., & Highhouse, S. (1999). Decoy effects and attribute-level inferences. *Journal of Applied Psychology*, *84*(5), 823–828. https://doi.org/10.1037/0021-9010.84.5.823

Spektor, M. S., Kellen, D., & Hotaling, J. M. (2018). When the Good Looks Bad: An Experimental Exploration of the Repulsion Effect. *Psychological Science*, *29*(8), 1309–1320. https://doi.org/10.1177/0956797618779041

Spektor, M. S., Kellen, D., & Klauer, K. C. (2022). The repulsion effect in preferential choice and its relation to perceptual choice. *Cognition*, *225*, 105164. https://doi.org/10.1016/j.cognition.2022.105164

Stephens, R. G., Matzke, D., & Hayes, B. K. (2020). State-trace analysis—misrepresented and misunderstood: Reply to ashby (2019). *Journal of Mathematical Psychology*, *96*, 102342.

Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences*, *54*(2), 407–411. https://doi.org/10.1073/pnas.54.2.407

Team, T. O. D. (2019). Gnu octave. https://www.gnu.org/software/octave/

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*(2), 273–286.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Trueblood, J. S. (2012). Multialternative context effects obtained using an inference task. *Psychonomic Bulletin & Review*, *19*(5), 962–968. https://doi.org/10.3758/s13423-012-0288-9

Trueblood, J. S., Brown, S. D., & Heathcote, A. (n.d.). The Multiattribute Linear Ballistic Accumulator Model of Context Effects in Multialternative Choice.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, *24*(6), 901–908.

Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, *125*(3), 329–362. https://doi.org/10.1037/rev0000089

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

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management science*, *39*(10), 1179–1189.

Usher, M., & McClelland, J. L. (2004). Loss Aversion and Inhibition in Dynamical Models of Multialternative Choice. *Psychological Review*, *111*(3), 757–769. https://doi.org/10.1037/0033-295X.111.3.757

van den Enden, G., & Geyskens, K. (2021). Attract the best: The attraction effect as an effective strategy to enhance healthy choices. *Plos one*, *16*(11), e0259521.

Wollschläger, L. M., & Diederich, A. (2012). The 2N-ary Choice Tree Model for N-Alternative Preferential Choice. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00189

Yearsley, J. M., Pothos, E. M., Barque-Duran, A., Trueblood, J. S., & Hampton, J. A. (2022). Context effects in similarity judgments. *Journal of Experimental Psychology: General*, *151*(3), 711–717. https://doi.org/10.1037/xge0001097