

Context Dependencies in Decision Making



Tsvetomira Dumbalska

Pembroke College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2021

Abstract

Irrelevant information should not affect our decisions. Yet, our choices are often swayed by contextual input that is entirely unrelated to the decision at hand. This is true for perceptual judgments pertaining to the appearance of a visual stimulus, as well as for more abstract evaluations of economic prospects. My thesis examines three specific instances of context-dependent behavior across the realms of perception and economics. Across all studies, I charted the behavioral signatures of context-dependent choices via careful experimental manipulation. I combined this behavioral data with simulations of mathematical models to arbitrate between different theoretical accounts of the mechanisms underlying context-dependent decision making. My first study investigated how the context provided by an additional (decoy) alternative sways economic choice between two target options. It mapped decoy influence across the full attribute space of possible decoy definitions. My second study examined the influence that an irrelevant distractor stimulus wields on perceptual decisions. It combined psychophysical measurement with regression-based and reverse correlation analytic approaches to shed light on the functional form of distractor effects. Finally, my third study probed the mechanisms driving context-dependent categorization decisions. It measured the contributions of putative drivers of contextual influence combining novel analytic techniques with experimental design controlling for potential confounds. Across the three studies, normalization-based information processing schemes emerged as a common theme. Viewed through the normative lens of efficient neural coding, normalization provides a compelling account for the observed patterns of context-dependent choice behavior.

Contents

List of Figures	v
1 Introduction	1
1.1 Optimal Decisions	2
1.2 Context Dependencies	6
1.3 The Underpinnings of Context-Dependent Choice	19
1.4 Thesis Outline	24
2 A map of decoy influence in multialternative economic choice	25
2.1 Introduction	26
2.2 Experiment	29
2.3 Computational Modeling	40
2.4 Discussion	55
2.5 Conclusion	58
2.6 Acknowledgements	58
3 How do perceptual distractors distract?	59
3.1 Introduction	60
3.2 Experiment 1	64
3.3 Experiment 2	78
3.4 Experiment 3	86
3.5 Computational Modeling	95
3.6 Discussion	98
3.7 Conclusion	102
3.8 Acknowledgements	103
4 Context-dependent categorization	104
4.1 Introduction	105
4.2 Experiment 1	110
4.3 Experiment 2	122

Contents

4.4	Discussion	134
4.5	Conclusion	138
5	Discussion	139
5.1	Overview of Results	140
5.2	Interpretation	141
5.3	Zooming Out	142
5.4	Conclusion	145
	References	147

List of Figures

1.1	Psychometric function	3
1.2	Color adaptation	9
1.3	Color constancy	11
2.1	Decoy effects	28
2.2	Experimental design	32
2.3	Conventional decoy analyses	36
2.4	Map of decoy influence	37
2.5	Map decomposition	38
2.6	Map simulations	42
2.7	New decoy relationship	44
2.8	Model compression approximation	49
2.9	T-SNE plot of model maps	51
2.10	Normalization with compressed inputs	54
3.1	Experiment 1, Trial structure	66
3.2	Reverse correlation analysis, Stimulus energy	69
3.3	Reverse correlation analysis, Decision kernels	71
3.4	Experiment 1, Stepwise regression results	73
3.5	Experiment 1, Regression results	74
3.6	Experiment 1, Decision kernel results	75
3.7	Experiment 1, Kernal decomposition results	77
3.8	Experiment 2, Trial structure	81
3.9	Experiment 2, Regression results	82
3.10	Experiment 2, Decision kernel results	83
3.11	Experiment 2, Kernel decomposition results	84
3.12	Experiment 3, Trial structure	87
3.13	Experiment 3, Attention manipulation check	89
3.14	Experiment 3, Regression results	90

List of Figures

3.15	Experiment 3, Decision kernel results	92
3.16	Experiment 3, Kernel decomposition results	93
3.17	Experiment 3, Results per cueing condition	94
3.18	Computational model, Simulations	96
3.19	Computational model, Fits to human data	97
4.1	Context-dependent categorization across domains	106
4.2	Sources of context dependence	108
4.3	Experiment 1, Trial structure	114
4.4	Experiment 1, Psychometric functions, part 1	118
4.5	Experiment 1, Psychometric functions, part 2	119
4.6	Experiment 1, Threshold values	120
4.7	Experiment 1, Regression of responses on stimulus history	121
4.8	Experiment 2, Trial structure	126
4.9	Experiment 2, Psychometric functions and perceptual scales	131
4.10	Experiment 2, Model results	133

1

Introduction

Contents

1.1	Optimal Decisions	2
1.1.1	Optimality in Perception	2
1.1.2	Optimality in Economics	4
1.2	Context Dependencies	6
1.2.1	Perceptual Decisions	6
1.2.2	Economic Decisions	12
1.2.3	Parallels	16
1.3	The Underpinnings of Context-Dependent Choice . . .	19
1.3.1	Neural Constraints & Efficiency	20
1.3.2	Behavioral Relevance	22
1.4	Thesis Outline	24

What makes a good decision? A common theme among the various answers provided over the centuries by writers, philosophers, and social scientists argues that a good decision accurately evaluates all of the relevant information, ignoring any irrelevant or distracting input. Yet, human choices often deviate from this compelling principle.

Our judgments about a stimulus are often swayed by irrelevant information in the context in which the decision occurs. In some settings, we might consistently judge the stimulus as large or valuable, while in others, that same stimulus might appear small and unappealing. How do these stereotyped patterns of context-dependent

1. Introduction

choice behavior come about? My thesis addresses this question across the realms of perception and economics. This chapter outlines normative views about optimal decisions from both disciplines, reviews established patterns of context-dependent decision making, and considers the underpinnings of contextual choice biases.

1.1 Optimal Decisions

Optimal decisions in perception maximize the likelihood of a correct evaluation of a sensory feature; optimal decisions in economics maximize the decision maker's utility. Despite these differences in definition, classic normative theories across both disciplines posit that optimal choice behavior should produce a monotonic function linking the relevant feature(s) of a target stimulus and the probability of choosing it.

1.1.1 Optimality in Perception

In perception research, human participants are typically presented with sensory stimuli, such as weighted objects, tilted lines or auditory tones, and are asked to compare or categorize them. In those psychophysical experiments, there is usually a clear-cut correct answer – the experimenter can measure the objective sensory value of each stimulus and determine the larger, more tilted or higher-pitched option. Human choices are usually stochastic, or randomly variable, in a way that is dependent on the true stimulus magnitude. The mapping of feature magnitude on participant choice is quantified in the psychometric function. While perfect performance would produce a step function, empirically measured psychometric functions are typically smooth (Fig. 1.1). That is, there is no threshold magnitude above which a stimulus would always be judged as big, and below which as small.

This characteristic stochasticity of perception posed a puzzle for earlier theories in the field; the advent of signal detection theory (Green and Swets 1966) ushered in modern thinking about it. Smooth psychometric curves can be explained by the assumption that perceptual choices are based on an internal representation $\hat{x}$ of the sensory stimulus x that is corrupted by encoding noise. As suggested by

1. Introduction

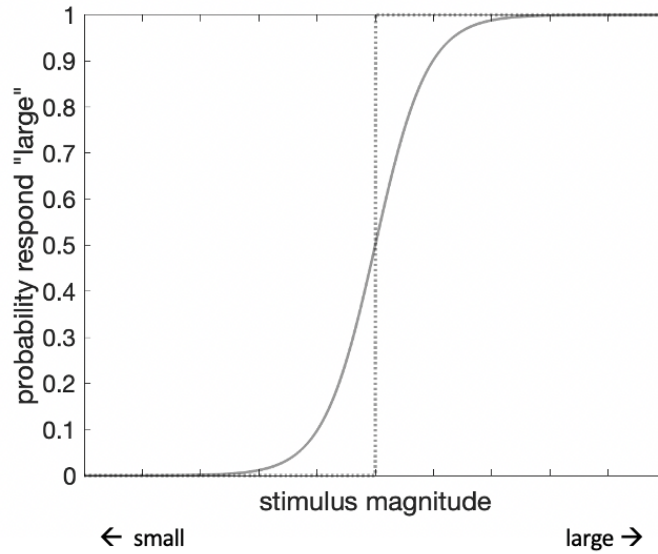


Figure 1.1: The psychometric function. As an example, an observer may be presented with series of objects of different sizes and asked to categorize each of them as large or small. The x-axis maps the size of each object in the experiment. The y-axis tracks the probability that the observer would categorize a given object as large. The dotted line illustrates a step-wise psychometric function, where all objects above a certain size are always categorized as large and all objects below that size are always categorized as small. By contrast, the solid curve illustrates a typical psychometric function estimated from empirical data.

its name, encoding noise arises early on in the processing of a sensory stimulus owing to the intrinsic randomness of neural processes or intrinsic limits on the observability of the stimulus. It is often assumed that encoding noise follows a Gaussian distribution. Consequently, sensory representations may be captured by a random draw from a normal distribution whose mean corresponds to some monotonically increasing function of the magnitude of the true sensory feature and some measure of the variability in the representation:

$$\hat{x} \sim N(f(x), \nu^2) \quad (1.1)$$

The shape of the transformation $f(x)$ is classically assumed to follow a logarithmic compression (Weber-Fechner’s law for sensory systems), whereby sensitivity to changes in sensory inputs diminishes with the absolute magnitude of stimulation. This phenomenon is ubiquitous across perceptual modalities: the addition of 50g

1. Introduction

would be easy to notice if we were holding an object weighing 100g; however, we would fail to detect it if the object weighed 1kg. Similarly, we might easily notice the difference between the first and second spoonful of sugar in a cup of tea; noticing the difference between the ninth and tenth would, however, pose a challenge. This characteristic logarithmic compression of sensory magnitudes is an efficient strategy for the neural coding of percepts under the biological capacity constraints of the neural system (Sun et al. 2012; Wei and Stocker 2017).

Ideal observer models treat observer choices as inferences about the true stimulus magnitude based on the noisy internal representation ($\hat{x}$). An efficient decision rule is a likelihood ratio test against a decision criterion (Neyman-Pearson Lemma). This test compares the ratio of the conditional probability of the observed internal state occurring, given the possible true stimulus magnitudes, against a threshold value.

1.1.2 Optimality in Economics

In economic analysis, human behavior is explained via the incentives presented by the environment. However, the incentive values of different economic prospects or consumer goods depend on the idiosyncratic preferences of each individual. Thus, compared with psychophysical perception experiments, it is less straightforward to define what the right answer is in any choice situation. Nevertheless, rational choice theory posits that an individual's preferences ought to be transitively ordered and internally consistent.

It is classically assumed that economic decisions elicit decision makers' preferences, which are a fixed function of stimulus characteristics. In the case of uncertain economic prospects (e.g. gambles with probabilistic outcomes), the expected value is determined by the product of the probability of the event and its magnitude (the monetary outcome). This idea, first introduced by Pascal in the 17th century, provides a benchmark for optimal choice behavior. However, in certain situations, expected value theory leads to choices that appear irrational – take for instance a lottery ticket offering a 50% chance of winning £1 million and a 50% chance of winning nothing. The expected value of this gamble is £500 000 ($50\% \cdot £1$

1. Introduction

million), and so if someone were to offer to purchase the lottery ticket from us for £499 999, expected value theory dictates we should refuse. Intuitively, however, this choice seems unappealing. In terms of practical changes to our quality of life, the difference between having £499 999 and having nothing (can I afford basic necessities?) is much starker than the difference between having £499 999 and having £1 million (can I afford additional luxuries?).

This idea was formalized by Bernoulli with the example of the St Petersburg paradox (another hypothetical gamble, Dutka 1988), which demonstrates that the *utility* ($u(x)$, i.e. subjective value) we attach to economic prospects is a concave function of their expected value (e.g. the von Neumann-Morgenstern utility function, Von Neumann and Morgenstern 1944). Thus, the marginal utility of an additional pound decreases with the total amount of money we have. Expected utility theory has been further refined since to capture the specific rate of decrease of marginal utility as well as risk preferences of individuals (e.g. Prospect theory, Kahneman and Tversky 1979). Recent work has demonstrated that the shape of these distortions of expected value constitutes an efficient reward-maximizing strategy under the biological constraints to computational precision in the human brain (Juechems et al. 2021).

These biological constraints are further reflected in the stochastic nature of empirically derived utility functions (e.g. Luce 1959; Block and Marschak 1960; McFadden 1981). While traditionally, empirical studies on individuals' choice behavior in economics have used one-off hypothetical scenarios or vignettes (e.g. Kahneman and Tversky 1979), studies where incentives are varied parametrically have mirrored findings about the stochasticity of choice behavior from perception research (e.g. Mosteller and Nogee 1951). Therefore, the empirically measured mapping of stimulus features on subjective utility is typically a smooth psychometric curve. The observed smooth, S-shaped psychometric functions can be captured by a stochastic utility representation, where the subjective value function is randomly variable like in sensory perception, $\hat{u}(x)$ (as in random utility models, McFadden 1981). It is usually assumed that, unlike in sensory systems, the source of noise does

1. Introduction

not stem from early processing stages (although noise might affect the encoding of both value, Bhui and Gershman 2018; and probability Steiner and Stewart 2016), but rather, from inaccurate retrievals from memory of the true stimulus utility to the individual (e.g. Polania et al. 2018). That is, there is no inaccuracy in the encoding of the identity and characteristics of the option under consideration. Instead, the mapping of the characteristics of the option onto choice is subject to random variability.

1.2 Context Dependencies

While these normative principles dictate that choices about a target stimulus should be a function of that stimulus alone, human decisions, across perception and economics alike, are often influenced by the context in which they occur.

1.2.1 Perceptual Decisions

Human perception is astonishingly context-dependent. A long tradition in perception research has considered that percepts, and the underlying neural responses forming them, are not simply a function of the focal input. Rather, they are a function of the focal input as well as the current background and recent history of stimulation (Webster 2015). The Weber-Fechner law described earlier offers one compelling example of the ubiquity of context dependence. Objectively the exact same stimulus can be more difficult to detect depending on the context in which it occurs. Consider the stars, for instance – they are practically invisible to us in the daytime under the bright illumination of sunlight, yet at night, in the relative darkness of moonlight, perceiving stars is much easier. Similarly, light pollution in cities, driven for instance by streetlights, makes stargazing a challenge compared to the darker nighttime settings available in the wilderness. Across all those scenarios our perception of the target stimulus (stars) is dramatically affected by the context (illumination provided by sunlight, moonlight or artificial lighting). Below I consider other instances of the influence of temporal or spatial context on our perception.

1. Introduction

Temporal Context

The visual system continuously adapts to the temporal context to ensure that it is maximally sensitive to changes in the inputs that are most likely to occur (i.e. those close to the mean levels of environmental stimulation). *Adaptation* to contextual levels of stimulation occurs continuously, at multiple time-scales and across multiple domains. Adaptation to light, for instance, can occur at the timescale of seconds and minutes, when moving from outdoors on a sunny day to a dimly lit room, but it also occurs across seasons as the colors dominating our natural environment change (Welbourne et al. 2015), and over the course of the lifetime as the increasing pigment density of the lens in the human eye modifies the spectral composition of light reaching the retina (Webster 2015). The visual system adapts to local low-level stimulus features such as lightness, color and tilt, as well as to more complex, spatially extended characteristics, such as spatial frequency and contrast, and even to composite characteristics, processed at higher levels of the visual hierarchy such as facial attributes and expressions (Webster 2011).

Adaptive effects are canonically repulsive. As the sensory system adapts to the contextual level of stimulation (i.e. the adaptation level), our experience of incoming stimuli is repelled away from it. Thus, the term repulsion refers to the relationship between our experience of the current stimulus and the properties of context, rather than the interaction between context and stimulus. This independent effect of context means that the same stimulus may be experienced as brighter (or larger) if it were presented following dim (or small) stimuli relative to how it would be experienced when presented following bright (or large) stimuli. Contextual repulsion is typically accompanied by a stimulus-context interaction. The more similar the stimulus to the current adaptation level, the higher the sensitivity of our visual system to changes in the stimulus. That is, small changes in stimulus level close to the contextual expectation are exaggerated.

Computationally, adaptation can be expressed as a shift in the transducer function, linking sensory inputs to neural responses, towards the mean of contextual stimulation (Fig. 1.2). The horizontal translation of the transducer illustrates the

1. Introduction

repulsive influence of context (i.e. the independent effect of context). This shift, however, also means that the slope of the transducer is steepest around the current environmental expectation, so the neural system is most sensitive to those signals that are most likely to occur (i.e. the stimulus-context interaction effect). This is an efficient strategy for neurons with limited dynamic range, as it prevents response saturation and maximizes the neural resources dedicated to the processing of the most common signals (Carandini and Heeger 2012).

Standard place-coding process models of adaptation appealing to this normative principle propose that adaptation impacts the tuning curves of feature-selective populations of neurons (channels), diminishing neural responses to features at (and close to) the adaptation level (e.g. for perceptual properties including color contrast, spatial frequency or viewpoint, Webster 2011). These diminished responses effectively equate sensitivity across the different channels, such that their average responses are similar within the current context. This outcome can also be achieved through a rate code, where the level of neural responsiveness conveys the stimulus identity (e.g. for perceptual properties including saturation, blurriness or facial distortion, Webster 2011). Rate-based models of adaptation propose that context neutralizes the intensity of the response, shifting the stimulus level which corresponds to the norm or null point neural response.

A similar *temporal repulsion effect* is also observed in studies where participants are asked to categorize stimuli according to a perceptual feature (e.g. as “big” or “small” according to their size, or as “heavy” or “light” according to their weight). Regardless of whether participants are given an explicit category boundary (or criterion) against which to compare stimuli (e.g. Lages and Treisman 1998; Olkkonen et al. 2014) or not (e.g. Helson 1947; Morgan et al. 2000; Levvari et al. 2018), their choices drift over the course of the experimental session such that the empirical decision boundary effectively tracks the central tendency of the stimulus set. Thus, the same stimulus can be categorized as big or small depending on what other (irrelevant) stimuli the decision maker has recently been exposed to. Beyond the mean, other characteristics of the distribution of stimulus magnitudes,

1. Introduction

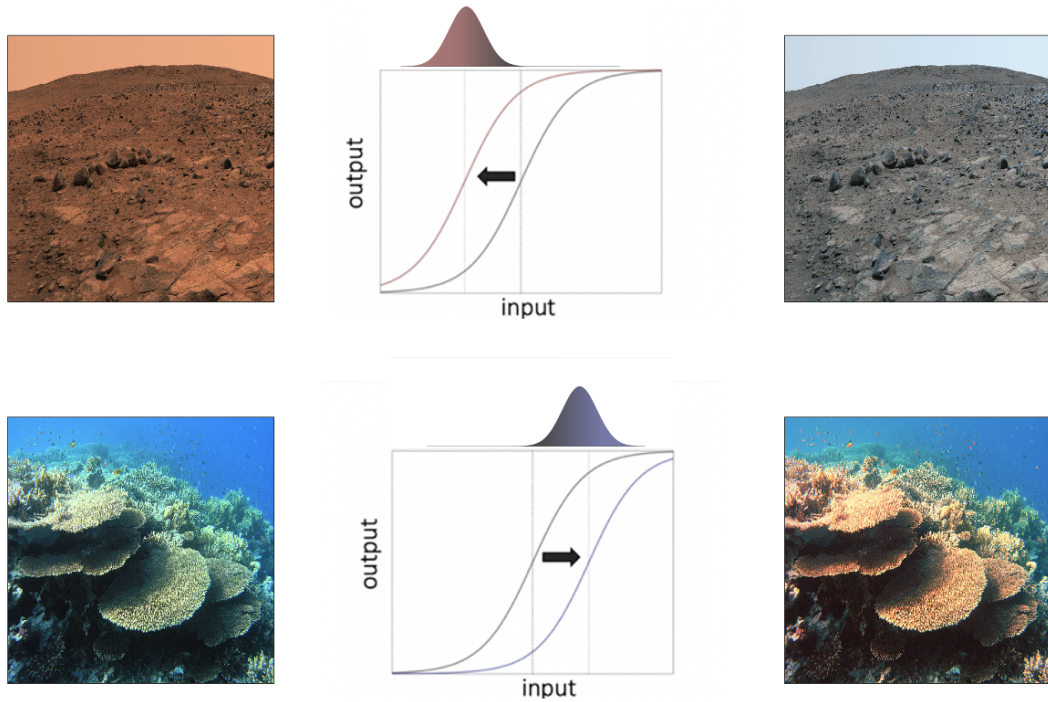


Figure 1.2: Color adaptation. The rightmost column illustrates how the surface of Mars and the sea would appear to a human observer adapted to the landscape of Earth. In the Martian scene, the landscape is dominated by red; in the sea scene – by blue. The middle column illustrates a neural processing scheme which adaptively shifts the transducer function to track the mean of contextual stimulation (that is, red distribution for Mars; blue distribution for the sea). The leftmost column simulates the appearance of both scenes following adaptation. Both scenes appear more chromatically neutral (adaptive repulsion). Following adaptation to the Martian context, the visual system is fine-tuned to distinguish different shades of red; following adaptation to the sea context – sensitivity is boosted for change detection in blues. Figure concept adapted from Webster 2015.

such as its range and skewness, can also sway judgments about sensory stimuli (Parducci 1965). Computational accounts of the repulsive influence of context on categorization judgments have appealed to temporal adaptation (Helson 1964; Parducci 1965), whereby neural responses to incoming inputs over time are dampened by normalization by the temporal context. An alternative account likens the observed effect to a *contraction bias* in working memory (also known as central tendency bias, Hollingworth 1910) affecting the decision maker’s reference of the category boundary. This theoretical view proposes that categorization decisions are biased by a drift in the representation of the category boundary towards the prior (e.g. midpoint of the stimulus set, contextual expectation, etc., Treisman and Williams 1984;

1. Introduction

Jou et al. 2004; Olkkonen et al. 2014). This attractive contextual bias produces a repulsive pattern of behavior, as it affects the categorization criterion against which imperative stimuli are evaluated.

Temporal context effects are not necessarily always repulsive. Indeed, a large literature centers on attractive *serial dependencies* of choice (Kiyonaga et al. 2017). In psychophysical experiments, decisions about a focal stimulus are attracted towards the features of the immediately preceding stimulus, especially if there is a delay between stimulus presentation and response or perceptual ambiguity in the stimulus (Akaishi et al. 2014). This effect is modulated by the degree of similarity between the two sequential stimuli; the more similar the previous stimulus to the current one, the stronger the attraction (Fischer and Whitney 2014). Normative explanations of this effect appeal to its adaptive role promoting the stability of perception given the statistics of the natural environment (Kiyonaga et al. 2017). The visual world is typically stable over time, so sensory inputs to a given area of the retina are autocorrelated. Thus, in the presence of uncertainty, a Bayesian optimal observer would use previous inputs to resolve the ambiguity (Ashourian and Loewenstein 2011; Olkkonen and Allred 2014). In support of this interpretation, previous work has shown that the stability of environmental statistics can reverse the sign of the sequential dependency effect. For more stable characteristics, such as recognizing the gender of a face, choices are positively associated with the gender of the preceding face, maximizing the stability of perceptions. Conversely, for more changeable characteristics, such as facial expression, choices are negatively associated with the expression of the preceding face, maximizing change detection (Taubert et al. 2016).

Spatial Context

Some of the most compelling demonstrations of the influence of spatial context come from the domain of color perception, where colors in the spatial vicinity of a focal stimulus can dramatically alter its color appearance. Fig. 1.3 illustrates how the same physical color (i.e. spectrum of wavelengths emanated from a given region) can be perceived differently depending on what colors surround it. This dependence

1. Introduction

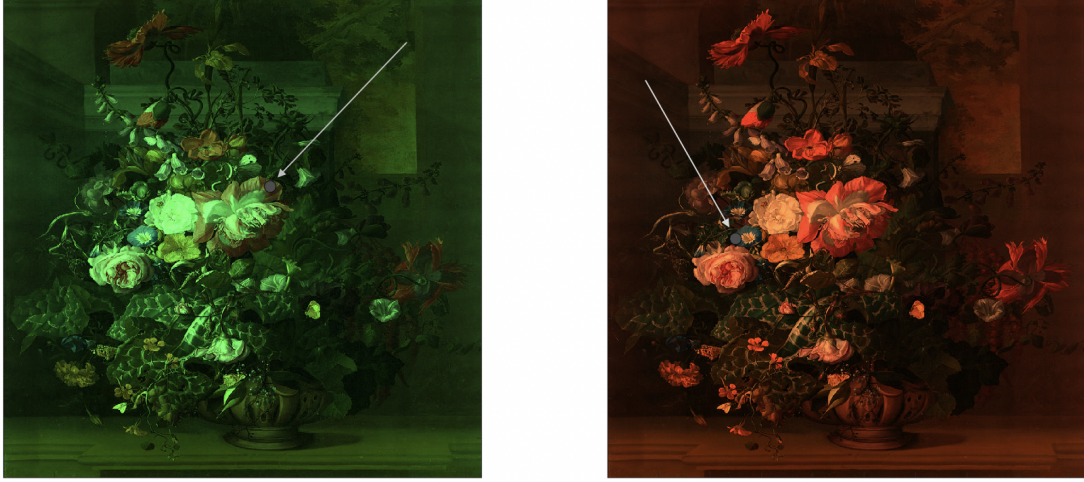


Figure 1.3: Color constancy. On the painting on the left, the circled region on the demarcated flower petal appears red, but is actually gray. On the painting on the right, the circled region on the demarcated flower petal appears blue, but is also gray. Painting by Rachel Ruysch, ca. 1680.

on spatial context has an adaptive purpose – it gives rise to *color constancy* or our ability to experience the color of an object (i.e. the perceptual correlate of spectral reflectance) as unchanging despite changes in the spectral composition of the illuminant. Thus, processing color signals in a relative manner, rather than in isolation, allows the neural system to solve the otherwise mathematically intractable task of interpreting the light signals reaching the retina in a way that separates an object’s surface color from the color of lighting (Smithson 2005). While the specific computational mechanisms underlying this phenomenon are still the topic of debate, a popular theme across proposals suggests that cone cell excitation signals in the retina are normalized by a quantity set by the context (Land 1983; Foster 2011).

Spatial dependencies in visual processing have been documented across a variety of other domains, including the perception of orientation and contrast. A classic example is the *tilt illusion* (Blakemore et al. 1970) where the tilt of a central grating is repulsed from the tilt of flanking gratings. Thus, a central horizontally oriented stimulus would appear tilted clockwise, when surrounded by anticlockwise flankers. Similarly, spatial context repels contrast perception. A central texture appears of lower contrast when surrounded by a high contrast texture compared to when

1. Introduction

surrounded by a uniform background (Chubb et al. 1989; Xing and Heeger 2001). The repulsive influences of context across visual attributes have been successfully modeled in a normalization framework, whereby neural responses to target inputs are normalized by the features of context (Carandini and Heeger 2012).

Another example of spatial dependencies in perception comes from the literature on *lateral interactions* (Stemmler et al. 1995). Psychophysical and electrophysiological evidence converges on a pattern of *contextual facilitation* in visual processing whereby the presence of proximal stimuli lowers thresholds for the detection of (Kapadia et al. 1995; Freeman et al. 2001; Freeman and Driver 2005) and boosts neural responses to (Kapadia et al. 1995; Ito and Gilbert 1999) a collinear target stimulus. Crucially, this contextual effect can reverse its direction. When target stimuli are close to detection thresholds, flanking stimuli often facilitate processing (Polat and Sagi 1993). For high-contrast targets, on the other hand, lateral interactions are, generally, of a suppressive nature (Stemmler et al. 1995). These characteristics of lateral interactions have been linked with the behavioral relevance of visual context in the real world. Context can be helpful to identify and interpret stimuli near detection thresholds (Gilbert et al. 2000). For example, neighboring visual inputs may be useful when the visual system is trying to make sense of blurry or interrupted lines on a faded photograph. Ambiguity is, however, lower for suprathreshold stimuli, so the visual system need not rely on contextual clues for stimulus identification. Instead, *contextual inhibition* may aid the quick detection of changes in the visual environment through pop out effects for inconsistent stimuli (Stemmler et al. 1995).

1.2.2 Economic Decisions

Over the past 40 years, the field of behavioral economics has cataloged instances of “suboptimal” context-dependent economic behavior (Thaler and Ganser 2015). From the standpoint of classical economics, these phenomena constituted “anomalies” from rational behavior. Despite the many commonalities between context dependence in perception and economics, the interpretations of the observed patterns of contextual

1. Introduction

influence have been starkly different between disciplines. In perception research, context dependencies have generally been interpreted in an adaptive light, allowing the perceptual system to take advantage of regularities in the environment to enhance efficiency in sensory processing (e.g. normalization-based neural coding) and resolve perceptual ambiguity (e.g. Bayesian inference). In economics, they have predominantly been considered as instances of irrationality (Kahneman et al. 1991) and normative explanations of context-dependent choice behavior have largely been overlooked till fairly recently (Loewenstein et al. 2008).

Temporal Context

Economic choices regarding monetary rewards are characterized by repulsion effects akin to the temporal context biases documented in perceptual categorization discussed above. When the reward context (as defined by the mean of rewards experienced in previous trials) changes, so too does a participant’s subjective valuation of rewards (as estimated by their choice behavior) shift. In a high-reward context (e.g. average reward = £5), participants generally value the same monetary sum (e.g. £5) less than they do in a low-reward context (e.g. when average reward = £3, Rigoli et al. 2016). In line with the observed *context-dependent valuation* behavior of humans, electrophysiological evidence from non-human primates (Tremblay and Schultz 1999; Padoa-Schioppa 2009) and fMRI evidence from humans (Cox and Kable 2014) in tasks where the reward context is manipulated suggests that value is not encoded in an absolute fashion. Rather, neural value signals are consistent with an encoding scheme whereby stimulus value is processed relative to the recent reward history (Rangel and Clithero 2012). Context-dependent subjective value decisions are affected by the mean (Rigoli et al. 2016), variability (Rigoli et al. 2016), and skewness (Stewart et al. 2006) of the distribution of previously experienced rewards in a direct parallel with context-dependent perceptual categorization (Parducci 1965). Accordingly, popular computational accounts of this effect appeal to normalization, whereby target values are divisively normalized by a summary characteristic of the distribution of

1. Introduction

contextual levels of stimulation, such as the average (as in divisive normalization, Louie et al. 2013), average and variability (as in the logistic model of subjective value, Rigoli 2019), or range (as in range normalization, Padoa-Schioppa 2009). A more extreme theory of relative value, decision by sampling, argues that the subjective value of a stimulus is entirely driven by the stimulus’ ranked order within the context (Stewart et al. 2006; Heng et al. 2020).

Outside of the lab, this temporal repulsion effect has been observed in the prices of properties rented by individuals who have recently moved to a new city. In line with reference-dependent valuation, those arriving from relatively cheaper cities rent cheaper properties compared to those arriving from more expensive cities, regardless of movers’ objective wealth (Simonsohn and Loewenstein 2006). In the marketing literature, a related concept is the so called *perceptual contrast* (Lynch et al. 1991; Tormala and Petty 2007; Lee and Suk 2010), where a consumer’s evaluation of a target good is repulsed away from their evaluation of a decoy good with which they were presented immediately before this. For instance, a retail agent might first show a prospective client a very unattractive property in order to boost the subjective value of a subsequent target property for the client (Cialdini 1987). Naive home buyers are far from the only victims of this bias. Experienced financial professionals are also swayed by contrast effects, such that investors’ evaluations of today’s earnings are inversely proportional to their balance from the previous day (Hartzmark and Shue 2018).

Conversely, *anchoring* refers to the systematic assimilation of subjective evaluations of a target stimulus to an immediately preceding, unrelated and arbitrary reference magnitude (Tversky and Kahneman 1974). The induction of an anchoring effect typically requires a comparison between the target evaluation and the anchor. For instance, Ariely and colleagues (2003) first asked a cohort of MBA students if they would be willing to purchase a variety of consumer goods, such as wine bottles, computer accessories or books, for a price equal to the last two digits of their social security number (the anchor). Then, the researchers asked participants to report the maximum price they would be willing to pay for each good. Value

1. Introduction

estimates were strongly influenced by the arbitrary anchor; higher anchors inflated the reported price, while lower anchors diminished it. Anchoring has been observed across a variety of domains, including economic valuation, general knowledge and factual estimates, probability estimates, legal judgments, and negotiations (Furnham and Boo 2011).

Spatial Context

A classic example of the influence of spatial context in economics is the *decoy effect*. It refers to the finding that consumer choices between two equally preferred options, which vary along two dimensions, can be swayed by the addition of a third, generally inferior, alternative. The introduction of a dispreferred, and hence termed decoy, option should not in theory affect the choice between two better alternatives (Block and Marschak 1960). However, value judgments are systematically affected by decoys. First described in 1982 (Huber et al. 1982), the *asymmetric dominance* (or attraction) effect is produced by a decoy good which is dominated asymmetrically, that is, it is objectively worse than one of the two target options, but not the other. Consumer choice is attracted towards the dominating alternative. For instance, when choosing between a cheap but slow computer and a pricey but fast one, the addition of an alternative which is slightly more expensive than the cheap one yet just as slow boosts preference for the cheap computer.

In a similar vein, a decoy which is similar to one of the two alternatives repels choice away from it (*similarity effect* Tversky 1972b) and a decoy which is more extreme than one of the options, but equally preferred, sways choice towards it (*compromise effect*, Simonson 1989). All three classic decoy effects have been replicated in the domain of perceptual decisions, using judgments about the area of rectangle stimuli of varying height and width (Trueblood et al. 2013). This finding suggests that the decoy effect is likely driven by a general principle in decision processing across stimulus domains. In line with this explanation, the most recent computational accounts of the decoy effect have proposed normalization in the neural coding of decision inputs as the driver of the observed context

1. Introduction

dependencies (e.g. Rigoli et al. 2017; Bushong et al. 2021; Daviet and Webb 2021; Landry and Webb 2021).

Another finding in multialternative choice concerns *value distraction* (Louie et al. 2013). It occurs when choosing between two items, one of which is preferred by the decision-maker, such that $v(A) > v(B)$. The value of a third option, inferior to the two original alternatives ($v(D) < v(B) < v(A)$), affects how often the option with the subjectively highest value is chosen. The probability that both humans and non-human primates would select option A over B falls as the value of the distractor option grows (in speeded choices as in Louie et al. 2013) and reaction times rise (in slower settings in humans as in Gluth et al. 2020). This context dependence can be explained in a normalization framework, where choice values are divisively normalized by the (weighted) sum of context. As the distractor value rises, so does the numerator in the normalization procedure, lowering sensitivity to value differences between A and B (Louie et al. 2013). An alternative account, drawing on a sequential sampling tradition, proposes that high value distractors capture the decision maker’s attention and thus bias behavior (Gluth et al. 2020).

1.2.3 Parallels

Perceptual and economic decisions have largely been the focus of two distinct research streams and this separation is reflected in divergences in methodology. Perception research often employs weak or noisy stimuli, striving to measure (changes to) an observer’s threshold for stimulus detection or discrimination. Perceptual decision making studies typically involve a long, repetitive sequence of choice trials. The large amount of data from each participant allows researchers to parametrically investigate how different levels of an input of interest affect the choices of the observer. By contrast, economic decisions have traditionally been studied via one-shot hypothetical scenarios or vignettes, where each participant provides one (or very few) data points; although recently, psychophysical investigations have become more common in economics.

1. *Introduction*

Beyond these differences in methodology, by definition, economic and perceptual judgments concern qualitatively vastly different inputs and operations. Perceptual decisions target some existing property of the world, such as the appearance of an object; by contrast, economic decisions deal with the decision maker’s preferences, such as their subjective valuation of the object. The sources of uncertainty are therefore different. In perceptual choice, it is largely the inputs that are uncertain (what is there?); in economic choice, it’s the value function (how much do I value it?).

Nevertheless, there are also commonalities. Across both domains, the goal of the decision maker is, generally, to compare the available options and select a course of action. The rich, high-dimensional and noisy evidence for each alternative must be integrated and translated into a discrete low-dimensional outcome – the decision (Summerfield and Tsetsos 2020). Indeed, a growing literature attempts to draw bridges between perceptual and economic decisions (Summerfield and Tsetsos 2012; Woodford 2020) highlighting commonalities in their behavioral signatures, neural underpinnings and computational models thereof. Another set of parallels between perceptual and economic choices emerges in the ways in which they can be affected by context.

First, the law of diminishing marginal utility strikingly mirrors the Weber-Fechner law of perception. As the absolute magnitude grows, small changes in magnitude become increasingly difficult to notice. These logarithmic compressions, across both perception and valuation, have been linked with the limited capacity of the neural system (e.g. Wei and Stocker 2017; Bhui and Gershman 2018). Calibrating sensitivity to the relevant range of magnitudes in the context makes the most efficient use of the limited dynamic range of neural firing. This parallel suggests that there might be a common principle guiding the processing of decision-relevant information across the two domains (Woodford 2020; although interestingly, Max Weber himself was skeptical of this idea, Weber and Schneider 1975). It also gives rise to symmetric contextual biases in the evaluation of magnitudes, effectively scaling percepts and subjective valuations of changes in an imperative stimulus according to its absolute magnitude.

1. Introduction

Judgments across perception and economics also exhibit analogous and robust reference dependence. Information in the temporal context of a decision sets the reference against which incoming inputs are judged. In a context dominated by large (or valuable) objects, a focal stimulus may be perceived (or evaluated) as small (or unappealing). That same stimulus can be judged as large (or valuable) in a context of small (or unappealing) objects. Quantitatively, this pattern of results is captured by a drifting psychometric function which tracks the contextual expectation. Across both domains, other features of the statistics of context affect judgments, including the shape, skewness and dispersion of the distribution of feature magnitudes present (Parducci 1965; Rigoli 2019). Reference-dependent percepts and valuations have commonly been modeled as the result of an adaptive process, whereby neural signals are normalized by the contextual levels of stimulation (Webster 2015; Rigoli et al. 2016). In line with this theoretical view, across the processing hierarchy neural responses to visual stimuli and economic prospects exhibit contextual adaptation (e.g. Enroth-Cugell and Shapley 1973; Cox and Kable 2014).

Reference-dependent judgments are typically repelled *away* from context. However, another class of contextual biases sways choices across perception and economics *towards* context. In the domain of perception, judgments about a focal stimulus can be attracted towards the stimulus immediately preceding it if the two stimuli are sufficiently similar and located nearby to one another (Liberman et al. 2014; St. John-Saaltink et al. 2016; Barbosa and Compte 2020). In the domain of economics, a value estimate of a prospect can be attracted towards a preceding irrelevant quantity if a direct comparison has been drawn between the two. This can be done, for instance, by asking if the subjective value of the prospect exceeds the anchor quantity (Tversky and Kahneman 1974). Sequential dependencies across both perceptual and economic judgments can be strengthened in the presence of noise, including increased stimulus-response intervals and scalar variability (Papadimitriou et al. 2016; Lee and Morewedge 2021).

The influence of spatial context also encompasses repulsive and attractive effects, whose strength and directionality typically interact with the imperative stimulus.

1. Introduction

For instance, the decoy effect has been documented in multialternative choices about economic prospects and visual stimuli alike (Trueblood et al. 2013). The decoy stimulus can sway choices towards different alternatives depending on the interplay of attribute values of all available options. Similarly, the effect of spatial context on perceptual detection judgments can reverse its direction depending on the characteristics of the focal stimulus (Stemmler et al. 1995). This pattern of results suggests that inputs are processed differently depending on their interrelationship with other information present in the spatial context. This idea has also been formalized in normalization frameworks where neural responses are modulated by context (Xing and Heeger 2001; Daviet and Webb 2021).

The influence of context on human information processing and decision making extends beyond the realms of perception and economics. As early as the 1960s, research documented the dynamics of contextual repulsion and attraction in human affect, motivation, learning, cognition and interpersonal behavior (Helson 1964). A popular account in perception, adaptation-level theory posits that the human brain processes information across all manner of input domains, including low-level sensory stimulus features and more high-level abstract concepts and characteristics, in a relative, context-dependent manner. The existing parallels in context dependencies across perceptual and economic decision domains lend support to the view that there might be a domain-general information processing mechanism which gives rise to contextual influence.

1.3 The Underpinnings of Context-Dependent Choice

What drives context dependencies in choice? In biology, there is an important distinction between proximate and ultimate explanations. Proximate causes capture the mechanisms directly driving a behavior (e.g. observed patterns of neural activity or inferred information processing mechanisms); ultimate causes refer to the evolutionary origins that have selected for and shaped the behavior (Davies et al.

1. Introduction

2012). Normative perspectives of decision making serve as clues for the selective pressures shaping choice behavior. For this, I will turn back to normative views of decision making to explore the ultimate causes underlying context-dependent choice. This section redefines what constitutes an optimal decision by incorporating assumptions about constraints to the decision making process, including the limited capacity of the brain (Sims 2003) and the inherent noisiness of neural computation (Tsetsos et al. 2016).

1.3.1 Neural Constraints & Efficiency

The discussion in section 1.1 regarding what constitutes an optimal decision assumed that neural systems are, generally, free from processing constraints. While acknowledging that neural representations are noisy, the models described in sections 1.1.1 and 1.1.2 neglect the limited capacity of the human nervous system and the metabolic costs associated with information processing. In cognitive science, these constraints would be reflected in the cost function of the processing system, which is minimized in efficient calculations. These efficiency considerations are reflected in the *efficient coding* hypothesis (Barlow et al. 1961; Simoncelli 2003).

An efficient strategy for the encoding of sensory information postulates that neural representations should follow the environmental statistics (Girshick et al. 2011), resulting in the observed overrepresentation of, for instance, cardinal orientations in early visual cortex (Furmanski and Engel 2000; Li et al. 2003). This strategy, whereby the distribution of possible neural states matches the structure of the environment, ensures that sensitivity is maximized for the features which are most likely to occur.

While the concept of efficient coding is generally associated with evolutionary and developmental neural adaptations to the natural environment (Barlow et al. 1961), similar processes are operating at much shorter time-scales, adapting the neural system to the levels of stimulation available in the local context. Light adaptation offers an example of this – adaptation to the contextual level of light dedicates the limited neuronal range (typically ~ 2 orders of magnitude) to those light intensities

1. Introduction

which are present in one’s immediate environment, and ensures that visual sensitivity is relatively stable despite changes in light levels of a factor of over 10 billion (or ~ 10 orders of magnitude, Stockman et al. 2006). Adaptation, thus, is an efficient strategy given the biological constraints of the visual system; it does, however, also lead to context dependence in our percepts (Webster 2015). Theoretic approaches from economics, such as bounded rationality (Simon 1955) and rational inattention (Sims 2003), have similarly appealed to the constraints of human information processing to explain the mechanisms driving context-dependent economic choice.

Normalization has been put forth as a “canonical operation” adapting the neural system to current levels of stimulation and promoting the efficiency of neural computation across the processing hierarchy (Carandini and Heeger 2012). Generally, normalization theories involve the multiplicative normalization of neural signals by a common factor drawn from, for instance, the activity of a neuronal population. The specific mathematical formulation of the normalization factor is the topic of debate (e.g. the weighted sum of inputs as in Carandini and Heeger 2012; Louie et al. 2013; or the range of inputs as in Soltani et al. 2012). However, a commonality across specific accounts is that normalization concentrates the available range of neural responses around some estimate of what inputs are most likely to occur. This, in turn, ensures that various efficiency considerations are met, including maximizing sensitivity to changes in inputs, boosting sensitivity for behaviorally relevant stimuli, avoiding neural saturation and reducing neural redundancy. Divisive normalization occurs across many different stimulus domains and computational operations, including sensory processing, attention modulation, the encoding of value and multisensory inference (Carandini and Heeger 2012).

Implementation-wise, normalization has been documented across various scales of neural processing. For instance, on the single neuron level, neural responses to a constant stimulus are adapted via inhibiting intrinsic ionic currents based on the neuron’s own activity (inhibitory feedback loop, Benda 2021). Normalization-based inhibition can also occur via lateral cortical connections, whereby stimuli outside the classic receptive field of a neuron can adjust that neuron’s response

1. Introduction

to stimulation within its receptive field center (Stemmler et al. 1995). Similar dynamics have also been observed on the population level, for instance, pooled inhibition in the posterior parietal cortex, which may play a role in decision making pertaining to categorical choices (Wang 2002).

1.3.2 Behavioral Relevance

Beyond the case for computational efficiency, relative decision making may also confer some behavioral advantages. The idea that evaluating stimuli relative to the local context is an efficient and evolutionary adaptive strategy is a central tenet of *optimal foraging theory* (Stephens and Krebs 1986; Kolling et al. 2012). An animal should choose whether to exploit a given resource or explore for other resources, depending on an estimate of the current resource value relative to the expected environmental reward. Relative perception and valuation can similarly be argued to serve an adaptive behavioral function.

Context can also help resolve ambiguities. A popular normative principle for decision making which incorporates the contextual expectation is *Bayesian decision theory* (Mamassian et al. 2002; Yuille and Kersten 2006). According to this theoretic account, an ideal observer should evaluate a stimulus taking into account not only the likelihood of the true stimulus magnitude given the noisy internal representation ($\hat{x}$), but also the prior expectations about the likelihood of this magnitude occurring. These prior expectations are shaped by context and follow the distribution of signals that are currently present or recently experienced. Thus, signal inferences for the true magnitude of x will be biased towards the contextual expectation. A similar theoretical framework underpins dominating accounts of the contraction bias, appealing to Bayesian inference to explain the finding that the influence of the prior grows with the presence of noise in the judgment process (Ashourian and Loewenstein 2011; Olkkonen et al. 2014).

In signal detection theory, prior expectations may be incorporated in the parameter value of the decision criterion, which sets a threshold that needs to be exceeded for a detection judgment to occur (Green and Swets 1966). While this

1. Introduction

can make detection judgments more liberal in a context where the signals of interest dominate, the criterion is classically considered stable over time. *Criterion setting theory* (Treisman and Williams 1984; Lages and Treisman 2010) is an extension of signal detection theory which instead argues that the location of the decision criterion drifts over time according to two opposite mechanisms which both work to further the behavioral goals of the observer. The probability tracking mechanism is driven by the autocorrelated nature of inputs in the natural environment: if a signal was present at $t - 1$, then it is more likely to be present than not at time t . Hence, it leads to cross-trial attractive serial dependencies. The stabilization mechanism, on the other hand, shifts the criterion towards the central tendency of context in order to ensure that positive and negative detection judgments occur equally often, maximizing information transmission. This, in turn, leads to reference-dependent judgments that are repelled from the statistics of local context.

The attractive and repulsive dynamics of contextual influence can, thus, be traced back to the interplay of efficiency considerations and the behavioral relevance of context. In the real world, our visual (and conceptual) environment is stable. The neural system explores this regularity to promote efficient computation and increase sensitivity to small changes in the environment (Mattar et al. 2018). This efficient adaptation of neural coding produces *repulsive* influences of context (Webster 2015; Fritsche et al. 2020). The spatially and temporally correlated nature of inputs, however, also gives rise to *attractive* contextual effects. A Bayesian observer would exploit the stability of the environment when decoding noisy sensory inputs, with contextual information guiding the prior probability of target signals. Thus, judgments would be attracted towards the statistics of context (van Bergen and Jehee 2019; Fritsche et al. 2020). Taken together, these two opposite efficient strategies can explain the rich (and contradictory) patterns of contextual influence outlined here: negative context dependence is optimal in encoding information, but positive context dependence is optimal when decoding neural signals (Fritsche et al. 2020).

1.4 Thesis Outline

Human choices often exhibit stereotyped patterns of context dependence. This is true across decisions pertaining to the evaluation of low-level sensory features, such as visual or auditory properties, and high-level features, such as subjective value, propositional statements and abstract concepts. The parallels between context-dependent choices in perception and economics reviewed in this chapter suggest that there might be a common, domain-general decision mechanism underlying these phenomena (Helson 1964). The neural underpinning of such a mechanism could stem from neural efficiency considerations. State-of-the-art psychophysical, electrophysiological and neuroimaging evidence suggests that normalization is a feasible candidate domain-general computational scheme.

This thesis explores specific instances of context-dependent choice behavior. Chapter 2 charts the full parametric space of decoy influence in economic decisions and arbitrates between different formalizations of neural normalization. Chapter 3 investigates and models the functional form of the effect of a distractor stimulus in a perceptual discrimination task. Chapter 4 arbitrates between opposing theoretical accounts of reference-dependent perceptual categorization.

This chapter is based on the published work:
Dumbalska, T., Li, V., Tsetsos, K., & Summerfield, C. (2020). A map of decoy influence in human multialternative choice. Proceedings of the National Academy of Sciences, 117(40), 25169-25178.

2

A map of decoy influence in multialternative economic choice

Contents

2.1	Introduction	26
2.2	Experiment	29
2.2.1	Methods	29
2.2.2	Results	35
2.2.3	Interim Discussion	38
2.3	Computational Modeling	40
2.3.1	Adaptive Gain Model	40
2.3.2	Model Comparisons	45
2.3.3	Charting Normalization Models of Decoy Influence	47
2.3.4	Interim Discussion	55
2.4	Discussion	55
2.5	Conclusion	58
2.6	Acknowledgements	58

When choosing between two alternatives, the addition of a third inferior option should not affect choices. In practice, however, even though the dispreferred decoy good is rarely chosen, it can systematically sway preference towards one of the two original alternatives. Previous work has identified three classic decoy biases, known as the attraction, similarity and compromise effects, which arise during choices between economic alternatives defined by two attributes. However, the

2. A map of decoy influence in multialternative economic choice

reliability, interrelationship, and computational origin of these three biases has been controversial. Here, a large cohort of human participants made incentive-compatible choices among assets that varied in price and quality. Instead of focusing on the three classic effects, we sampled decoy stimuli exhaustively across bidimensional multiattribute space and constructed a full map of decoy influence on choices between two otherwise preferred target items. Our analysis revealed that the decoy influence map was highly structured even beyond the three classic biases. We identified a very simple model that can fully reproduce the decoy influence map and capture its variability in individual participants. This model reveals that the three decoy effects are not distinct phenomena but are all special cases of a more general principle, by which attribute values are repulsed away from the context provided by rival options. The model helps understand why the biases are typically correlated across participants and allows us to validate a new prediction about their interrelationship. This work helps clarify the origin of three of the most widely studied biases in human decision-making.

2.1 Introduction

Optimal decisions should be driven solely by information that is relevant for the choice. When deliberating among more than two options (“multialternative” choices) this means ignoring those alternatives that are inferior or unavailable. Hence, the choice between two consumer goods should not be affected when an unaffordable third option is introduced. Similarly, voting preferences between two electoral candidates should not be changed when a third contender with more dubious merit enters the race. This normative principle, which is enshrined in the axiom of regularity (Block and Marschak 1960; Rieskamp et al. 2006), has been of significant and long-standing interest to behavioral scientists as it is robustly violated by a variety of animals including humans (Tversky 1972a; Huber et al. 1982; Simonson 1989), monkeys (Parrish et al. 2015), amphibians (Lea and Ryan 2015), invertebrates (Shafir 1994), and even unicellular organisms (Latty and Beekman 2011). The existing empirical evidence has indicated that where choice alternatives

2. A map of decoy influence in multialternative economic choice

are characterized by two value dimensions (e.g. the price and the quality of a product or the likability and competence of a political candidate), the introduction of an irrelevant distractor item to the choice set leads to rich and stereotyped biases in decision-making. A key research goal in the fields of psychology and economics has been to identify a simple and elegant computational principle that can explain the biases provoked by an irrelevant “decoy” stimulus (Turner et al. 2018).

The existing literature has focused on three decoy effects that can arise during ternary (three-way) choice among alternatives characterized by two independent and equally weighed attributes. The phenomena are illustrated in Fig. 2.1. Consider for example a consumer who is choosing between three products that are each characterized by dimensions (attributes) of quality and economy. The axes in Fig. 2.1 are scaled such that these attributes are perfect substitutions in that the consumer will forego one unit of one attribute for one unit of the other. Two target items, A and B, lie on the line of isopreference which is perpendicular to the identity line. In other words, A is less expensive but lower quality than B, such that the consumer should be indifferent between these options. The empirical phenomena describe how preferences may be biased towards either A or B as a function of a third “decoy” item D that lies on or below the isopreference line. The consensus view holds that a bias towards A can be provoked by the inclusion of a decoy D_a that it dominates, that is where A (but not B) is equivalent or superior on both dimensions (the attraction effect); that a bias towards A occurs in the presence of a more extreme decoy D_c which is superior in quality but yet more expensive than A, making A the “compromise” option (the compromise effect); and that a bias towards A is incurred by a decoy D_s which is similar to B in price and quality (the similarity effect, Fig. 2.1).

These three phenomena have been a major object of study in psychology and behavioral economics for several decades, with research interest reinvigorated in the last couple of decades after a landmark study that proposed the first unified computational account of the three classic decoy effects (Roe et al. 2001). Since then, a variety of empirical results have been produced and a plethora of computational

2. A map of decoy influence in multialternative economic choice

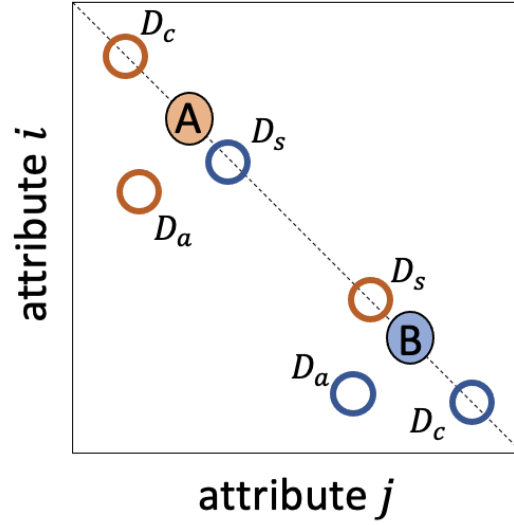


Figure 2.1: Illustration of the attraction, compromise and similarity effects. A and B denote two equally preferred stimuli; A is strong on attribute i but weak on attribute j and vice versa. The introduction of decoy stimuli (rings; denoted D_a , D_c and D_s) can bias preferences towards either A or B. The color of each ring signals the direction of the bias, e.g. for orange rings, A is preferred. Stimuli falling on the dashed line are equally preferred.

models have emerged. These include models that rely on loss aversion (Tsetsos et al. 2010), pairwise normalization (Landry and Webb 2021), attentional weighing (Hotelling et al. 2010; Tsetsos et al. 2012; Bhatia 2013; Trueblood et al. 2014), lateral inhibition (Hotelling et al. 2010), associative biases (Bhatia 2013), power-law transformation of attribute values (Bhatia 2013), sampling from memory (Noguchi and Stewart 2014; Bhui and Gershman 2018), or various other forms of reference-dependent computation (Soltani et al. 2012; Li et al. 2018; Natenzon 2019; Rigoli 2019). However, there has been a notable lack of consensus about the computational principles that give rise to decoy effects (Turner et al. 2018). There are a number of potential reasons for this, but here we focus on one limitation of past studies: most have tested for decoy effects by selecting fixed attribute values for D_a , D_c and D_s and calculating for each the relative choice share (RCS) for target items A and B, with relative deviations from choice equilibrium signalling a bias indicative of the successful detection of a decoy effect. However, reducing the dimensionality of the data in this way (i.e. to 6 data points) makes it harder to distinguish theoretical accounts, as many models may mimic one another in successfully capturing the

2. *A map of decoy influence in multialternative economic choice*

phenomena, so that comparisons among models are reduced to questions of a priori plausibility and parsimony. Relatedly, what defines a “decoy” of each class is typically left largely to the discretion of the researcher, who is free to choose a priori the values for D_a , D_c and D_s – i.e. the space over which the attraction, compromise or similarity might occur. This issue, coupled with the fact that the effects are often studied in small participant cohorts, using diverse stimulus materials – consumer choices (Josiam and Hobson 1995), text-based vignettes (Yang and Lynn 2014), or perceptual judgments (Trueblood et al. 2013) – has led to disagreement over the provenance and reliability of the three effects (Trueblood 2012; Frederick et al. 2014; Yang and Lynn 2014; Trueblood et al. 2015).

In this chapter, we address these issues by reporting a large-scale ($n > 200$), incentive-compatible study. We systematically mapped the decoy influence across attribute space, calculating the relative choice share RCS_{ij} for each decoy with attribute values i and j . This allowed us to explore the dimensionality of the data, with a view to asking whether a single principle can explain the ensemble of observed decoy effects. We found that a remarkably simple model, which draws on a computational framework based on the principle of divisive normalization, can capture the full decoy influence (RCS) map. Importantly, the model suggests that the three canonical decoy effects are not in fact distinct phenomena but rather fall naturally out of previously described dynamics of attraction and repulsion of decision values towards and away from a reference value given by the mean of available options.

2.2 Experiment

2.2.1 Methods

Participants

A total of 358 US-based participants were recruited via the platform Amazon Mechanical Turk to participate in a three-phase online study. All participants took part in the first phase (rating task), and those who passed a performance

2. A map of decoy influence in multialternative economic choice

threshold ($n = 231$; see below) were invited to join the second and third parts of the study in separate testing sessions (choice task). Of these, 189 met our criteria for inclusion in the analysis, namely $p < 0.001$ of responding randomly during the choice task (binomial test). Phases 2 and 3 (choice task) were identical; phase 3 simply allowed us to gather more data ($n = 149$ completed both phases 2 and 3). Data were collected in two distinct batches. In the first batch, we paid participants \$4 for completing each phase, in addition to a performance-based bonus of up to \$20 for the second and third part of the study (a maximum payment of \$32). To reduce the dropout rate, in the second batch the base payment was increased to \$5 and the bonuses increased to \$12 and \$18 in the second and third phases respectively (a maximum payment of \$45). The study received ethical approval from the Central University Research Ethics Committee at the University of Oxford (approval reference number: R50750/RE001).

Experimental Tasks

Rating Task The first phase of the study (rating task, Fig. 2.2a) was introduced as a “property rental price guessing game”. The task involved estimating the market rental price of residential real estate by viewing an image of the exterior of a series of properties. On each trial, an image of a property was shown along with a horizontal slider for a maximum of 60 seconds. The task was to guess the market rental price of the property, i.e. the dollar amount that an average person would be willing to pay per month to rent it, and to indicate it on a slider scale. The slider ranged from \$0 to \$2500 and the initial position of the slider was randomized on each trial. There were a total of 250 unique house images, each presented twice in randomized order (for a total of 500 trials per participant). The 250 houses had been selected to have the lowest average choice variability in a pilot study involving 30 distinct participants and a larger set of properties ($n = 450$), which we conducted before the main experiment. We used participant ratings from the pilot dataset to include/exclude participants. After phase 1, we correlated the 250 ratings for each participant against the average ratings obtained from the pilot

2. *A map of decoy influence in multialternative economic choice*

study. Participants with a Spearman’s rank correlation of $\rho < 0.7$ were excluded; the rest ($n = 231$) were invited to progress to the choice task.

Choice Task We introduced phases 2 and 3 as a “best-value property hunting game” (choice task, Fig. 2.2c). Here, participants were told to imagine that they were a real estate agent recommending to a client the best value house we, a fictitious real estate company, have on offer. On each trial, 3 properties (i.e. choice alternatives) were displayed for a maximum of 60 seconds on left, central and right positions on the screen. Underneath each image we displayed an allocated rental cost (in dollars) and a number of stars. The number of stars was proportional to the value given by the participant in the ratings task, and merely served as a reminder. In piloting, we found that this improved choice consistency. Participants were informed that the property images were a subset of those that they had viewed in phase 1, and that the stars were related to the ratings they themselves had reported. The task asked participants to press three keyboard buttons (left, down, and right arrow buttons) to indicate their ordered preference from the best-valued house to the worst-valued house. Participants were explicitly instructed that the best-valued house was the one with the highest market value but the lowest allocated rental price. At the end of each block, participants were told how many trials’ recommendations were correct, given their initial ratings. The bonus payment at the end of each phase was proportional to their accuracy.

Unbeknownst to participants, the options were carefully selected for each participant individually to allow us to test our hypotheses of interest (Fig. 2.2b). First, for each participant we filtered out the 90 properties with the highest rating variability, i.e. the highest absolute deviance between the two ratings. Second, the remaining 160 properties were binned into quality deciles (attribute i) on the basis of each participant’s ratings. The binned properties could then be associated with an allocated rental cost that was drawn uniformly from within the range of dollar values that defined each decile (attribute j). This allowed us to select, on each trial, three stimuli that differed on two dimensions: two targets A and B, and a decoy

2. A map of decoy influence in multialternative economic choice



Figure 2.2: Experimental design. **A:** Participants first played a “property price guessing game”. On each trial they estimated the monthly rental value (in dollars) of a residential property, using a sliding scale. **B:** After discarding properties with inconsistent responses, ratings were sorted into deciles for each participant. These bins were used to select stimuli for targets A and B (deciles 3 and 8 of estimated ratings; red and blue squares), and decoy stimuli. Each choice task stimulus was created by matching a property with a given decile estimated value (quality; attribute j) to a new rental price (economy; attribute i) on a 10x10 grid. Eight property/price combinations were generated for each cell in the grid that lay below the diagonal (white cells), and 2 property/price combinations for each cell above the diagonal (grey cells). **C:** Participants then played a “best value property hunting game” in which they were asked to rank 3 stimuli according to their economy/quality trade off. Two of the stimuli corresponded to the targets A and B, the third stimulus was the decoy. A star rating system was used as a reminder of their previous price estimation judgment.

2. A map of decoy influence in multialternative economic choice

stimulus. Target A was always a property drawn from the 3rd decile of quality (i.e. participant rating) and the 8th decile of economy (i.e. the 3rd decile of cost); target B was always drawn from the 8th decile of quality and the 3rd decile of economy (i.e. the 8th decile of cost). The decoy stimulus was sampled exhaustively from the full attribute space (any of the 10 quality x 10 economy bins). Thus, targets A and B were equally valued options (A being low quality and low cost, and B being high quality and high cost) and the decoy stimulus could be both superior or inferior in value. Participants completed a total of 530 trials in the second part of the study. The third part constituted an additional 530 trials of the same task.

Analyses

We calculated relative choice share for the “low” item A over the “high” item B by calculating the proportion of trials for which A was chosen before B across all trials in each decoy decile bin. We were able to do this for both inferior and superior decoys, as participants provided us with ranked preference for all three items. Thus, even if on a given trial they chose the decoy, they still reported whether they preferred A over B or vice-versa. This allowed us to plot the relative preference for A over B across the full attribute space (RCS_{ij}).

Conventional Decoy Analyses We first adopted a standard approach from previous studies that have focused on estimating the effects of the three classic decoy effects, that is, calculating RCS for D_a , D_c and D_s . To this end, we defined portions of the influence map that corresponded to the traditionally defined positions of attraction, compromise and similarity decoys (Fig. 2.3a). We also included an additional decoy set that we called “repulsion” decoys (D_r): these were mirror-symmetric to the attraction decoys but located in the upper triangle of the influence map where the decoy was the objectively best option (i.e. a set of “superior” decoys). We then calculated the RCS for D_a , D_c , D_s and D_r , each defined with respect to target A and target B. The strength of each effect was defined as the difference in RCS for each decoy set defined with respect to targets A and B. We tested each effect against zero for statistical significance via a single sample t test.

2. *A map of decoy influence in multialternative economic choice*

Interrelationship of Effects To investigate the associations between the three classic decoy effects in our cohort, we calculated Pearson’s correlation coefficients. We estimated those pairwise for each of the three possible combinations of effects (D_c - D_a , D_c - D_s , and D_a - D_s), based on the strength of the effects for each participant in our sample.

Map of Decoy Influence Our major goal for this project was to go beyond conventional analyses and chart decoy influence across the attribute space. The ordered RCS for option A over B in each decoy decile bin across all possible attribute combinations constitutes the map of decoy influence. In addition to the decoy map, as a sanity check, we also plotted preferences for the two targets *over* the decoy item. We did this by calculating the RCS for option A over the decoy D (i.e. in what proportion of trials did the participant choose A before D?), as well as the RCS for option B over the decoy D in each of the 10x10 decoy decile bins.

Map Decomposition The resulting map of decoy influence allowed us to not only ask how the three classic decoy effects are interrelated in our cohort of participants, but to seek structure beyond these three locations in attribute space. Using an exhaustive range of decoy locations allowed us to use dimensionality reduction approaches to examine the (potentially distinct) factors from which the map of decoy influence is composed. We used singular value decomposition (SVD) to identify factors contributing to the map of preference for $A > B$ and calculate the variance explained by these factors. To this end, we flattened the 10x10 decoy map for each participant into a vector of values and constructed a large matrix of RCS, where each row indexed an individual participant and each column – the decoy location in attribute space. SVD produces a set of basis functions, or factors, and allows us to examine the amount of variance explained by them in the empirical data. This, in turn, illustrates the structure of decoy effects across participants. Knowing this structure can give us clues as to whether distinct mechanisms drive

2. A map of decoy influence in multialternative economic choice

decoy influence across different locations in attribute space, or alternatively, if the general structure of RCS can be explained by a single factor.

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/u6br3/>) for this project.

2.2.2 Results

Conventional Decoy Analyses

Fig. 2.3b illustrates the difference in choice share driven by decoys in the 4 locations associated with D_c , D_a , D_s and D_r (defined in Fig. 2.3a). As a first general observation, despite the careful sampling of targets that were matched in price/quality ratio according to participants' responses in the valuation phase, and despite the incentives offered for consistent responding, participants exhibited a bias towards the high item B (average RCS ~ 0.7) over the low item A. Despite this additive bias, however, decoys had a clear and robust influence on choice. Participants exhibited clear attraction ($t_{188} = 4.74$, $p < 0.001$) and compromise ($t_{188} = 6.31$, $p < 0.001$) effects, which were statistically significant and in the expected direction. There was also empirical support for a repulsion effect ($t_{188} = 3.45$, $p < 0.001$) from superior decoys. On average, the presence of attraction, compromise or repulsion decoys shifted preferences from A to B by about 3-5%. However, we failed to replicate the similarity decoy effect ($t_{188} = 0.41$, $p = 0.68$).

Interrelationship of Effects We observed a positive association between the attraction and compromise effects ($r = 0.72$, $p < 0.001$) and a negative relation between the similarity effect and both compromise ($r = -0.59$, $p < 0.001$) and attraction ($r = -0.46$, $p < 0.001$) effects. Note that the latter correlations were observed despite the fact that in our data, the similarity effect was on average non-significant. These are shown in Fig. 2.3c.

2. A map of decoy influence in multialternative economic choice

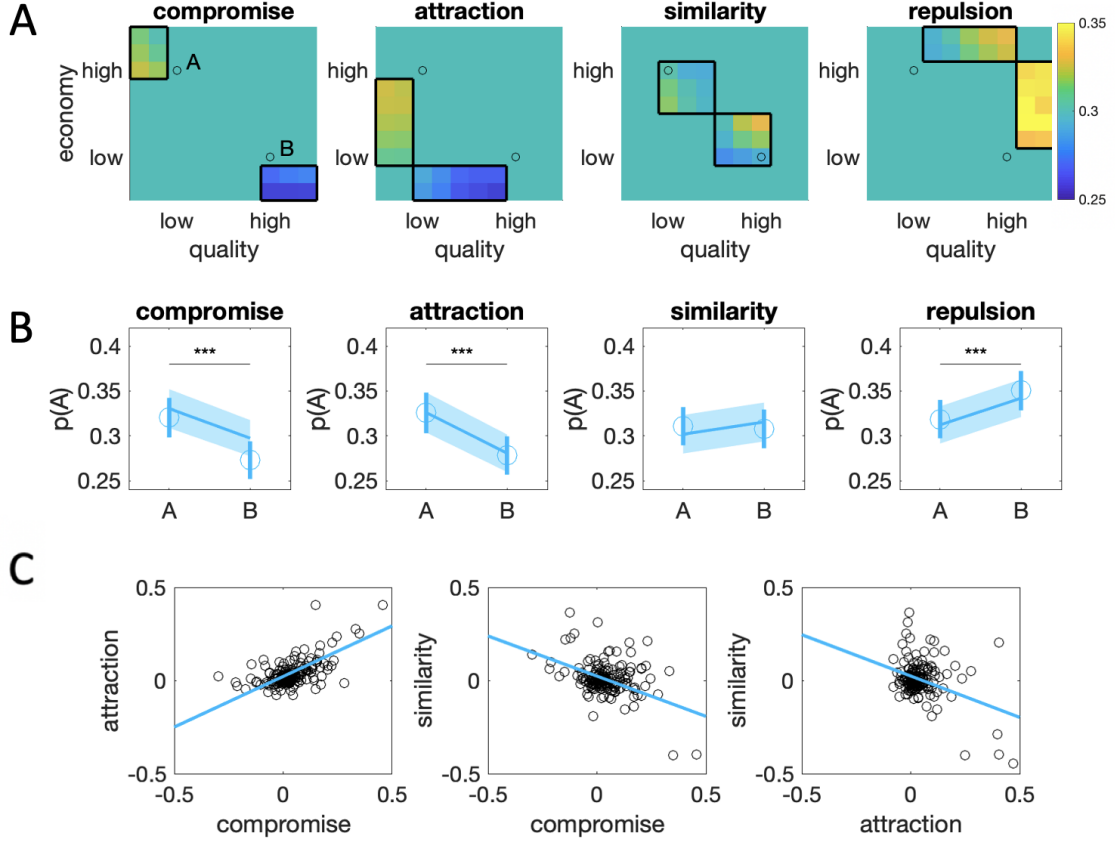


Figure 2.3: Conventional decoy analyses. **A:** Each panel illustrates the chosen locations in decoy space for compromise (D_c), attraction (D_a), similarity (D_s) and repulsion (D_r) decoys (boxes). The blue-yellow color scale illustrates relative preference for target A over B (warmer colors) or vice versa (colder colors) at each location. Black circles indicate the locations of targets A and B. **B:** Average choice share for target A as a function of decoy location across D_c , D_a , D_s and D_r . Blue dots are human data, shaded lines are model fit of adaptive gain model (see below). Bars/shaded area signal $M \pm \text{SEM}$ *** indicates $p < 0.001$. **C:** Correlations between the attraction, compromise and similarity effects. Each dot is a single participant; the decoy estimate is calculated as the difference between the RCS for given decoy with respect to targets A and B.

Map of Decoy Influence The full map of decoy influence RCS_{ij} is shown in Fig. 2.4a. Visual inspection reveals that the map has rich structure beyond the traditional decoy locations. Relative preferences for A and B seem to be driven by a dynamic of attraction and repulsion that depends on the position of the decoy with respect to each target stimulus. Robust attraction effects (whereby the presence of a decoy that is dominated by A shifts preferences towards A) were mirrored by strong repulsion effects (whereby a decoy that dominates A shifts preferences towards B). Attraction and repulsion were observed for both targets in approximate symmetry.

2. A map of decoy influence in multialternative economic choice

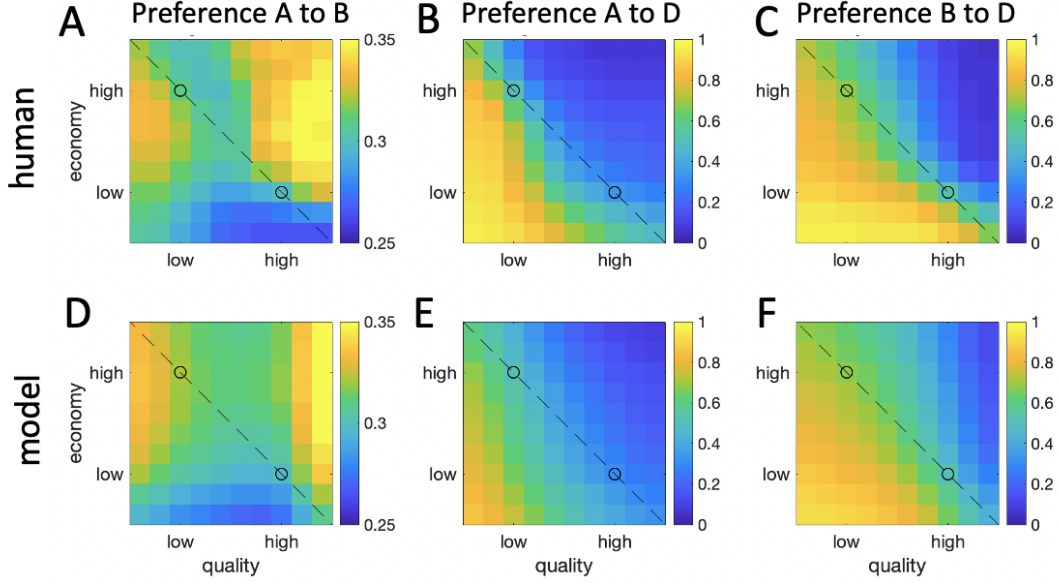


Figure 2.4: Map of decoy influence. **A:** Decoy influence map showing RCS_{ij} for A over B (left panels), A over D (middle panels) and B over D (right panels). Upper plots are the human data and lower plots are the same data for the simulated adaptive gain model (see below). The dashed line signals isopreference and the black circles are the targets A and B.

The map charting preference for $A > D$ (Fig. 2.4b) is in line with the expected pattern, whereby inferior decoys, i.e. those under the left diagonal, or line of theoretical isopreference, are chosen less often (warmer colors) than superior decoys (colder colors). Similarly, the map charting preference for $B > D$ (Fig. 2.4c) adheres to this general pattern. In theory, if A and B were equally preferred, these two maps would be identical. However, in line with the observed preference bias for the high value item B, we see that overall, preferences for $A > D$ appear weaker compared to preferences for $B > D$.

Map Decomposition Next, we sought to examine the structure of the map of decoy influence (Fig. 2.4a) via dimensionality reduction. The first 5 factors identified by SVD are visualized in Fig. 2.5a (top row). The first factor accounted for 95% of the variance in the data, suggesting that there is a single explanatory variable that drives decoy effects across participants (Fig. 2.5c, right panel, note the logarithmic scale on the y-axis). This finding is consistent with the correlation

2. A map of decoy influence in multialternative economic choice

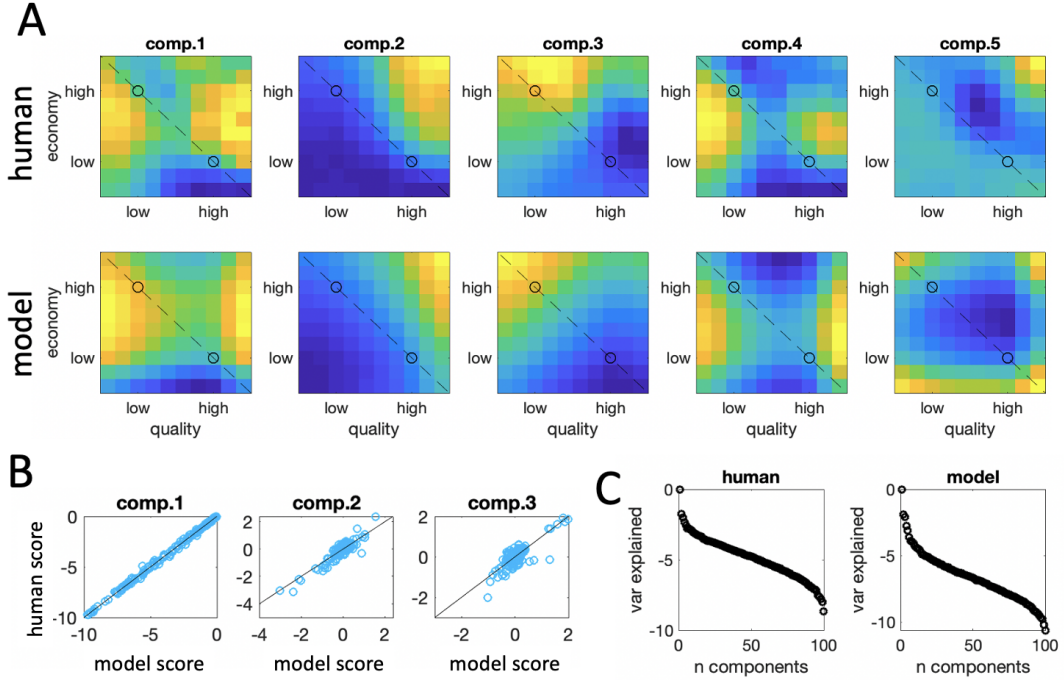


Figure 2.5: Map decomposition. **A:** First five components obtained from singular value decomposition (SVD) of the RCS for A vs. B. Upper plots are the human data and lower plots are the model. **B:** Correlation between component scores for components 1-3 between the human and the best-fitting model. For components 1-3, this correlation was very high. **C:** The variance explained by each component obtained by SVD for the humans (left panel) and the model (right panel). Note that the y-axis is on a log scale; the data is dominated by the first component in both cases.

results reported above, showing the robust interrelationships between the three classic effects in our participants.

2.2.3 Interim Discussion

Conventional decoy analyses of our data replicated the previously extensively documented compromise and attraction effects (Huber et al. 1982; Simonson 1989). However, we failed to replicate the similarity decoy effect in this dataset, consistent with weak or absent similarity effect from other studies (Noguchi and Stewart 2014; Berkowitsch et al. 2014). Indeed, while the attraction effect tends to be highly robust and consistent across participants, the compromise effect and similarity effects tend to be more idiosyncratic, with a high proportion of participants showing effects which are inverted with respect to the canonical form. For example, in previous studies

2. A map of decoy influence in multialternative economic choice

only a minority of participants show all three effects (numerically) in the expected direction (e.g. only 23% in Trueblood et al. 2015, we find a comparable figure of 22%).

Despite the fact that the similarity effect was on average non-significant in our dataset, its strength inversely correlated with the strength of the compromise and attraction effects. Those two effects were in turn positively associated with one another. This pattern of interrelationships of the decoy effects mirrors previous reports (Noguchi and Stewart 2014) and is further supported by our results on the structure of the full map of decoy influence. In particular, we found that a single explanatory factor is sufficient to capture 95% of the variability of decoy influence in our dataset.

One interpretation of this data is that the attraction, compromise and similarity effects are not distinct phenomena, but rather they are all driven by a single common computational principle. This principle ensures that decision values of stimuli A and B are repulsed away from proximal decoys on a line which lies perpendicular to the line of isopreference, giving rise firstly to what is traditionally known as the “attraction” effect (a potentially confusing nomenclature, given that decision values for a target are repulsed away from those of the decoy) and secondly to its converse identified here for superior decoys, which we have dubbed the “repulsion” effect. The compromise and similarity effects can arise because these repulsive effects may vary in strength for targets A and B, and/or for inferior and superior decoys, such that repulsive effects “spill over” asymmetrically towards the extremes of the isopreference line (compromise effect) or into the center of the bivariate decoy map (similarity effect). It is the fact that compromise and similarity effects are really driven by asymmetric attraction and repulsion that gives rise to the strong correlations in effects across the cohort observed in this paper and previously (Berkowitsch et al. 2014).

2.3 Computational Modeling

These findings suggest that it should be possible to identify a simple model that can reproduce the full decoy influence map with parameters that control the putative repulsive principle, as well as additional degrees of freedom that allow for asymmetric scaling of the attribute values (quality and economy). Previous simulation work demonstrated that a gain-control process can provide a unifying explanation for similar choice biases (Li et al. 2018). In this theoretic account, the gain with which alternatives are encoded is adaptively adjusted according to the context provided by other available options. Recent proposals of models for decoy phenomena have also appealed to a similar adaptive framework, whereby decision signals are normalized by context (Rigoli et al. 2017; Daviet 2018).

2.3.1 Adaptive Gain Model

The normalization principle in the adaptive gain model assumes that for each attribute (e.g. quality or economy), a *relative* value estimate is computed that reflects the distance between each item and the mean of the available choice set. We thus begin by defining $r(A)$, which is the value of target A relative to other items, including the decoy, on a given attribute i :

$$r(A_i) = v(A_i) - v(avg_i^{ABD}) \quad (2.1)$$

where $v(avg_i^{ABD})$ corresponds to the average value of the three alternatives on attribute i :

$$v(avg_i^{ABD}) = \frac{v(A_i) + v(B_i) + v(D_i)}{3} \quad (2.2)$$

The model proposes that the subjective utility $u(A_i)$ of target A on attribute i is computed through a logistic transformation of the relative value, as follows:

$$u_i(A_i) = \frac{1}{1 + e^{-(r(A_i) - c_i) \cdot s^{-1}}} \quad (2.3)$$

2. A map of decoy influence in multialternative economic choice

Thus, for each attribute, the utility of each target is given by a logistic function with slope s whose inflection point is the mean value of all items plus an additive bias term c . The additive bias c can potentially vary across attributes.

The utility of target A is a weighed sum of its attributes i and j , and the final decision is made by passing the utilities of all three rival stimuli through a softmax function to make a ternary choice:

$$u(A) = w \cdot u_i(A_i) + (1 - w) \cdot u_j(A_j) \quad (2.4)$$

$$p(A) = \frac{e^{\tau u(A)}}{e^{\tau u(A)} + e^{\tau u(B)} + e^{\tau u(D)}} \quad (2.5)$$

In addition to the softmax temperature τ , the model potentially has 4 free parameters of interest: the slope s , and inflection points c_i and c_j of the logistic function Eq. (2.4), and the weighing parameter w in Eq. (2.5).

Model Simulations

We explore the effects of manipulating these parameters on the predicted decoy influence map in Fig. 2.6; a full description of the parameters used is available in Table 2.1. This figure shows how the model can systematically account for not only the pattern observed in the current experiment, but also those from previous (and potentially contradictory) reports. In Fig. 2.6a we show the effect of manipulating the parameter w . This simply shows how we can tip the balance of preferring A over B according to the relative weight given to each attribute.

In Fig. 2.6b, with w now fixed to 0.5 (equal weighing of the two attributes), we show how the decoy effects grow in strength with s . Above each plot the relative positive or negative strength of the compromise (D_c), attraction (D_a), similarity (D_s) and repulsion (D_r) effects are shown in a bar plot. As can be seen, the attraction and repulsion effects grow as s grows, including a weak compromise effect but no similarity effect.

Fig. 2.6c shows the influence of varying $c_i = -(c_j)$ while s and w are fixed. This has the effect of shifting the relative strength of the attraction/repulsion effect for targets A and B . For example, when $c_i > c_j$ the attraction/repulsion effects

2. A map of decoy influence in multialternative economic choice

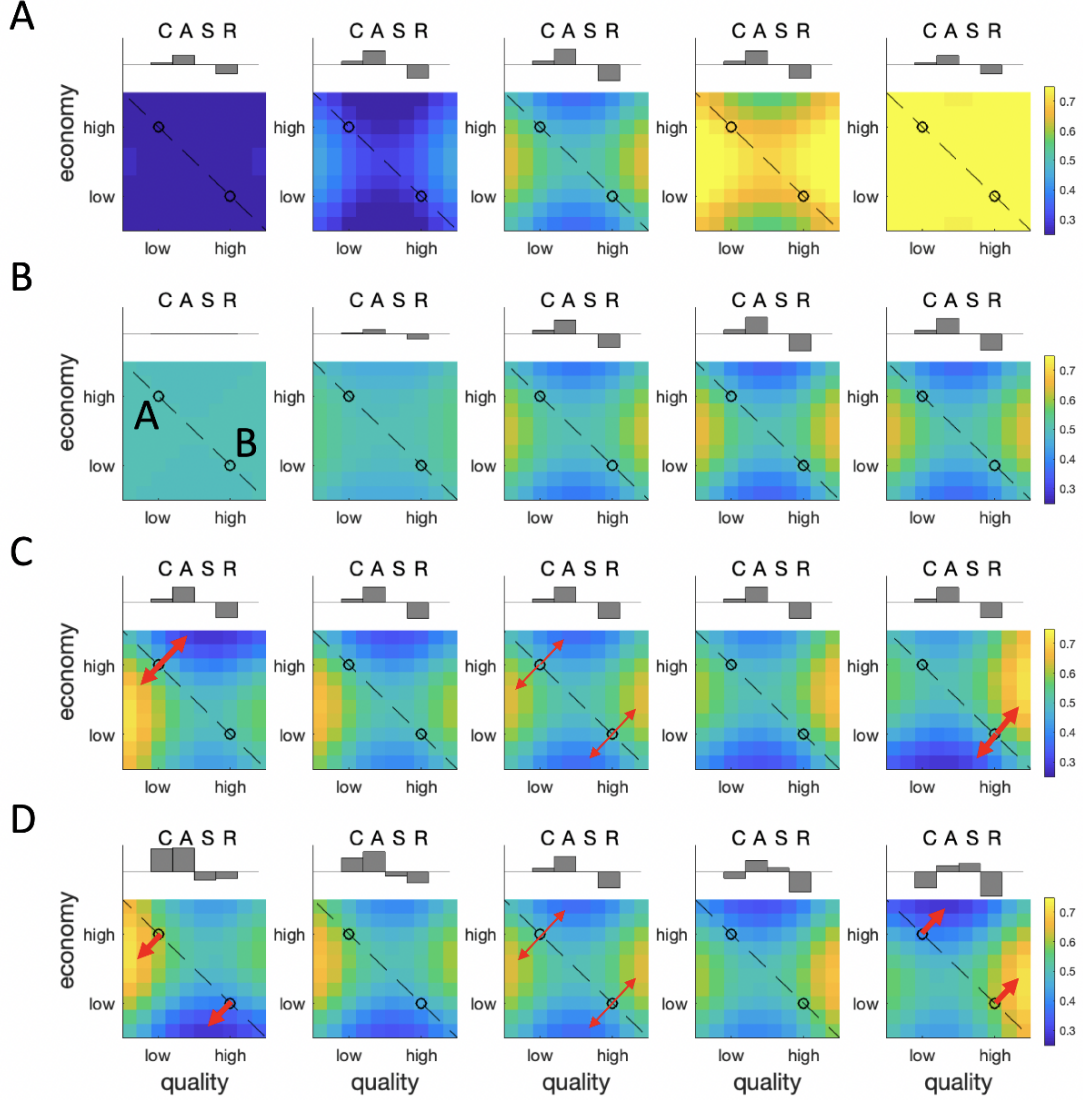


Figure 2.6: Simulated maps of decoy influence. **A:** Effect of varying the parameter w from low (left panels) to high (right panels). This parameter controls the relative preference for low price/quality to high price/quality items. **B:** The effect of varying the parameter s from high to low. This parameter controls whether A and B are equally preferred, or whether there is decoy-like distortion. **C:** The effect of varying the difference of bias terms $c_i = -c_j$ from negative (left panels) to positive (right panels). Varying this difference alters whether the maximal distortion occurs proximal to target A (left) or target B (right). **D:** The effect of varying the sum of bias terms $c_i = c_j$ from negative (left panels) to positive (right panels). Varying this difference alters whether the maximal distortion occurs for inferior decoys (left) or superior decoys (right). Red arrows highlight directions of repulsion, with arrow width schematically representing the strength of the effect.

2. A map of decoy influence in multialternative economic choice

Table 2.1: Parameters for map simulation.

Panel	Constants	Column 1	Column 2	Column 3	Column 4	Column 5
A	$s=0.1; \tau = 0.1$ $c_i = 0; c_j = 0$	$w = 0.4$	$w = 0.45$	$w = 0.5$	$w = 0.55$	$w = 0.6$
B	$w = 0.5; \tau = 0.1$ $c_i = 0; c_j = 0$	$s = 0.01$	$s = 0.0325$	$s = 0.055$	$s = 0.0775$	$s = 0.1$
C	$w = 0.5; \tau = 0.1$ $s = 0.1$	$c_i = -0.04$ $c_j = 0.04$	$c_i = -0.02$ $c_j = 0.02$	$c_i = 0$ $c_j = 0$	$c_i = 0.02$ $c_j = -0.02$	$c_i = 0.04$ $c_j = -0.04$
D	$w = 0.5; \tau = 0.1$ $s = 0.1$	$c_i = -0.04$ $c_j = -0.04$	$c_i = -0.02$ $c_j = -0.02$	$c_i = 0$ $c_j = 0$	$c_i = 0.02$ $c_j = 0.02$	$c_i = 0.04$ $c_j = 0.04$

are strongest for the target A, whereas when $c_j > c_i$ they are strongest for B (red arrows). However, because these effects cancel out symmetrically, this does not affect the overall RCS for D_c , D_a , D_s or D_r .

Finally, varying $c_i = c_j$ (Fig. 2.6d) brings about an asymmetric distortion whereby either attraction effects are stronger (i.e. below the isopreference line) or repulsion effects are stronger (in the superior decoy portion of space). In addition to varying the relative strength of attraction and repulsion, this allows the compromise effect to vary from positive to negative, as described in previous studies; it allows a weak similarity effect to emerge. Combinations of all of these factors give the model systematic flexibility to account for a wide range of observed effects.

Interestingly, the simulations shown in Fig. 2.6d allow us to make a new prediction about the human data. As seen in the bar plots accompanying each predicted influence map, when c_i and c_j are both negative, the compromise effect is positive and attraction is stronger than repulsion. By contrast, when c_i and c_j are both positive, the compromise effect is negative and the repulsion effect is stronger than attraction. The model thus predicts that on average, in the human data there will be a correlation between the (signed) compromise effect and the relative magnitude of attraction vs repulsion. This is plotted in Fig. 2.7, and as can be seen this prediction holds for the data we collected ($r = 0.57$, $p < 0.001$). Of note, this effect was driven both by a positive correlation between the compromise effect and the strength of repulsion ($r = 0.43$, $p < 0.001$), as well as the correlation between compromise and attraction effects described above.

2. A map of decoy influence in multialternative economic choice

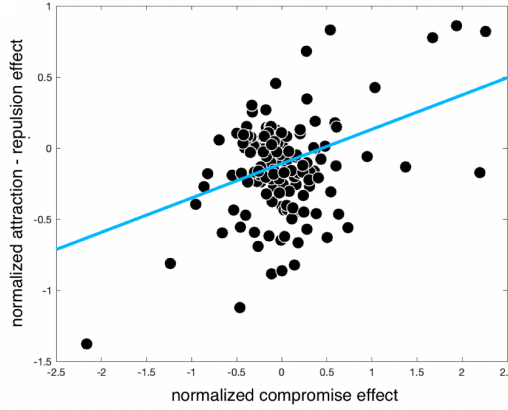


Figure 2.7: Correlation between the compromise effect and the relative strength of attraction vs repulsion in the human data. Each dot is a participant; the blue line is the best fitting linear trend.

Model Fitting

Next, we estimated the model parameters which best fit the data for each participant in our experiment to assess whether the model was able to capture the results at the individual as well as aggregate level.

Model Fitting Methods The model provided predictions about probabilities of choosing option A over B, which allowed us to compute model likelihoods on each trial. Those likelihoods were then used for model fitting and comparison. We estimated the model for each participant individually using gradient descent with the `globalsearch` function from the MATLAB Optimization Toolbox.

Model Fitting Results Fitting this 5-parameter (τ, s, c_i, c_j, w) model to human data, we can fully recreate the decoy effects observed in this study using both conventional (Fig. 2.3b) and novel (Fig. 2.4d-f) analysis methods. Specifically, the model captured almost exactly the pattern of traditional decoy effects, in terms of the relative impact on RCS of D_a , D_c and D_s , as well as the repulsion decoy D_r (blue shaded lines in Fig. 2.3b). The model reproduced the pattern of preferences for target A > target B qualitatively and quantitatively across the decoy space, and when we applied SVD to the model data generated under the best-fitting

2. *A map of decoy influence in multialternative economic choice*

parametrization for each participant, the first five components that emerged were nearly identical to those for humans, and the first model component explaining 97% of the variance (Fig. 2.5c). When we plotted the estimated singular values for the first three components for humans and the best-fitting model, we found them to be very tightly correlated (Fig. 2.5b). The model also displayed the same pattern of positive association between attraction and compromise effect ($r = 0.86$) and negative association between the similarity and attraction ($r = -0.85$) and similarity and compromise effects ($r = -0.94$). In other words, the model captures the human data very closely, both at the individual and the aggregate level.

2.3.2 Model Comparisons

The adaptive gain model belongs to a class of models which describe contextual biases as arising from normalization among populations of neurons. Normalization models share commitment to the idea that information is divisively normalized by the local context. The precise mathematical formulation of this procedure, however, differs between theoretical accounts. Some models, for instance, use measures of the central tendency of the features present in the local context (e.g. the average of attribute values, Louie et al. 2013), while other appeal to measures of the dispersion of features (e.g. the range of attribute values, Soltani et al. 2012). Similarly, models might encode features linearly (e.g. Bushong et al. 2021) or compress them prior to normalization (e.g. via a logistic operation as in the model described here and in Rigoli 2019). Below we highlight several prominent normalization models in the literature. Our aim was to compare our model to a broad space of alternative accounts that similarly assume the value of each target (on each attribute) is encoded relative to its competitors. Thus, we began our model comparison exercise with contrasts with those well established competitors.

The classic formulation of the divisive normalization theoretical tradition (henceforth vanilla divisive normalization, VDN) can be broadly encapsulated

2. A map of decoy influence in multialternative economic choice

in the following expression (Carandini and Heeger 2012; Louie et al. 2013):

$$u_i(A_i)^{VDN} = \frac{v(A_i)}{v(avg_i^{ABD}) + c_i} \quad (2.6)$$

where c_i is a small regularization constant. The model assumes that decision inputs are divided by the current contextual expectation, that is, the average of all current options.

In one variant of this model, which has been recently and successfully used to account for decoy effects, this normalization is also “recurrent”, i.e. it overweighs the focal item’s contribution to normalization (Daviet 2018; Webb et al. 2021):

$$u_i(A_i)^{RDN} = \frac{v(A_i)}{v(A_i) + v(avg_i^{ABD}) + c_i} \quad (2.7)$$

Note that the adaptive gain model (Eq. (2.3)) is equivalent to a form of the recurrent divisive normalization model in which the values are exponentiated prior to normalization:

$$u_i(A_i)^{AG} = \frac{\exp(\frac{v(A_i)}{s})}{\exp(\frac{v(A_i)}{s}) + \exp(\frac{v(avg_i^{ABD}) + c_i}{s})} \quad (2.8)$$

Alongside these two models, we also considered a class of contextual normalization, which uses the range (rather than the average) of attribute values being encoded on a given trial to normalize an imperative stimulus:

$$u_i(A_i)^{RN} = \beta \frac{v(A_i)}{v(rng_i^{ABD})} \quad (2.9)$$

where $v(rng_i^{ABD})$ corresponds to the difference between highest and lowest values of attribute i across all stimuli (target or decoy) in the trial and β is a scaling term (Soltani et al. 2012; Bushong et al. 2021).

Methods

We compared the fit of each of the candidate normalization accounts to adaptive gain in a direct model comparison exercise. To achieve this, we used Bayesian model selection on cross-validated model evidence. Cross-validation involved estimating model parameters from one half of trials (by comparing fits to preferences

2. A map of decoy influence in multialternative economic choice

between target items A and B, as well as preferences between target and decoys) and computing log-likelihoods from the held-out trials. We implemented this using the `globalsearch` function from the MATLAB Optimization Toolbox. For Bayesian model selection, we used the Statistical Parametric Mapping Toolbox (Stephan et al. 2009).

Results

Comparisons between cross-validated model likelihoods revealed that the exceedance probability for the adaptive gain model over each of the three competitor models (vanilla divisive normalization, recurrent divisive normalization and range normalization) was 0.99, providing decisive evidence for the former over each of the latter.

2.3.3 Charting Normalization Models of Decoy Influence

Grandmother Model

To go beyond those established models and compare our model to a broader range of competitors, we also devised and fit a more flexible model which encompassed a large space of possible normalization schemes. This “grandmother” model could capture the encoding scheme proposed by the 4 models introduced above, along with a number of other “hybrid” models. Thus, exploring the parameter space of the grandmother model allows us to explore the space of putative normalization schemes, including those that are yet to be described in the literature, and chart how different model features translate to predicted patterns of decoy influence.

The model had the general form:

$$u_i(A_i) = \frac{1}{\beta_1 + e^{-(v(A_i)^k - \mu_i^k)(s \cdot k)^{-1}}} \quad (2.10)$$

where

$$\mu_i = c_i + \beta_2 \cdot v(avg_i^{ABD}) + (1 - \beta_2) \cdot v(rng_i^{ABD}) \quad (2.11)$$

We focus on free parameters β_1 , β_2 , and k , which respectively encode the tendency to engage in asymmetric weighing of the evaluated (recurrent) item (β_1),

2. A map of decoy influence in multialternative economic choice

the tendency to normalize with respect to the mean vs range (β_2) and the extent to which inputs are compressed before being transduced (k).

Parameter k facilitated comparison between the adaptive gain model (which incorporates a logistic transformation) and other normalization models (which do not), by leveraging the following limit:

$$\lim_{k \rightarrow 0} \frac{x^k - 1}{k} = \log_e x, \forall x > 0 \quad (2.12)$$

Thus, substituting x above for our variables of interest, $v(A_i)$ and μ , we obtain:

$$\log v(A_i) - \log \mu_i = \lim_{k \rightarrow 0} \frac{v(A_i)^k - 1}{k} - \frac{\mu_i^k - 1}{k} = \lim_{k \rightarrow 0} \frac{v(A_i)^k - \mu_i^k}{k} \quad (2.13)$$

Thus, k thus corresponds to the extent to which inputs are compressed before being transduced, with $k \sim 0$ signifying a logarithmic compression, and $k = 1$ signifying linear input. Simulations revealed that implementing a value of k as high as 0.001 satisfactorily approximates the natural logarithm (Fig. 2.8).

Grandmother Model Derivations

The adaptive gain, vanilla and recurrent divisive normalization, and range normalization models were nested within the grandmother model. This section shows how each of them can be derived from the general grandmother model equation.

Leveraging the limit from Eq. (2.12), when we plug in the parameter values for vanilla divisive normalization specified in Table 2.2 into the general grandmother model, we arrive at the formulation of vanilla divisive normalization (as per Eq. (2.6)):

$$u_i(A_i) = \frac{1}{e^{-(v(A_i)^{0.001} - \mu_i^{0.001})(0.001)^{-1}}} = \frac{1}{e^{-(\log v(A_i) - \log \mu_i)}} = \frac{1}{\frac{\mu_i}{v(A_i)}} = \frac{v(A_i)}{\mu_i} = \frac{v(A_i)}{v(avg_i^{ABD}) + c_i} \quad (2.14)$$

Similarly, we may simplify the grandmother model into recurrent divisive normalization (Eq. (2.8)) by plugging in the relevant parameter values specified in Table 2.2:

2. A map of decoy influence in multialternative economic choice

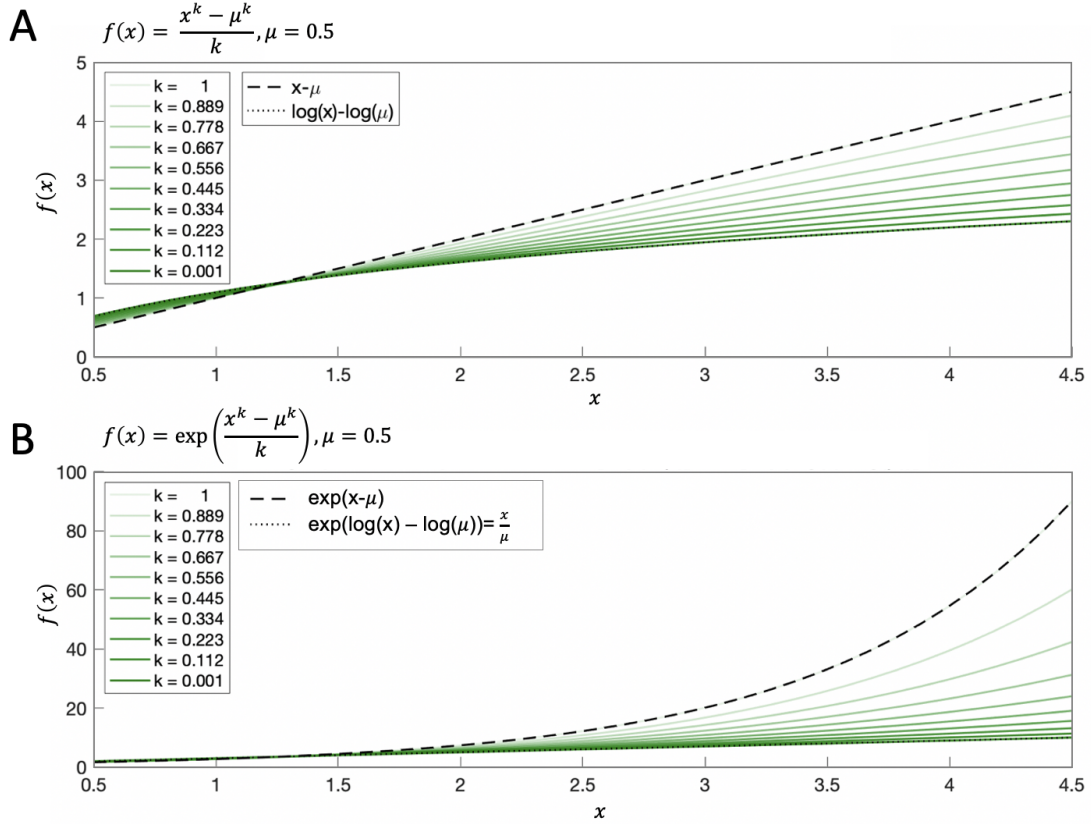


Figure 2.8: Illustration of variations in parameter k . **A:** Using parameter k to interpolate between a linear expression and its logarithmic compression. Note that the darkest green curve ($k = 0.001$) lies on the dotted line (logarithmic expression), suggesting that implementing a value of k as high as 0.001 satisfactorily approximates the natural logarithm. **B:** Exponentiating the function from panel A, as in the logistic formulation of the grandmother model. Note that the darkest green curve ($k = 0.001$) lies on the dotted line (linear expression, $\exp(\log(x) - \log(\mu)) = \frac{x}{\mu}$). This demonstrates that we can recover both normalization models that rely on logistic compression and normalization models that do not.

Table 2.2: Parametrizations of grandmother model corresponding to computational schemes proposed in the literature.

Model	c_i	s	β_1	β_2	k	τ
VDN	<i>free</i>	1	0	1	0.001	<i>free</i>
RDN	<i>free</i>	1	1	1	0.001	<i>free</i>
AG	<i>free</i>	<i>free</i>	1	1	1	<i>free</i>
RN	0	1	0	0	0.001	<i>free</i>

2. A map of decoy influence in multialternative economic choice

$$u_i(A_i) = \frac{1}{1 + e^{-(v(A_i)^{0.001} - \mu_i^{0.001})(0.001)^{-1}}} = \frac{1}{1 + e^{-(\log v(A_i) - \log \mu_i)}} = \frac{1}{1 + \frac{\mu_i}{v(A_i)}} = \frac{v(A_i)}{v(A_i) + \mu_i} = \frac{v(A_i)}{v(A_i) + v(avg_i^{ABD}) + c_i} \quad (2.15)$$

Note that if we allow parameter s to vary freely, it assumes the role of a power transform for inputs:

$$u_i(A_i) = \frac{1}{1 + e^{-(v(A_i)^{0.001} - \mu_i^{0.001})(s \cdot 0.001)^{-1}}} = \frac{1}{1 + e^{-(\log v(A_i) - \log \mu_i)s^{-1}}} = \frac{1}{1 + \frac{\mu_i^{s^{-1}}}{v(A_i)^{s^{-1}}}} = \frac{v(A_i)^{s^{-1}}}{v(A_i)^{s^{-1}} + \mu_i^{s^{-1}}} = \frac{v(A_i)^{s^{-1}}}{v(A_i)^{s^{-1}} + (v(avg_i^{ABD}) + c_i)s^{-1}} \quad (2.16)$$

Similarly, we may obtain the formulation of range normalization (Eq. (2.9)) by plugging in the relevant parameter values from Table 2.2 into the general grandmother model:

$$u_i(A_i) = \frac{1}{e^{-(v(A_i)^{0.001} - \mu_i^{0.001})(0.001)^{-1}}} = \frac{1}{e^{-(\log v(A_i) - \log \mu_i)}} = \frac{1}{\frac{\mu_i}{v(A_i)}} = \frac{v(A_i)}{\mu_i} = \frac{v(A_i)}{v(rng_i^{ABD})} \quad (2.17)$$

Finally, we may reduce the grandmother model to adaptive gain (Eq. (2.3)) by plugging in the relevant parameter values from Table 2.2:

$$u_i(A_i) = \frac{1}{1 + e^{-(v(A_i) - \mu_i) \cdot s^{-1}}} = \frac{1}{1 + e^{-(v(A_i) - \mu_i) \cdot s^{-1}}} = \frac{1}{1 + e^{-(v(A_i) - v(avg_i^{ABD}) - c_i) \cdot s^{-1}}} \quad (2.18)$$

Note that along with those specific models, which have been described in the literature, there exist many more parametrizations of the grandmother model which correspond to “hybrid” normalization schemes, combining specific features of the above described accounts.

2. A map of decoy influence in multialternative economic choice

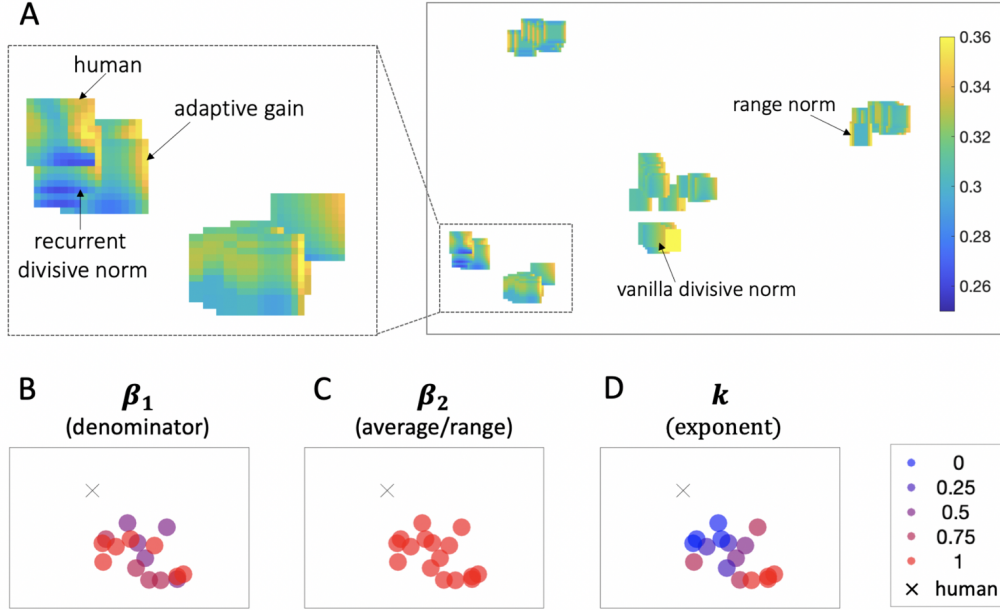


Figure 2.9: Charting decoy models. **A:** Embedding space for normalization models of decoy effects. A t -distributed stochastic neighbour visualization of the maps of decoy influence produced by different variants of the grandmother model. Each map represents a variant of the grandmother model positioned in 2D space such that models with similar decoy influence patterns are nearby, while models with more different decoy patterns are further apart. Heat maps illustrate decoy influence. The rectangle on the left shows a zoomed in version of the denoted subset of embedding space. **B-D:** These plots depict the zoomed in subset of embedding space presented above. Each model-produced decoy map is denoted as a dot and color coded to indicate parameter value: β_1 (panel B), β_2 (panel C), or k (panel D). Human data is represented with a cross.

Fitting the Grandmother Model

Methods We fit 125 variations of the grandmother model, by varying β_1 and β_2 in 5 steps between 0 and 1, and k in 5 steps between 0.001 and 1, in addition to the four constrained parametrizations of the grandmother model that result in the vanilla divisive normalization, recurrent divisive normalization, adaptive gain and range normalization models (Table 2.2). We used the t -distributed stochastic neighbor embedding (t-SNE) visualization technique (Van Der Maaten and Hinton 2008) to calculate relations among the resulting decoy influence maps from the 129 parametrizations. We specified the t-SNE hyperparameter perplexity, which controls the number of expected close neighbors, following guidelines in the

2. A map of decoy influence in multialternative economic choice

literature to balance the trade-off between perplexity and the Kullback–Leibler divergence (Cao and Wang 2017).

Results Fitting the grandmother model to human data revealed the different patterns of decoy influence predicted by variations of the general coding scheme. The embedding plot in Fig. 2.9a shows the simulated decoy maps for each of the 129 parametrizations of the grandmother model. Neighboring maps reflect models that produce relatively similar patterns of decoy influence (and vice versa for distant points). In Fig. 2.9b-d the points are colored according to levels of β_1 , β_2 and k , revealing the human data is neighbored by maps generated by models with high values of parameters β_1 and β_2 , i.e. those that resemble the adaptive gain model.

We note that the precise value of parameter k , which interpolates between models implementing logistic and linear divisive normalization, is less important for fitting human data. This implies that the model is equally well fit with an exponentiated implementation of recurrent divisive normalization ($k = 1$, such as the adaptive gain model) and with a linear implementation of recurrent divisive normalization ($k \sim 0$), where inputs are transformed with a power parameter s prior to normalization ($s \neq 0$, see Eq. (2.16)). This latter formulation of recurrent divisive normalization (as in Eq. (2.16)) is conceptually analogous to a form of recurrent divisive normalization in which values are power transformed (as in Daviet 2018):

$$u_i(A_i)^{PRDN} = \frac{v(A_i)^\alpha}{v(A_i)^\alpha + v(avg_i^{ABD})^\alpha + c_i} \quad (2.19)$$

Indeed, Bayesian model selection reveals that this model fits the data equally well as the adaptive gain model (exceedance probability for PRDN = 0.49), suggesting that our dataset cannot arbitrate between an exponential and power transform of attribute values.

Why is the nonlinear compression of attribute values important? To explore the role of this procedure, we compared the maps of decoy influence produced by an implementation of the grandmother model incorporating a logistic (i.e. exponential) transform of attribute values (equivalent to the adaptive gain model) and an

2. A map of decoy influence in multialternative economic choice

implementation of the grandmother model incorporating attribute values linearly (equivalent to recurrent divisive normalization without a power transform) in Fig. 2.10a. The nonlinearity qualitatively changes the form of the decoy map. The model featuring linearly coded attribute values produces stronger contextual effects for decoys of higher attribute values. This result can be traced back to the shape of the transducer function under linear attribute values (Fig. 2.10b, red curve), which is a non-symmetric concave function. Consequently, the curve deviates more from the identity line along the higher end of the abscissa and this leads to higher distortions in the translation of higher compared lower attribute values.

This is in contrast to what happens in the presence of an exponential nonlinearity as in the adaptive gain model. This compression of attribute values produces an S-shaped transducer (Fig. 2.10b, blue curve), akin to the sigmoidal shapes of response functions ubiquitous in neural coding. Notably, this transducer is symmetric around the inflection point (which corresponds to the mean of attribute values of the context when $c_i = 0$) and thus results in equal levels of distortion across the low and high end of attribute space. The bias term c_i can tip the balance in either direction: it can produce stronger effects for higher attribute values when $c_i > 0$ and stronger effects for lower attribute values when $c_i < 0$ (see Fig. 2.6d for simulations). Of note, the sigmoidal shape of the transducer here can be approximated by power transforming inputs prior to normalizing them in the recurrent divisive normalization scheme (i.e. $s > 0$ and $k \sim 0$, Fig. 2.10d lightest blue curve). This result helps clarify why this operation can capture the human data equally well as the exponential nonlinearity in the adaptive gain model.

Finally, we asked whether the adaptive gain model fits better than the full grandmother model after appropriate penalization for complexity. A failure to do so would imply the existence of a “hybrid” normalization solution that fits the human data even better, presumably involving some combination of parameters that has yet to be described in the literature. To assess this, we performed Bayesian model selection on complexity-penalized model fit metrics (Bayesian Information Criterion, BIC) which revealed that the exceedance probability for the normalization scheme

2. A map of decoy influence in multialternative economic choice

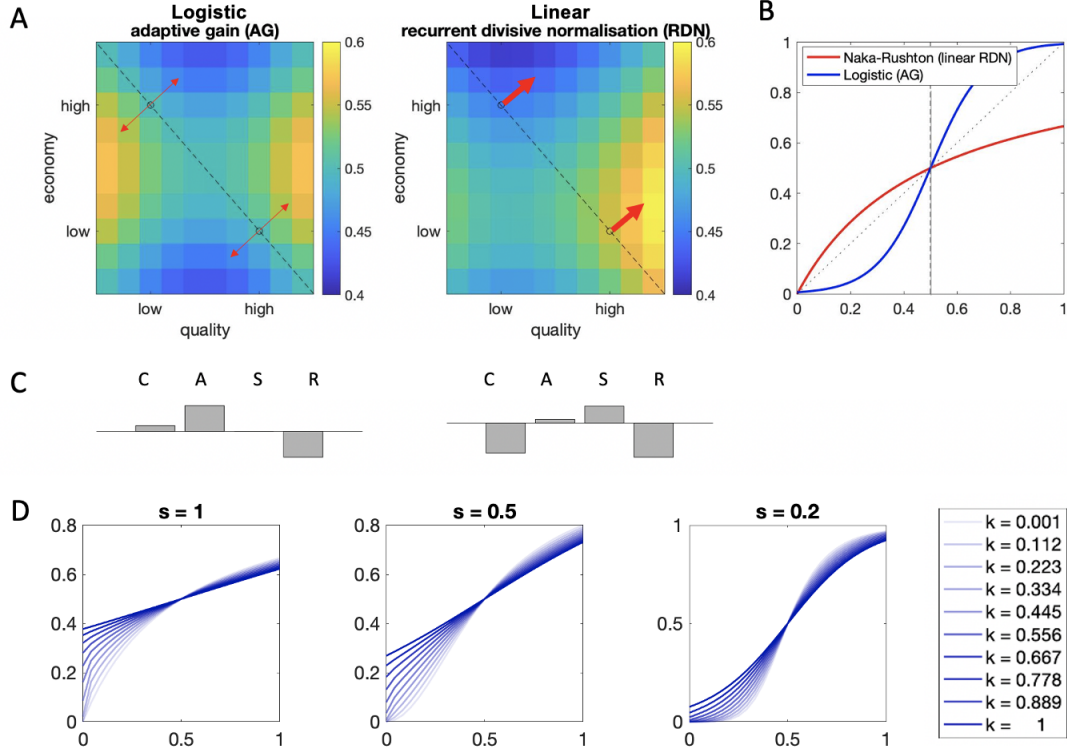


Figure 2.10: The role of the compressive nonlinearity. **A:** Simulated decoy maps for the adaptive gain and recurrent divisive normalization models. It is noteworthy that the adaptive gain model (with bias terms $c_i = c_j = 0$) produces symmetric regions of repulsion and attraction around the line of isopreference. By contrast, "linear" recurrent divisive normalization (i.e. RDN without a power transform applied to values prior to transduction) produces stronger repulsion and attraction for superior decoys (i.e. decoys above rather than below the isopreference line). **B:** The asymmetric pattern of decoy influence seen for linear RDN occurs because the transducer is a (decelerating) Naka-Rushton function. This function is concave, meaning that the derivative is always higher (i.e. curve steeper) for low than high attribute values irrespective of the value of x . By contrast, the transducer for AG is sigmoidal and it is thus symmetric around the midpoint (which is itself adjusted to the context; this is the "adaptive" part). This panel plots the transductions applied to inputs by the Naka-Rushton function in red and by the Sigmoidal function in blue, with $v((avg)_i) = 0.5$ (dashed line), $s = 0.1$. The recurrent divisive normalization framework (without additional nonlinearity) implies that those attribute values which are always relatively smaller (closer to zero) are processed with higher gain. By contrast, the adaptive gain model implies that resources are allocated preferentially to the mean of a context, exaggerating binary distinctions which potentially straddle that midpoint (e.g. "low" vs. "high" value). **C:** The relative strengths of the compromise, attraction, similarity and repulsion effects per the two decoy maps. **D:** Illustration of the transfer function of the grandmother model, with $c_i = c_j = 0$, $\beta_1 = \beta_2 = 1$, and $v((avg)_i) = 0.5$, across different values of k and s . Note that with this parametrization, the transducer function is equivalent to a logistic function when $k = 1$, and to a Naka-Rushton function when $k \sim 0$ and $s = 1$. Power transforming inputs (by setting $s \neq 1$) in the Naka-Rushton definition approximates the sigmoidal shape of the adaptive gain transducer.

2. A map of decoy influence in multialternative economic choice

avored by our empirical data, the adaptive gain model, over the grandmother model is 0.97, offering evidence against a hybrid solution.

2.3.4 Interim Discussion

Altogether, the results of our modeling suggest that a very simple computational scheme appealing to the relative encoding of attribute values may capture the full pattern of decoy influence in our dataset. We identified a 5-parameter model, previously termed adaptive gain, which fit the data exceptionally well, both at a fine scale (at the level of effects in individual participants) and on the aggregate level (the average map of decoy influence). We compared this model to related theoretical accounts in the literature, which appeal to a contextual normalization of attribute values, first, via conventional model selection techniques and second, through a comprehensive model comparison exercise for which we devised a flexible “grandmother” model encompassing a wide space of putative normalization schemes. The results of our computational work indicated that the rich pattern of decoy influence is best accounted for by information processing schemes that compressively transduce inputs, appealing to normalization by the central tendency of context and recurrently overweighing the contribution of the target input to the normalization.

2.4 Discussion

Decoy effects have been studied for decades, but substantial controversy has surrounded their replicability, interrelationship, and computational origins. This chapter sheds new light on these debates by gathering and analyzing a large-scale dataset that systematically maps the influence of a decoy stimulus across both the inferior and superior locations of multiattribute space. Conducting our analysis in a conventional fashion, we broadly replicate past studies, in that we find strong attraction effects, strong compromise effects, and a weaker and more variable similarity effect (not significant in our dataset). As in previous studies, the three decoy effects are correlated across the cohort, with a positive relationship observed

2. A map of decoy influence in multialternative economic choice

between attraction and compromise, and a negative relationship between those two and the strength of the similarity effect (Berkowitsch et al. 2014). This finding implies that three decoy phenomena have a single cause, and indeed previous dynamic models (in which information is accumulated over time) have been able to capture the three discrete effects with a single set of parameters (Roe et al. 2001; Usher and McClelland 2004; Bhatia 2013; Trueblood et al. 2014; Bhui and Gershman 2018). Here, we use dimensionality reduction on the full decoy influence map to confirm that indeed, there is a single component that explains the vast majority ($\sim 95\%$) of the variance in decoy influence, suggestive of a single computational origin for these biases.

To understand the computational origin of decoy effects, we chose to model our data with a framework based on divisive normalization. We made this choice because the normalization model offers a simple, parsimonious account of contextual biases in decision-making based on a rich, neurobiologically grounded tradition in the cognitive sciences (Carandini and Heeger 2012; Louie et al. 2013; Rigoli 2019; Landry and Webb 2021; Webb et al. 2021). In particular, it allowed us to systematically measure the influence of various candidate computational steps on the predicted decoy map, providing an interpretable mapping from model to data (Fig. 2.6). On this basis, we were able to establish (for example) that normalization occurs relative to the average of the available values (via a sigmoidal gain function) rather than to the lower end (via a concave gain function) as proposed in some previous models (Fig. 2.10). This characteristic sigmoidal shape of the transfer function may be approximated by transforming inputs via a power term ($\alpha > 1$) in recurrent divisive normalization (Daviet 2018; Webb et al. 2021).

Overall, out of the models tested here, evidence favored a model that has previously been described in the literature as the *adaptive gain model* (Cheadle et al. 2014; Li et al. 2018). This account is closely related to other models involving recurrent divisive normalization, especially those proposing that values are nonlinearly transformed beforehand, as well as being very similar to another model known as the logistic model of subjective value (Rigoli 2019). There is a

2. A map of decoy influence in multialternative economic choice

close correspondence between qualitative features of model-simulated and human performance displayed in Fig. 2.4, and in particular, a close correspondence achieved after decomposition of the decoy map into linear components using singular value decomposition (Fig. 2.5). The adaptive gain model even predicted a new and potentially counter-intuitive relationship between the decoy effects: that when the compromise effect is positive, attraction should dominate over repulsion (and vice versa), a prediction that was satisfied in the data.

Under the adaptive gain control framework described here, decoy effects occur because of contextual biases arising when each target item is transduced via a logistic function whose inflection point lies at the mean of all three items including the decoy. For example, the “attraction” effect thus occurs because when the decoy is lower in value than item A, the inflection point is lower than item A, and so A lies at the steepest portion of the sigmoidal gain function and is thus “overvalued” or repulsed away from this mean point. The precise converse occurs when the decoy is higher in value than A, as well as for B. Previous work has demonstrated that exactly this mechanism can in principle account for a range of decision biases arising in the presence of distractors, across perceptual, cognitive and economic domains (Li et al. 2018). The brain may have evolved the type of normalization scheme proposed here because it promotes efficient neural coding (Summerfield and Tsetsos 2015; Summerfield and Tsetsos 2020).

Whereas the attraction effect tends to be highly robust and consistent across participants, the strength and directionality of the compromise effect and similarity effects tend to be more idiosyncratic. Indeed, the similarity effect did not reach statistical significance in our dataset. In our model, the compromise and similarity effects occur when attractive and repulsive processes are asymmetric due to differential weighing or biasing of the two attributes, causing attraction effects (and their converse for superior decoys) to warp and/or “spill over” into locations where compromise and similarity decoys are typically tested. In other words, the fragile nature of the compromise and similarity effects might be at least in part due to heterogeneity in the asymmetric way each attribute is coded or transformed, which

2. *A map of decoy influence in multialternative economic choice*

in turn might (for example) be due to differing choice concerning stimulus materials. A systematic unpicking of ways in which different classes of stimulus material (e.g. numerical values in distinct ranges, perceptual stimuli such as rectangles, and vignettes) are encoded, and thus why decoy effects may or may not have emerged in previous studies (Frederick et al. 2014; Huber et al. 2014; Spektor et al. 2021), is beyond the scope of our research project here. However, our simulations suggest that a relatively low-dimensional encoding model may be sufficient to capture this variation and thus to pinpoint the source of variation in previous studies.

2.5 Conclusion

Adding a “decoy” alternative to a choice set can shift preferences in multialternative economic decisions. A long research tradition focusing on this contextual bias has provided evidence for three classic decoy effects – the compromise, attraction and similarity effects – but their replicability, interrelationship and computational origins have sparked controversy. To address these, we carried out a large-scale incentive-compatible study, charting the full two dimensional attribute space of decoy effects. Our analyses indicated that a single explanatory factor can account for the rich pattern of decoy influence we observe. Indeed, a simple process model, appealing to the principle of divisive normalization, can capture our results both at the individual and aggregate level.

2.6 Acknowledgements

This chapter is based on the published work *Dumbalska, T., Li, V., Tsetsos, K., & Summerfield, C. (2020). A map of decoy influence in human multialternative choice. Proceedings of the National Academy of Sciences, 117(40), 25169-25178.*

The human data for this chapter was procured by Dr Vickie Li. Dr Vickie Li and Prof Christopher Summerfield jointly designed and Dr Vickie Li built the experiment and administered data collection. Dr Konstantinos Tsetsos provided feedback on code for the modeling simulations.

This chapter is based on a manuscript under peer review:

Dumbalska, T., Rudzka, K., Smithson, H. E., & Summerfield, C. (2021). How do (perceptual) distractors distract?

3

How do perceptual distractors distract?

Contents

3.1	Introduction	60
3.2	Experiment 1	64
3.2.1	Methods	64
3.2.2	Results	73
3.2.3	Interim Discussion	78
3.3	Experiment 2	78
3.3.1	Methods	79
3.3.2	Results	81
3.3.3	Interim Discussion	85
3.4	Experiment 3	86
3.4.1	Methods	86
3.4.2	Results	89
3.4.3	Interim Discussion	94
3.5	Computational Modeling	95
3.6	Discussion	98
3.7	Conclusion	102
3.8	Acknowledgements	103

When a target stimulus occurs in the presence of distractors, decisions are less accurate. But how exactly do distractors affect choices? This chapter explores that question using measurement of human behavior, psychophysical reverse correlation and computational modeling. We contrasted two models: one in which targets and distractors have an independent influence on choices (independent model) and one

3. How do perceptual distractors distract?

in which distractors modulate choices in a way that depends on their similarity to the target (interaction model). Across three experiments, participants were asked to make fine orientation judgments about the tilt of a target grating presented adjacent to an irrelevant distractor. We found strong evidence for the interaction model, in that decisions were more sensitive when target and distractor were consistent relative to when they were inconsistent. This consistency bias occurred in the frame of reference of the decision, that is, it operated on decision values rather than on sensory signals. Further, it was independent of spatial attention. A normalization framework, where target features are normalized by the expectation and variability of the spatial context, successfully captures the observed pattern of results.

3.1 Introduction

Visual scenes in the real world are typically cluttered, and so perceptual decisions necessarily occur in the context of irrelevant information. For instance, when adjusting a crooked picture frame, a decision maker must focus on its offset from vertical, ignoring the tilts of surrounding paintings, as those are irrelevant to the task. A long line of research across psychology and neuroscience has investigated humans' ability to focus on an imperative stimulus in the context of distracting input. A classic laboratory analogy of the above scenario asks participants to judge a cued item in a multi-element array. For example, participants might be asked to discriminate the tilt of a (cued) target grating presented among one or more irrelevant gratings (distractors). Empirical work has focused on understanding how the cue impacts behavior and modulates neural signals, informing models of how target processing is prioritized in the face of distraction (Reynolds and Chelazzi 2004; Carrasco 2011). These models have, for example, produced detailed predictions about how tuning functions in visual neurons vary according to whether the neuron's preferred stimulus is attended or unattended (Moran and Desimone 1985; McAdams and Maunsell 1999; Treue and Trujillo 1999; Reynolds and Heeger 2009).

3. How do perceptual distractors distract?

However, we know remarkably little about how distraction itself influences decisions. When (a cued) target and (an uncued) distractor occur together, how exactly does the distractor affect choices? This question lies at the intersection of several distinct literatures in psychology and neuroscience which are rarely discussed together. Firstly, it is sometimes argued the effects of spatial cueing can be understood through the normative lens provided by Bayesian inference. The cue provides probabilistic information about which of two locations is decision-relevant; an ideal observer will combine this with noisy signals arising from target and distractor locations, weighted by their likelihoods (according to the probabilistic cue, Dayan and Zemel 1999; Eckstein et al. 2009). In other words, the distractor and target hold independent sway over decisions, with the job of the observer being to weigh them appropriately according to prior information. Similar models have sometimes tried to subsume the literature on attention within the normative framework given by decision theory (Anderson 2011).

A related view of distraction can also be found in a popular theoretical framework proposing that stimuli effectively compete for neural resources. Attention serves as controller that can bias processing towards percepts which are more relevant for ongoing behaviors (Desimone and Duncan 1995). While this view does not clearly translate onto decision theory, it nevertheless implies that decision variables are derived from admixtures of independent (relevant and irrelevant) sensory signals, with their relative weighing determined by the strength of control, for example arising in prefrontal cortex (Miller and Cohen 2001; Egner and Hirsch 2005). Alternative but related views posit that relevant and irrelevant features race independently to drive decisions (Bundesen 1990) or contribute to the differential weighing of mutually inhibitory sources of evidence in an accumulation-to-bound model, a framework known as decision field theory (Busemeyer and Townsend 1993). One specific instance of this class of theory argues that when humans choose between a preferred and a less preferred item, inputs to the decision process are given by a subtractive mixture of the two stimulus values, with a higher weight given to the currently fixated item (Krajovich et al. 2010).

3. How do perceptual distractors distract?

What the aforementioned models share is the central idea that target and distractor contribute independently to choice. Broadly, we can view decisions as driven by a process such as:

$$y^{ind} = \beta_0 + \beta_1\theta_T + \beta_2\theta_D + \epsilon \quad (3.1)$$

where y^{ind} is the decision variable, θ_T and θ_D are the features of the target and distractor respectively, β_1 and β_2 are their respective weights, and ϵ is an error term. We can consider the application of the weights to be outside voluntary control, so that distraction is at least in part inevitable. Here, we call this the “independent” model of distraction. For the most part, assuming that accuracy is above chance, we can expect $\beta_1 > \beta_2$. Some theories posit special downstream mechanisms that attempt to eliminate the effect of β_2 altogether, known as distractor suppression (Geng 2014; Chelazzi et al. 2019).

However, a different proposal is that target and distractor do not wield independent sway over y but rather:

$$y^{int} = \beta_0 + \beta_1\theta_T + \beta_2 \cdot f(\theta_T|\theta_D) + \epsilon \quad (3.2)$$

We use the term $f(\theta_T|\theta_D)$ in this equation to refer to some unspecified process by which target and distractor interact. For instance, this could be a straightforward multiplicative process or it could involve rectification or discretization operations; e.g. target processing could be stronger or weaker when the distractor is similar/dissimilar (Blakemore et al. 1970) or congruent/incongruent (Eriksen and Eriksen 1974). Importantly, what this implies is that the influence of the distractor on the decision occurs only by virtue of how the distractor modulates or interacts with the target. Thus, we call this the “interaction” model of distraction.

This class of model has also been popular at various times and in various guises. For instance, one well-known body of literature discusses the phenomenon of contextual facilitation, whereby in the extrastriate visual cortex, neuronal responses to the receptive field center are biased by stimulation of the surround (Krause and Pack 2014). At the neural level, this area has been amply explored but fewer studies

3. *How do perceptual distractors distract?*

have examined how spatial context may modulate behavior. Still, previous work has established that detection thresholds for a grating are lowered if it is flanked by collinear gratings (Polat and Sagi 1994; Kapadia et al. 1995; Ito and Gilbert 1999; Gilbert and Li 2013). Here, the distractors influence the decision by virtue of their similarity to the target. Another example involving visual stimuli is the tilt illusion whereby the tilt of a central grating is perceived as repulsed away from a ring of flanking distractors (Blakemore et al. 1970). Again, the distractors do not influence the decision directly; for instance, decisions about the target are less repulsed if the distractors are more dissimilar, contrary to the predictions of the independent model. These sensory level interaction effects have been successfully modeled in a divisive normalization framework, where individual neural responses are scaled by the population response (Carandini and Heeger 2012).

Indeed, the family of normalization models discussed in the previous chapter, which successfully capture decoy effects in decisions about economic prospects, also constitute specific instances of the interactive model. In this class of models, known as divisive normalization, $f(\theta_T|\theta_D)$ may be viewed as a divisive operation, following the general form $y \sim \frac{\theta_T}{r+\theta_T+\theta_D}$, where r is a regularizer (Louie et al. 2013; Landry and Webb 2021). Through the years, alternative explanations of the phenomenon have sometimes instead appealed to the independent account (Busemeyer and Townsend 1993). However, it is difficult to reconcile those theories with findings that the modulation strength of a decoy depends on its proximity in conceptual space to the target (Soltani et al. 2012; Landry and Webb 2021), echoing results from the perceptual literature about the effect of target-distractor similarity on choice.

Relatedly, in the cognitive domain, perception of quantities (e.g. the number of African nations) or value (the price of bottle of wine) can be anchored by irrelevant prior information (Ariely et al. 2003). However, anchoring by an irrelevant quantity also seems to diminish with distance between the anchor and the target value (Wegener et al. 2001), indicative of an interaction between target and distractors. Together, these effects from diverse literatures imply a view whereby distractors and targets interact. Indeed, a recent paper showed how a version of the interaction model

3. How do perceptual distractors distract?

can explain a range of distractor-mediated phenomena in perceptual, cognitive, and value-based decision tasks (Li et al. 2018).

This chapter describes three experiments in which participants made a fine discrimination judgement about a target grating that is presented concurrently with an irrelevant distractor grating. We combined computational modeling and reverse correlation analysis to probe how the distractor influences choices. This approach allowed us to directly compare variants of the independent and interaction models as explanations of our data. Our experiments addressed three related questions. Firstly, how does the distractor influence decisions? Secondly, does it do so at the perceptual or decision level? And thirdly, how does spatial attention mitigate (or otherwise) the effect of distraction?

3.2 Experiment 1

To investigate the influence distractors wielded on choice, we presented participants with brief flashes of a pair of Gabor patches embedded in smoothed noise. We analyzed participants' reports of the tilt of the target grating on each trial, using information about the features of the orientation signal of both target and distractor stimuli, as well as information about the signal-like fluctuations in the noise of each stimulus.

3.2.1 Methods

Participants

Twenty-six participants (aged 25.33 ± 4.38) took part in Experiment 1. Two participants were excluded from the main analyses as the accuracy of their responses for one of the possible stimulus locations was at chance level ($|\text{accuracy} - 50\%| < 5\%$). Our exclusion strategy took into account that 75% accuracy (the staircase target and approximately the average performance in the experiment) could be achieved if participants only focused on one side of the screen (and achieved near 100% accuracy for stimuli there) while guessing on the other side (achieving near 50% accuracy for stimuli there). To ensure that such a strategy was not driving the

3. How do perceptual distractors distract?

observed level of accuracy, we excluded participants who performed at chance level for stimuli on either side of the screen. The study received ethical approval from the Central University Research Ethics Committee at the University of Oxford (approval reference number: R51752/RE001). All participants provided written informed consent and were compensated £10 per hour for their time.

Apparatus

Participants were seated in a dark room approximately 60cm away from a computer monitor (120Hz refresh rate, 1280x1024 resolution, 21" CRT) with linearized output light intensities. Visual stimuli were created and presented with Psychophysics Toolbox Version 3 (PsychToolbox-3, Brainard 1997) for MATLAB.

Stimuli

The experimental stimuli were noisy gratings, each of which consisted of a Gabor pattern and smoothed Gaussian noise. The Gaussian envelope of the Gabor patterns had a standard deviation of 1° of visual angle. The phase of each Gabor pattern was sampled from a uniform distribution; orientation was drawn from a uniform distribution with range -10° to 10° rotation from a vertical decision boundary. Noise was generated independently for each stimulus by sampling pixel values from a Gaussian distribution and passing the resulting pixel values through a smoothing Gaussian filter. The dimension of the filter, 0.083° of visual angle, was chosen to maximize the trial-to-trial variability of the convolution between the smoothed noise and Gabor signal (Wyart et al. 2012). Stimuli were presented at full contrast. The ratio of the contrast of the Gabor pattern to the contrast of the noise of a stimulus was adaptively determined for each participant (see procedure for staircase details). The average value of the contrast ratio produced by the staircase across participants was 0.35 (i.e. 35% of the contrast of the stimulus was dedicated to the signal and 65% to the smoothed noise). Stimulus size spanned 4° visual angle and stimulus position was centered at 4° of visual angle to the left and right of fixation.

3. How do perceptual distractors distract?

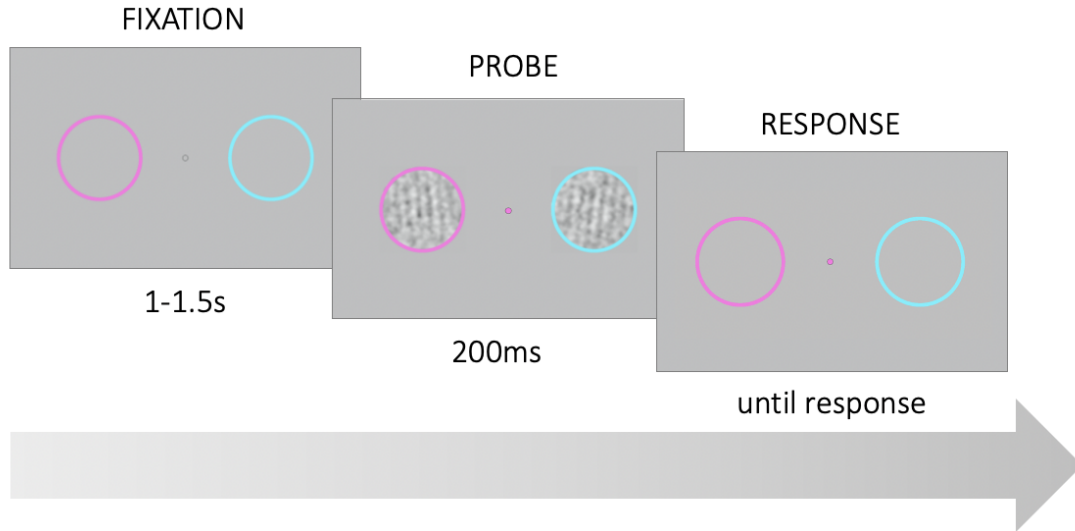


Figure 3.1: Trial structure. Participants were asked to fixate on a point in the center of the screen. Two noisy gratings appeared briefly within the colored rings on the left and right of the fixation point. At stimulus onset, the fixation point changed color and served as a probe. Participants had to report the tilt (CW or CCW) of the probed grating relative to vertical.

Experimental Procedure

Each participant completed the experiment in a single hour-long session. The experimental session consisted of 2 training blocks and up to 9 test blocks (100 trials each). The training blocks served to familiarize participants with the task and calibrate participant performance. To this end, during training we employed an adaptive staircase procedure, accelerated stochastic approximation (Lu and Doshier 2013), to titrate the contrast ratio for the noise in the stimuli for a target participant performance of 75% correct. The training and test trials followed the same structure; the only difference was in the contrast ratio of the stimuli. During training blocks, the contrast ratio was adjusted in each trial according to the staircase, which varied the step size of the change in contrast adaptively after each response. The training blocks had a variable trial count. Each training block ended after the staircase converged and took on average between 5-10 minutes. We used the contrast ratio estimate from the second training block to generate stimuli for the test blocks.

Each test block lasted approximately 5 minutes and participants were invited to take short breaks between blocks. During the experiment, participants were

3. *How do perceptual distractors distract?*

asked to fixate their gaze at the center of the screen on a fixation point between two colored rings (Fig. 3.1). The rings were located at 4° of visual angle to the left and right of center. The left ring was magenta and the right ring was cyan. In each trial, after a uniformly variable interval between 1-1.5s, two noisy gratings appeared within the colored rings for 200ms. At stimulus onset, the fixation point assumed the color of one of the two rings until the end of the trial. The color of the fixation point served as the probe indicating which of the two stimuli was the target. Participants had to report whether the target was tilted clockwise or counterclockwise (CW or CCW) relative to a vertical decision boundary. Participants reported their responses via keyboard presses (left and right arrow keys) and instantaneously received fully informative auditory feedback (correct: a high 880Hz tone, incorrect: a low 440 Hz tone).

Analyses

In our analyses, we asked how the distractor stimulus affects participant choices. We addressed this question with both regression-based and reverse correlation analytic approaches.

Stepwise Regression We first took a regression-based approach to identify the contributions of the target and distractor to choice. We defined the decision-relevant features of the target θ_T and distractor θ_D as their respective angular offsets from the decision boundary on each trial. We assumed that choice probabilities were a logistic function of a decision variable y , computed from both target and distractor features:

$$p(CW) = \frac{1}{1 + e^{-y}} \quad (3.3)$$

where $p(CW)$ corresponds to the probability to respond clockwise on a given trial.

In order to ascertain the nature of the relationship between distractor features and choice, we utilized a forward and backwards stepwise regression approach for including predictors. Due to the way we have framed the question above, there is some liberty in specifying the nature of the target-distractor interaction. The

3. How do perceptual distractors distract?

hierarchical factor knock-in and knock-out approach of the stepwise regression allowed us to consider a number of possible definitions of the interactive relationship $f(\theta_T|\theta_D)$ and empirically evaluate each of them. The predictor variables we considered included:

- θ_T , the independent effect of the target angular offset
- θ_D , the independent effect of the distractor angular offset
- $|\theta_D|$, the absolute value of the distractor angular offset
- *congruency*, a binary indicator whether the angular offsets of the target and distractor fall on the same side of the category boundary
- $\theta_T \cdot \theta_D$, the (unsigned) multiplicative interaction between the target and distractor angular offsets
- $\theta_T \cdot |\theta_D|$, the multiplicative interaction between the target angular offset and the absolute value of the distractor angular offsets
- $\theta_T \cdot |\theta_T - \theta_D|$, the interaction between target angular offset and the consistency between the target and the distractor angular offsets

We estimated separate models for each individual participant. We used the **stepwise** algorithm from the MATLAB Statistics and Machine Learning Toolbox for the backward and forwards stepwise regression analysis and the Statistical Parametric Mapping Toolbox for Bayesian model selection. We used Bayesian model selection (Stephan et al. 2009) to compare the cross-validated log likelihoods of the model identified with the stepwise regression against nested competitor models.

To further explore the effect of the distractor on choices, we used logistic regression to estimate target sensitivity across different values of the distractor. We sorted distractors into 6 bins, each bin spanning 3.33° , and regressed choices within each distractor bin on target offset, θ_T :

$$y = \beta_0 + \beta_1 \theta_T \quad (3.4)$$

Running this regression model separately for each distractor bin allowed us to estimate the influence of θ_T given different values for θ_D and detect distractor-dependent changes to target sensitivity.

3. How do perceptual distractors distract?

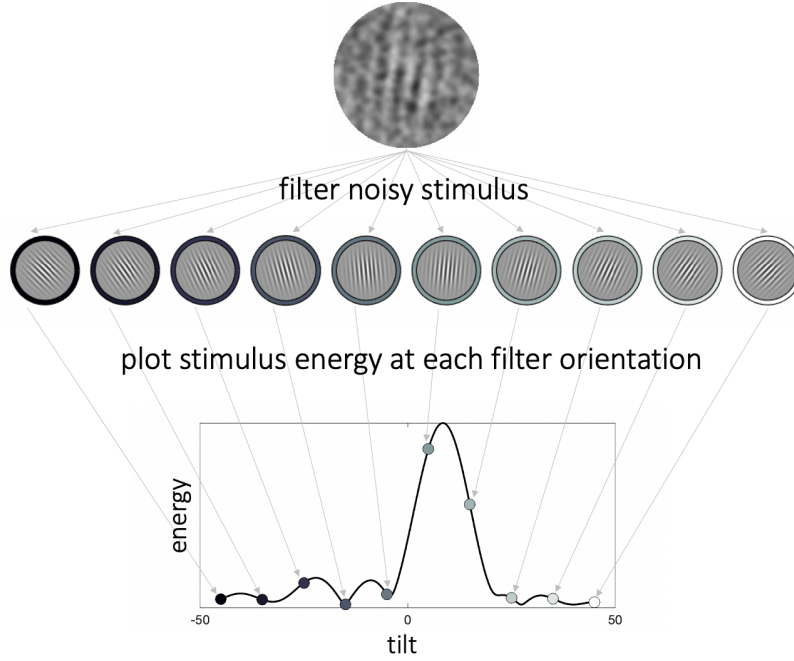


Figure 3.2: Stimulus energy profiles. We filtered each noisy stimulus through a pool of oriented Gabor patches to calculate energy across orientation space. The higher the stimulus energy at a given orientation, the more the noisy stimulus resembles a Gabor patch of that orientation. The collection of energies of a given stimulus across orientation space constitutes the energy profile of that stimulus.

Reverse Correlation The regression analysis describes each target and distractor with a scalar quantity that indexes the disparity between its orientation signal and the boundary. However, the target and distractor stimuli constituted images containing multidimensional information about a full range of orientations. In particular, the energy at each orientation varied from trial to trial for both target and distractor because of the smoothed noise we applied. This allowed us to carry out a reverse correlation analysis to examine how the signal-like fluctuations in signal energy at each orientation influence choices, and thus to plot decision kernels in orientation space for both target and distractor.

Stimulus Energy Profiles To quantify the relationship between trial-to-trial signal-like fluctuations in the noise of each stimulus and participant decisions, we first characterized the filter energy of each stimulus S across orientation space using

3. How do perceptual distractors distract?

an ordered set of Gabor filters $F_{\phi_n}^\theta$ (Fig. 3.2). The spatial frequency of the filters matched the spatial frequency of the Gabor patterns. Filter orientation (θ) ranged between -45° and 45° in increments of $.5^\circ$. To correct for differences between the phase of the stimuli (sampled uniformly between $[0, 1] \cdot 2\pi$) and the filters, we specified 5 filter phases ϕ_n ($[0.1, 0.9] \cdot 2\pi$, in steps of $0.2 \cdot 2\pi$). We calculated stimulus response to each filter $R(S|F_{\phi_n}^\theta)$ as the amount of variance in the filter that was explained by the covariance of the filter and stimulus. We then estimated the filter phase ϕ^{max} (which can take any value between $[0, 1] \cdot 2\pi$) at which the highest response would be obtained for that orientation, by fitting the following mathematical model:

$$R(S|F_n^\theta) = E(S|\theta) \cdot \cos(\phi_n - \phi^{max}) \quad (3.5)$$

where $E(S|\theta)$ corresponds to the energy of the stimulus S at orientation θ , and ϕ_n to the phase of filter n (n ranges from 1 to 5). Thus, the model assumes that the relationship between filter response and filter phase follows a cosine function. We can use gradient descent to identify the value of ϕ^{max} that would produce the strongest filter response, as well as predict the magnitude of that response, i.e. signal energy at that orientation, $E(S|\theta)$.

Calculating the energy profiles in this manner, that is, via convolution with Gabor filters of the same parametric specification as the target signal, ensures that we are obtaining orientation energy at the spatial scale of the target. This operation is more specific than, for instance, a Fourier transform in that it takes into account the expectation for the target (and distractor) signals. Thus, the energy profile lends itself to an intuitive interpretation – the more a noisy stimulus resembles a signal grating with a given orientation, the higher the filter energy of the stimulus at that orientation.

Decision Kernels Next, we related fluctuations in energy at each orientation in each trial to participant choices to calculate the decision kernel, i.e. estimate of the latent function mapping stimulus features on choice (Fig. 3.3). We did this via a series of parallel binomial regressions, one per each filter orientation, with

3. How do perceptual distractors distract?

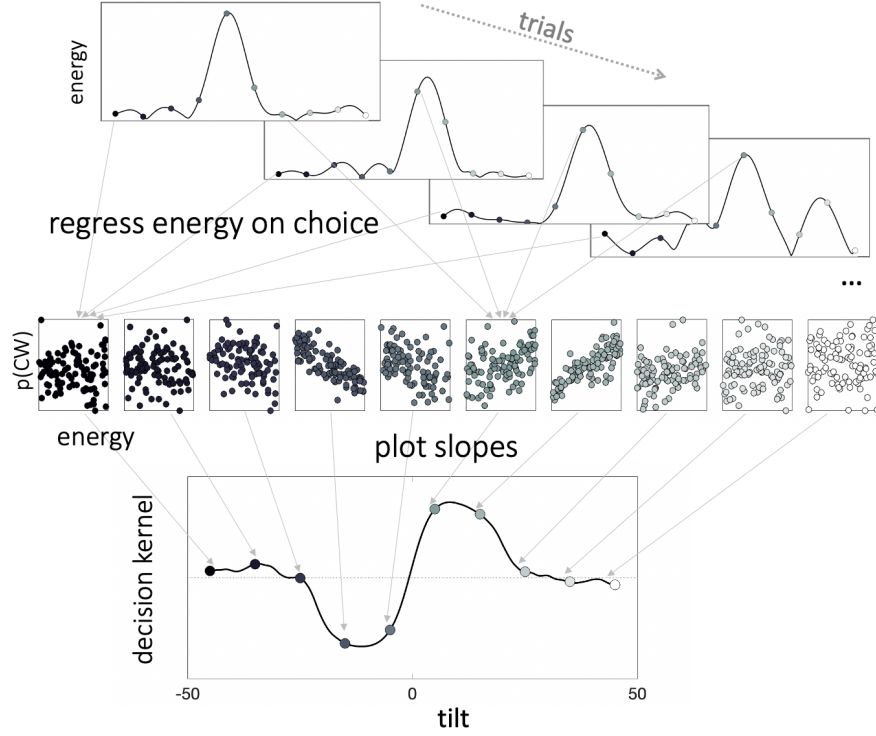


Figure 3.3: Decision kernels. We regressed choices ($p(CW)$) on z -scored stimulus energy via separate binomial regressions in each orientation bin. The resulting beta coefficients make up the decision kernel. Thus, the decision kernel indexes the relationship between energy at a given orientation and choice. Positive kernel values suggest that the more a stimulus resembles a Gabor patch of that orientation, the more likely is a participant to make a clockwise choice.

participant choice (CW or CCW) as the outcome variable and target energy $E(T_i|\theta)$, standardized within each orientation bin as a z score, as the predictor variable. The resulting beta coefficients capture the relationship between choice and energy at each orientation, and when plotted together constitute the decision kernel. The reverse correlation approach we describe has been used previously to investigate the effects of feature expectation and attention, as well as spatial attention to target stimuli on choice sensitivity (Wyart et al. 2012; Barbot et al. 2014; Cheadle et al. 2015).

In addition to the kernel quantifying the relationship between target energy and choice, we calculated a kernel quantifying the direct effect of the distractor energy on choice (as in the independent model). We estimated the distractor influence by adding distractor energy as a competitive regressor in the model above. We assessed the independent effect of the distractor on choices by testing whether the average

3. How do perceptual distractors distract?

absolute values of the distractor kernel deviate from zero at different angles relative to the category boundary. To reduce the number of multiple comparisons, we divided orientation space into angles near the decision boundary ($|\theta| < 22.5^\circ$) and angles far from the boundary ($|\theta| > 22.5^\circ$) and averaged the absolute value of the decision kernel within those two bins. Our choice of median split of orientation space was guided by simulations of the theoretical decision kernels (i.e. simulated responses based on ground truth tilt corrupted by Gaussian noise). We compared the resulting kernel values in the two bins across participants to zero via a single sample t test.

Kernel Decomposition The decision kernel is smooth, so statistical tests conducted at different orientation bins are not independent. To tackle this issue, we reduced the dimensionality of the data via singular value decomposition (SVD). This approach allowed us to derive a set of basis functions (components) from the data itself, and to assess their relative weighing in driving choices. We carried out SVD on a matrix containing the energy profiles of all stimuli, both targets and distractors, from the experiment. The rows of the matrix indexed individual stimuli and the columns indexed orientation bins. The SVD analysis produced a set of components and assigned component scores for each stimulus, which correspond to the variance in the stimulus explained by a given component.

Next, we identified the (number of) components V^i which explained 95% of the variance in the data and assessed which of them contained information about the tilt of the stimulus relative the decision boundary via visual inspection. We then used the component scores U^i for the tilt-informative component(s) of the target and distractor stimuli as predictor variables in a regression model. The model was analogous to the regression model identified through the stepwise algorithm. The only difference was that we used the component scores for the target U_T^i and distractor U_D^i instead of their angular offsets, θ_T and θ_D , as predictor variables.

3. How do perceptual distractors distract?

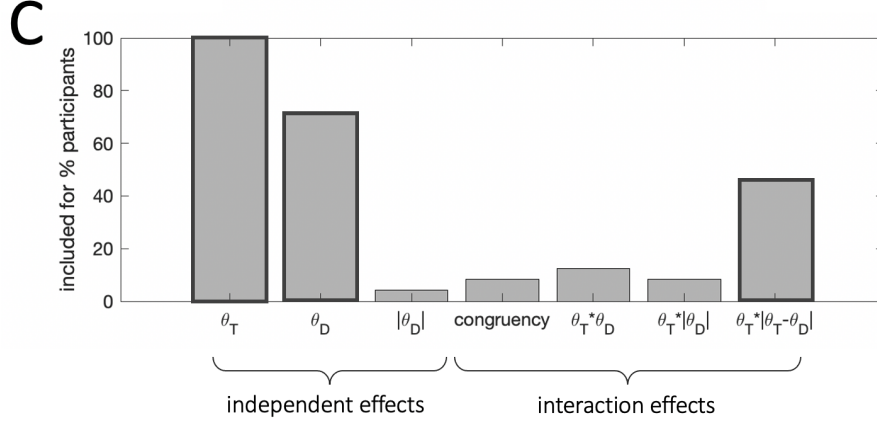


Figure 3.4: Results from stepwise regressions estimated individually for each participant in the experiment. Bar height indicates the percentage of participants for whom a parameter was included in the final model. Darker bar outlines indicate predictors we included in the final model.

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/54rf2/>) for this project.

3.2.2 Results

Regression Results

Fig. 3.4 illustrates the proportion of participants whose data was best explained by a model including each of the predictor variables we considered. The model variant that best fit the human data among the candidates we considered was the following:

$$y = \beta_0 + \beta_1\theta_T + \beta_2\theta_D + \beta_3(\theta_T \cdot |\theta_T - \theta_D|) \quad (3.6)$$

This model provided a better explanation of human choices compared to models that included a different functional form for the interaction $f(\theta_T|\theta_D)$. It also provided a better fit to models that omitted either the independent θ_D or interaction effect $f(\theta_T|\theta_D)$ of the distractor. Bayesian model selection on cross-validated model log likelihoods produced exceedance probabilities of $p = 0.80$ and $p = 0.99$ for this model over an independent effect only model (i.e. $\beta_3 = 0$) and over an interaction effect only model (i.e. $\beta_2 = 0$), respectively.

3. How do perceptual distractors distract?

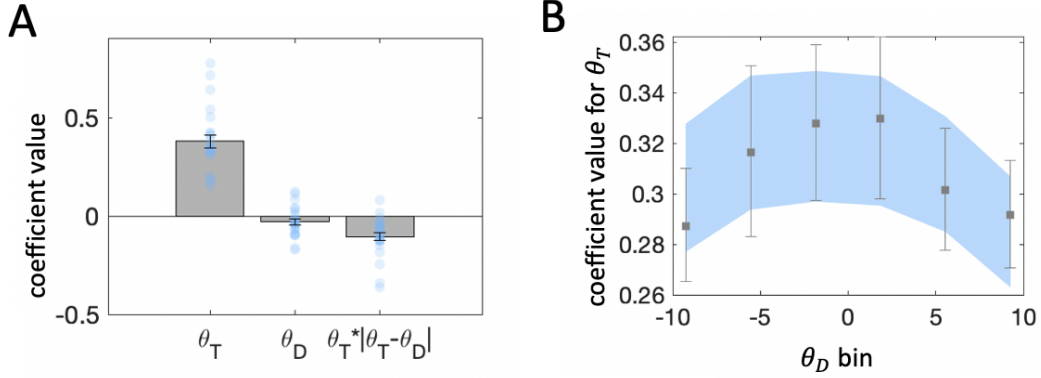


Figure 3.5: **A:** Beta coefficients from regression of choices on target and distractor orientation. Note that the interaction term is divided by 10, to ensure that unit changes in each predictor variable are reported in consistent terms. Grey bars denote averages, error bars denote standard error of the mean, blue dots represent individual participants. **B:** Beta coefficients for target orientation from binned regressions. Distractor orientation was divided in 6 equally spaced bins. Within each bin we estimated a binomial regression of choices on target orientation. Grey squares denote regression with human choice as the outcome variable, and error bars denote standard error of the mean. Blue region denotes standard error from the same regression, but with choice estimates derived from the regression model (from panel A) fit to human data.

In Eq. (3.6) the coefficients β_1 and β_2 encode the influence (weight) of the target and distractor on choice. The third coefficient β_3 encodes the influence of the target-distractor interaction. The functional form of the interaction term indicates a consistency bias – the influence of the target on choices depends on the degree of consistency between the target and distractor features. That is, choice sensitivity to θ_T varies with the similarity between θ_T and θ_D .

The best-fitting coefficients β_1 , β_2 and β_3 are shown in Fig. 3.5a. Firstly, θ_T holds significant sway over decisions ($t_{23} = 11.41$, $p < 0.001$). This is expected because participants were instructed to discriminate the target item. Secondly, the effect of consistency, indexed by β_3 , was statistically significant at the group level ($t_{23} = -5.09$, $p < 0.001$). The negative sign of the coefficient means that the influence of the target on choices was greater when the target-distractor consistency was greater (i.e. a positive consistency bias).

Thirdly, the effect of θ_D was more variable. It appears that the direction of influence of the distractor varied from individual to individual ($\beta_2 = -0.03 \pm 0.075$)

3. How do perceptual distractors distract?

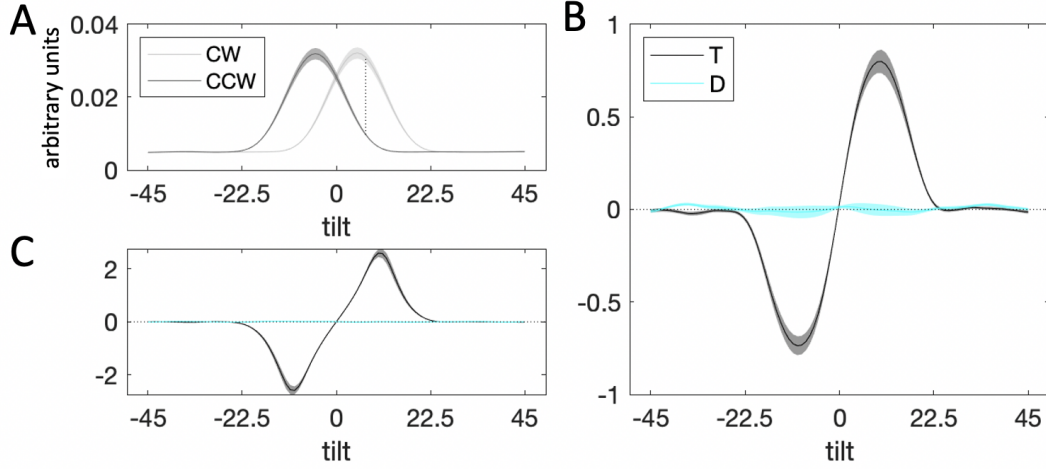


Figure 3.6: Decision kernel results. **A:** Average stimulus energy profiles. Dark grey curve denotes CCW stimuli, light grey curve – CW stimuli. Shaded regions correspond to $M \pm \text{SEM}$. The dashed line denotes maximum signal-to-noise ratio between the distribution for CW and CCW stimuli. **B:** Decision kernel based on participant choices. Black curve denotes the kernel for the target stimulus, cyan curve – the kernel for the distractor. Shaded regions correspond to $M \pm \text{SEM}$. **C:** Decision kernel based on ground truth stimulus tilts. Black curve denotes the kernel for the target stimulus, cyan curve – the kernel for the distractor. Shaded regions correspond to $M \pm \text{SEM}$.

and a t test conducted on β_2 at the group level failed to reach significance at an α level of .05, ($t_{23} = -1.99$, $p = 0.06$). Although this coefficient was nonsignificant at the group level, the fits of the regression model to data from individual participants suggested that the inclusion of θ_D was warranted by the variance it explained.

Fig. 3.5b illustrates the results of the binned regression analysis estimating the influence of the target on choices across different distractor values. Sensitivity to the target was reduced for more extreme distractor orientations, across both clockwise and counterclockwise orientations. This finding is expected from the consistency effect, because target and distractor are necessarily more different on average when one of them is at the extreme. This pattern of results is also captured by the predicted choices from the stepwise regression-identified model (Eq. (3.6)).

Reverse Correlation

Fig. 3.6b illustrates the decision kernel for the target (black curve) and distractor (cyan line) stimuli averaged over all trials of the experiment. Positive kernel values

3. How do perceptual distractors distract?

indicate that an increase in energy at a given orientation is associated with higher probability of responding CW and negative kernel values – with responding CCW. The peak of the target kernel is on the CW side of orientation space, close to the maximum signal-to-noise ratio of the CW and CCW energy distributions (Fig. 3.6a, dashed line). Note that the kernel tapers towards zero beyond the range of -10° and 10° degrees from which stimulus orientations were sampled. A statistical test confirmed that while the average absolute value of the target kernel was significantly different from zero at orientations close to the decision boundary (i.e. $|\theta| < 22.5^\circ$, $t_{23} = 15.57$, $p < 0.001$), it was not so at orientations far from the decision boundary (i.e. $|\theta| > 22.5^\circ$, $t_{23} = 1.78$, $p = 0.09$). This pattern is also present in a decision kernel calculated based on the ground truth (whether the true stimulus orientation is CW/CCW) rather than participant responses (Fig. 3.6c, near the boundary: $t_{23} = 17.20$, $p < 0.001$, far from the boundary: $t_{23} = 0.99$, $p = 0.33$).

In Fig. 3.6b, it can be seen that the decision kernel for the distractor was completely flat throughout orientation space. A statistical test on the average of the absolute values of the distractor kernel near the decision boundary ($t_{23} = 0.16$, $p = 0.87$) and far from the boundary ($t_{23} = -0.54$, $p = 0.59$) confirmed that the kernel for the distractor did not deviate from zero. This is consistent with the fact that on its own, the angular offset of the distractor had an inconsistent impact on choices (β_2 in the Eq. (3.6) regression).

To assess the interaction effect of the distractor, we next turn to the SVD results. The top 4 components derived from the target and distractor stimuli on each trial, which collectively accounted for 96% of the variance in filter energy (the first component captured 75% and the second – 16%), are shown in Fig. 3.7a. Of these 4 components, only component 2 had a different profile for the CW and CCW side of orientation space. Components 1, 3 and 4 were symmetric around the decision boundary. To confirm this intuition, we examined the component scores for each of the 4 components. A negative component score means that the variance in the stimulus is best captured by a mirror image of the component, flipped around the horizontal axis. In line with our visual inspection observations, on average,

3. How do perceptual distractors distract?

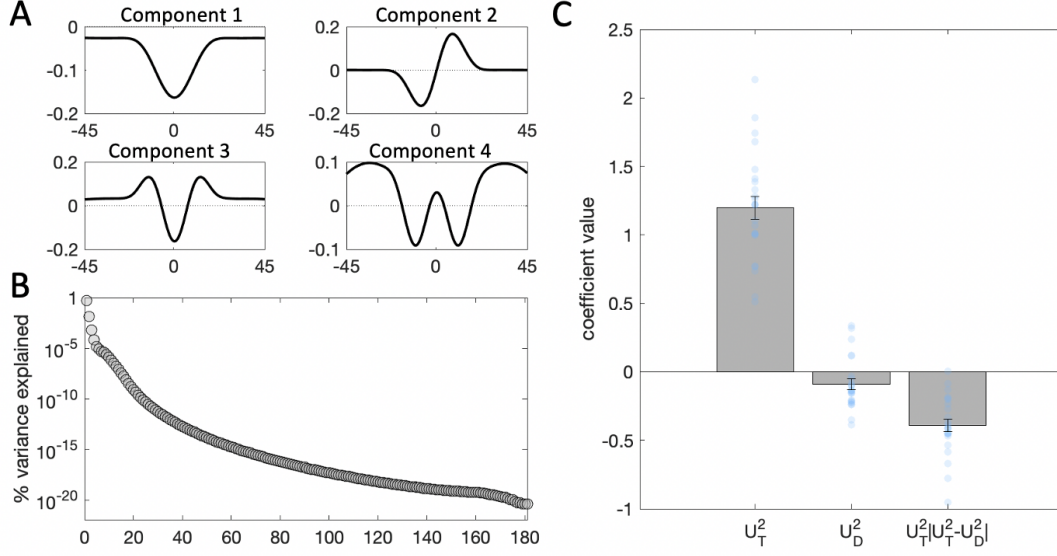


Figure 3.7: Kernel decomposition results. **A:** Illustration of the first four components yielded by singular value decomposition (SVD). **B:** Percentage of variance in the stimulus energy profile explained by each of the SVD components. Note that the y-axis is in logarithmic scale. **C:** Beta coefficients from a binomial regression of participant choices on stimulus scores for the second component and interactions with distractor signals. Grey bars denote averages, error bars denote standard error of the mean, blue dots represent individual participants.

stimulus scores associated with components 1, 3 and 4 did not differ across CW and CCW oriented stimuli. By contrast, the stimulus scores for the second component appear to contain information pertaining to the tilt of the stimulus – they were predominantly positive for CW stimuli, and predominantly negative for CCW stimuli. Thus, we chose to use component 2 for our kernel regression analyses.

We regressed participant choices on the stimulus scores for the second component for the target U_T^2 and distractor U_D^2 on each trial, following the Eq. (3.6) regression model:

$$y = \beta_0 + \beta_1 U_T^2 + \beta_2 U_D^2 + \beta_3 (U_T^2 \cdot |U_T^2 - U_D^2|) \quad (3.7)$$

Thus, the coefficient β_1 reflects the independent effect of the target energy profile (as captured by component 2) on choices, β_2 – the independent effect of the distractor profile, and β_3 – the modulatory (interaction) effect of the distractor on sensitivity to the target, based on the consistency between target and distractor

3. *How do perceptual distractors distract?*

profiles. As expected, the main effect of the target profile on choices was strongly positive ($t_{23} = 14.67$, $p < 0.001$). In this more sensitive analysis, however, distractor profiles now exerted a small direct repulsive effect on choices ($t_{23} = -2.34$, $p = 0.03$). Moreover, as expected from the analyses above, the distractor profile modulated the relationship between target profiles and choices ($t_{23} = -8.80$, $p < 0.001$), such that the more consistent the target and distractor profile, the stronger the relationship between target profile and choice.

3.2.3 Interim Discussion

Taken together, the results from both our classic and reverse correlation analyses suggest that the effect of distraction is primarily interactive, i.e. distractors modulate choices by virtue of the way they interact with targets. There may also be a more idiosyncratic independent effect of the distractor on choice. However, this effect was much weaker than the interactive effect and indeed nonsignificant on the group level in our classical regression analysis. The independent distractor effect identified in the reverse correlation analysis was repulsive, meaning that it pushed choices away from the orientation indicated by the distractor. We take these data to mean that the primary way in which distractors influences choices is interactive rather than independent.

3.3 Experiment 2

We next asked whether the interactive effect of distraction operates at the perceptual level or at the decision level. That is, is the contextual influence of the distractor driven by its sensory features or by its decision information? The interaction model has been proposed to operate across both sensory domains (e.g. in visual cortex, Kapadia et al. 1995) and in more abstract decision processes (e.g. about value-based information, Louie et al. 2013). Our task features a low level visual property (orientation), suggesting that interactions between raw sensory information could be driving the effect. However, sensory level interactions have been documented at

3. How do perceptual distractors distract?

comparably closer locations in retinotopic space (Polat and Sagi 1994) relative to our task here, which positioned target and distractor in opposite visual hemifields. Thus, the effects we observe might instead be driven by interactions at later stages of processing of decision inputs.

To disentangle the perceptual and decision level information in our task, we orthogonalized the decision boundaries for target and distractor stimuli. Thus, in Experiment 2, target and distractor stimuli that share the same tilt relative to their respective boundary would be offset by 90° in terms of their raw tilts. We reasoned that if the contextual influence of the distractor occurred at the perceptual level and thus depended on similarity in terms of raw tilt, then the interaction effect would be null in Experiment 2 where this similarity is all but destroyed. Hence, any interaction effect that survives this manipulation must be due to encoding of information at the decision level (i.e. similarity of tilt relative to the boundary).

We analyzed the data from Experiment 2 in the frame of reference of the decision boundary. That is, we used stimulus tilt relative to the respective boundary rather than raw stimulus orientation in all our analyses. This allowed us to first, avoid wrap around effects in sensory similarity due to the circular nature of orientation space, and second, compare directly the magnitude of the *decision level* interaction effects between Experiments 1 and 2. If we see equivalent consistency effects in both experiments, this would indicate that the influence of the distractor occurs at the decision level.

3.3.1 Methods

Participants

Thirty-two participants (aged 24 ± 4.75) took part in Experiment 2. Eight participants were excluded from the main analyses as the accuracy of their responses for one of the possible stimulus locations was at chance level (defined as in Experiment 1). The study received ethical approval from the Central University Research Ethics Committee at the University of Oxford (approval reference number:

3. How do perceptual distractors distract?

R51752/RE002). All participants provided written informed consent and were compensated £10 per hour for their time.

Apparatus

Participants were seated in a dark room approximately 60cm away from a computer monitor (60Hz refresh rate, 1080x1290 resolution, 23" Dell LCD) with linearized output light intensities. As in Experiment 1, visual stimuli were created and presented with Psychophysics Toolbox Version 3 (PsychToolbox-3, Brainard 1997) for MATLAB.

Stimuli

Experimental stimuli were as in Experiment 1; the only difference was in stimulus orientation. In one of the two stimulus locations, the decision boundary was vertical like in Experiment 1. In the other location, the decision boundary was horizontal. In line with this, Gabor pattern orientation was drawn from a uniform distribution with range -10° to 10° offset from vertical or horizontal for the two respective stimulus locations. The average value of the contrast ratio for the noise in the stimuli produced by the staircase across participants was 0.42. This relatively higher signal contrast in Experiment 2 (versus 0.32 in Experiment 1) indicates that, on average, participants found the task harder when using two different decision boundaries. This is in line with the relatively higher number of participants excluded due to chance-level performance in this experiment (8 versus 2 in Experiment 1).

Experimental Procedure

The experimental procedure followed that of Experiment 1 (Fig. 3.8). The only difference between the two experiments was in decision boundaries. In Experiment 2, for one of the colored rings the decision boundary was vertical, and for the other – horizontal. The decision boundaries were fixed for the duration of the experiment and ring-boundary combinations were counterbalanced between participants. Participants were reminded of the relevant decision boundary by two dots (oriented vertically or horizontally) overlayed over the target ring on each trial.

3. How do perceptual distractors distract?

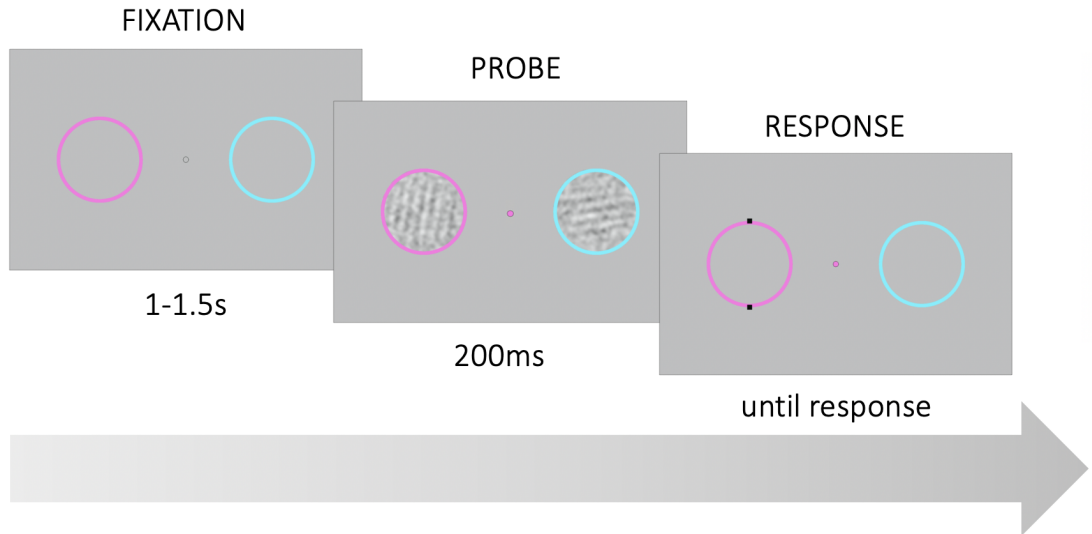


Figure 3.8: Trial structure followed Experiment 1. However, the two colored rings had different reference orientations. In the example here, the decision boundary for the pink ring was vertical, and horizontal for the cyan ring.

Analyses

We followed the same analysis plan as for Experiment 1. As stimulus features were coded as angular offset from the relevant decision boundary, all analyses were conducted in decision space rather than in the raw sensory orientation space.

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/54rf2/>) for this project.

3.3.2 Results

Regression Results

Fig. 3.9a illustrates the results of the stepwise regression on participant choices. Interestingly, while the independent effect of the target and the interaction effect of distraction mirrored the findings from Experiment 1, the independent effect of the distractor explained a sufficient amount of variance to be included in the final model for only 3 out of 24 participants. Thus, the best fitting model for

3. How do perceptual distractors distract?

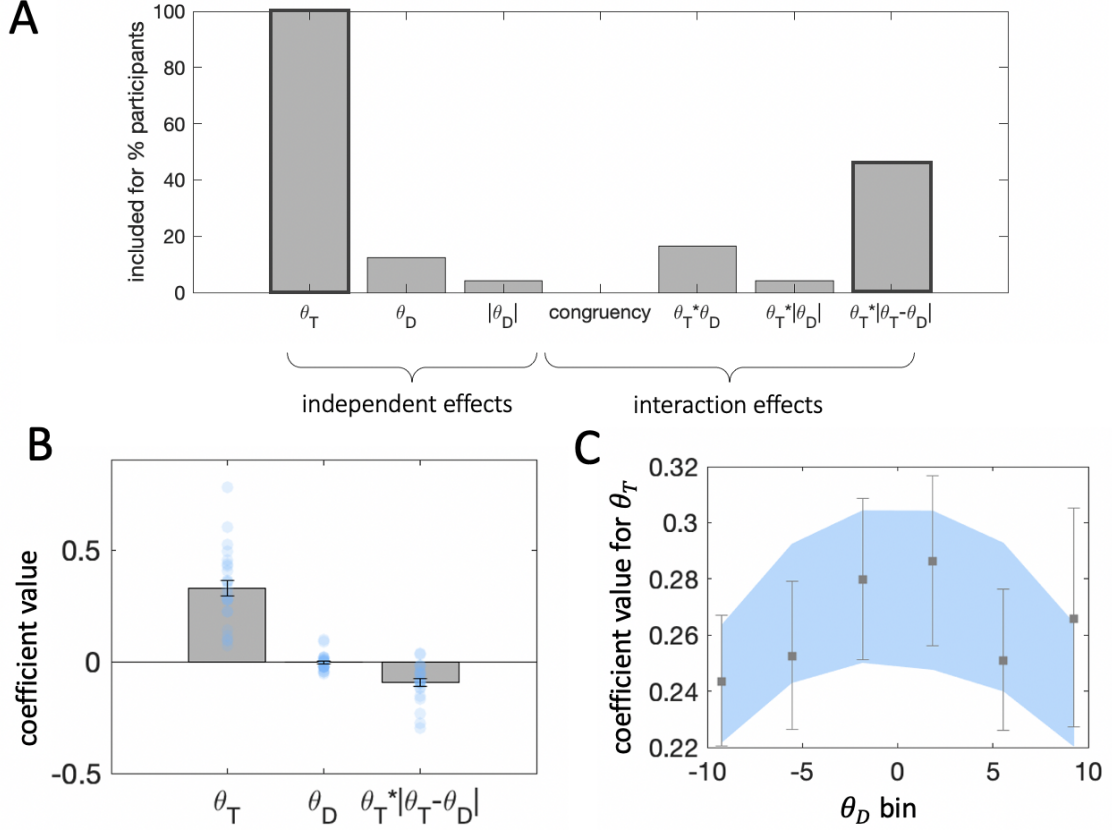


Figure 3.9: Regression results. Plots are as in Experiment 1. **A:** Results from stepwise regressions estimated individually for each participant in the experiment. Darker bar outlines indicate predictors we included in the final model. **B:** Beta coefficients from regression of choices on target and distractor orientation. **C:** Beta coefficients for target orientation from binned regressions.

the choice data in Experiment 2 was:

$$y = \beta_0 + \beta_1 \theta_T + \beta_3 (\theta_T \cdot |\theta_T - \theta_D|) \quad (3.8)$$

As expected, the target had a strong positive influence on choices ($t_{23} = 9.89$, $p < 0.001$) and the coefficient for the interaction effect of distraction was negative ($t_{23} = -5.13$, $p < 0.001$). For consistency with Experiment 1, we also report the results for the Eq. (3.6) model (Fig. 3.9b). The results for the target ($t_{23} = 9.40$, $p < 0.001$) and the interaction effect of distraction ($t_{23} = -5.43$, $p < 0.001$) follow the same pattern. The coefficient for the independent effect of the distractor failed to reach statistical significance ($t_{23} = -0.51$, $p = 0.62$). Thus, in this experiment, the direction of the independent influence of the distractor again varied from individual

3. How do perceptual distractors distract?

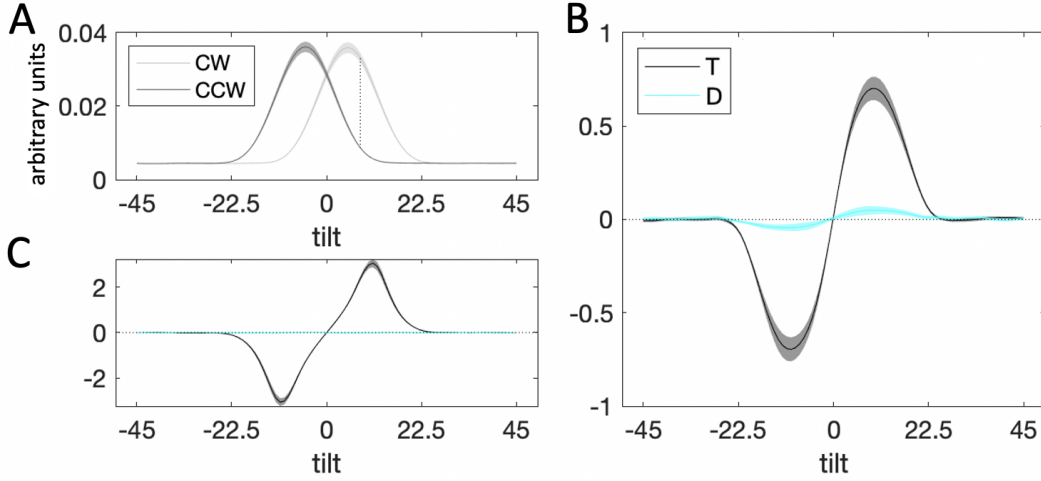


Figure 3.10: Decision kernel results. Plots are as in Experiment 1. **A:** Average stimulus energy profiles. **B:** Decision kernel based on participant choices. **C:** Decision kernel based on ground truth stimulus tilts.

to individual ($\beta_2 = -0.004 \pm 0.037$), however, its inclusion in the model was not warranted by the amount of variance it explained in participant choices.

As can be seen from Fig. 3.9c, the results of the binned regression analysis followed closely those from Experiment 1. Across both participant and regression model-predicted choices, sensitivity to the target was reduced for more extreme distractor orientations, across both clockwise and counterclockwise orientations. This pattern of results is in line with the observed consistency bias.

Reverse Correlation

Fig. 3.10b illustrates the decision kernel for the target and distractor. The target kernel follows the same pattern as in Experiment 1, with a peak on the CW side of feature space and a trough on the CCW side. Again, it tapered off to zero for angles away from the decision boundary (close to the boundary: $t_{23} = 13.06$, $p < 0.0001$, far from the boundary: $t_{23} = 1.27$, $p = 0.22$). Interestingly, the distractor kernel appears slightly peaked in a way that follows the target kernel. Indeed, the average absolute value of the kernel at orientations close to the category boundary was significantly different from zero ($t_{23} = 2.47$, $p = 0.02$), but further from the category boundary it was nonsignificant ($t_{23} = 0.17$, $p = 0.86$). This

3. How do perceptual distractors distract?

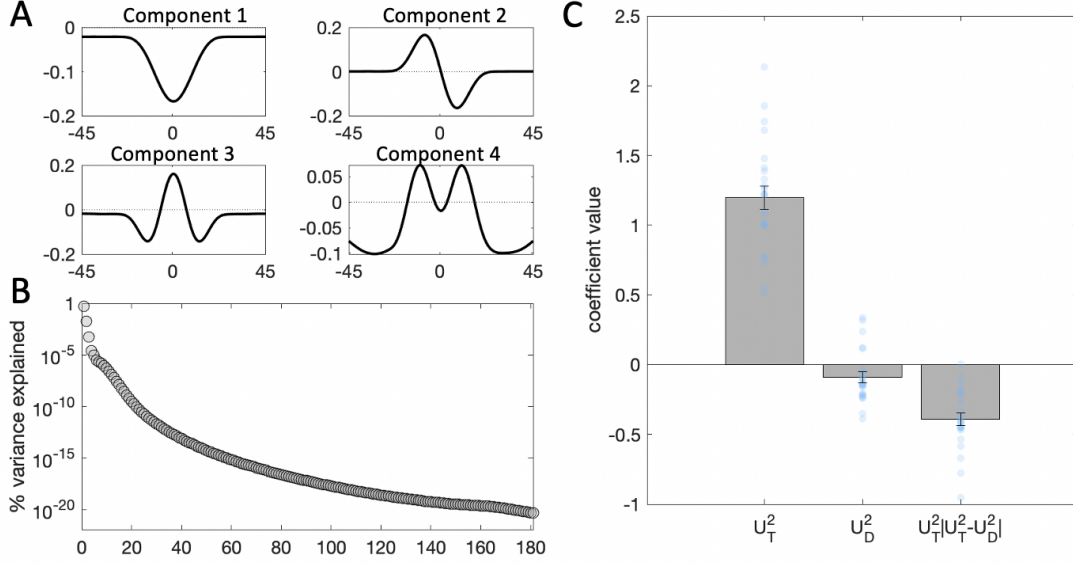


Figure 3.11: Kernel results. Plots are as in Experiment 1. **A:** Illustration of the first four components yielded by SVD. Note that the second component is a mirror image of the component resulting from Experiment 1. For ease of interpretability and consistency between the two experiments, we flipped the sign of the component scores for component 2 in the following analyses. **B:** Percentage of variance in the stimulus energy profile explained by each of the SVD components. **C:** Beta coefficients from a binomial regression of participant choices on stimulus scores for the second component and interactions with distractor signals. Grey bars denote averages, error bars denote standard error of the mean, blue dots represent individual participants.

suggests that participant choices were attracted, albeit rather weakly, towards the signal-like energy in the distractor.

The SVD regression analysis yielded analogous components as in Experiment 1 (Fig. 3.11a-b). This is perhaps unsurprising, as we followed the same stimulus generation procedure. The second component, however, was flipped compared to the second component derived from the stimuli in Experiment 1, suggesting that here, positive component scores would be associated with CCW stimuli, and negative component scores with CW stimuli. Thus, for ease of interpretability and consistency with Experiment 1, we multiplied U^2 by (-1) prior to fitting the Eq. (3.7) regression model (Fig. 3.11c). The results followed those from Experiment 1. We found a positive independent effect of the target ($t_{23} = 10.84$, $p < 0.001$) and a negative interaction effect of target-distractor consistency ($t_{23} = -7.94$, $p < 0.001$). The independent effect of the distractor failed to reach statistical significance in this

3. How do perceptual distractors distract?

analysis, where, unlike in the simple t test above, we controlled for the interaction effect of the distractor ($t_{23} = -1.19$, $p = 0.24$).

3.3.3 Interim Discussion

Overall, the results of Experiment 2 replicate the main findings from Experiment 1. The effect of target-distractor consistency was similar across the two experiments despite the fact that here the decision boundaries for target and distractor were orthogonal (horizontal and vertical). This implies that the relevant features for target and distractor interact at the decision rather than at the perceptual level, i.e. at a stage in which they have been placed in a frame of reference that is relative to the decision boundary. The independent effect of the distractor was more elusive. Unlike Experiment 1, here the independent effect of the distractor was not included in the best fitting model. It also produced nonsignificant coefficients in regression analyses based on distractor tilt and distractor energy. Interestingly, while in Experiment 1 there were suggestions that the distractor may exert a weak repulsive influence on choices, here a small attraction effect was apparent in the distractor kernel.

Perhaps counter-intuitively, this pattern of results is actually in line with the direct repulsive influence of the distractor reported in Experiment 1 if the direct effect of the distractor occurs at the perceptual level. As analyses for Experiment 2 were carried out on the tilt offset of the two stimuli relative to two orthogonal boundaries, inputs that are close in decision space (e.g. tilt offsets 10° and $9^\circ = 1^\circ$ distance in decision space) are actually further in sensory space (respectively 10° and $99^\circ = 89^\circ$ distance in orientation space) compared to inputs that are more distanced in decision space (e.g. tilt offsets 10° and $-9^\circ = 19^\circ$ distance in decision space, respectively 10° and $81^\circ = 71^\circ$ distance in orientation space). Thus, to interpret these results in the same frame of reference as the findings from Experiment 1, we should swap the sign of the effect. The relatively weaker direct effect here could be owing to the fact that target-distractor tilt pairs used in Experiment 2 were much more distant in orientation space compared to those in Experiment 1.

3.4 Experiment 3

Focusing attention on a decision-relevant location promotes target processing at the expense of distractors, and leads to heightened accuracy and faster reaction times (Posner et al. 1980; Hawkins et al. 1990; Carrasco 2011). This naturally raises the question of how the interactive effect of distraction described here changes when targets are, or are not, the focus of spatial attention. In Experiment 3, thus, we manipulated spatial attention by probabilistically cueing participants to focus on either the target or distractor stimulus in order to assess interactions with distraction.

3.4.1 Methods

Participants

Twenty participants (aged 24.3 ± 3.69) took part in the Experiment 3. The study received ethical approval from the Central University Research Ethics Committee at the University of Oxford (approval reference number: R64927/RE001). All participants provided written informed consent and were compensated £10 per hour for their time.

Apparatus

Participants were seated in a dark room approximately 60cm away from a computer monitor (60Hz refresh rate, 1024x768 resolution, 17" LCD) with linearized output light intensities. Visual stimuli were created and presented with Psychophysics Toolbox Version 3 (PsychToolbox-3, Brainard 1997) for MATLAB. Throughout the experiment, participants' eye gaze position was monitored via eye-tracking using SR Research EyeLink 1000.

Stimuli

Experimental stimuli were generated as described in Experiment 1.

3. How do perceptual distractors distract?

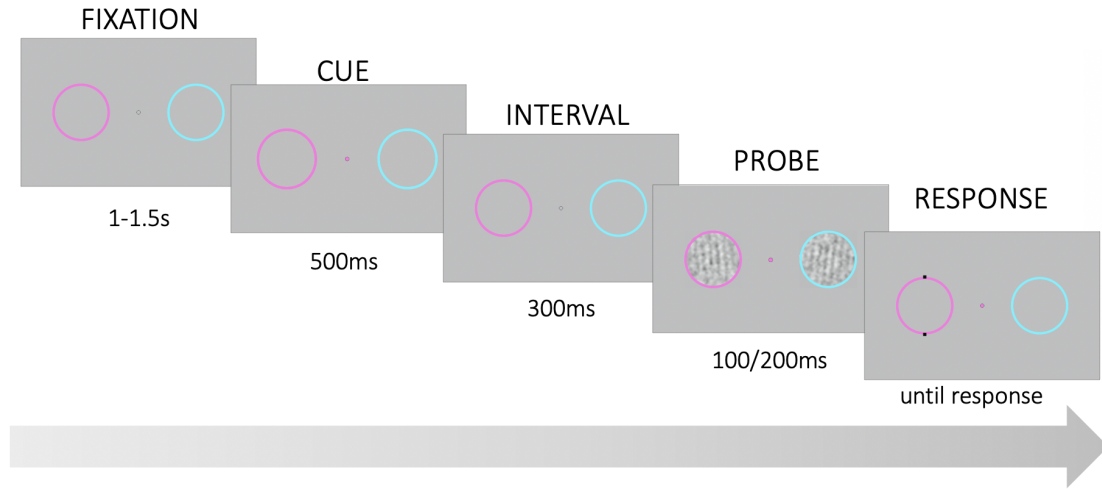


Figure 3.12: Trial structure followed Experiment 1. However, here, 800ms prior to stimulus presentation, the fixation point assumed the color of the ring which was more likely to be probed for 500ms. Note that on each trial stimulus duration was randomly determined as either 100ms or 200ms.

Experimental Procedure

Each participant completed the experiment in a single session, lasting approximately three hours. The experimental session consisted of 2 training mini-blocks of 50 trials each, an adaptive staircase and 9 test blocks of 200 trials each with a short break midway through each test block. The 2 training mini-blocks served to familiarize the participant with the task and the staircase followed the same procedure as in Experiments 1 and 2.

We adjusted the trial sequence in the test blocks to manipulate spatial attention via a spatial cue component (Fig. 3.12). The cue provided information about which stimulus was more likely to be the target. In 2 of the test blocks, the cue was neutral, and thus, trials were in practice equivalent to Experiment 1 (other than some small differences in trial timing). Participants were informed at the start of the neutral blocks that the cue would not be informative and the color of the fixation point (white) reflected this. In the remaining 7 blocks, the cue assumed the color of the ring which was more likely (70%) to be probed.

Each trial began with the probabilistic cue – the fixation point assumed the color of one of the two rings (or white in the neutral condition) for 500ms. 300ms

3. How do perceptual distractors distract?

following cue offset, two noisy gratings appeared within the colored rings for either 100ms or 200ms. Stimulus timings were randomized across trials to maximize the effect of cue validity on choice. At stimulus onset, the fixation point assumed the color of one of the two rings until the end of the trial. The color of the fixation point served as the probe indicating which of the two stimuli was the target. At stimulus offset, two dots were overlaid over the target ring. These dots corresponded to the vertical decision boundary relative to which participants had to report the tilt of the target. Participants reported their responses via keyboard presses (left and right arrow keys) and instantaneously received fully informative auditory feedback (correct: high tone, 880Hz and incorrect: low tone, 440Hz). Responses were not accepted on trials where the participant failed to maintain fixation, defined as gaze location further than 2° in radius from fixation. The experimenter remained in the same room as the participant to re-calibrate the eye tracker in between experimental blocks. Auditory feedback was delivered to the participant via headphones.

Analyses

We adapted the analytic strategy from Experiments 1 and 2, to assess the influence of spatial attention and its interaction with the effect of the distractor.

Manipulation checks First, we conducted manipulation checks to confirm that our spatial attention manipulation was successful. We coded each trial in the experiment into three cueing conditions: 1) valid, where the probabilistic cue correctly signaled the target stimulus, 2) invalid, where the probabilistic cue incorrectly signaled the distractor stimulus, and 3) neutral, where the cue provided no information. We compared accuracy rates (converted to log odds to avoid violating the normality assumption) and response times (log transformed to again avoid violating the normality assumption) across these three cueing conditions via a GLM, controlling for the nonindependence of data points from the same observer. We followed up the results of the GLM with pairwise comparisons via t tests.

3. How do perceptual distractors distract?

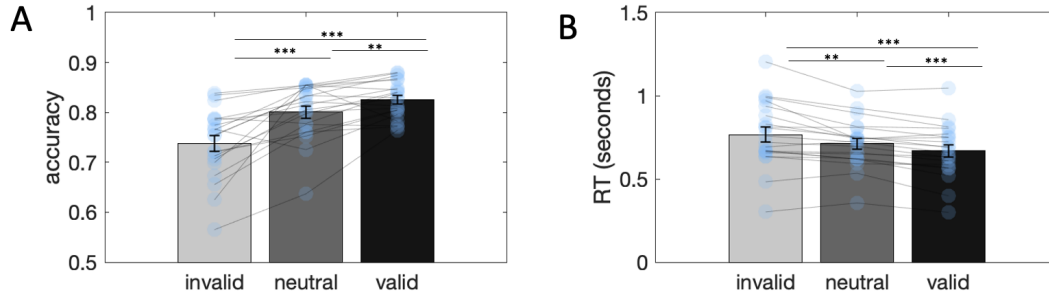


Figure 3.13: Attention manipulation check. **A:** Accuracy across invalid, neutral and valid trials. Statistical tests were performed on odds ratio transformed measures. **B:** Response times across invalid, neutral and valid trials. Statistical tests were performed on log transformed measures. Stars denote statistical significance: ** = $p < 0.01$, *** = $p < 0.001$.

Regression Analyses Collapsing across cueing condition, we followed the analytic pipeline from Experiments 1 and 2. To assess interactions with spatial attention, in addition to the predictors identified via the stepwise approach, we added cueing condition as an interaction term to each of the predictors. Valid, neutral and invalid cueing conditions were coded as 1, 0 and -1 respectively.

Reverse correlation We repeated all reverse correlation analyses from Experiments 1 and 2, collapsing across cueing condition. To assess interactions with attention, we plotted kernels for each cueing condition separately. For the SVD regression analyses, we followed the same strategy as above – we included cueing condition as an interaction term to each of the predictors in the regression model.

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/54rf2/>) for this project.

3.4.2 Results

Manipulation checks

Fig. 3.13 shows accuracy rates and response times across the three cueing conditions. As expected, accuracy was highest for valid trials and lowest for invalid

3. How do perceptual distractors distract?

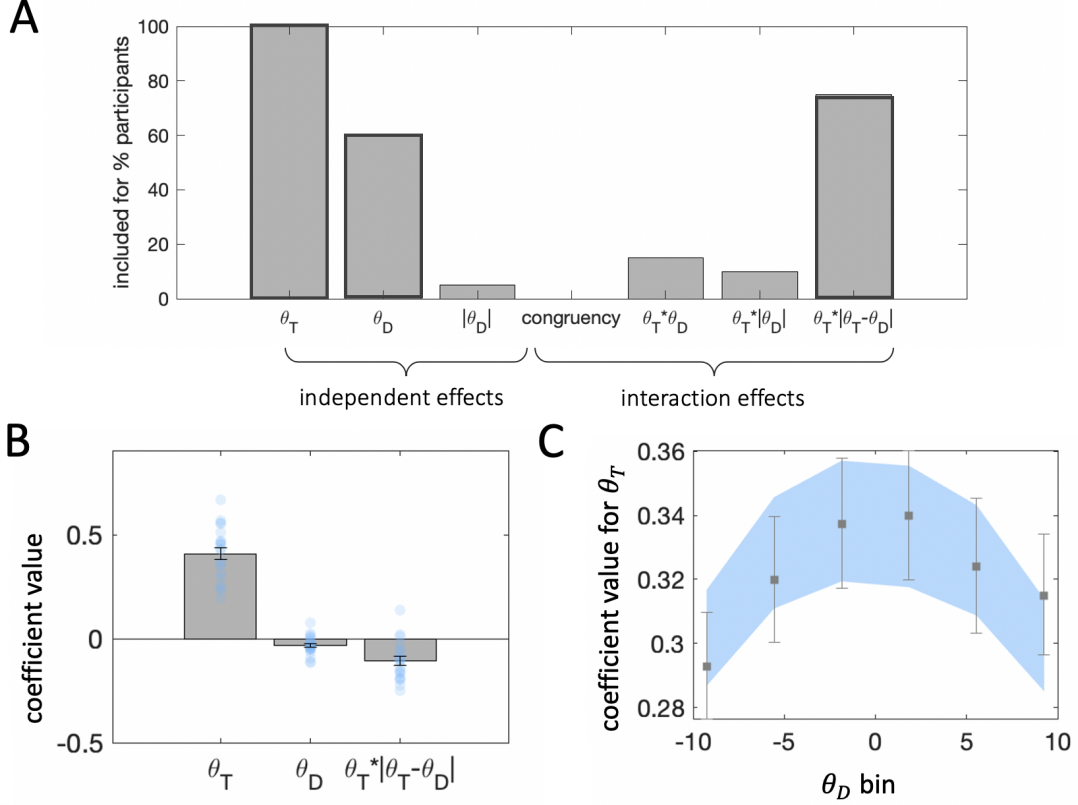


Figure 3.14: Regression results collapsing across cueing condition. Plots are as in Experiment 1. **A:** Results from stepwise regressions estimated individually for each participant in the experiment. Darker bar outlines indicate predictors we included in the final model. **B:** Beta coefficients from regression of choices on target and distractor orientation. **C:** Beta coefficients for target orientation from binned regressions.

trials, while response times followed the opposite pattern, indicating that our attention manipulation was successful. Log-transformed response time showed a valid < neutral < invalid pattern (GLM coefficient $\beta = -0.07$, $p < 0.0001$; pairwise comparisons: valid < invalid: $t_{19} = -8.76$, $p < 0.0001$, valid < neutral: $t_{19} = -4.34$, $p < 0.0001$, neutral < invalid: $t_{19} = -2.61$, $p < 0.02$) and accuracy (transformed into log odds) showed a valid > neutral > invalid pattern (GLM coefficient $\beta = 0.26$, $p < 0.0001$; valid > invalid: $t_{19} = 6.07$, $p < 0.0001$, valid > neutral: $t_{19} = 3.00$, $p < 0.01$, neutral > invalid: $t_{19} = 4.82$, $p < 0.0001$).

3. How do perceptual distractors distract?

Regression Results

First, we attempted to replicate the regression analyses from Experiments 1 and 2 collapsing across cueing condition. As can be seen from Fig. 3.14a, the stepwise regression identified the same predictors as in Experiment 1. Fitting Eq. (3.6) to the choice data from Experiment 3 replicated the strong sway of targets on choice ($t_{19} = 14.17$, $p < 0.001$) and the target-distractor consistency effect ($t_{19} = -5.00$, $p < 0.001$). Here, the distractor had a statistically significant repulsive effect ($t_{19} = -3.17$, $p = 0.005$). The binned regressions produced the same pattern of results as in Experiments 1 and 2. Target sensitivity in participant and model-produced choices was reduced at more extreme distractor orientations, in line with the observed consistency bias.

Next, to assess the impact of target and distractor (and their interaction) under different attentional cueing conditions, we again used logistic regression (Fig. 3.17b), but now including interaction terms with attention (cueing) condition:

$$y = \beta_0 + \beta_1\theta_T + \beta_2\theta_D + \beta_3(\theta_T \cdot |\theta_T - \theta_D|) + \beta_4 A \cdot \theta_T + \beta_5 A \cdot \theta_D + \beta_6 A \cdot (\theta_T \cdot |\theta_T - \theta_D|) \quad (3.9)$$

This expression looks a little complicated but in fact builds naturally from Eq. (3.6). The first three terms (associated with β_1, β_2 and β_3) are identical to those in Eq. (3.6). The final three terms are new predictors that encode the extent to which θ_T , θ_D and their interaction influence choices as a function of cueing validity. In Eq. (3.9) thus A is an indicator variable coding cueing validity.

The results for $\beta_{(1-3)}$ restate those described above (β_1 : $t_{19} = 13.57$, $p < 0.001$, β_2 : $t_{19} = -2.68$, $p = 0.02$, β_3 : $t_{19} = -4.69$, $p < 0.001$). Attention modulated sensitivity to target orientations (β_4 : $t_{19} = 4.80$, $p < 0.001$), as expected from previous work in the attention literature. Surprisingly, however, attention did not impact the direct effect of the distractor (β_5 : $t_{19} = -1.46$, $p = 0.16$), nor its modulatory influence on target sensitivity (β_6 : $t_{19} = 0.71$, $p = 0.49$).

3. How do perceptual distractors distract?

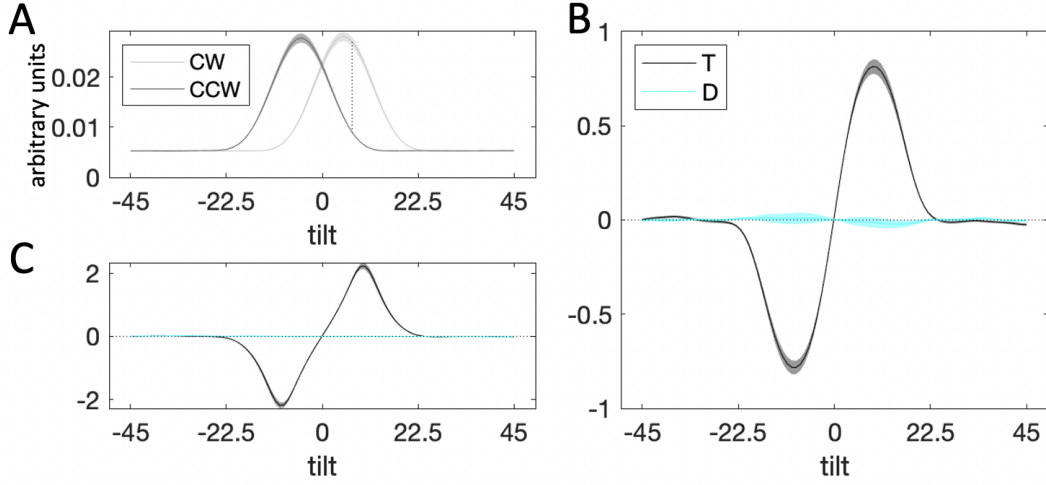


Figure 3.15: Decision kernel results collapsing across cueing condition. Plots are as in Experiment 1. **A:** Average stimulus energy profiles. **B:** Decision kernel based on participant choices. **C:** Decision kernel based on ground truth stimulus tilts.

Reverse Correlation Results

Again, we started by attempting to replicate the reverse correlation results from Experiments 1 and 2 collapsing across cueing condition. The kernel results were consistent with those above (Fig. 3.15). The target kernel significantly drove choices close to the category boundary ($t_{19} = 23.80$, $p < 0.001$) and tapered off to zero for orientations further away from the boundary ($t_{19} = -1.43$, $p = 0.17$). The distractor kernel was flat throughout orientation space (near the category boundary: $t_{19} = -0.55$, $p = 0.59$, far from the category boundary: $t_{19} = 0.09$, $p = 0.92$).

The SVD results (Fig. 3.16) and the regression based on component 2 replicated the findings from Experiment 1. The direct effect of the target profile on choices was strongly positive ($t_{19} = 18.72$, $p < 0.001$) and we again observed the interaction effect of the consistency of target and distractor profiles on choice ($t_{19} = -8.80$, $p < 0.001$). Like in Experiment 1, we also observed a small direct repulsive effect of the distractor on choices ($t_{19} = -3.29$, $p = 0.004$).

Next, we considered these effects under different cueing conditions. Fig. 3.17a shows that there appears to be an attention-induced enhancement of target sensitivity in the decision kernels for the three different attention conditions. To quantitatively assess this, we re-ran the attention-inclusive regression model from

3. How do perceptual distractors distract?

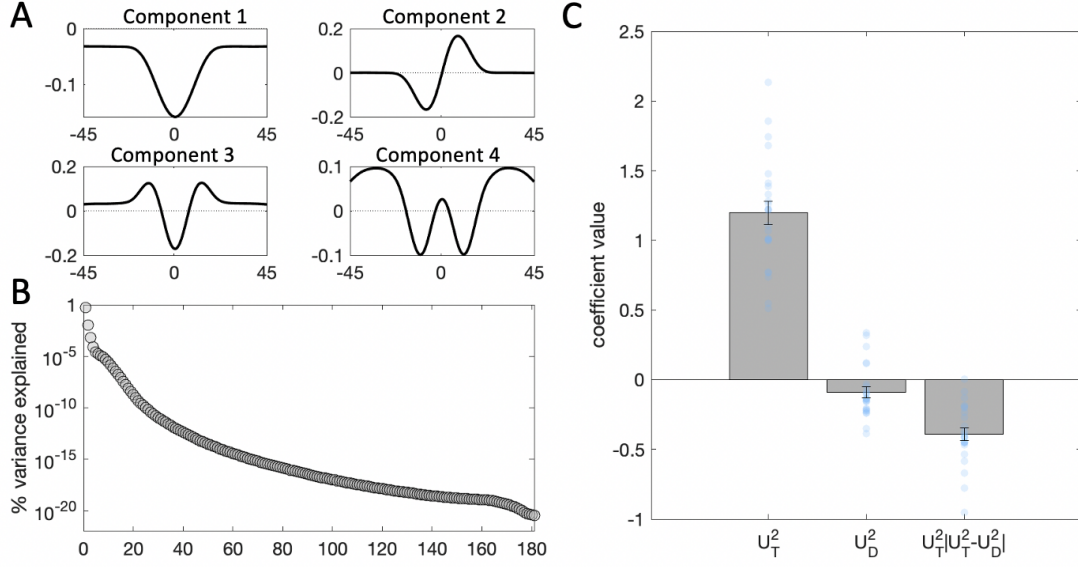


Figure 3.16: Kernel decomposition results collapsing across cueing condition. Plots are as in Experiment 1. **A:** Illustration of the first four components yielded by SVD. **B:** Percentage of variance in the stimulus energy profile explained by each of the SVD components. **C:** Beta coefficients from a binomial regression of participant choices on stimulus scores for the second component and interactions with distractor signals. Grey bars denote averages, error bars denote standard error of the mean, blue dots represent individual participants.

Eq. (3.9) using the scores for component 2 as the predictor variables:

$$y = \beta_0 + \beta_1 U_T^2 + \beta_2 U_D^2 + \beta_3 (U_T^2 \cdot |U_T^2 - U_D^2|) + \beta_4 A \cdot U_T^2 + \beta_5 A \cdot U_D^2 + \beta_6 A \cdot (U_T^2 \cdot |U_T^2 - U_D^2|) \quad (3.10)$$

The results of this model replicated the results from the model based on angular offset θ (Eq. (3.9)). The target profile held a significant sway on decisions (β_1 : $t_{19} = 17.44$, $p < 0.001$), the distractor profile exerted a repulsive influence (β_2 : $t_{19} = -2.73$, $p = 0.01$) as well as a negative interactive effect driven by its consistency with the target profile (β_3 : $t_{19} = -9.63$, $p < 0.001$). Crucially, cueing condition modulated sensitivity to the target (β_4 : $t_{19} = 5.12$, $p < 0.001$), but had no impact on the direct effect of the distractor (β_5 : $t_{19} = -1.66$, $p = 0.14$), nor its modulatory influence on target sensitivity (β_6 : $t_{19} = -1.15$, $p = 0.27$).

3. How do perceptual distractors distract?

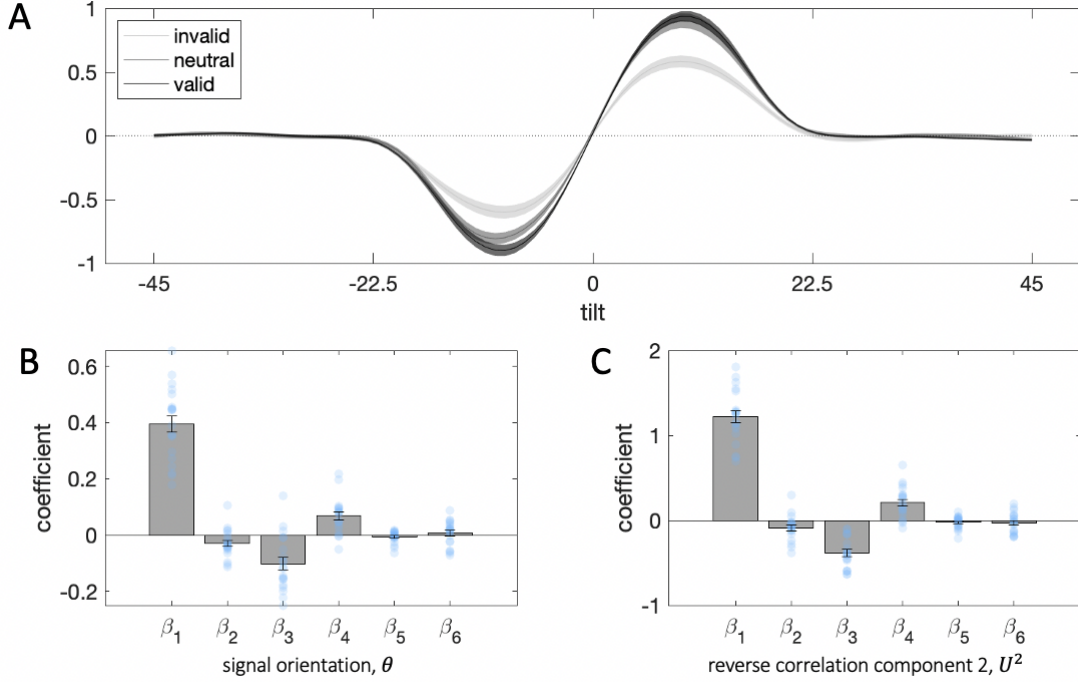


Figure 3.17: Results per cueing condition. **A:** Decision kernels for the target kernel in invalid, neutral and valid trials. Lightest grey denotes invalid, mid-grey – neutral and darkest grey – valid trials. **B:** Beta coefficients from a binomial regression on signal orientation. Grey bars denote averages, error bars denote standard error of the mean, blue circles denote individual participants. β_1 is the coefficient for the target stimulus, β_2 – the distractor, β_3 – the interaction between target and the consistency between target and distractor, β_4 – the interaction between target and cueing condition (coded as -1=invalid, 0=neutral, 1=valid), β_5 – the interaction between the distractor and cueing condition, β_6 – the three-way interaction between target, target-distractor consistency and cueing condition. **C:** As in B, but dependent variables in the regression correspond to the score for SVD component 2 from the reverse correlation analysis.

3.4.3 Interim Discussion

Collapsing across cueing condition, the results of Experiment 3 again replicate the interactive influence of the distractor. Here, we also observe a direct repulsive influence of distractors on choice, which we detected in only some of the analyses in Experiment 1. The observed contextual effects of irrelevant information do not differ across valid, invalid and neutral trials. This implies that directing attention heightens sensitivity to the imperative target stimulus, but, perhaps counter-intuitively, does not impact the contextual effects of irrelevant information.

3.5 Computational Modeling

The regression models above (Eq. (3.6)-(3.10)) were an analytic tool designed to partition variance in choice between independent predictors (θ_T and θ_D) and an interaction term $f(\theta_T|\theta_D)$. However, there is a perhaps more elegant way to model the data, and one that, although closely related, makes more direct links to models of attention and contextual biasing that rely on normalization (Reynolds and Heeger 2009):

$$y = \frac{(x - \rho \cdot \mu)}{r + \tau \cdot \sigma} \quad (3.11)$$

Here, y is the decision variable (as in Eq. (3.3)). A stimulus x is additively/subtractively modulated by the contextual biasing term μ (in the attention literature this is sometimes known as a contrast gain effect, especially where the task involves detection) and multiplicatively/divisively normalized by a context variability term σ (similar to a multiplicative response gain effect). The degree of each type of gain control is respectively modulated by ρ and τ , and r is a regularizer; these 3 parameters are free to vary. Mapping this framework onto our experiment, the target feature $x = \theta_T$, the context mean $\mu = \frac{\theta_T + \theta_D}{2}$ and the context variability $\sigma = |\theta_T - \theta_D|$, i.e. it is given by the range of the two targets. We chose this measure of contextual dispersion over the more commonly used standard deviation, as our experimental manipulation comprised only two stimuli (target and distractor).

This compact formulation of the model allows us to visualize how changes to context affect the (sigmoidal) transducer function linking decision inputs to choice (see Eq. (3.3)). Changes in contextual variability adjust the slope of the transducer – the more consistent the target and distractor, the steeper the slope (Fig. 3.18a), and consequently, the higher the gain of processing (Fig. 3.18c). The parameter τ controls the degree of this change. As Fig. 3.18c shows, as τ grows, the interaction effect of the distractor (target-distractor consistency) increases. By contrast, the contextual expectation shifts the transducer function along the abscissa (Fig. 3.18b). This impact translates into an independent effect of the distractor – for negative

3. How do perceptual distractors distract?

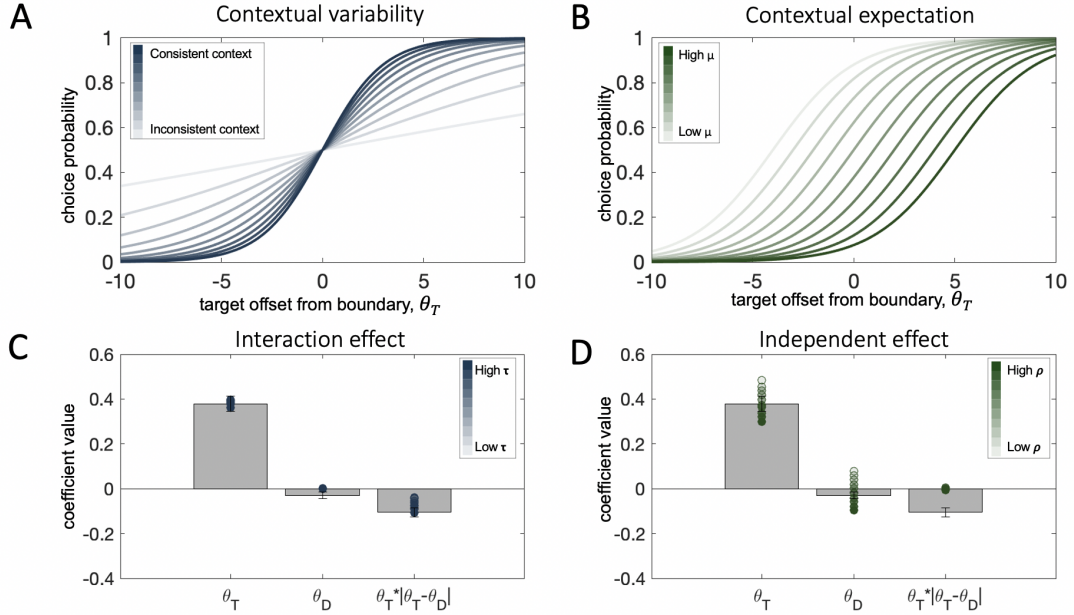


Figure 3.18: Model simulations of the effects of context on the transducer function translating decision inputs into choice probabilities (top) and on the distractor effects in choice behavior (bottom). **A:** The consistency of context changes the slope of the transducer. Lower contextual variability leads to steeper slope (darker blue curves) and vice versa for high variability (lighter blue curves). **B:** The contextual expectation changes the location of the inflection point of the transducer. A high contextual average leads to a transducer shifted rightwards along the x-axis (darker green curves) and vice versa for low contextual average. **C:** Contextual variability leads to an interaction effect of the distractor (target-distractor consistency). Bars correspond to coefficients from a regression on human choices from Experiment 1, filled circles correspond to an analogous regression on model simulations where we varied parametrically the value of τ , while keeping ρ and r constant. The higher the value of the multiplicative free parameter τ , the stronger the consistency effect (darker circles). **D:** The contextual expectation leads to an independent effect of the distractor. Bars correspond to coefficients from a regression on human choices from Experiment 1, filled circles correspond to an analogous regression on model simulations where we varied parametrically the value of ρ , while keeping τ and r constant. The higher the value of the multiplicative free parameter ρ , the stronger the repulsion from the distractor (darker circles). Note that manipulations to ρ do not lead to an interaction effect (overlapping circles on last bar).

3. How do perceptual distractors distract?

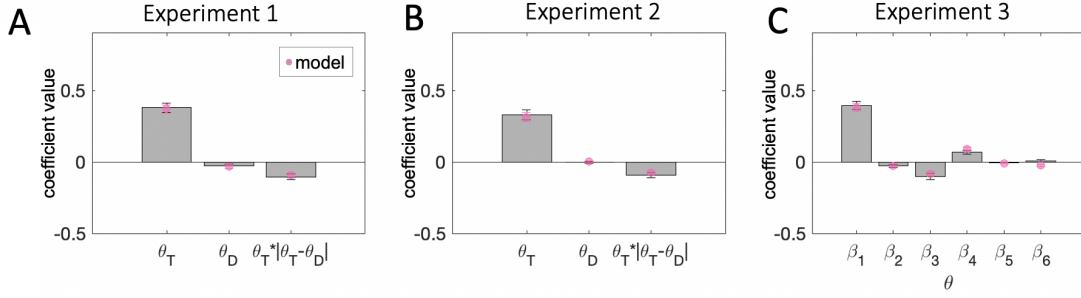


Figure 3.19: Model fits to human data. **A:** Experiment 1. **B:** Experiment 2. **C:** Experiment 3. Across all three experiments bars correspond to coefficients from a regression on human choices and pink circles correspond to model-simulated choices with free parameters estimated to best fit human data. Error bars correspond to standard errors of the mean. Coefficient labels in Experiment 3 are as in the regression analysis plot.

values of ρ we observe an attraction towards the distractor, and for positive values of ρ we observe a repulsion (Fig. 3.18d). This rich pattern of distractor influence maps onto the variability of the independent effect observed in human participants across the three experiments (the coefficient estimates for β_2 in Eq. (3.6)).

We fit this 3-parameter normalization model (r , τ , and ρ) to human data from the first two experiments through gradient descent at the single trial level. Best fitting parameters were estimated for each participant individually by minimizing negative log likelihood. The model captures the pattern observed in the human data remarkably well (Fig. 3.19a-b). It recreates both the independent and interaction effect of the distractor. It is also quantitatively favored over a static model (where $\tau = \rho = 0$) by Bayesian model selection on cross-validated model likelihoods (exceedance probabilities: Exp.1: $p = 0.99$, Exp.2: $p = 0.99$).

To account for the observed effect of attention in Experiment 3, we allowed the parameter r , which captures the baseline slope of the transducer above and beyond the impacts of contextual consistency, to vary freely across cueing conditions. Thus, in our process model, the spatial attention manipulation does not affect the strength of contextual normalization indexed by parameters τ and ρ . Rather, it changes the overall sensitivity to the target, as evidenced by the estimated parameter values: the slope of the transducer was steepest (i.e. lowest value for parameter r) in valid

3. *How do perceptual distractors distract?*

trials and shallowest (i.e. highest value for parameter r) in invalid trials (valid < invalid: $t_{19} = 3.94$, $p < 0.001$, valid < neutral: $t_{19} = 3.79$, $p = 0.001$, neutral < invalid: $t_{19} = 2.46$, $p = 0.02$). Indeed, incorporating attention in this manner in the normalization model reproduced the observed changes in sensitivity to target orientation (Fig. 3.19c - β_4). Moreover, it captured the attention-independent effects of the distractor (β_2 and β_3), as well as the lack of interaction of those with cueing condition (β_5 and β_6).

3.6 Discussion

Across all three experiments, we found consistent evidence that distractors mainly influence choices by modulating the influence that targets have on choices. Our experiment involved a classic perceptual task – discrimination of noisy gratings – and our stimuli were spatially separated across the two hemifields, minimizing retinotopic visual interactions in the sensory encoding of target and distractor. Yet, our modelling suggests that targets have a stronger effect on choices when they are similar to distractors. This result was consistent using both conventional regression-based modeling and an approach based on reverse correlation analysis across all three experiments. This finding does not necessarily contradict, but does potentially nuance, previous accounts of distraction. For example, a standard model assumes that targets and distractors compete independently to influence choice, with the job of control processes being to upweigh targets at the expense of distractors.

In a different study, which is notable for applying a very rigorous modelling approach to a large body of data gathered from tasks that broadly resemble our own, the authors (Shen and Ma 2019) characterize the effect of distractor as being a generalized decrement to precision, or an increase in overall lapses of judgment (i.e. guess rate). However, their analysis does not attempt to ask how the features of target and distractor impact choices, but rather compares experimental conditions that have different levels of distraction, which might have led to their overlooking any interaction between target and distractor that is present in that data.

3. How do perceptual distractors distract?

Relatedly, this finding is not consistent with a familiar framework for understanding decisions based on multiple, potentially noisy, sources of information with variable decision-relevance – that which is formalized by Bayesian inference (Dayan and Zemel 1999; Eckstein et al. 2009). In our experiment, a Bayesian ideal observer would weigh the target and distractor according to how relevant they were judged to be for the decision. How this relevance was determined would have to depend on some assumptions about the likely sources of error in the task, but in general, this class of explanation would point to decisions being determined exclusively by a weighted mixture of independent target and distractor features. This is not what we observed.

The form of the interactive effect we did observe seems related to a phenomenon has been previously described as a consistency bias (Cheadle et al. 2014). A consistency bias occurs when stimuli that resemble the local context are processed with heightened gain, and so have greater impact on choices. Previous work has reported a bias for decisions to be consistent with the recent temporal context, which is captured by a normalization model analogous to the one we discuss here (Cheadle et al. 2014). A closely related paper showed that a range of phenomena relating to spatial context – from across the subfields of perceptual, cognitive and economic decision-making – can be explained with a gain control model that also resembles the one described here (Li et al. 2018). Thus, seemingly unrelated phenomena such as the tilt illusion (Blakemore et al. 1970), congruency effects in a Flanker task (Eriksen and Eriksen 1974) and the distracting effect of decoy food items (Louie et al. 2013) can all be captured by a closely related interactive model in which items that are similar to the context are processed with higher gain.

We observed an interaction between target and distractor decision values (tilt relative to the boundary) rather than perceptual values (raw tilt). This implies that the distractor, as well as the target, is placed in a frame of reference that expresses tilt relative to the category boundary. Consistent with this finding, a previous study using EEG found evidence from neural encoding that both target and distracting signals are placed in a decision-based frame of reference (Wyart et al. 2015), although of course it is likely that the extent to which selection occurs early

3. How do perceptual distractors distract?

or late depends on both the nature of the task and the available resources (Lavie and Tsal 1994). The data described here suggest that the interaction between target and distractor occurs at this later stage, beyond immediate sensory perception. This perhaps surprising result is in line with previous reports on the consistency bias (Cheadle et al. 2014) which also found that this effect occurred at a later stage of processing by dissociating sensory and decision inputs (via asking participants to make a decision based on the cardinality or diagonality of grating orientations). The finding that the contextual influence of perceptual distractors operates on decision level signals could perhaps also be anticipated from the many reports of contextual effects for stimuli that share an abstract property (such as value) but not a low level property (such as tilt), for example, in the literature on decoy effects discussed in the previous chapter.

We found limited evidence that the distractor impacted choices directly. The independent effect of the distractor appeared to largely vary idiosyncratically between participants both in terms of strength and direction. On the group level, this independent effect seemed to be reliable in Experiment 3, but was weaker in Experiment 1. Of note, these are the tasks for which the category boundary is the same for target and distractor, so it seems likely that any such effect (unlike the contextual consistency bias) operates at a perceptual, not a decision level. It is notable that the direction of the effect is repulsive, meaning that the more clockwise the distractor is, the less likely participants are to respond “clockwise”. The observed repulsion of choices from the sensory features of the distractor across Experiments 1 and 3 is also consistent with the weak independent effect we found in Experiment 2 (see section 3.3.3).

Indeed, contextual effects in sensory processing, such as adaptation over time (Webster 2015) and the tilt illusion (Blakemore et al. 1970), typically follow a repulsive pattern. The tilt illusion features grating stimuli similar to the ones we use here, however, the distracting flankers are closer to the target compared to in our set up. The higher degree of spatial separation can account for the weak and inconsistent nature of the sensory level effects we observe. Similar spatial

3. How do perceptual distractors distract?

interactions are documented in contrast perception (Chubb et al. 1989; Xing and Heeger 2001), where a central texture appears of lower contrast when surrounded by a high context texture compared to when surrounded by a uniform background. These psychophysical effects have previously been modeled in a normalization framework (Carandini and Heeger 2012). Here, we also demonstrate that the observed repulsive effect of the distractor can be captured by a model in which decision inputs are normalized by the contextual expectation. The direction of this effect cannot, however, be explained by a model in which target and distractor compete for attention and/or processing resources.

Finally, our findings also speak to the large literature which has explored the effect of attention on decisions made about a target stimulus in the face of distraction. In Experiment 3, we included a cueing manipulation that allowed us to test how attention might be deployed to mitigate the effect of distraction. Although we saw that spatial attention enhanced target processing – as demonstrated by both conventional and reverse correlation approaches – it did not seem to have any impact on the consistency bias. In other words, both target and distractor were elaborated into the frame of reference of the decision (Experiment 2) and then apparently interacted in a way that drove choices, but at no point (as far as we can see) did spatial attention intervene to facilitate or inhibit this process (Experiment 3). This finding seems hard to reconcile with physiological work with non-human primates, which shows that when spatial attention is directed towards a target stimulus, contextual modulation of neural responses by irrelevant stimuli in extra-striate area V4 is reduced by approximately 50% (Sundberg et al. 2009). This discrepancy might be related to the fact that the consistency bias appears to operate on the decision level rather than on the raw sensory signals likely captured by electrophysiological work in visual areas. In our data, spatial attention appeared to multiplicatively increase the gain with which a target stimulus is encoded (Carandini and Heeger 2012), but it did not modulate the independent or contextual effect of the distractor.

This pattern of results parallels findings on feature-based attention, which is known to operate broadly and independently of spatial attention (Liu et al. 2007),

3. How do perceptual distractors distract?

and which has been associated with a feature-similarity gain control process not dissimilar from the consistency bias reported here. For example, when attention is cued towards a visual feature – such as downwards motion – in one part of the visual field, neurons sensitive to downwards visual motion (but with receptive fields at different locations) show heightened responses (Treue and Trujillo 1999). In attention studies, feature gain similarity also affects performance, albeit as potentially nonmonotonic function of target-distractor feature disparity (Störmer and Alvarez 2014; Liu 2019). However, in our study there was no overt cueing of features. One possibility is that there is an automatic feature-similarity gain process that depends on the context provided by the distractor. Future work can shed light on the relationship between the phenomena reported here, and the established literature on feature-based attention.

3.7 Conclusion

Perceptual decisions often occur in the presence of irrelevant information. Studies on attention have contributed greatly to our understanding of the behavioral signatures and neural underpinnings of focusing attention onto a target stimulus and away from irrelevant distractors. However, we know remarkably little about the influence distractors wield on choice. This chapter examined the functional relationship between the features of a distractor stimulus and perceptual decisions. Across three experiments, we found consistent evidence that distractors affect perceptual decisions primarily through an interactive effect, whereby consistency between target and distractor features boosts sensitivity to the target. This effect appears to operate on the level of the decision, rather than on raw sensory inputs, and is independent of attention. A simple normalization framework, where the dispersion of features present in the local context determines the gain of processing of decision-relevant information, can capture the behavioral signatures of distraction in our results, across all three experiments.

3.8 Acknowledgements

This chapter is submitted for publication and available as a preprint *Dumbalska, T., Rudzka, K., Smithson, H. E., & Summerfield, C. (2021). How do (perceptual) distractors distract?*.

Experiment 3 formed part of Katarzyna Rudzka’s final year undergraduate research project. Data collection for this experiment was carried out in its entirety by Katarzyna Rudzka. I am grateful for constructive feedback on an early draft of this chapter from my transfer of status assessors, Prof Jill O’Reilly and Dr MaryAnn Noonan. I also thank Prof Marisa Carrasco, members of her research group, and the participants in the Symposium on Biology of Decision Making (2019) for comments on presentations of this work.

4

Context-dependent categorization

Contents

4.1	Introduction	105
4.2	Experiment 1	110
4.2.1	Methods	111
4.2.2	Results	117
4.2.3	Interim Discussion	121
4.3	Experiment 2	122
4.3.1	Methods	123
4.3.2	Results	130
4.3.3	Interim Discussion	133
4.4	Discussion	134
4.5	Conclusion	138

Categorization decisions are repelled away from the characteristics of temporal context. This pattern of results is consistent across the realm of perception, cognition and economics; yet the underlying computational mechanisms that produce it remain elusive. This chapter arbitrates between putative sources of contextual bias in categorization decisions based on the lightness, size or numerosity of visual stimuli, across two experiments. Our experimental design allowed us to rule out response bias as a driver of the effect. We considered two alternative explanations: contextual shifts in the category boundary (contraction bias) and contextual changes in our encoding and experience of the current stimulus (adaptation). We used changes in

4. Context-dependent categorization

suprathreshold sensitivity to estimate the contribution of adaptation. Our results suggest that adaptive changes to the appearance of the to-be-categorized stimulus can account for up to half of the contextual bias in lightness categorization, and for ~10% and ~30% of the contextual bias in categorizations of stimulus size and numerosity.

4.1 Introduction

Sensory information is continuous, yet the choices we make based on it are necessarily discrete. In daily life, we seamlessly sort sensations into meaningful categories to guide our actions and behaviors. We do this by drawing boundaries through sensory space that divide the continuum of our experiences into distinct classifications. The location of those boundaries, however, is not static. It shifts over the course of our lifetime, starting with early cognitive development (Gopnik and Meltzoff 1987; Slawinski and Fitzgerald 1998), through our lived experiences (Pollak and Kistler 2002) and domain-specific expertise (Chi et al. 1981; Shafto and Coley 2003). As we learn about the world, or a particular corner or field within it, our standards about what constitutes a particular category evolve, perhaps becoming more fine-tuned and consistent (Slawinski and Fitzgerald 1998; Shafto and Coley 2003).

Interestingly, category boundaries may also shift on a much finer time-scale, depending on our most recent experiences. In the domain of visual perception, for instance, the same square stimulus can be categorized as small or large depending on the size of the squares that precede it (Fig. 4.1a). If it happens to occur after several large squares, it would, on average, be more often categorized as small than if it happens to occur after several small squares. In the laboratory, this phenomenon has been examined by asking participants to categorize (or label) series of visual stimuli, presented one after the other. The experimenter would manipulate the distribution of stimulus characteristics: in some blocks, the participant may be presented with predominantly large stimuli; in others, they may be presented with predominantly small stimuli. The location of the category boundary in each experimental condition can then be quantified based on participant

4. Context-dependent categorization

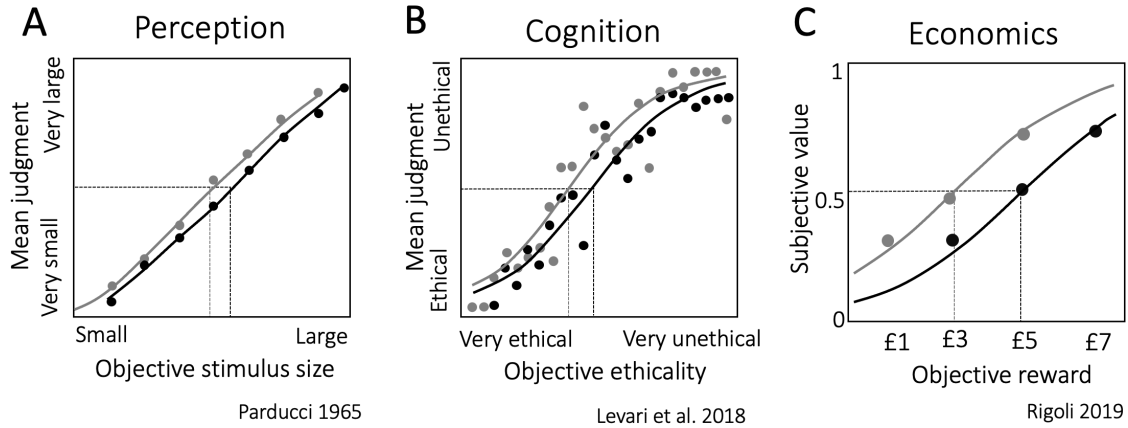


Figure 4.1: Context-dependent categorization across domains. Across all panels, the x-axis indexes the objective stimulus property and the y-axis – categorization decisions. **A:** Perception. Participants were asked to rate square stimuli on a scale between very large and very small. The psychometric curves depict participants’ categorization decisions across two contexts. The grey curve tracks judgments in a context dominated by objectively smaller stimuli, the black curve tracks judgments in a context dominated by objectively larger stimuli. Data from Parducci 1965, Fig. 7. **B:** Cognition. Participants were asked to accept or reject proposals for scientific studies that varied on a continuum of very ethical to very unethical (according to an independent group of raters). The shading of the curve again tracks context (grey: low prevalence of unethical studies, black: high prevalence of unethical studies). Data from Levari et al. 2018, Fig. 3b. **C:** Economics. The psychometric functions here depict participants’ subjective values estimated from choice behavior. Shading again tracks context (grey: low value context, black: high value context). Data from Rigoli 2019, Fig. 3a.

responses by estimating what stimulus size would lead to an equal proportion of “large” and “small” categorizations (known as the threshold or point of subjective equivalence/indifference point). The results of these experiments indicate the dynamic malleability of our categorization decisions. The location of the category boundary shifts predictably, tracking the central tendency of the distribution of stimulus characteristics present in the temporal context.

This pattern of results is robust. It generalizes across tasks that feature different types of visual stimuli (e.g. spatial frequency, Lages and Treisman 1998; distance, Morgan et al. 2000; hue, Olkkonen et al. 2014; facial expression, Levari et al. 2018) or engage other sensory modalities (e.g. weight perception, Helson 1947; interval timing, Jazayeri and Shadlen 2010). It holds regardless of whether participants are asked to use their “internal concept” of the category (Levari et al. 2018), whether

4. Context-dependent categorization

they are provided with an explicit or implicit standard for categorization, or with feedback on the accuracy of their judgments (Morgan et al. 2000). It also extends beyond perception, when we need to abstract away from sensory information to evaluate, for instance, scenarios based on their ethicality (Fig. 4.1b, Levari et al. 2018) or economic prospects based on their subjective value (Fig. 4.1c, Rigoli 2019). Temporal context sways judgments across perception, cognition and economics in a predictable fashion. Category boundaries track the frequency and dispersion of recently experienced stimulus properties (Parducci 1965; Rigoli 2019), and so categorizations are repulsed away from recent experience. The smaller the squares we just saw, the more likely we would be to categorize a medium square as large. The more unethical the scenarios we just heard, the more likely we would be to judge an ambivalent scenario as ethical. The smaller the financial rewards we just experienced, the more likely we would be to perceive a moderate sum as valuable.

What drives this pattern of context dependence? One possibility is that it is driven by a response bias, whereby we calibrate our categorization judgments, such that the relative frequency of our actions remain stable across contexts. For instance, if we were faced with the choice of whether to accept or reject a proposal based on its ethicality, we might accept about half the proposals we encounter. But if the ethicality of proposals is particularly low in a given context (as it would be in the above described experimental manipulations) and we are compelled to accept half, this would naturally lead us to exhibit a lower threshold for (acceptable) ethicality. This is particularly likely in the laboratory, where response biases abound (Poulton and Poulton 1989; Bonnet 1990), as participants bring their priors (how likely is this category?), expectations of experiment demands (how would the experimenter like me to categorize this?, Rosenthal 1976), and misperceptions of randomness (can chance produce this sequence of categories?, Bar-Hillel and Wagenaar 1991). All of these factors may give rise to a response bias, producing the observed context-dependent behavior. It is difficult to rule out this possibility based on the empirical data, as the ground truth response frequency is necessarily confounded with the

4. Context-dependent categorization

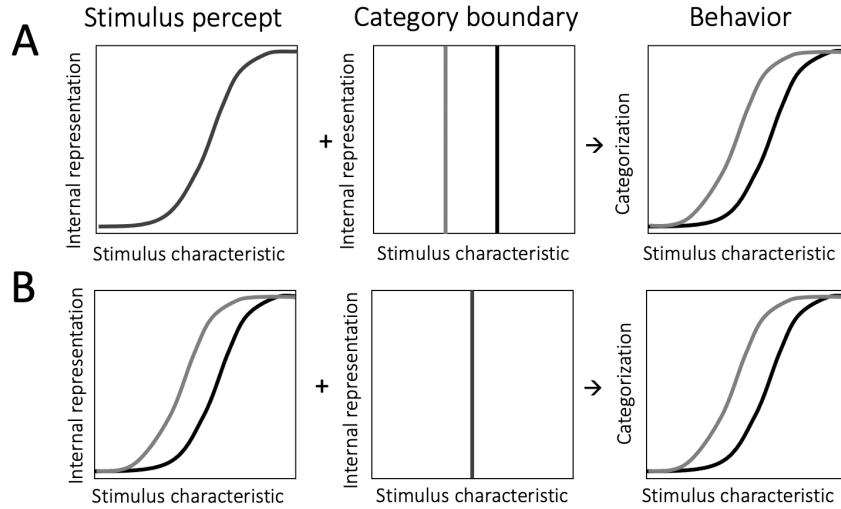


Figure 4.2: Possible sources of context dependence. **A:** Participants’ internal representation (left panel) of the stimulus is stable across contexts. Their representation of the category boundary (middle panel), however, changes such that it is lower in a context dominated by low stimulus property values (grey line) and higher in a context dominated by high stimulus property values (black line). This may be driven by a Bayesian process incorporating the prior distribution of stimulus characteristics in a noisy representation of the category boundary. It can lead to the observed context dependence in categorization decisions (right panel). **B:** Participants’ internal representation of stimulus property space shifts across contexts (left panel). This change may be driven by neural adaptation to contextual levels of stimulation. Participants’ representation of the category boundary, however, remains stable (middle panel). This can also lead to the observed context dependence in categorization decisions (right panel). Note that both of these two processes can occur at the same time and influence categorizations together.

experimental manipulation of context. Response bias can, however, be minimized by careful experimental design; we took this route in the present study.

An alternative account, popular in the literature, argues that the internal representation of the standards we use to draw category boundaries is noisy. To make categorization judgments, we draw on a working memory reference for those standards. The precision of the reference, however, decays in memory (Jou et al. 2004; Ashourian and Loewenstein 2011; Olkkonen and Allred 2014). To resolve the uncertainty about its precise location in sensory (or conceptual) space, the neural system uses information available in the temporal context. It draws on the distribution of stimuli that have been recently experienced to form a prior; this prior then informs the standard. As a result, the category boundary differs across

4. Context-dependent categorization

contexts – it is biased towards the central tendency of context (contraction bias), and so, behaviorally, our categorizations exhibit the observed context dependence (schematic illustration in Fig. 4.2a). A Bayesian account of this process, whereby the prior weighs more heavily in the presence of noise, is consistent with data showing that increasing the noisiness of the representation increases the contextual bias in categorizations. This can be done by, for instance, introducing cognitive demands (Ashourian and Loewenstein 2011; Allred et al. 2016) or manipulating delays between the reference and the to-be-categorized stimulus or the presentation quality of the reference (Olkkonen et al. 2014).

While this proposal puts the emphasis on our internal representation of the category boundary, context may also affect how we experience the current stimulus. Neural systems adapt to the levels of stimulation present in the environment. Visual adaptation to light, for instance, changes our sensitivity to stimulus brightness. After some minutes in darkness, we might be able to detect a faint light that would be indistinguishable to us in a bright room. These adaptive processes help calibrate the range of possible neural responses to the distribution of behaviorally relevant characteristics of the environment (Webster 2015) through normalization of each neural response by the pooled neural responses to context (Carandini and Heeger 2012). Adaptation can lead to changes in the appearance of stimuli, such that the neutral point (i.e. the stimulus that would be considered neutral) is shifted towards the central tendency of context. For instance, adapting to a face exhibiting an angry facial expression, would alter how we perceive subsequent faces: objectively neutral faces will cease to feel neutral and objectively angry facial expressions will appear less angry and more neutral (Webster 2011). This process may be illustrated as a recentering of the internal representation of sensory space towards the characteristics that dominate the temporal context (Fig. 4.2b). This recentering would lead to the same stimulus being encoded and experienced differently depending on the temporal context in which it is encountered. Thus, even if the internal representation of the category boundary remains stable, adaptation can capture the observed context-dependent categorization behavior.

4. Context-dependent categorization

In addition to changes to the category boundary, normalization-based adaptive processes also bring about changes in sensitivity. The sigmoid shape of the transducer function, mapping sensory space onto an internal representation, is ubiquitous in neural systems; sensitivity is highest near the inflection point of the transducer. If the transducer shifts between contexts, tracking the distribution of features, then sensitivity would necessarily change too – it would be highest for those stimulus characteristics that dominate the temporal context. Temporal context in adaptation, however, does not only encompass preceding task-relevant information. While the prior in a Bayesian process (as the one which might affect the location of the category boundary) is shaped by previously experienced instances of task-relevant stimuli, contextual influences on adaptation are less specific. Light adaptation, for instance, is driven by the overall light levels in the context, regardless of what light source is producing them (e.g. light from a lamp + moonlight = overall level of light). By contrast, our prior for the brightness of moonlight would be driven by moon brightness on previous nights, and not by our previous experiences with lamplight.

In this chapter, we set out to examine the underlying mechanisms driving context-dependent categorizations across two experiments. First, we disentangled response frequencies from the frequencies of stimulus characteristics in our contextual manipulation to rule out response bias as a driver of context-dependent categorization (Experiment 1). Second, we used the specificity of contextual influences on categorizations (Experiment 1) and changes in sensitivity (Experiment 2), to probe whether and how much adaptation contributes to the observed context dependencies across categorizations based on sensory (Experiments 1 & 2) and more abstract (Experiment 2) stimulus features.

4.2 Experiment 1

Experiment 1 measured the impact of temporal context on categorizations, controlling for biases in response frequencies introduced by the experimental manipulation. We did this by asking participants to complete one of two possible tasks on each

4. Context-dependent categorization

trial. On some trials, participants had to categorize the lightness of a circle stimulus; on other trials, participants had to instead categorize its size. We manipulated the size context on trials asking about lightness categorizations only (and vice versa for lightness context manipulations). This ensured that participants were exposed to a biased set of stimulus sizes, but that they would not have to respond to disproportionately many objectively large stimuli in one context, and to disproportionately many objectively small stimuli in another. This eliminated the possibility that context-dependent categorizations could be driven by a bias, whereby participants might be striving for an equal number of “large” and “small” categorizations across contexts. At the same time, it also made it so that any such bias would act in opposition of the contextual effect. That is, in our design, responding “large” as equally often as “small” would lead to a stable categorization boundary across the different contexts, as size trials are composed of an equal number “large” and “small” ground truth categorizations in all contexts.

Experiment 1 also manipulated the timing of trials to examine two additional research questions. First, is the contextual bias in categorization decisions stronger when participants have had the opportunity to experience (and potentially, adapt to) each stimulus for longer? To address this question, we compared the contextual effect on category boundaries between a short (200ms) and long (1000ms) stimulus presentation condition. And second, how specific is the contextual effect? Is it just the context provided by task-relevant stimuli that impacts categorizations, or does the visual environment more generally also influence the location of the category boundary? The timing manipulation allowed us to examine this as participants’ visual environment, and consequently exposure to task-irrelevant properties, differed between the two stimulus timing conditions.

4.2.1 Methods

Participants

Ten participants (aged 22.3 ± 4.37) with normal or corrected-to-normal vision took part in Experiment 1. The study received ethical approval from the Central

4. Context-dependent categorization

University Research Ethics Committee at the University of Oxford (approval reference number: R58531/RE001) and took part in the Perception Laboratory at the Department of Experimental Psychology. All participants provided written informed consent and were compensated with either course credit or £10 per hour for their time.

Apparatus

Participants were seated in a dark room approximately 175cm away from a computer monitor (1080x1080 resolution, 32" Display++ LCD monitor) with linearized output light intensities. Visual stimuli were created and presented with ViSaGe, CRS Toolbox (Cambridge Research Systems) for MATLAB.

Stimuli

Stimuli in the experiment consisted of grayscale circles presented on a light yellow background (CIE $X=.372$, $Y=.400$, $Z=76.1\text{cd/m}^2$). We followed previous work in lightness perception in choosing a chromatic background to minimize background-anchoring effects (e.g. to the minimum lightness available – black background, the maximum – white background, mean – half-contrast grey, etc.). For the contextual manipulation, we constructed 3 types of stimulus sets (low, neutral and high) for 2 stimulus property dimensions (lightness and size). For each of the two stimulus dimensions we first generated an array of 16 possible property values. Those values corresponded to the radius of the circle in pixels, and the lightness of the circle in 10-bit pixel values. We then sorted those values into two biased sets (low set: the lowest 10 values, high set: the highest 10 values) and a neutral set (the middle 10 values).

We constructed the full arrays of property values using a perceptual difference scale of lightness and size, which we estimated with a pilot group of participants ($n=6$, following same ethics protocol R58531/RE001) via maximum likelihood difference scaling (MLDS, Maloney and Yang 2003; Knoblauch, Maloney, et al. 2008, see Experiment 2 for details on this method). We took this approach to ensure that the stimuli we employed in the experiment would all be equally discriminable.

4. Context-dependent categorization

The size array ranged from 90 to 210 pixels in steps of 8 pixels (corresponding to approximately 1.09° to 2.54° of visual angle at the viewing distance of the display). Note that linear step sizes in terms of radius are quadratic in terms of circle area. The lightness array ranged from 24 to $75\text{cd}/\text{m}^2$ in steps of $3.4\text{cd}/\text{m}^2$. The linear perceptual spacing of lightness we observed in piloting mirrors past results for this luminance range, (e.g. Wiebel et al. 2017; Rogers et al. 2016).

Experimental Procedure

Each participant completed the experiment in a single hour-long session. The experiment consisted of 2 practice blocks and 30 test blocks of 60 trials each. Each block started with an example circle stimulus whose lightness and size corresponded to the median of the overall stimulus set. Participants were advised that they could use this circle as a reference for neutral (or medium) size and lightness.

The first practice block served to familiarize participants with the experiment. Stimuli were presented for up to 4 seconds, until a categorization response was registered. The block continued until the participant achieved an average response time of 1.5 seconds or below. The length of the first practice block was constrained to be of a minimum of 80 trials and a maximum of 150 trials. The second practice block comprised 100 trials and introduced a shorter stimulus duration – 200ms.

In each trial (Fig. 4.3a), the participant first heard a computer-generated voice (Google Text-to-Speech API, gender neutral voice) announce the type of trial, by producing either the word “lightness” or “size”. 200ms after the onset of the voice, a circle appeared on the screen. In size trials, participants had to categorize the circle as either large or small; in lightness trials, participants had to categorize the circle as either light or dark. Responses were accepted until the end of the trial, 1.5 seconds post stimulus onset. There was a 200ms interval between trials.

The experiment followed a 2×5 factor design. The first factor was timing – in half of the blocks, stimuli were presented for 200ms (brief condition), and in the other half of the blocks – for 1000ms (long condition). The timing manipulation allowed us to test whether the contextual effect is stronger when the participant

4. Context-dependent categorization

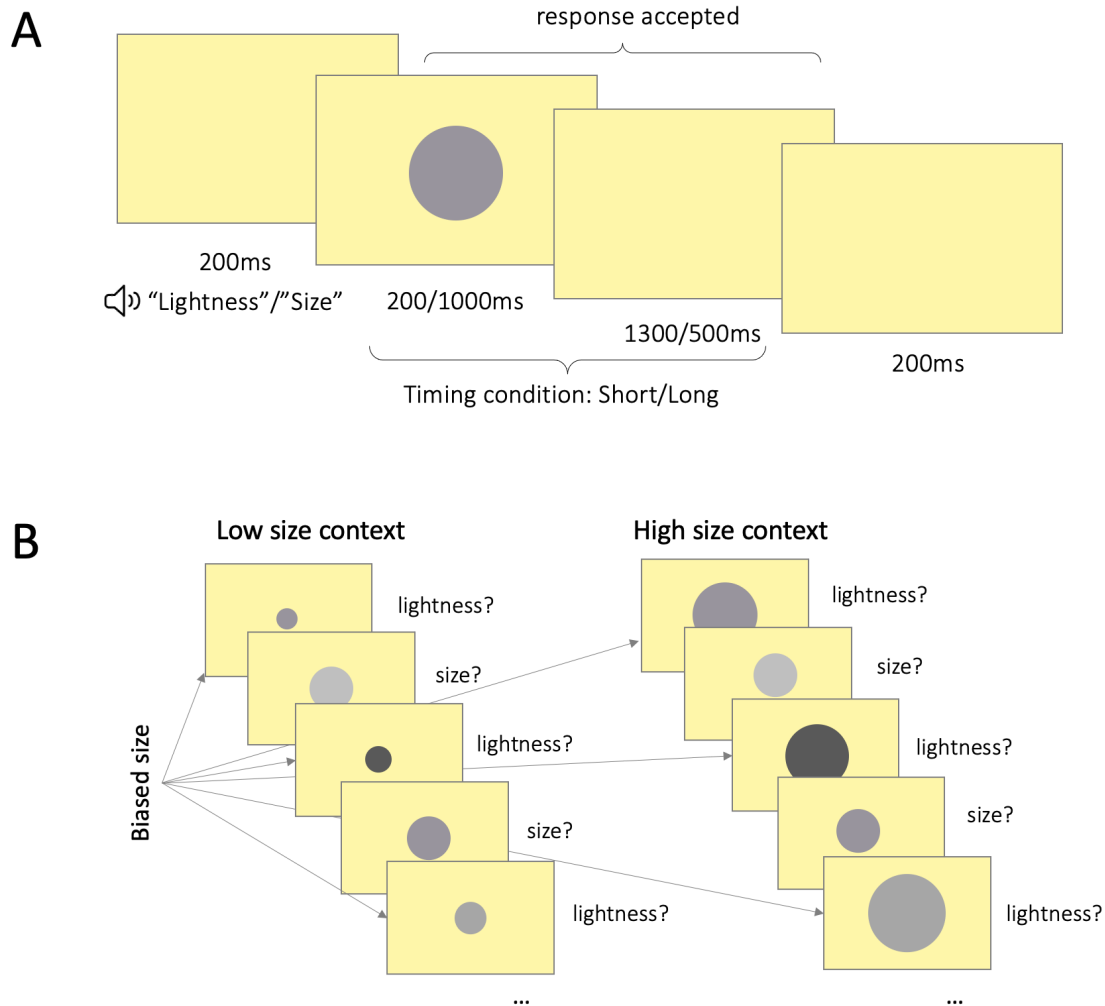


Figure 4.3: A: Trial structure. Participants first heard a computer generated voice announce the trial type. 200ms following this, a circle stimulus was presented in the center of the screen for either 200ms or 1000ms. In size trials, participants had to categorize the stimulus as small or large; in lightness trials, as light or dark. Responses were accepted for 1.5s starting from stimulus onset. The stimulus remained on the screen for the full stimulus duration even if the participant responded prior to stimulus offset. There were 200ms breaks between trials. **B:** Contextual manipulation. The left panel illustrates a sequence of 5 trials in a low size context, the right panel – in a high size context. Across both panels, in trials asking about size, stimulus size was drawn from a neutral set, such that half the stimuli would be objectively small and half the stimuli – objectively large. In trials asking about lightness, stimulus size was drawn from biased sets, such that stimulus size would on average be lower in the low context, and higher in the high context.

4. Context-dependent categorization

has had the opportunity to adapt to the stimulus set for longer. That is, is there an interaction between stimulus duration and the contextual manipulation? The timing manipulation also allowed us to test another question – how specific is the contextual effect? Are categorization decisions impacted only by the context provided by previous *task-relevant stimuli*, or do other, task-irrelevant inputs also affect categorizations? In the brief condition, participants were exposed to the background for longer than in the long stimulus presentation condition. As the luminance of the background (76cd/m^2) was higher than that of the stimuli ($24\text{--}75\text{cd/m}^2$), in the brief condition, participants were exposed to a lighter visual environment compared to in the long condition. If participants' categorization decisions were influenced by the properties of the task-relevant stimuli only, lightness categorizations should not differ between the brief and long conditions (since the relative frequencies of stimulus lightness values are matched across the long and short conditions). But if the contextual effects were driven by a more general adaptive process which is not specific to the imperative stimuli, we would expect lightness categorizations to differ between the two timing conditions (with higher thresholds in the brief condition). For size, as participants were seated in a dark room, the only obvious size reference available in the gaps between stimulus presentations was the size of the screen itself. By contrast, during stimulus presentations, the size environment included both the size of the monitor and the (relatively smaller) size of the current stimulus. Thus, the size environment was on average smaller in the long condition (where stimuli were presented for longer) than in the brief condition. We reasoned that, according to the adaptation hypothesis, this might lead to the same pattern of contextual bias for size as for lightness – higher categorization thresholds in the brief condition.

The second experimental factor was the contextual manipulation. In each block, we manipulated the context for only one stimulus dimension by using a biased set (i.e. either the low or high stimulus set) in the trials that asked about the *other* stimulus dimension (Fig. 4.3b). For instance, in a “high size” block, stimulus size was drawn from the high size stimulus set only on trials that asked participants

4. Context-dependent categorization

to categorize the lightness of the circle; on trials where the participant had to categorize the size of the circle, stimulus size was drawn from the neutral size stimulus set. This ensured that our contextual manipulation would not introduce a confounding response bias, i.e. it would not skew the ground truth number of ‘large’ responses. In each block, we manipulated the context for only one stimulus dimension, the context for the other stimulus dimension remained neutral (neutral stimulus set). This resulted in 4 combinations of context conditions (size-lightness: high-neutral, low-neutral, neutral-high, neutral-low), to which we added a fifth neutral condition (neutral-neutral).

Responses were registered via button presses on a hand-held number keyboard. Participants were instructed to hold the keyboard with both hands and respond with their left thumb on size trials, and right thumb on lightness trials (or vice versa, counterbalanced across participants). Categorizing a stimulus as “large” or “light” was associated with the press of an upper button; categorizing a stimulus as “small” or “dark” was associated with the press of a down button. As the timings of the trials were fixed, nothing changed on the screen upon participant response; instead, participants heard a click to confirm their choice was registered.

Analyses

First, we estimated psychometric functions separately for each participant and each condition in the experiment, using the Psignifit toolbox for MATLAB (Schütt et al. 2015). We tested whether the contextual manipulation biased participants’ categorization decisions on the group level with a general linear model (GLM) of participant thresholds (normalized as a z score). The GLM quantified the impact of the contextual manipulation (coded as -1, 0, 1 for low, neutral and high), timing condition, and their interaction, controlling for the nonindependence of threshold measurements from the same observer. We followed up the results of the GLM with pairwise comparisons with paired samples t tests.

As a complementary approach for probing the effect of temporal context, we regressed participant choices on each trial on the relevant property of the stimulus

4. Context-dependent categorization

on the current trial, as well as on the value of that same property on the previous 3 trials. Thus, for each participant we built two logistic regression models, one for lightness and one for size. In the model for size, the outcome variable comprised the responses to all trials asking the participant to categorize the size of the stimulus; the predictor variables included the size of the stimulus on the current trial t , as well as the size of the stimuli on the immediately preceding 3 trials, $t - 1$, $t - 2$ and $t - 3$ (all predictors were z -scored). This approach allowed us to quantify the effect of stimulus history (size on previous trials, $t - 1$, $t - 2$ and $t - 3$), controlling for the effect of the imperative stimulus (size on current trial, $t - 1$). Note that as trial types (lightness or size categorizations) were interleaved, the predictor variables in the regression included both types of trials. For instance, in a regression on lightness categorizations, we included as predictors the lightness of stimuli on the preceding three trials; on some of those trials, the participant was likely asked to judge the size (and not lightness) of the stimulus. We assessed the statistical significance of the beta coefficients from the regression on the group level via one sample t tests.

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/vxh3n/>) for this project.

4.2.2 Results

The psychometric functions for each participant are presented in Fig. 4.4 and Fig. 4.5 where each column represents one participant. While there is considerable variability in performance, there is also a clear pattern of dissociation of psychometric functions between the different context conditions. Generally, the psychometric function for the low condition (light grey) is shifted to the left and the psychometric function for the high condition (black) is shifted to the right relative to the psychometric function for the neutral condition (dark grey). These horizontal shifts in the psychometric function are captured by the parameter *threshold*, which tracks the stimulus property value that corresponds to an equal probability of categorizing a

4. Context-dependent categorization

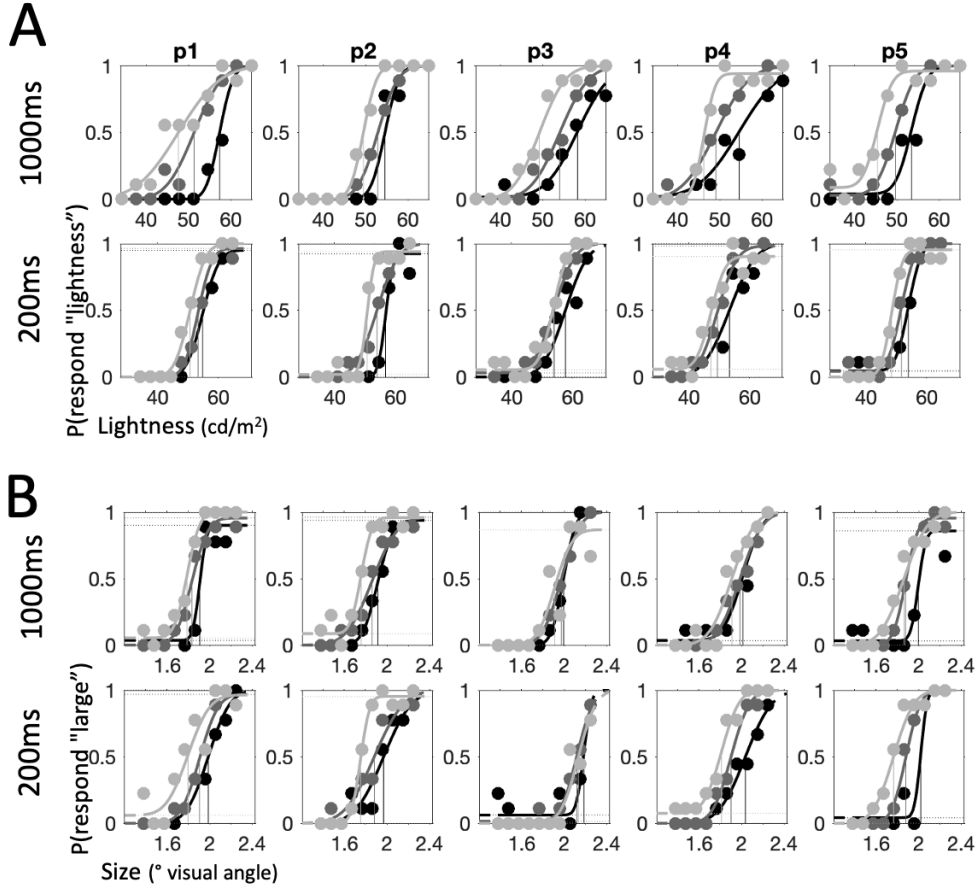


Figure 4.4: Psychometric functions for each individual participant: first half of participants, continued in the next figure. Across all panels, columns depict individual participants (p1-p10). Shading tracks context (light grey: low, dark grey: neutral, black: high). Points show raw data, curves – cumulative Gaussian functions fit to data. Vertical lines show threshold value estimates for each condition. **A:** Psychometric functions for lightness trials. Top row shows long condition (1000ms) and bottom row – short condition (200ms). **B:** Psychometric functions for size trials. Top row shows long condition (1000ms) and bottom row – short condition (200ms).

stimulus as “large” or “small” (/“light” or “dark”). Threshold values are visualized in Fig. 4.4 as vertical lines projecting downwards from the inflection point of each psychometric function.

The average values of the thresholds across participants are presented in Fig. 4.6. We quantitatively assessed the influence of context and stimulus timing on threshold values via a GLM, controlling for individual influences on thresholds. The contextual manipulation significantly influenced categorizations (lightness: $\beta = 0.81$, $p < 0.0001$, size: $\beta = 0.63$, $p < 0.0001$), such that thresholds increased

4. Context-dependent categorization

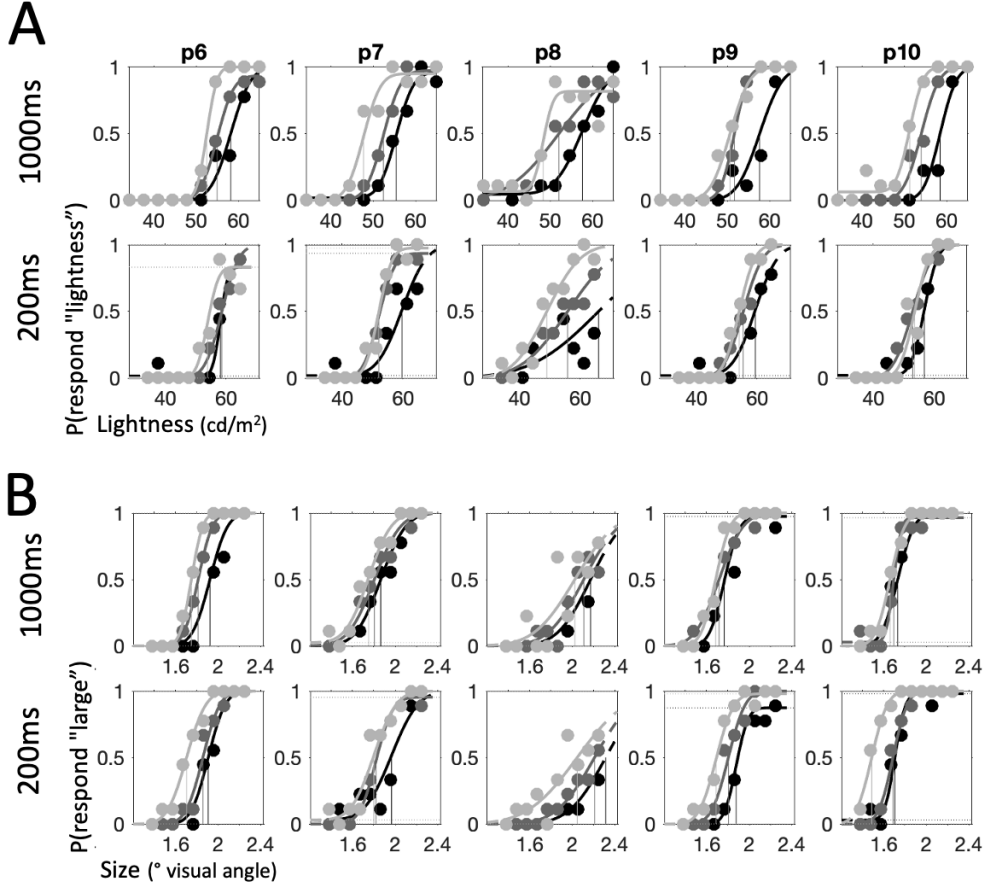


Figure 4.5: Psychometric functions for each individual participant: second half of participants, continued from previous figure

from low to neutral, and neutral to high conditions. Pairwise comparisons confirmed this pattern of results (lightness low < neutral: $t_9 = 6.78$, $p < 0.001$, lightness neutral < high: $t_9 = 8.66$, $p < 0.001$; size low < neutral: $t_9 = 7.02$, $p < 0.001$, size neutral < high: $t_9 = 7.97$, $p < 0.001$).

Stimulus timing impacted lightness thresholds ($\beta = -0.45$, $p < 0.001$), with higher thresholds for lightness judgments in the brief presentation condition. Shorter stimulus timing also increased thresholds for size judgments ($\beta = -0.21$, $p = 0.01$), albeit the effect was weaker than for lightness. This pattern of results suggests that categorization decisions are subject to contextual influences going beyond the immediately task-relevant stimuli. In the brief condition, the relatively high lightness of the background provided a lighter visual environment across all context conditions, resulting in higher categorization thresholds for lightness.

4. Context-dependent categorization

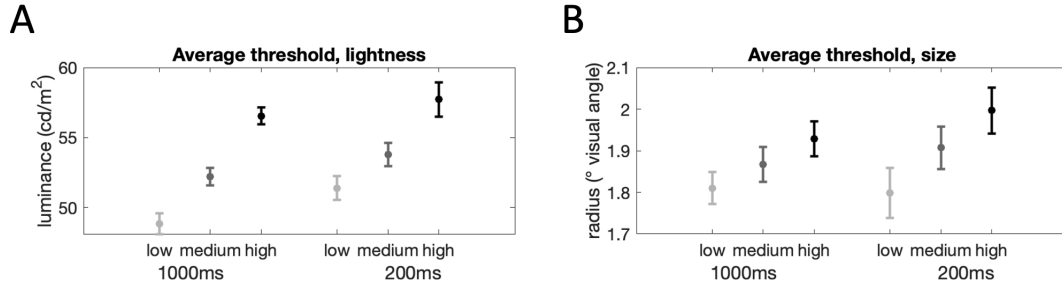


Figure 4.6: Threshold values for each context condition (light grey: low, dark grey: neutral, black: high) and timing condition (left: long, right: short). **A:** Lightness judgments. **B:** Size judgments.

Similarly, the brief condition provided, on average, a larger size reference, leading to increased categorization thresholds for size. This result suggests that the visual environment (beyond the immediately task-relevant inputs) contributes to the contextual bias in categorization decisions.

Finally, the interaction between timing and context was not statistically significant for lightness ($\beta = 0.18$, $p = 0.20$). This null result indicates that the contextual effect for lightness did not increase when participants were exposed to stimuli for longer (1000ms vs 200ms). The interaction between stimulus duration and context for size, however, was significant ($\beta = -0.26$, $p = 0.01$) indicating that the contextual effect *decreased* when participants were exposed to stimuli for longer (1000ms vs 200ms). Note that this effect goes in the opposite direction of the adaptation prediction; an adaptive process should lead to stronger context-dependencies as stimulus duration increases.

We explored the effect of temporal context further by quantifying the influences of recent stimulus history on choice via a logistic regression. The resulting beta coefficients for the decision-relevant stimulus property on the most recent 4 trials are available in Fig. 4.7. As expected, the lightness ($t_9 = 14.31$, $p < 0.001$) and size ($t_9 = 14.95$, $p < 0.001$) of the stimulus on the current trial were both strongly associated with a higher likelihood of categorizing it as light and large, respectively. Categorizations were, however, repulsed away from the properties of the stimulus on the preceding trials. Higher stimulus lightness on the previous trial(s) led to a lower likelihood of categorizing the stimulus on the current trial as light (trial $t - 1$:

4. Context-dependent categorization

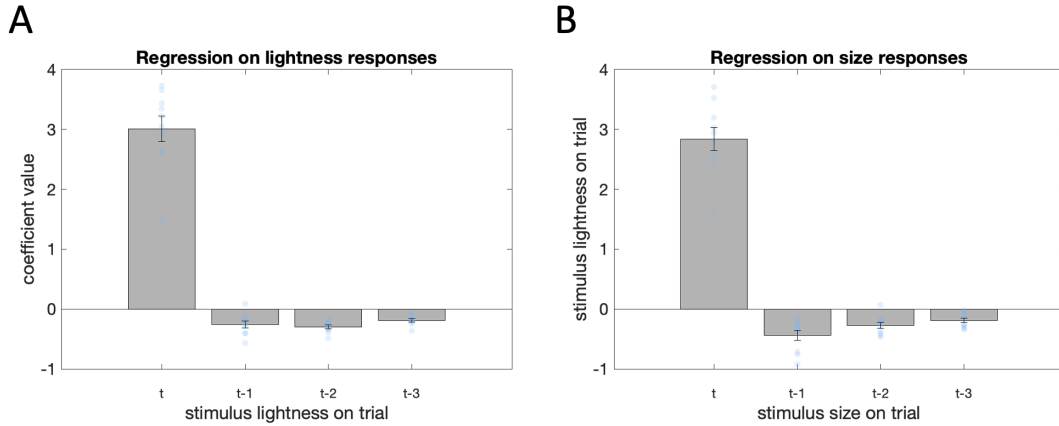


Figure 4.7: Regression on stimulus history. Across both panels, the height of the bar depicts the average standardized coefficient value for the stimulus property value on the current trial (first bar: trial t), the stimulus property value on the immediately preceding trial (second bar: trial $t-1$), the stimulus property value on trial before last (second bar: trial $t-2$), and the stimulus property value three trials ago (second bar: trial $t-3$). Error bars depict SEM, blue dots show coefficient values for individual participants. **A:** Coefficients for lightness responses. **B:** Coefficients for size responses.

$t_9 = -5.00$, $p < 0.01$; trial $t - 2$: $t_9 = -10.10$, $p < 0.001$; trial $t - 3$: $t_9 = -7.24$, $p < 0.001$). Similarly, higher stimulus size on the previous trial(s) led to a lower likelihood of categorizing the stimulus on the current trial as large (trial $t - 1$: $t_9 = -5.31$, $p < 0.001$; trial $t - 2$: $t_9 = -5.39$, $p < 0.001$; trial $t - 3$: $t_9 = -4.80$, $p = 0.001$). Exploratory analyses (by adding more predictor variables to the logistic regression) indicated that the effect of stimulus history extended as far back as to the preceding 13 trials for lightness (trial $t - 13$: $t_9 = -3.17$, $p = 0.01$), and the preceding 8 trials for size (trial $t - 9$: $t_9 = -4.13$, $p < 0.01$).

4.2.3 Interim Discussion

Temporal context yielded a significant influence on categorizations in Experiment 1. Category boundaries for size and lightness were higher when the temporal context was dominated by larger and lighter stimuli; categorization judgments were also repulsed from the task-relevant characteristics of the stimuli on preceding trials. The smaller the size of the preceding stimuli, the more likely the current stimulus to be categorized as large. Our experimental design allowed us to rule out response

4. *Context-dependent categorization*

bias as a driver of the observed effects. In fact, temporal context biases decisions even when this would lead to a distortion in the relative frequency of actions. This is because in our design, the relative frequency of ground truth “high” and “low” judgments was carefully counterbalanced. Thus, context-dependent categorization necessarily skews the distribution of “high” vs “low” responses.

We did not find evidence that the contextual effect increases when each stimulus is presented for 1000ms versus 200ms. We predicted that longer stimulus durations would allow for additional adaptation to the stimulus set. Contrary to our prediction, for size, we found that *shorter* stimulus durations led to stronger contextual biases. This finding may speak to the literature on the effect of noise on context-dependencies; shorter stimulus durations made the task more demanding. A Bayesian perspective dictates that this should lead to heavier reliance on the prior and consequently, stronger contextual biases. Interpreted in this light, our results appear to provide support for a Bayesian contraction bias mechanism driving the observed context dependence.

However, we did also find evidence that the contextual effect is not isolated to task-relevant information. Contrary to the predictions of a Bayesian process, which incorporated the distribution of recently experienced stimulus properties in a prior, we found that participants were also biased by task-irrelevant visual information. Categorization thresholds were higher in the brief condition where the task-irrelevant visual environment was lighter and larger. This suggests that a more automatic, sensory visual adaptation process, acting on our perception of the current to-be-categorized stimulus, likely contributes to the observed contextual dependencies. In Experiment 2, we follow up on these findings and assess more directly the contributions of adaptation.

4.3 Experiment 2

In Experiment 2, we aimed to disentangle the two putative sources of contextual influence – changes to our representation of the category boundary (Fig. 4.2a) versus

4. Context-dependent categorization

changes to our experience of the current stimulus (Fig. 4.2b). We took advantage of the fact that adaptation processes, which impact our perception of incoming input, are characterized by changes in sensitivity to the dominating characteristics in the environment. Thus, we estimated the contribution of adaptation to context-dependent categorization judgments by comparing participants' sensitivity to stimulus characteristics across two temporal contexts. We did this separately for the two visual properties from Experiment 1, size and lightness, as well as for a more abstract stimulus property – numerosity. Computing numerical magnitudes engages later stages of processing compared to low-level visual properties, such as lightness, contrast or orientation (Heng et al. 2020) and has been linked with various higher-level cognitive processes, which rely on an ‘approximate number system’ (Piazza et al. 2007; Nieder and Dehaene 2009). Imprecision in the estimation of numerical magnitudes has also been put forth as the driver of biases in value-based decisions about economic prospects (Woodford 2020). Consequently, including numerosity in this experiment allows us to link the current chapter more closely with context dependencies in economic decisions.

4.3.1 Methods

While Experiment 1 was conducted in the Perception Laboratory at the Department of Experimental Psychology, the pandemic prevented onsite data collection for Experiment 2. Thus, all data for this experiment were collected remotely.

Participants

Twenty participants (aged 28.3 ± 8.63) with normal or corrected-to-normal vision took part in Experiment 2. The study received ethical approval from the Central University Research Ethics Committee at the University of Oxford (approval reference number: R55652/RE005). All participants provided written informed consent and were compensated with £10 per hour for their time.

4. Context-dependent categorization

Aparatus

Participants completed the experiment on their personal computers. The study was programmed using PsychoPy (Peirce 2007) and PsychoJS and was hosted on Pavlovia. The output (combined RGB) light intensities of all monitors were calibrated with a perceptual calibration method (half tone pattern matching, Xiao et al. 2011) or with a photometer (ColorCal, Cambridge Research Systems). Participants were instructed to set their monitors to the lowest brightness setting and to disable any automatic brightness adjustment or light filtering applications on their computers. We chose to conduct the perceptual calibration procedure at low brightness for two reasons. First, increasing the brightness settings on a monitor, might introduce a brightness offset (i.e. requested 0% lightness does not correspond to black pixel output, but produces some light), which would render the gamma model ($y = x^\gamma$) of the electro-optical transfer function of the monitor inaccurate. Second, human visual sensitivity to changes in brightness decreases with absolute brightness levels, and so, we can minimize variability in brightness matches by conducting them at a low brightness setting. This in turn ensures that our perceptual estimate of the electro-optical transfer function of the monitor, which is based on the brightness matches, is reliable.

Viewing distance was determined individually for each participant. We asked participants to adjust the size of a rectangle on their screen to match the size of a standard bank card (Yung et al. 2015). We used the size of the rectangle in pixels to calculate the required visual distance such that 1 pixel would subtend 57'' of visual angle; this ensured that the size of the stimuli in visual angle would be approximately the same for all participants.

Stimuli

The experiment consisted of three parts, each of which followed the same structure, but concerned a different stimulus property – size, lightness or numerosity. In the size part of the study, the stimuli comprised grayscale circles presented against a black background. In the lightness part of the study, the stimuli comprised

4. Context-dependent categorization

grayscale circles presented against a yellow background (as in Experiment 1). In the numerosity part of the study, the stimuli comprised clouds of black and white dots of varying sizes against a grey (half contrast) background. The color (white or black) was randomly selected for each dot, ensuring that the brightness of the stimulus would not be informative for the numerosity judgment. Following previous work, we also constrained the average dot radius or the total area covered by the dots to be the same across dot clouds in each trial (Izard and Dehaene 2008; Van den Berg et al. 2017; Heng et al. 2020). For each trial, we randomly chose whether to hold dot radius or total area constant. This manipulation ensured that stimulus size neither on the level of individual dots nor on the level of the entire dot cloud would be informative for the numerosity judgment.

As in Experiment 1, for each of the three properties, we constructed an array of possible property values. The full arrays consisted of 10 possible values; the lowest 7 of those constituted the low stimulus set, and the highest 7 – the high stimulus set. For size, the values ranged from 69 to 195 pixels in steps of 4 (corresponding to approximately 1.1° to 3.1° of visual angle). For lightness, the pixel values ranged between 0.35 and 0.65 in steps of 0.033 (assuming 0-1 range on an 8-bit monitor with linearized output intensities). For numerosity, the number of dots ranged from 8 to 52 in logarithmic steps, using a multiplier of 1.232.

Experimental Procedure

The experiment took approximately 3 hours, which participants were invited to complete over three shorter sessions. Links to the online experiments are available in the OSF repository (<https://osf.io/vxh3n/>) for this study. We did not counterbalance the order of the parts across participants, as the subjective difficulty of the tasks in the three parts increased progressively from size through lightness to numerosity.

Across all parts of the experiment, participants had to complete two types of tasks: categorization and difference comparisons. Participants' difference comparisons allowed us to estimate suprathreshold sensitivity. In categorization trials,

4. Context-dependent categorization

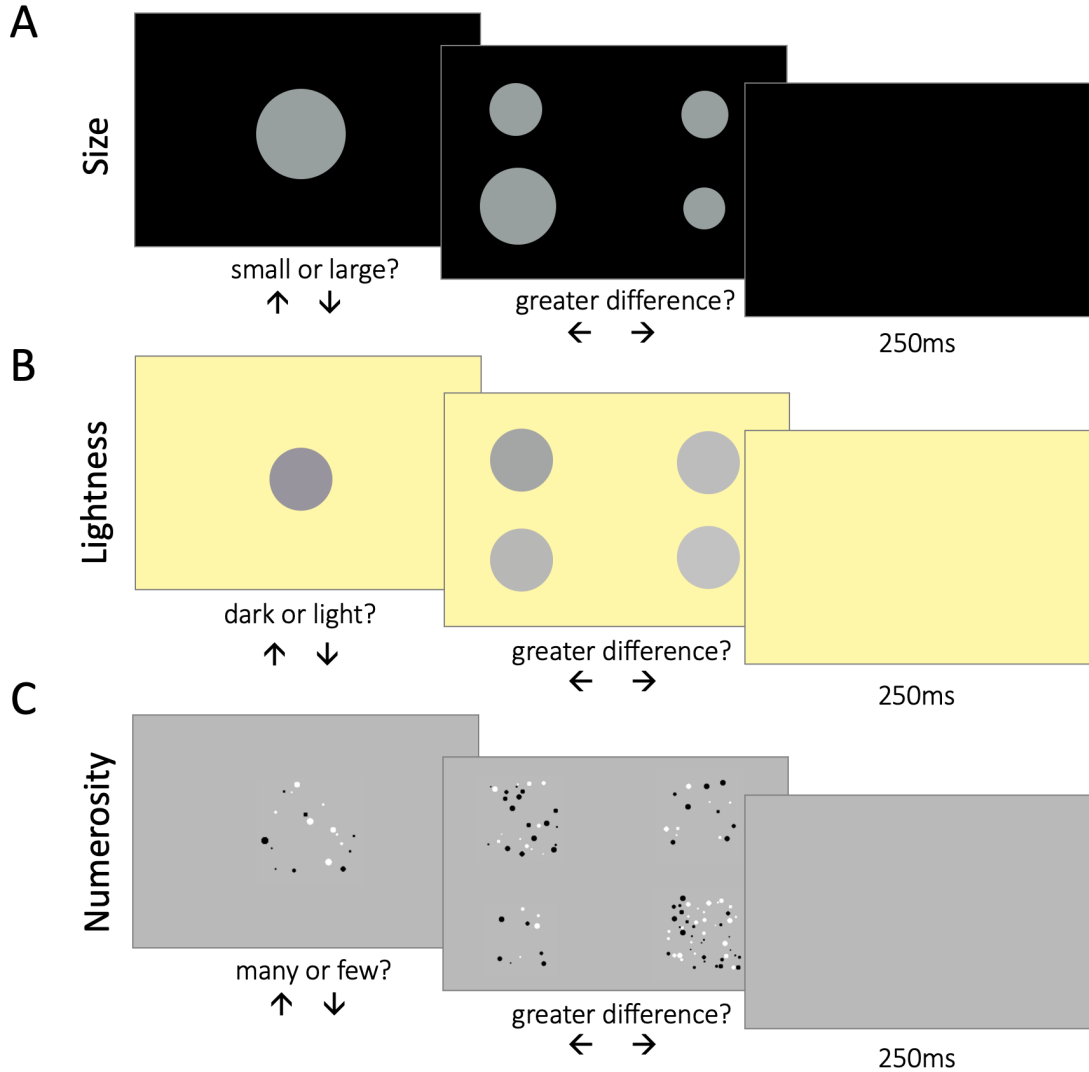


Figure 4.8: Trial structure. Across all parts of the experiment, categorization and difference comparison trials were interleaved. On categorization trials, participants saw a single stimulus in the center of the screen. On difference comparison trials, participants saw two pairs of stimuli, two stimuli on the left and two on the right side of the screen. **A:** Part 1, size task. In categorization trials, participants had to categorize circle stimuli as large or small. They responded by pressing the up or down arrow key. In difference comparison trials, participants had to compare the difference in size between the two pairs of stimuli. If they considered the difference between the stimuli on the left larger than the difference between the stimuli on the right, they responded by pressing the left arrow key (and vice versa for right). **B:** Part 2, lightness task. Participants had to categorize circle stimuli as light or dark, and compare the difference in lightness between the two pairs of stimuli. **C:** Part 3, numerosity task. Participants had to categorize clouds of dots as many or few, and compare the difference in number between the two pairs of clouds.

4. Context-dependent categorization

participants saw a single stimulus in the center of the screen. In the size part, their task was to categorize the circle as large or small; in the lightness part – to categorize the circle as light or dark; and in the numerosity part – to categorize the number of dots in the cloud as many or few. Participants registered their choice via button press (up or down arrow). In difference comparison trials, participants saw two pairs of stimuli on the left and right side of the screen. Participants had to compare the two stimuli on the left, and decide whether the difference between them was larger than the difference between the two stimuli on the right. Thus, participants had to indicate which pair of stimuli was most different from one another. Again, they registered their choice via button press (left or right arrow). The stimuli remained on the screen for the full duration of the trial. The trial ended only when the participant provided a response.

Each part of the experiment comprised 4 blocks. Each block began with a series of 10 categorization trials, followed by 210 comparison and 210 categorization trials, presented in an interleaved fashion. We implemented the contextual manipulation during the categorization trials. In half of the blocks, the stimulus in categorization trials was drawn from the high stimulus set, and in the other half – from the low stimulus set. This resulted in two contextual bias conditions – low and high. Across all blocks, the stimuli in the difference comparison trials were drawn from the full stimulus set (i.e. the 10 value array).

Analyses

First, we analyzed categorization trials. We estimated psychometric functions for the categorization decisions of each participant for each condition (low and high context) and each stimulus type (size, lightness and numerosity). As a sanity check, we confirmed that the contextual manipulation impacted categorizations, by comparing the threshold values for the low and high context conditions via paired samples t test.

Next, we analyzed difference comparison trials. We estimated sensitivity in the two conditions via maximum likelihood difference scaling (MLDS, Maloney

4. Context-dependent categorization

and Yang 2003; Knoblauch, Maloney, et al. 2008). MLDS constitutes a stochastic model of suprathreshold perceptual differences. The model assigns each of the 10 stimuli in the experiment a value on a latent perceptual scale, ψ_1 to ψ_{10} , such that the absolute differences between scale values (corresponding to the stimuli on a given trial) maximize the likelihood of the participant’s response (the reported difference comparison). Since any linear transformation of the perceptual scale produces the same pattern of results, the model constrains the scale to span values between 0 and 1, (i.e. $\psi_1 = 0$ and $\psi_{10} = 1$). Thus, there are 8 free ψ parameters (ψ_2 to ψ_9) plus one additional free scaling term, σ , which tracks the variability of participant responses. To assess whether the temporal context manipulation led to changes in suprathreshold sensitivity, we compared the perceptual scales estimated based on responses in the low condition (ψ_2^l to ψ_9^l) and responses in the high condition (ψ_2^h to ψ_9^h). More specifically, an adaptive process should lead to higher sensitivity for low stimulus magnitudes in the low context condition and higher sensitivity for high stimulus magnitudes in the high context condition. We tested this prediction on the group level with a GLM, which quantified the impact of the contextual manipulation on participants’ scale values (ψ_{2-9}^l and ψ_{2-9}^h), controlling for the nonindependence of scale values from the same observer and for the effect of stimulus level (2-9). To ascertain whether the effect of context was localized to a portion of stimulus space (e.g. the lower or higher end of stimulus values), we also assessed the statistical significance and direction of the coefficient for the interaction between context and stimulus level via a GLM.

Computational Modeling

Finally, to bridge the results from categorical and difference comparison trials, we built one common model encompassing both tasks. At the heart of the model is the assumption depicted in Fig. 4.2. That is, the model assumes that categorization decisions are based on the participant’s representation of the stimulus (Fig. 4.2b) and the participant’s representation of the category boundary (Fig. 4.2a), both

4. Context-dependent categorization

of which might differ between contexts:

$$r_{cat}^l = \frac{1}{1 + \exp(-((x - c^l) - b^l) \cdot s^{-1})}$$

$$r_{cat}^h = \frac{1}{1 + \exp(-((x - c^h) - b^h) \cdot s^{-1})}$$
(4.1)

where r_{cat} refers to the categorization decision, which is a logistic function of the stimulus property x ; the superscripts h and l denote the relevant context (high and low). The model comprises 2 free parameters tracking changes to the stimulus representation (c^l and c^h), and 2 free parameters tracking changes to the category boundary (b^l and b^h). The former two parameters determine the location of the transducer function along stimulus space, the latter two free parameters determine the location of the category boundary along stimulus space. A further free parameter, s , determines the slope of the transducer which may differ between participants.

If the model were fit to categorization decisions only, the parameters tracking the contextual shift in the transducer would trade off with the parameters tracking the contextual shift in category boundary, and so, it would not be possible to reliably estimate both. However, difference comparison trials are based on the participant's representation of the stimulus *only*, and not on the representation of the category boundary:

$$y_i^l = \frac{1}{1 + \exp(-(x_i - c^l) \cdot s^{-1})}$$

$$y_i^h = \frac{1}{1 + \exp(-(x_i - c^h) \cdot s^{-1})}$$
(4.2)

$$r_{diff}^l = |y_1^l - y_2^l| - |y_3^l - y_4^l|$$

$$r_{diff}^h = |y_1^h - y_2^h| - |y_3^h - y_4^h|$$

where r_{diff} refers to the difference comparison decision, y_i refers to the representation of stimulus x_i and the subscript i tracks each of the four stimuli on a difference comparison trial. Thus, feeding information about participants' difference comparisons

4. Context-dependent categorization

into the model allowed us to disentangle contextual changes to the representation of the stimulus (c^l versus c^h) from contextual changes to the representation of the category boundary (b^l and b^h).

We fit the model to empirical data from the experiment individually for each participant and each condition via gradient descent, implemented using the `globalsearch` algorithm from the MATLAB Optimization Toolbox. The best fitting parameter values allowed us to estimate the relative contributions of the two putative sources of influence to the contextual bias in categorizations. We did this by, first, calculating the contextual difference in the representation of the category boundary ($\Delta(b) = b^h - b^l$) and the contextual difference in the representation of the stimulus ($\Delta(c) = c^h - c^l$), as well as their sum ($\Delta(\text{total}) = \Delta(b) + \Delta(c)$); we then estimated their relative strengths ($\frac{\Delta(c)}{\Delta(\text{total})}$).

Data & Code Availability Statement

All data and code to reproduce the analyses are available in the OSF repository (<https://osf.io/vxh3n/>) for this project.

4.3.2 Results

Psychometric functions for categorization decisions are available in Fig. 4.9a; there is clear separation between low and high context conditions. Paired samples t tests on participants' threshold values for the two contexts confirm that categorization boundaries were higher in the high context across size ($t_{18} = 10.18$, $p < 0.0001$), lightness ($t_{18} = 7.41$, $p < 0.0001$), and numerosity ($t_{18} = 5.95$, $p < 0.0001$) judgments (Fig. 4.9b). We excluded one participant from the main analyses, as visual inspection of their psychometric function indicated that they responded randomly on the categorization task for numerosity. We confirmed this by estimating the impact of the numerosity of each stimulus on that participant's categorization judgment via logistic regression (coefficient in the low condition $b = 0.04$, $p = 0.49$, coefficient in the high condition: $b = 0.05$, $p = 0.32$).

4. Context-dependent categorization

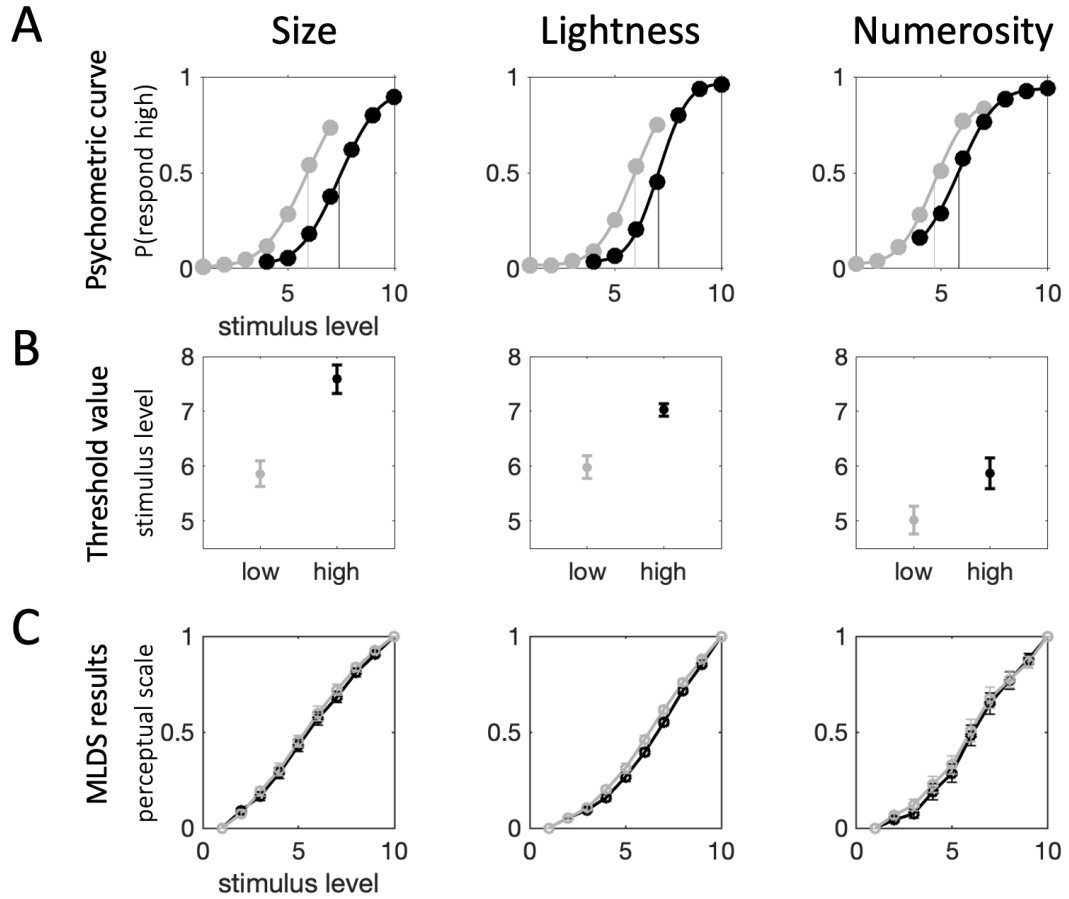


Figure 4.9: Experiment 2 results. Columns show results for each of the three different parts (left: size, middle: lightness and right: numerosity). **A:** Psychometric functions based on aggregate data collapsed across participants. Points show raw data, curves – cumulative Gaussian functions fit to data. Vertical lines show threshold value estimates for each condition. Shading indexes context condition (grey: low, black: high). Note that these average curves are for illustration purposes only, they are not necessarily representative of any one participant in the experiment. **B:** Average threshold values (across all participants) for the two conditions (grey: low, black: high). Error bars depict SEM. **C:** Perceptual scales for the two conditions (grey: low, black: high). Dots depict average scale value across participants, error bars depict SEM.

4. Context-dependent categorization

Perceptual scales estimated from participants' difference comparison decisions are depicted in Fig. 4.9c. The plots show the average estimates of perceptual scale values of all participants, error bars track SEM, shading denotes context. The distance between scale values (on the same perceptual scale) illustrates the estimated sensitivity – the further away two points are, the more easily the participant can distinguish the stimuli they refer to. Thus, if the perceptual scale lay on the identity line (i.e. all points were equidistant along the y-axis) then all stimuli would be equally discriminable, and so, participants would be equally sensitive to changes along stimulus space. Distortions to the perceptual scale, on the other hand, illustrate differential sensitivity. For instance, the more convex the perceptual scale, the higher the relative sensitivity for higher values of the stimulus property compared to lower values of the stimulus property. This means that the directionality of the dissociation between perceptual scales in the two contexts is important. An adaptive process should lead to higher sensitivity to high stimulus property values in the high context, and to low stimulus property values in the low context. Consequently, we would expect to see a more convex perceptual scale for the high context (i.e. black curve below the grey curve).

The trends depicted by the average perceptual scales (Fig. 4.9c) seem to, on the whole, conform to this pattern. However, the effect appears small. Statistical tests confirmed that the contextual effect on sensitivity was reliable on the group level across all three stimulus types (size: $\beta = 0.10$, $p < 0.0001$, lightness: $\beta = 0.08$, $p < 0.0001$, numerosity: $\beta = 0.10$, $p < 0.0001$). We did not find evidence that this effect was localized to the lower or higher end of stimulus space; the interaction between context and stimulus level was not statistically significant (size: $\beta = -0.004$, $p = 0.20$, lightness: $\beta = -0.005$, $p = 0.09$, numerosity: $\beta = 0.006$, $p = 0.22$).

Finally, we estimated the relative contributions of the two putative sources of influence on context-dependent categorizations using our common model for the two types of tasks. Fig. 4.10 shows the results of the model fitting exercise. The three values plotted on each panel depict the average estimates for the difference in parameter values for the internal representation of the stimulus (diamond dot:

4. Context-dependent categorization

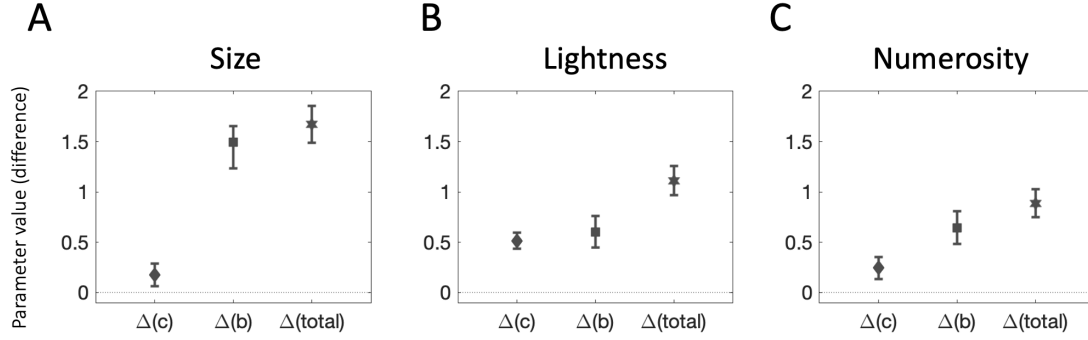


Figure 4.10: Model results. Each panel depicts the difference in parameters between the high and low context conditions. The diamond-shaped dot corresponds to the difference between the parameters for the representation of the stimulus ($\Delta(c)$). The square dot corresponds to the difference between the parameters for the representation of the category boundary ($\Delta(b)$). The hexagonal dot corresponds to the total difference between parameters in the two conditions ($\Delta(total)$). Error bars track SEM. **A:** Size task. **B:** Lightness task. **C:** Numerosity task.

$\Delta(c) = c^h - c^l$, the difference in parameter values for the internal representation of the category boundary (square dot: $\Delta(b) = b^h - b^l$), and their sum (hexagonal dot: $\Delta(total) = \Delta(c) + \Delta(b)$), which tracks the contextual difference in categorization behavior between the two contexts.

Generally, model estimates are consistent with the results reported above. As can be seen from the plots in Fig. 4.10, for size and numerosity, the total contextual difference in categorization judgments ($\Delta(total)$) was driven primarily by changes to the representation of the category boundary ($\Delta(b)$). On average, differences in the representation of the category boundary explained 89% and 73% of the difference in categorization judgments for size and numerosity stimuli, respectively. For lightness, changes in the representation of the category boundary ($\Delta(b)$) contributed approximately 54% of the difference in categorization judgments, suggesting that adaptation accounts for about half of the contextual bias in lightness categorizations.

4.3.3 Interim Discussion

Experiment 2 also provided evidence for the robust influence of temporal context on categorization judgments. The effects held across all three properties we tested. When the context was dominated by larger, lighter and more numerous

4. Context-dependent categorization

stimuli, category boundaries were pushed towards higher levels of size, lightness and numerosity. We found evidence that adaptation contributes to the observed context dependencies. On average, participants exhibited changes in sensitivity; however, those changes were modest in magnitude and differed between stimulus types. For lightness judgments, approximately half of the contextual bias can be captured by adaptation. For size and numerosity judgments, adaptation may account for 11% and 28% of the contextual bias. This pattern of results is consistent with the extensive literature on visual adaptation to light levels (Reuter 2011; Barlow 1972), and the sparser literature on adaptation to size (Blakemore and Sutton 1969; Zimmermann and Fink 2016) and higher level “cognitive adaptation” processes (Palumbo et al. 2017). These results suggest that an adaptive encoding process adjusting the perception of a to-be-categorized stimulus contributes to the observed context dependencies in categorization.

4.4 Discussion

The temporal context in which a stimulus is situated affects how we choose to categorize it. Across two experiments, we replicate the robust contextual biases in categorization documented in the literature. When the temporal context is dominated by larger, lighter and more numerous stimuli, participants are more likely to categorize an objectively neutral stimulus as small, dark and low in number. We probed the computational origins of this bias.

The results of Experiment 1 demonstrated that context-dependent categorizations cannot be explained by a response bias driving an equal distribution of actions across contexts. Both experiments provide evidence that adaptation contributes to context dependencies in categorization decisions. Experiment 1 shows that contextual judgments are influenced by the visual environment beyond immediately task relevant inputs, consistent with the predictions of visual adaptation. Experiment 2 provides evidence that sensitivity to changes in stimulus characteristics shifts in a direction predicted by a normalization-based adaptive encoding process.

4. *Context-dependent categorization*

The magnitude of the observed adaptation varies – it is strongest for lightness, accounting for about half of the contextual bias in categorization decisions, and more modest in magnitude but significant for size and numerosity.

This adaptation process, however, does not lead to stronger contextual biases when each stimulus is experienced for 1000ms versus 200ms. This result speaks to the rich literature on the time courses of adaptation. Visual adaptation, for instance, has been documented to occur on timescales ranging between milliseconds to years over the course of an individual’s lifetime (Webster 2015). The dynamics of adaptation in the mammalian visual system vary with the past history of stimulation; the discriminability of changes to the stimulus characteristics as well as their statistical properties can also change the observed timescale of adaptation (measured over the course of seconds and minutes, Wark et al. 2009). It is likely that all of these factors contribute to the adaptation effects we observe here. In Experiment 1, we manipulate context on the level of an experimental block (lasting ~5 minutes), yet each stimulus is presented on the screen for a much briefer duration ($\leq$ a second). Crucially, stimulus duration also varied on the block level, allowing the visual system to adjust to the timings of incoming stimuli and update the timescale of adaptation.

We found that shorter stimulus presentation led to *more* context-dependent performance for size judgments. This finding is in line with a Bayesian hypothesis (Olkkonen et al. 2014), where reducing the duration of the stimulus increases task demands – the participant has to remember the to-be-categorized stimulus as well as the category boundary in order to respond. This would lead participants to rely more heavily on the prior and so the contextual bias would be stronger, as in our size results. In Experiment 2, we estimated that contextual changes to the category boundary, akin to a Bayesian drift towards the stimulus characteristics dominating the temporal context, account for approximately the half of the context dependence observed in lightness judgments and the majority of the contextual influence in size and numerosity decisions.

Somewhat counter-intuitively, this Bayesian process produces an anti-Bayesian pattern of behavior. Typically, we might expect Bayesian influences on behavior to

4. *Context-dependent categorization*

lead to judgments which are positively associated with the features present in the temporal context. In fact, this is exactly what happens in the commonly documented phenomenon of serial dependence, whereby choices concerning a typically noisy stimulus are attracted towards (or positively influenced by) the properties of the preceding stimulus (Fischer and Whitney 2014; Kiyonaga et al. 2017). This effect makes sense in our temporally autocorrelated sensory (and conceptual) environment. Input on time t is typically strongly correlated with input on time $t - 1$. For instance, visual scenes tend to be relatively stable over time, so if there is uncertainty about the identity of an input, previous stimulation would serve as a useful cue as to its identity. In the laboratory, however, trials are randomized and carefully counterbalanced, and so, any such adaptation effect would lead to suboptimal choice behavior biased towards an objectively unrelated and irrelevant stimulus on the preceding trial.

Here, we see the opposite pattern. In our experiments, categorization decisions are repulsed away from the properties of the stimuli on preceding trials (Fig. 4.7) as well as from the properties dominating the temporal context more generally (Fig. 4.6). This discrepancy in results is driven by a difference in task demands. In tasks where participants exhibit serial dependence, there is typically some uncertainty concerning the imperative stimulus – e.g. it has been presented prior to a short delay and so participants need to rely on a memory of the sensory experience, rather than on the sensory experience directly (Hollingworth 1910; Kiyonaga et al. 2017). To resolve that uncertainty, the neural system uses information from the temporal context like, for instance, the identity of the immediately preceding stimulus or some prior distribution of stimulus properties shaped by recent experience. The task we consider here is different. Participants need to categorize (or label) a stimulus that is currently available to them. Thus, there is no uncertainty about the imperative stimulus, the uncertainty lies in the category boundary – i.e. what should be the criterion for categorization? Thus, the prior distribution is incorporated in the standard *against* which participants judge the stimulus, and so, the observed effects on behavior are repulsive.

4. Context-dependent categorization

The reach of these context dependencies extends far beyond the laboratory. This is because we base our actions on how we evaluate or categorize a given stimulus or situation. If our judgments are affected by information in the temporal context, then so too will our actions exhibit robust context dependencies. According to behavioral ecology, context dependence in judgments relating to sustenance maximizes an organism's chance for survival (Stephens and Krebs 1986; Kolling et al. 2012). When deciding whether to stick to the current resource or forage for something better, it only makes sense to expend valuable energy on foraging if it is likely that better resources would be available in the environment. Thus, our definition of an acceptable banana should change depending on context. If the available bananas are predominantly unripe, we might happily eat a banana that is a bit on the green side. Conversely, a spotted, overripe banana might make the cut in a context dominated by rotting, brown bananas.

Context, however, impacts our judgments far beyond these evolutionarily relevant realms where context dependence may serve an adaptive, fitness-maximizing purpose. For instance, an ethics review committee should maintain stable standards for the ethicality of research proposals. Yet, we know that when student volunteers are faced with this task, their decisions to accept or reject a research proposal are significantly impacted by the levels of ethicality currently dominating the context (Levari et al. 2018). In the realm of political science, the context dependence of human judgment is encapsulated in the Overton window, a term coined in the 1990s by political commentator Joseph Overton. It refers to the range of ideas that are acceptable to the public, and it can shift and expand with the evolution of public opinion. Recently, analyses have highlighted how manipulating what information dominates the public sphere can be used to alter what the general public considers acceptable (Greenhill 2018; Bobric 2021). Research into the mechanisms underlying the context dependence of human judgments, such as our study described here, bears implications for this line of work. Cognitive neuroscience can help inform the social sciences about when and how context affects behavior and what can be done to counteract it.

4.5 Conclusion

Context consistently sways categorization decisions across perception, cognition and economics. This chapter examined the computational mechanisms underlying this effect. We found that context dependencies in categorization behavior cannot be explained by a response bias, aiming to maintain the distribution of actions stable across contexts. Instead, our results suggest that context dependencies are driven by a combination of two complementary processes. The first process affects our representation of the category boundary, pushing it towards the central tendency of the context (Bayesian contraction bias). The second process affects our perception of the to-be-categorized stimulus, repelling it away from the properties of temporal context (adaptation). These two processes appear to contribute equally to the context dependence of lightness categorizations; however, the role of adaptation is weaker, yet significant, for judgments of stimulus size and numerosity.

5

Discussion

Contents

5.1	Overview of Results	140
5.2	Interpretation	141
5.3	Zooming Out	142
5.3.1	Context Dependence in the Wild	142
5.3.2	Context Dependence as a Tool	144
5.4	Conclusion	145

The optimality of our actions is constrained by the accuracy of our judgments. Yet, our percepts and evaluations often appear inaccurate. Depending on the context, the exact same alternatives might lead to different choices – seemingly irrelevant contextual information biases our judgments in a consistent and systematic manner. This robust context dependence, thus, poses a puzzle. Why has evolution sculpted the neural system to process information in a manner that reduces the apparent accuracy of our judgments? I argued for the adaptive role of such a mechanism. Adjusting neural responses based on contextual information boosts the efficient encoding and processing of information; contextual information also helps resolve ambiguities in the face of noise and uncertainty.

In this discussion, I will summarize the findings from my doctoral work and highlight their implications for and connections with phenomena outside the scope

5. *Discussion*

of perception and economics, and outside the laboratory. My thesis considered three different instances of context-dependent choice across the realms of perception and economics. Each study combined careful experimental investigation of human behavior with simulations using mathematical models of hypothesized computational mechanisms. This approach allowed me to first, chart the parametric influence of context on human choices, and second, arbitrate between different theoretical accounts of the drivers of the observed context dependencies. On the behavioral level, the studies revealed rich patterns of contextual influence spanning interactive, attractive and repulsive effects. On the algorithmic level, my computational work provided evidence that normalization schemes of neural processing can account for many of the observed contextual effects.

5.1 Overview of Results

Chapter 2 introduced decoy effects in economic multialternative choice. It went beyond existing work in the field, which has traditionally focused on three well established decoy effects (attraction, compromise and similarity), and mapped the strength and directionality of decoy influence across the full two-dimensional attribute space. Computational simulations revealed that the rich structure of decoy influence can be produced by a simple, general principle of relative encoding, where attribute values are repulsed away from the context of rival options. In the chapter, I further explored the space of putative normalization schemes and charted their predicted behavioral signatures in the multialternative choice task. Evidence from the choice behavior of a large cohort of human volunteers converged with normalization theories that compressively transduce inputs, appealing to normalization by the central tendency of context and recurrently overweighing the contribution of the target input to the normalization.

Chapter 3 considered the role of this information processing scheme in the domain of perceptual choices along a single visual dimension – orientation. Evidence from three psychophysical experiments indicates that perceptual distractors exert an

5. Discussion

interactive influence on choices, biasing the accuracy of judgments according to their consistency with target input. This result is in stark contrast with the classic model of distractor influence, which proposes that distractors compete with targets for processing resources and attract choices as a function of the observer’s attention. Instead, I find that focusing attention enhances sensitivity to target stimuli, but does not ameliorate the contextual influence of the distractor. This counter-intuitive finding is consistent with an attention-independent normalization coding scheme, where the gain with which target signals are processed is boosted in consistent contexts, a mechanism which presumably lies outside of voluntary control.

Finally, chapter 4 delved into the computational mechanisms of context-dependent categorization decisions. Through rigorous experimental design, I provide evidence that the observed contextual bias is driven by the combined influence of a normalization-based information processing scheme, continuously adapting to the temporal context of the decision maker, and a Bayesian process, pushing the categorization standards towards the central tendency of context. This finding holds across three distinct stimulus dimensions – lightness, size and numerosity – highlighting the domain-generalty of the mechanisms underlying the observed context dependence, and bridging our results with related work on context dependence in evaluations of propositional statements and subjective value.

5.2 Interpretation

Taken together, the results of the three studies in this thesis contribute a unifying account of context dependencies across perceptual and economic decisions. Across all chapters, normalization emerged as a common theme, offering a compelling computational account of the behavioral results. It is perhaps somewhat counter-intuitive that a common neural mechanism would drive context dependencies across all the scenarios I consider here. They do, after all, concern very different stimulus characteristics, spanning visual appearance and subjective value, and pose distinct task demands, including comparing, categorizing and ordering alternatives. Yet,

5. Discussion

divisive normalization has been identified as a “canonical neural computation”, repeated across brain regions and modalities (Carandini and Heeger 2012). It might, thus, form part of an arsenal of operations that the neural system has evolved to apply across the processing hierarchy for a variety of problems.

The existence of a common computational mechanism underlying relative judgment has long been hypothesized in the psychological literature. Adaptation-level theory (ALT, Helson 1964), for instance, posits that the human brain processes all manner of information in a relative manner. According to ALT, inputs are evaluated against the current adaptation level of the individual, which is shaped by the stimulation present in the temporal and spatial context (residual and background stimuli, respectively, in ALT). ALT strove to provide a unified account of context dependencies across perception, affect, motivation, learning, cognition and interpersonal behavior. Research on contextual effects across these areas has since, however, splintered off into distinct and largely theoretically unrelated research streams, like for instance, work on adaptation and perceptual constancy in perception (Smithson 2005; Webster 2015) and work on the hedonic treadmill in well-being research (Diener et al. 2009). The parallels between the underpinnings of context dependencies in perceptual and economic choices highlighted in this thesis signal that there is merit in considering the relativity of human judgment holistically and renewing the search for a domain-general computational mechanism of contextual influence.

5.3 Zooming Out

5.3.1 Context Dependence in the Wild

The instances of context dependence I considered here differ in their definitions of context. Chapters 2 and 3 considered the influence of information present in the spatial context of a decision. In chapter 3, this information was completely irrelevant to the choice problem; the contextual input constituted a distractor stimulus which the decision maker was asked to ignore. In chapter 2, this information formed

5. Discussion

part of the choice set; the decision maker was asked to evaluate the contextual (decoy) input and assign a preference rank to it. Chapter 4, on the other hand, considered temporal context. The contextual input consisted of stimuli on previous trials. The task demands in the laboratory rendered context “irrelevant” for the evaluation of the imperative stimulus across all three of these definitions. Yet, outside the artificial laboratory setting, where stimuli and trials are independently and randomly sampled, the irrelevance of context is not so clear cut.

Context may provide useful cues for the interpretation of information. The Kuleshov effect, for instance, refers to a phenomenon whereby the same (ambiguous) facial expression can be interpreted as showing different and potentially contradictory emotions depending on the context in which it is encountered. Popularized by filmmaker Lev Kuleshov over a century ago, the effect emerged when the same clip of an actor was stitched together with oppositely valenced scenes, a scene of a funeral or a scene of a child playing. In the former case, the actor’s emotional state was interpreted as melancholy, in the latter – as happiness (Mobbs et al. 2006). It is easy to see why using context as a cue in this situation would be compelling. In daily life, the emotions of the people around us do usually covary with the context in which they are situated. It is primarily in artificial settings, like the psychology laboratory, that facial expressions may be divorced from their surroundings and need to be interpreted following *unrelated* emotionally charged clips (Mobbs et al. 2006) or photographs (Mullennix et al. 2019).

Context is, thus, useful to make sense of the world. In fact, its role in helping us interpret inputs is so crucial that it features prominently in attempts to construct artificial brains. Context dependencies are inbuilt in various state-of-the-art artificial neural networks. Convolutional networks (CNNs) incorporate information from the spatial context of inputs to analyze and classify images. The base operations of CNNs largely mirror the structure of the visual system, with convolutional and pooling layers serving an analogous computational role as simple and complex cells (Lindsay 2021). Similarly, recurrent networks (RNNs) incorporate information from the temporal context of inputs to interpret information sequences. Analogous processes,

5. Discussion

where signals are dynamically fed back to processing units, are commonly found in biological brains (Goulas et al. 2021). Context dependence is, hence, ubiquitous and crucial across both biological and artificial information processing operations.

5.3.2 Context Dependence as a Tool

Reliance on context is a compelling strategy in the real world, but it can also lead to inconsistent and inaccurate behavioral outcomes. The same processes that I describe in this thesis – decoy (Wu and Cosguner 2020) and repulsive “contrast” effects (Levin 2002) – are often applied by marketing professionals to influence consumer choices. Artificially constructing choice contexts that can tip consumer preferences towards a desired alternative is a lucrative avenue for private companies; recent estimates show that decoy effects can increase sales by as much as 15% (Wu and Cosguner 2020). There is, thus, a clear financial incentive to understand and take advantage of contextual influences on decisions.

Beyond purchasing choices, unintended contextual biases can arise in various social settings, like for instance, the selection of candidates in hiring and admissions decisions (Highhouse 1996; Norton et al. 2004; Simonsohn and Gino 2013). Raising the stakes does not seem to eliminate the influence of context. Context dependencies sway even the (arguably) most consequential decisions in society concerning policy choices (Herne 1997) and political leadership (Chang et al. 2019). Archival electoral data and experimental surveys demonstrate that voting choices in American electoral races systematically violate the normative principle of independence of irrelevant alternatives (Sue O’Curry and Pitts 1995; Chang et al. 2019).

Over the past decade, a movement to harness these robust context dependencies “for good” has garnered momentum (Thaler and Sunstein 2008). The central idea of the movement is that policy makers can “nudge” individuals towards better decisions about, for instance, their retirement savings and health care, by designing choice contexts which bias people towards the desired outcomes. This approach is often called libertarian paternalism: manipulating the decision context (e.g. how choice alternatives are ordered or presented) would not affect decision makers’

5. Discussion

freedom to choose. It would, however, ensure that if the decision makers were to behave in a context-dependent manner, they would choose the intended option determined by a presumably benevolent entity.

Translating research on choice context dependencies into nudges has been the focus of governmental nudge units, such as the UK’s Behavioural Insights Team and (previously) the USA’s Social and Behavioral Sciences Team. Economically, nudging approaches are estimated to be several times more cost-effective than traditional public policy interventions, such as tax incentives and other financial inducements (Benartzi et al. 2017). In the domain of retirement savings, for instance, nudging has been implemented to boost enrollment rates, prompt individuals to choose more appropriate retirement plans and increase their savings contributions (Benartzi and Thaler 2007; Iyengar and Kamenica 2010). More recently, there has been an increased interest in applying contextual choice biases to prompt individual action in response to the climate crisis (Schubert 2017; Zaneva and Dumbalska 2020; Carlsson et al. 2021).

The success of green nudges and other similar choice interventions, however, is constrained by our understanding of context dependence in human choice. Despite their ever growing popularity, little is actually known about why and how nudging interventions work (Marteau et al. 2020). Rigorous methodological evaluations, which are largely missing from the field (Marchiori et al. 2017; Van Kleef and van Trijp 2018), and controlled experiments, like the ones presented in this thesis, are necessary to develop a theoretical understanding of the observed behaviors and mechanisms underlying them. Building that understanding is key for the optimization of choice architecture interventions and the maximization of their impact.

5.4 Conclusion

Why do people do what they do? The behavioral sciences aim to address this big question. The study of decision making is crucial for this endeavor, since our actions

5. Discussion

are, generally, guided by our decisions. It is particularly important to understand the decision process in those cases when our actions appear inconsistent or inaccurate. Oftentimes, inconsistencies can be traced back to reliance on irrelevant contextual information. Uncovering why context matters for our decisions is important not only as a theoretical venture (e.g. what are the neural mechanisms?, are humans rational?, etc.), but also has important translational implications. Context dependence can be wielded as a tool to protect our best interests, as is the goal of the nudge movement. It may, however, also be applied against those interests. It is, thus, crucial to understand and be aware of manipulations to the choice context.

This thesis contributes to a body of work that aims to further our collective understanding of precisely those processes. My doctoral work compiled evidence on context-dependent choice across the domains of perception and economics. Through careful experimental manipulations, I mapped the behavioral signatures of contextual effects. I combined this data with computer simulations to shed light on the information processing mechanisms underlying context-dependent choice. Taken together, these tools can help uncover the factors and processes driving context dependencies in decision making, bringing us one small step closer to answering the big question.

References

- Akaishi, Rei et al. (Jan. 2014). “Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia”. In: *Neuron* 81.1, pp. 195–206. DOI: [10.1016/j.neuron.2013.10.018](https://doi.org/10.1016/j.neuron.2013.10.018).
- Allred, Sarah R. et al. (Dec. 2016). “Working memory and spatial judgments: Cognitive load increases the central tendency bias”. In: *Psychonomic Bulletin & Review* 23.6, pp. 1825–1831. DOI: [10.3758/s13423-016-1039-0](https://doi.org/10.3758/s13423-016-1039-0).
- Anderson, Britt (2011). “There is no Such Thing as Attention”. In: *Frontiers in Psychology* 2. DOI: [10.3389/fpsyg.2011.00246](https://doi.org/10.3389/fpsyg.2011.00246).
- Ariely, Dan, George Loewenstein, and Drazen Prelec (2003). ““Coherent Arbitrariness”: Stable Demand Curves Without Stable Preferences”. In: *QUARTERLY JOURNAL OF ECONOMICS*, p. 33.
- Ashourian, Paymon and Yonatan Loewenstein (May 12, 2011). “Bayesian Inference Underlies the Contraction Bias in Delayed Comparison Tasks”. In: *PLoS ONE* 6.5. Ed. by Adrian G. Dyer, e19551. DOI: [10.1371/journal.pone.0019551](https://doi.org/10.1371/journal.pone.0019551).
- Bar-Hillel, Maya and Willem A Wagenaar (1991). “The perception of randomness”. In: *Advances in applied mathematics* 12.4. Publisher: Elsevier, pp. 428–454.
- Barbosa, Joao and Albert Compte (Dec. 2020). “Build-up of serial dependence in color working memory”. In: *Scientific Reports* 10.1, p. 10959. DOI: [10.1038/s41598-020-67861-2](https://doi.org/10.1038/s41598-020-67861-2).
- Barbot, Antoine, Valentin Wyart, and Marisa Carrasco (2014). “Spatial and feature-based attention differentially affect the gain and tuning of orientation-selective filters”. In: *Journal of Vision* 14.10. Publisher: The Association for Research in Vision and Ophthalmology, pp. 703–703.
- Barlow, Horace B et al. (1961). “Possible principles underlying the transformation of sensory messages”. In: *Sensory communication* 1.1.
- Barlow, Horace B (1972). “Dark and light adaptation: psychophysics”. In: *Visual psychophysics*. Springer, pp. 1–28.
- Benartzi, Shlomo and Richard H Thaler (2007). “Heuristics and Biases in Retirement Savings Behavior”. In: *Journal of Public Economics*, p. 47.
- Benartzi, Shlomo et al. (2017). “Should Governments Invest More in Nudging?” In: p. 15.
- Benda, Jan (2021). “Neural adaptation”. In: *Current Biology* 31.3. Publisher: Elsevier, R110–R116.
- Berkowitsch, N.A.J., B. Scheibehenne, and J. Rieskamp (2014). “Rigorously testing multialternative decision field theory against random utility models”. In: *Journal of Experimental Psychology: General* 143.3, pp. 1331–1348. DOI: [10.1037/a0035159](https://doi.org/10.1037/a0035159).
- Bhatia, S. (2013). “Associations and the accumulation of preference.” In: *Psychological review* 120.3, pp. 522–543. DOI: [10.1037/a0032457](https://doi.org/10.1037/a0032457).
- Bhui, R. and S.J. Gershman (2018). “Decision by sampling implements efficient coding of psychoeconomic functions”. In: *Psychological Review* 125.6, pp. 985–1001. DOI: [10.1037/rev0000123](https://doi.org/10.1037/rev0000123).

References

- Blakemore, Colin, Roger H. S. Carpenter, and Mark A. Georgeson (Oct. 1970). “Lateral Inhibition between Orientation Detectors in the Human Visual System”. In: *Nature* 228.5266, pp. 37–39. DOI: [10.1038/228037a0](https://doi.org/10.1038/228037a0).
- Blakemore, Colin and Peter Sutton (1969). “Size adaptation: A new aftereffect”. In: *Science* 166.3902. Publisher: American Association for the Advancement of Science, pp. 245–247.
- Block, H.D. and J. Marschak (1960). “Random orderings and stochastic theories of responses”. In: *Contributions to Probability and Statistics*, pp. 97–132.
- Bobric, George-Daniel (2021). “The Overton Window: A Tool for Information Warfare”. In: *ICCWS 2021 16th International Conference on Cyber Warfare and Security*. Academic Conferences Limited, p. 20.
- Bonnet, Claude (1990). “Psychophysical approaches, contextual effects and response bias”. In: *Advances in psychology*. Vol. 68. Elsevier, pp. 221–242.
- Brainard, David H (1997). “The psychophysics toolbox”. In: *Spatial vision* 10.4, pp. 433–436.
- Bundesen, C. (Oct. 1990). “A theory of visual attention”. In: *Psychol Rev* 97.4, pp. 523–47.
- Busmeyer, J. R. and J. T. Townsend (July 1993). “Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment”. In: *Psychol Rev* 100.3, pp. 432–59.
- Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein (Feb. 1, 2021). “A Model of Relative Thinking”. In: *The Review of Economic Studies* 88.1, pp. 162–191. DOI: [10.1093/restud/rdaa055](https://doi.org/10.1093/restud/rdaa055).
- Cao, Yanshuai and Luyu Wang (Aug. 10, 2017). “Automatic Selection of t-SNE Perplexity”. In: *arXiv:1708.03229 [cs, stat]*. arXiv: [1708.03229](https://arxiv.org/abs/1708.03229).
- Carandini, M. and D. J. Heeger (Jan. 2012). “Normalization as a canonical neural computation”. In: *Nature Reviews Neuroscience* 13.1, pp. 51–62. DOI: [10.1038/nrn3136](https://doi.org/10.1038/nrn3136).
- Carlsson, Fredrik et al. (2021). “The Use of Green Nudges as an Environmental Policy Instrument”. In: *Review of Environmental Economics and Policy* 15.2. Publisher: The University of Chicago Press Chicago, IL, pp. 216–237.
- Carrasco, M. (July 1, 2011). “Visual attention: the past 25 years”. In: *Vision Res* 51.13, pp. 1484–525. DOI: [10.1016/j.visres.2011.04.012](https://doi.org/10.1016/j.visres.2011.04.012).
- Chang, Linda W, Samuel J Gershman, and Mina Cikara (2019). “Comparing value coding models of context-dependence in social choice”. In: *Journal of Experimental Social Psychology* 85. Publisher: Elsevier, p. 103847.
- Cheadle, S. et al. (Mar. 2014). “Adaptive Gain Control during Human Perceptual Choice”. In: *Neuron* 81.6, pp. 1429–1441. DOI: [10.1016/j.neuron.2014.01.020](https://doi.org/10.1016/j.neuron.2014.01.020).
- Cheadle, S. et al. (Oct. 1, 2015). “Feature expectation heightens visual sensitivity during fine orientation discrimination”. In: *J Vis* 15.14, p. 14. DOI: [10.1167/15.14.14](https://doi.org/10.1167/15.14.14).
- Chelazzi, Leonardo et al. (Oct. 2019). “Getting rid of visual distractors: the why, when, how, and where”. In: *Current Opinion in Psychology* 29, pp. 135–147. DOI: [10.1016/j.copsyc.2019.02.004](https://doi.org/10.1016/j.copsyc.2019.02.004).
- Chi, Micheline TH, Paul J Feltovich, and Robert Glaser (1981). “Categorization and representation of physics problems by experts and novices”. In: *Cognitive science* 5.2. Publisher: Elsevier, pp. 121–152.
- Chubb, Charles, George Sperling, and Joshua A Solomon (1989). “Texture interactions determine perceived contrast”. In: *Proc. Natl. Acad. Sci. USA*, p. 6.

References

- Cialdini, Robert B (1987). *Influence*. Vol. 3. A. Michel Port Harcourt.
- Cox, Karin M. and Joseph W. Kable (Dec. 3, 2014). “BOLD Subjective Value Signals Exhibit Robust Range Adaptation”. In: *The Journal of Neuroscience* 34.49, pp. 16533–16543. DOI: [10.1523/JNEUROSCI.3927-14.2014](https://doi.org/10.1523/JNEUROSCI.3927-14.2014).
- Davies, Nicholas B, John R Krebs, and Stuart A West (2012). *An introduction to behavioural ecology*. John Wiley & Sons.
- Daviet, R. (2018). “Methods for statistical analysis and prediction of discrete choices”. In: *PhD/Doctoral dissertation*.
- Daviet, R. and R. Webb (2021). “A Double Decoy Experiment to Distinguish Theories of Dominance Effects”. In: *A double decoy experiment to distinguish theories of dominance effects*.
- Dayan, P. and R. Zemel (1999). “Statistical models and sensory attention”. In: London: International Conference on Artificial Neural Networks, pp. 1017–1022.
- Desimone, Robert and John Duncan (1995). “Neural mechanisms of selective visual attention”. In: *Annual review of neuroscience* 18.1. Publisher: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, pp. 193–222.
- Diener, Ed, Richard E Lucas, and Christie Napa Scollon (2009). “Beyond the hedonic treadmill: Revising the adaptation theory of well-being”. In: *The science of well-being*. Springer, pp. 103–118.
- Dutka, Jacques (1988). “On the st. petersburg paradox”. In: *Archive for History of Exact Sciences* 39.1. Publisher: Springer, pp. 13–39.
- Eckstein, M. P. et al. (June 2009). “Statistical decision theory to relate neurons to behavior in the study of covert visual attention”. In: *Vision Res* 49.10, pp. 1097–128. DOI: [10.1016/j.visres.2008.12.008](https://doi.org/10.1016/j.visres.2008.12.008).
- Egner, T. and J. Hirsch (Dec. 2005). “Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information”. In: *Nat Neurosci* 8.12, pp. 1784–90. DOI: [10.1038/nn1594](https://doi.org/10.1038/nn1594).
- Enroth-Cugell, Christina and R. M. Shapley (Sept. 1, 1973). “Adaptation and dynamics of cat retinal ganglion cells”. In: *The Journal of Physiology* 233.2, pp. 271–309. DOI: [10.1113/jphysiol.1973.sp010308](https://doi.org/10.1113/jphysiol.1973.sp010308).
- Eriksen, B. A. and C. W. Eriksen (Jan. 1974). “Effects of noise letters upon the identification of a target letter in a nonsearch task”. In: *Perception & Psychophysics* 16.1, pp. 143–149. DOI: [10.3758/BF03203267](https://doi.org/10.3758/BF03203267).
- Fischer, Jason and David Whitney (May 2014). “Serial dependence in visual perception”. In: *Nature Neuroscience* 17.5, pp. 738–743. DOI: [10.1038/nn.3689](https://doi.org/10.1038/nn.3689).
- Foster, David H. (Apr. 2011). “Color constancy”. In: *Vision Research* 51.7, pp. 674–700. DOI: [10.1016/j.visres.2010.09.006](https://doi.org/10.1016/j.visres.2010.09.006).
- Frederick, S., L. Lee, and E. Baskin (2014). “The limits of attraction”. In: *Journal of Marketing Research* 51.4, pp. 487–507. DOI: [10.1509/jmr.12.0061](https://doi.org/10.1509/jmr.12.0061).
- Freeman, Elliot and Jon Driver (May 2005). “Task-dependent modulation of target-flanker lateral interactions in vision”. In: *Perception & Psychophysics* 67.4, pp. 624–637. DOI: [10.3758/BF03193519](https://doi.org/10.3758/BF03193519).
- Freeman, Elliot, Dov Sagi, and Jon Driver (Oct. 2001). “Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers”. In: *Nature Neuroscience* 4.10, pp. 1032–1036. DOI: [10.1038/nn728](https://doi.org/10.1038/nn728).
- Fritsche, Matthias, Eelke Spaak, and Floris P De Lange (2020). “A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception”. In: *Elife* 9. Publisher: eLife Sciences Publications Limited, e55389.

References

- Furmanski, Christopher S. and Stephen A. Engel (June 2000). “An oblique effect in human primary visual cortex”. In: *Nature Neuroscience* 3.6, pp. 535–536. DOI: [10.1038/75702](https://doi.org/10.1038/75702).
- Furnham, Adrian and Hua Chu Boo (2011). “A literature review of the anchoring effect”. In: *The Journal of Socio-Economics* 40.1, pp. 35–42. DOI: <https://doi.org/10.1016/j.socec.2010.10.008>.
- Geng, Joy J. (Apr. 2014). “Attentional Mechanisms of Distractor Suppression”. In: *Current Directions in Psychological Science* 23.2, pp. 147–153. DOI: [10.1177/0963721414525780](https://doi.org/10.1177/0963721414525780).
- Gilbert, C. D. and W. Li (May 2013). “Top-down influences on visual processing”. In: *Nat Rev Neurosci* 14.5, pp. 350–63. DOI: [10.1038/nrn3476](https://doi.org/10.1038/nrn3476).
- Gilbert, Charles et al. (June 2000). “Interactions between attention, context and learning in primary visual cortex”. In: *Vision Research* 40.10, pp. 1217–1226. DOI: [10.1016/S0042-6989\(99\)00234-5](https://doi.org/10.1016/S0042-6989(99)00234-5).
- Girshick, Ahna R, Michael S Landy, and Eero P Simoncelli (July 2011). “Cardinal rules: visual orientation perception reflects knowledge of environmental statistics”. In: *Nature Neuroscience* 14.7, pp. 926–932. DOI: [10.1038/nn.2831](https://doi.org/10.1038/nn.2831).
- Gluth, Sebastian et al. (June 2020). “Value-based attention but not divisive normalization influences decisions with multiple alternatives”. In: *Nature Human Behaviour* 4.6, pp. 634–645. DOI: [10.1038/s41562-020-0822-0](https://doi.org/10.1038/s41562-020-0822-0).
- Gopnik, Alison and Andrew Meltzoff (1987). “The development of categorization in the second year and its relation to other cognitive and linguistic developments”. In: *Child development*. Publisher: JSTOR, pp. 1523–1531.
- Goulas, Alexandros, Fabrizio Damicelli, and Claus C Hilgetag (2021). “Bio-instantiated recurrent neural networks: Integrating neurobiology-based network topology in artificial networks”. In: *Neural Networks* 142. Publisher: Elsevier, pp. 608–618.
- Green, David Marvin and John A Swets (1966). *Signal detection theory and psychophysics*. Vol. 1. Wiley New York.
- Greenhill, Kelly M (2018). “How Trump Manipulates the Migration Debate”. In: *Foreign Affairs*.
- Hartzmark, Samuel M. and Kelly Shue (Aug. 2018). “A Tough Act to Follow: Contrast Effects in Financial Markets: A Tough Act to Follow”. In: *The Journal of Finance* 73.4, pp. 1567–1613. DOI: [10.1111/jofi.12685](https://doi.org/10.1111/jofi.12685).
- Hawkins, H. L. et al. (Nov. 1990). “Visual attention modulates signal detectability”. In: *J Exp Psychol Hum Percept Perform* 16.4, pp. 802–11.
- Helson, Harry (Jan. 1947). “Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data”. In: *The American Journal of Psychology* 60.1, p. 1. DOI: [10.2307/1417326](https://doi.org/10.2307/1417326).
- (1964). “Adaptation-level theory: an experimental and systematic approach to behavior.” In: Publisher: New York.
- Heng, Joseph A, Michael Woodford, and Rafael Polania (Sept. 15, 2020). “Efficient sampling and noisy decisions”. In: *eLife* 9, e54962. DOI: [10.7554/eLife.54962](https://doi.org/10.7554/eLife.54962).
- Herne, Kaisa (1997). “Decoy alternatives in policy choices: Asymmetric domination and compromise effects”. In: *European Journal of Political Economy* 13.3. Publisher: Elsevier, pp. 575–589.
- Highhouse, Scott (1996). “Context-dependent selection: The effects of decoy and phantom job candidates”. In: *Organizational Behavior and Human Decision Processes* 65.1. Publisher: Elsevier, pp. 68–76.

References

- Hollingworth, H. L. (1910). “The Central Tendency of Judgment”. In: p. 10.
- Hotaling, J.M., J.R. Busemeyer, and J. Li (2010). “Theoretical Developments in Decision Field Theory: Comment on Tsetsos, Usher, and Chater (2010)”. In: *Psychological Review* 117.4, pp. 1294–1298. DOI: [10.1037/a0020401](https://doi.org/10.1037/a0020401).
- Huber, J., J. W. Payne, and C. P. Puto (June 1982). “Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis”. In: *Journal of Consumer Research* 9.1, p. 90. DOI: [10.1086/208899](https://doi.org/10.1086/208899).
- Huber, Joel, John W Payne, and Christopher P Puto (2014). “Let’s be honest about the attraction effect”. In: *Journal of Marketing Research* 51.4. Publisher: SAGE Publications Sage CA: Los Angeles, CA, pp. 520–525.
- Ito, Minami and Charles Gilbert (Mar. 1999). “Attention Modulates Contextual Influences in the Primary Visual Cortex of Alert Monkeys”. In: *Neuron* 22.3, pp. 593–604. DOI: [10.1016/S0896-6273\(00\)80713-8](https://doi.org/10.1016/S0896-6273(00)80713-8).
- Iyengar, Sheena S. and Emir Kamenica (2010). “Choice proliferation, simplicity seeking, and asset allocation”. In: *Journal of Public Economics* 94.7, pp. 530–539. DOI: <https://doi.org/10.1016/j.jpubeco.2010.03.006>.
- Izard, Véronique and Stanislas Dehaene (2008). “Calibrating the mental number line”. In: *Cognition* 106.3. Publisher: Elsevier, pp. 1221–1247.
- Jazayeri, Mehrdad and Michael N Shadlen (Aug. 2010). “Temporal context calibrates interval timing”. In: *Nature Neuroscience* 13.8, pp. 1020–1026. DOI: [10.1038/nn.2590](https://doi.org/10.1038/nn.2590).
- Josiam, Bharath M and JS Perry Hobson (1995). “Consumer choice in context: the decoy effect in travel and tourism”. In: *Journal of Travel Research* 34.1. Publisher: Sage Publications Sage CA: Thousand Oaks, CA, pp. 45–50.
- Jou, Jerwen et al. (2004). “Contraction Bias in Memorial Quantifying Judgment: Does It Come from a Stable Compressed Memory Representation or a Dynamic Adaptation Process?” In: *The American Journal of Psychology* 117.4, p. 543. DOI: [10.2307/4148991](https://doi.org/10.2307/4148991).
- Juechems, Keno et al. (2021). “Optimal utility and probability functions for agents with finite computational precision”. In: *Proceedings of the National Academy of Sciences* 118.2. Publisher: National Acad Sciences.
- Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1991). “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias”. In: *Journal of Economic Perspectives*, p. 128.
- Kahneman, Daniel and Amos Tversky (Mar. 1979). “Prospect Theory: An Analysis of Decision under Risk”. In: *Econometrica* 47.2, p. 263. DOI: [10.2307/1914185](https://doi.org/10.2307/1914185).
- Kapadia, Mitesh K. et al. (Oct. 1995). “Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys”. In: *Neuron* 15.4, pp. 843–856. DOI: [10.1016/0896-6273\(95\)90175-2](https://doi.org/10.1016/0896-6273(95)90175-2).
- Kiyonaga, Anastasia et al. (July 2017). “Serial Dependence across Perception, Attention, and Memory”. In: *Trends in Cognitive Sciences* 21.7, pp. 493–497. DOI: [10.1016/j.tics.2017.04.011](https://doi.org/10.1016/j.tics.2017.04.011).
- Knoblauch, Kenneth, Laurence T Maloney, et al. (2008). “MLDS: Maximum likelihood difference scaling in R”. In: *Journal of Statistical Software* 25.2, pp. 1–26.
- Kolling, Nils et al. (Apr. 6, 2012). “Neural Mechanisms of Foraging”. In: *Science* 336.6077, pp. 95–98. DOI: [10.1126/science.1216930](https://doi.org/10.1126/science.1216930).

References

- Krajbich, I., C. Armel, and A. Rangel (Oct. 2010). “Visual fixations and the computation and comparison of value in simple choice”. In: *Nat Neurosci* 13.10, pp. 1292–8. DOI: [10.1038/nn.2635](https://doi.org/10.1038/nn.2635).
- Krause, Matthew R. and Christopher C. Pack (Nov. 2014). “Contextual modulation and stimulus selectivity in extrastriate cortex”. In: *Vision Research* 104, pp. 36–46. DOI: [10.1016/j.visres.2014.10.006](https://doi.org/10.1016/j.visres.2014.10.006).
- Lages, Martin and Michel Treisman (Feb. 1998). “Spatial frequency discrimination: Visual long-term memory or criterion setting?” In: *Vision Research* 38.4, pp. 557–572. DOI: [10.1016/S0042-6989\(97\)88333-2](https://doi.org/10.1016/S0042-6989(97)88333-2).
- (2010). “A Criterion Setting Theory of Discrimination Learning that Accounts for Anisotropies and Context Effects”. In: *Seeing and Perceiving* 23.5, pp. 401–434. DOI: [10.1163/187847510X541117](https://doi.org/10.1163/187847510X541117).
- Land, Edwin H (1983). “Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image.” In: *Proceedings of the national academy of sciences of the United States of America* 80.16. Publisher: National Academy of Sciences, p. 5163.
- Landry, Peter and Ryan Webb (Apr. 2021). “Pairwise normalization: A neuroeconomic theory of multi-attribute choice”. In: *Journal of Economic Theory* 193, p. 105221. DOI: [10.1016/j.jet.2021.105221](https://doi.org/10.1016/j.jet.2021.105221).
- Latty, T. and M. Beekman (2011). “Irrational decision-making in an amoeboid organism: Transitivity and context-dependent preferences”. In: *Proceedings of the Royal Society B: Biological Sciences* 278.1703, pp. 307–312. DOI: [10.1098/rspb.2010.1045](https://doi.org/10.1098/rspb.2010.1045).
- Lavie, Nilli and Yehoshua Tsal (Mar. 1994). “Perceptual load as a major determinant of the locus of selection in visual attention”. In: *Perception & Psychophysics* 56.2, pp. 183–197. DOI: [10.3758/BF03213897](https://doi.org/10.3758/BF03213897).
- Lea, A.M. and M.J. Ryan (2015). “Irrationality in mate choice revealed by túngara frogs”. In: *Science* 349.6251, pp. 964–966. DOI: [10.1126/science.aab2012](https://doi.org/10.1126/science.aab2012).
- Lee, Chang-Yuan and Carey K Morewedge (2021). “Noise Increases Anchoring Effects”. In: p. 55.
- Lee, Michelle P. and Kwanho Suk (Feb. 2010). “Disambiguating the Role of Ambiguity in Perceptual Assimilation and Contrast Effects”. In: *Journal of Consumer Research* 36.5, pp. 890–897. DOI: [10.1086/605299](https://doi.org/10.1086/605299).
- Levari, David E. et al. (June 29, 2018). “Prevalence-induced concept change in human judgment”. In: *Science* 360.6396, pp. 1465–1467. DOI: [10.1126/science.aap8731](https://doi.org/10.1126/science.aap8731).
- Levin, Aron M (2002). “Contrast and assimilation processes in consumers’ evaluations of dual brands”. In: *Journal of Business and Psychology* 17.1. Publisher: Springer, pp. 145–154.
- Li, Baowang, Matthew R. Peterson, and Ralph D. Freeman (July 2003). “Oblique Effect: A Neural Basis in the Visual Cortex”. In: *Journal of Neurophysiology* 90.1, pp. 204–217. DOI: [10.1152/jn.00954.2002](https://doi.org/10.1152/jn.00954.2002).
- Li, Vickie et al. (Sept. 18, 2018). “Gain control explains the effect of distraction in human perceptual, cognitive, and economic decision making”. In: *Proceedings of the National Academy of Sciences* 115.38, E8825–E8834. DOI: [10.1073/pnas.1805224115](https://doi.org/10.1073/pnas.1805224115).
- Liberman, Alina, Jason Fischer, and David Whitney (Nov. 2014). “Serial Dependence in the Perception of Faces”. In: *Current Biology* 24.21, pp. 2569–2574. DOI: [10.1016/j.cub.2014.09.025](https://doi.org/10.1016/j.cub.2014.09.025).
- Lindsay, Grace W (2021). “Convolutional neural networks as a model of the visual system: Past, present, and future”. In: *Journal of cognitive neuroscience* 33.10.

References

- Publisher: MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA
journals-info . . . , pp. 2017–2031.
- Liu, Taosheng (Oct. 2019). “Feature-based attention: effects and control”. In: *Current Opinion in Psychology* 29, pp. 187–192. DOI: [10.1016/j.copsyc.2019.03.013](https://doi.org/10.1016/j.copsyc.2019.03.013).
- Liu, Taosheng, Sean T. Stevens, and Marisa Carrasco (Jan. 2007). “Comparing the time course and efficacy of spatial and feature-based attention”. In: *Vision Research* 47.1, pp. 108–113. DOI: [10.1016/j.visres.2006.09.017](https://doi.org/10.1016/j.visres.2006.09.017).
- Loewenstein, George, Scott Rick, and Jonathan D. Cohen (Jan. 2008). “Neuroeconomics”. In: *Annual Review of Psychology* 59.1, pp. 647–672. DOI: [10.1146/annurev.psych.59.103006.093710](https://doi.org/10.1146/annurev.psych.59.103006.093710).
- Louie, K., M. W. Khaw, and P. W. Glimcher (Apr. 9, 2013). “Normalization is a general neural mechanism for context-dependent decision making”. In: *Proceedings of the National Academy of Sciences* 110.15, pp. 6139–6144. DOI: [10.1073/pnas.1217854110](https://doi.org/10.1073/pnas.1217854110).
- Lu, Zhong-Lin and Barbara Doshier (2013). *Visual psychophysics: From laboratory to theory*. MIT Press.
- Luce, R. Duncan (1959). “On the possible psychophysical laws.” In: *Psychological Review* 66.2, pp. 81–95. DOI: [10.1037/h0043178](https://doi.org/10.1037/h0043178).
- Lynch Jr., John G., Dipankar Chakravarti, and Anusree Mitra (Dec. 1991). “Contrast Effects in Consumer Judgments: Changes in Mental Representations or in the Anchoring of Rating Scales?” In: *Journal of Consumer Research* 18.3, p. 284. DOI: [10.1086/209260](https://doi.org/10.1086/209260).
- Maloney, Laurence T and Joong Nam Yang (2003). “Maximum likelihood difference scaling”. In: *Journal of Vision* 3.8. Publisher: The Association for Research in Vision and Ophthalmology, pp. 5–5.
- Mamassian, Pascal, Michael Landy, and Laurence T Maloney (2002). “Bayesian modelling of visual perception”. In: *Probabilistic models of the brain*, pp. 13–36.
- Marchiori, David R, Marieke A Adriaanse, and Denise TD De Ridder (2017). “Unresolved questions in nudging research: Putting the psychology back in nudging”. In: *Social and Personality Psychology Compass* 11.1. Publisher: Wiley Online Library, e12297.
- Marteau, Theresa M et al. (2020). “14 Changing Behavior by Changing Environments”. In: *The handbook of behavior change*. Publisher: Cambridge University Press, p. 193.
- Mattar, Marcelo G et al. (2018). “Adaptation decorrelates shape representations”. In: *Nature communications* 9.1. Publisher: Nature Publishing Group, pp. 1–9.
- McAdams, C. J. and J. H. Maunsell (Jan. 1, 1999). “Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4”. In: *J Neurosci* 19.1, pp. 431–441.
- McFadden, Daniel (1981). “Econometric models of probabilistic choice”. In: *Structural analysis of discrete data with econometric applications* 198272. Publisher: MIT Press.
- Miller, E. K. and J. D. Cohen (2001). “An integrative theory of prefrontal cortex function”. In: *Annu Rev Neurosci* 24, pp. 167–202. DOI: [10.1146/annurev.neuro.24.1.167](https://doi.org/10.1146/annurev.neuro.24.1.167).
- Mobbs, Dean et al. (2006). “The Kuleshov Effect: the influence of contextual framing on emotional attributions”. In: *Social cognitive and affective neuroscience* 1.2. Publisher: Oxford University Press, pp. 95–106.
- Moran, J and R Desimone (Aug. 23, 1985). “Selective attention gates visual processing in the extrastriate cortex”. In: *Science* 229.4715, pp. 782–784. DOI: [10.1126/science.4023713](https://doi.org/10.1126/science.4023713).

References

- Morgan, M.J, S.N.J Watamaniuk, and S.P McKee (Aug. 2000). “The use of an implicit standard for measuring discrimination thresholds”. In: *Vision Research* 40.17, pp. 2341–2349. DOI: [10.1016/S0042-6989\(00\)00093-6](https://doi.org/10.1016/S0042-6989(00)00093-6).
- Mosteller, Frederick and Philip Nogee (1951). “An Experimental Measurement of Utility”. In: p. 35.
- Mullennix, John, Jeremy Barber, and Trista Cory (2019). “An examination of the Kuleshov effect using still photographs”. In: *Plos one* 14.10. Publisher: Public Library of Science San Francisco, CA USA, e0224623.
- Natenzon, P. (2019). “Random choice and learning”. In: *Journal of Political Economy* 127.1, pp. 419–457. DOI: [10.1086/700762](https://doi.org/10.1086/700762).
- Nieder, Andreas and Stanislas Dehaene (2009). “Representation of number in the brain”. In: *Annual review of neuroscience* 32. Publisher: Annual Reviews, pp. 185–208.
- Noguchi, T. and N. Stewart (2014). “In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions”. In: *Cognition* 132.1, pp. 44–56. DOI: [10.1016/j.cognition.2014.03.006](https://doi.org/10.1016/j.cognition.2014.03.006).
- Norton, Michael I, Joseph A Vandello, and John M Darley (2004). “Casuistry and social category bias.” In: *Journal of personality and social psychology* 87.6. Publisher: American Psychological Association, p. 817.
- Olkkonen, M., P. F. McCarthy, and S. R. Allred (Sept. 5, 2014). “The central tendency bias in color perception: Effects of internal and external noise”. In: *Journal of Vision* 14.11, pp. 5–5. DOI: [10.1167/14.11.5](https://doi.org/10.1167/14.11.5).
- Olkkonen, Maria and Sarah R. Allred (Jan. 27, 2014). “Short-Term Memory Affects Color Perception in Context”. In: *PLoS ONE* 9.1. Ed. by Floris P. de Lange, e86488. DOI: [10.1371/journal.pone.0086488](https://doi.org/10.1371/journal.pone.0086488).
- Padoa-Schioppa, C. (Nov. 4, 2009). “Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex”. In: *Journal of Neuroscience* 29.44, pp. 14004–14014. DOI: [10.1523/JNEUROSCI.3751-09.2009](https://doi.org/10.1523/JNEUROSCI.3751-09.2009).
- Palumbo, Rocco, Stefania D’Ascenzo, and Luca Tommasi (2017). “Editorial: high-level adaptation and aftereffects”. In: *Frontiers in psychology* 8. Publisher: Frontiers, p. 217.
- Papadimitriou, Charalampos, Robert L. White, and Lawrence H. Snyder (Apr. 24, 2016). “Ghosts in the Machine II: Neural Correlates of Memory Interference from the Previous Trial”. In: *Cerebral Cortex*, bhw106. DOI: [10.1093/cercor/bhw106](https://doi.org/10.1093/cercor/bhw106).
- Parducci, Allen (1965). “Category judgment: A range-frequency model.” In: *Psychological Review* 72.6, pp. 407–418. DOI: [10.1037/h0022602](https://doi.org/10.1037/h0022602).
- Parrish, A.E., T.A. Evans, and M.J. Beran (2015). “Rhesus macaques (*Macaca mulatta*) exhibit the decoy effect in a perceptual discrimination task”. In: *Attention, Perception, and Psychophysics* 77.5, pp. 1715–1725. DOI: [10.3758/s13414-015-0885-6](https://doi.org/10.3758/s13414-015-0885-6).
- Peirce, Jonathan W (2007). “PsychoPy—psychophysics software in Python”. In: *Journal of neuroscience methods* 162.1. Publisher: Elsevier, pp. 8–13.
- Piazza, Manuela et al. (2007). “A magnitude code common to numerosities and number symbols in human intraparietal cortex”. In: *Neuron* 53.2. Publisher: Elsevier, pp. 293–305.
- Polania, Rafael, Michael Woodford, and Christian Ruff (June 29, 2018). “Efficient coding of subjective value”. In: DOI: [10.1101/358317](https://doi.org/10.1101/358317).
- Polat, U. and D. Sagi (Feb. 15, 1994). “Spatial interactions in human vision: from near to far via experience-dependent cascades of connections.” In: *Proceedings of the National Academy of Sciences* 91.4, pp. 1206–1209. DOI: [10.1073/pnas.91.4.1206](https://doi.org/10.1073/pnas.91.4.1206).

References

- Polat, Uri and Dov Sagi (May 1993). “Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments”. In: *Vision Research* 33.7, pp. 993–999. DOI: [10.1016/0042-6989\(93\)90081-7](https://doi.org/10.1016/0042-6989(93)90081-7).
- Pollak, Seth D and Doris J Kistler (2002). “Early experience is associated with the development of categorical representations for facial expressions of emotion”. In: *Proceedings of the National Academy of Sciences* 99.13. Publisher: National Acad Sciences, pp. 9072–9076.
- Posner, M. I., C. R. Snyder, and B. J. Davidson (June 1980). “Attention and the detection of signals”. In: *J Exp Psychol* 109.2, pp. 160–74.
- Poulton, E. C. and S. Poulton (1989). *Bias in quantifying judgements*. Taylor & Francis.
- Rangel, Antonio and John A Clithero (Dec. 2012). “Value normalization in decision making: theory and evidence”. In: *Current Opinion in Neurobiology* 22.6, pp. 970–981. DOI: [10.1016/j.conb.2012.07.011](https://doi.org/10.1016/j.conb.2012.07.011).
- Reuter, Tom (2011). “Fifty years of dark adaptation 1961–2011”. In: *Vision research* 51.21. Publisher: Elsevier, pp. 2243–2262.
- Reynolds, J. H. and L. Chelazzi (2004). “Attentional modulation of visual processing”. In: *Annu Rev Neurosci* 27, pp. 611–47. DOI: [10.1146/annurev.neuro.26.041002.131039](https://doi.org/10.1146/annurev.neuro.26.041002.131039).
- Reynolds, J. H. and D. J. Heeger (Jan. 29, 2009). “The normalization model of attention”. In: *Neuron* 61.2, pp. 168–85. DOI: [10.1016/j.neuron.2009.01.002](https://doi.org/10.1016/j.neuron.2009.01.002).
- Rieskamp, Jörg, Jerome R Busemeyer, and Barbara A Mellers (Aug. 1, 2006). “Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice”. In: *Journal of Economic Literature* 44.3, pp. 631–661. DOI: [10.1257/jel.44.3.631](https://doi.org/10.1257/jel.44.3.631).
- Rigoli, Francesco (Nov. 2019). “Reference effects on decision-making elicited by previous rewards”. In: *Cognition* 192, p. 104034. DOI: [10.1016/j.cognition.2019.104034](https://doi.org/10.1016/j.cognition.2019.104034).
- Rigoli, Francesco, Karl J. Friston, and Raymond J. Dolan (Nov. 2016). “Neural processes mediating contextual influences on human choice behaviour”. In: *Nature Communications* 7.1, p. 12416. DOI: [10.1038/ncomms12416](https://doi.org/10.1038/ncomms12416).
- Rigoli, Francesco et al. (2017). “A unifying Bayesian account of contextual effects in value-based choice”. In: *PLOS Computational Biology* 13.10, e1005769.
- Roe, R.M., J.R. Busemeyer, and J.T. Townsend (2001). “Multialternative decision field theory: A dynamic connectionist model of decision making”. In: *Psychological Review* 108.2, pp. 370–392. DOI: [10.1037/0033-295X.108.2.370](https://doi.org/10.1037/0033-295X.108.2.370).
- Rogers, Marie, Kenneth Knoblauch, and Anna Franklin (2016). “Maximum likelihood conjoint measurement of lightness and chroma”. In: *JOSA A* 33.3. Publisher: Optical Society of America, A184–A193.
- Rosenthal, Robert (1976). “Experimenter effects in behavioral research”. In: Publisher: Irvington.
- Schubert, Christian (2017). “Green nudges: Do they work? Are they ethical?” In: *Ecological Economics* 132. Publisher: Elsevier, pp. 329–342.
- Schütt, Heiko et al. (2015). “Psignifit 4: Pain-free Bayesian inference for psychometric functions”. In: *Journal of vision* 15.12. Publisher: The Association for Research in Vision and Ophthalmology, pp. 474–474.
- Shafir, S. (1994). “Intransitivity of preferences in honey bees: Support for ‘comparative’ evaluation of foraging options”. In: *Animal Behaviour* 48.1, pp. 55–67. DOI: [10.1006/anbe.1994.1211](https://doi.org/10.1006/anbe.1994.1211).
- Shafra, Patrick and John D Coley (2003). “Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge.” In:

References

- Journal of Experimental Psychology: Learning, Memory, and Cognition* 29.4. Publisher: American Psychological Association, p. 641.
- Shen, Shan and Wei Ji Ma (Jan. 2019). “Variable precision in visual perception.” In: *Psychological Review* 126.1, pp. 89–132. DOI: [10.1037/rev0000128](https://doi.org/10.1037/rev0000128).
- Simon, Herbert A (1955). “A behavioral model of rational choice”. In: *The quarterly journal of economics* 69.1. Publisher: MIT Press, pp. 99–118.
- Simoncelli, Eero P (Apr. 2003). “Vision and the statistics of the visual environment”. In: *Current Opinion in Neurobiology* 13.2, pp. 144–149. DOI: [10.1016/S0959-4388\(03\)00047-3](https://doi.org/10.1016/S0959-4388(03)00047-3).
- Simonsohn, Uri and Francesca Gino (2013). “Daily horizons: Evidence of narrow bracketing in judgment from 10 years of MBA admissions interviews”. In: *Psychological science* 24.2. Publisher: Sage Publications Sage CA: Los Angeles, CA, pp. 219–224.
- Simonsohn, Uri and George Loewenstein (Jan. 1, 2006). “Mistake: The Effect of Previously Encountered Prices on Current Housing Demand”. In: *The Economic Journal* 116.508, pp. 175–199. DOI: [10.1111/j.1468-0297.2006.01052.x](https://doi.org/10.1111/j.1468-0297.2006.01052.x).
- Simonson, Itamar (Sept. 1989). “Choice Based on Reasons: The Case of Attraction and Compromise Effects”. In: *Journal of Consumer Research* 16.2, p. 158. DOI: [10.1086/209205](https://doi.org/10.1086/209205).
- Sims, Christopher A. (Apr. 2003). “Implications of rational inattention”. In: *Journal of Monetary Economics* 50.3, pp. 665–690. DOI: [10.1016/S0304-3932\(03\)00029-1](https://doi.org/10.1016/S0304-3932(03)00029-1).
- Slawinski, Elzbieta B and Laurie K Fitzgerald (1998). “Perceptual development of the categorization of the/rw/contrast in normal children”. In: *Journal of Phonetics* 26.1. Publisher: Elsevier, pp. 27–43.
- Smithson, H.E (June 29, 2005). “Sensory, computational and cognitive components of human colour constancy”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1458, pp. 1329–1346. DOI: [10.1098/rstb.2005.1633](https://doi.org/10.1098/rstb.2005.1633).
- Soltani, Alireza, Benedetto De Martino, and Colin Camerer (July 19, 2012). “A Range-Normalization Model of Context-Dependent Choice: A New Model and Evidence”. In: *PLoS Computational Biology* 8.7. Ed. by Laurence T. Maloney, e1002607. DOI: [10.1371/journal.pcbi.1002607](https://doi.org/10.1371/journal.pcbi.1002607).
- Spektor, Mikhail S, Sudeep Bhatia, and Sebastian Gluth (2021). “The elusiveness of context effects in decision making”. In: *Trends in Cognitive Sciences*. Publisher: Elsevier.
- St. John-Saaltink, Elexa et al. (June 8, 2016). “Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex”. In: *The Journal of Neuroscience* 36.23, pp. 6186–6192. DOI: [10.1523/JNEUROSCI.4390-15.2016](https://doi.org/10.1523/JNEUROSCI.4390-15.2016).
- Steiner, Jakub and Colin Stewart (2016). “Perceiving prospects properly”. In: *American Economic Review* 106.7, pp. 1601–31.
- Stemmler, M, M Usher, and E Niebur (Sept. 29, 1995). “Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics”. In: *Science* 269.5232, pp. 1877–1880. DOI: [10.1126/science.7569930](https://doi.org/10.1126/science.7569930).
- Stephan, Klaas Enno et al. (July 2009). “Bayesian model selection for group studies”. In: *NeuroImage* 46.4, pp. 1004–1017. DOI: [10.1016/j.neuroimage.2009.03.025](https://doi.org/10.1016/j.neuroimage.2009.03.025).
- Stephens, David W and John R Krebs (1986). *Foraging theory*. Princeton University Press.
- Stewart, Neil, Nick Chater, and Gordon D.A. Brown (Aug. 2006). “Decision by sampling”. In: *Cognitive Psychology* 53.1, pp. 1–26. DOI: [10.1016/j.cogpsych.2005.10.003](https://doi.org/10.1016/j.cogpsych.2005.10.003).

References

- Stockman, A. et al. (Oct. 1, 2006). “Human cone light adaptation: From behavioral measurements to molecular mechanisms”. In: *Journal of Vision* 6.11, pp. 5–5. DOI: [10.1167/6.11.5](https://doi.org/10.1167/6.11.5).
- Störmer, Viola S. and George A. Alvarez (Sept. 2014). “Feature-Based Attention Elicits Surround Suppression in Feature Space”. In: *Current Biology* 24.17, pp. 1985–1988. DOI: [10.1016/j.cub.2014.07.030](https://doi.org/10.1016/j.cub.2014.07.030).
- Sue O’Curry, Yigang Pan and Robert Pitts (1995). “The attraction effect and political choice in two elections”. In: *Journal of Consumer Psychology* 4.1. Publisher: Wiley Online Library, pp. 85–101.
- Summerfield, Christopher and Konstantinos Tsetsos (2012). “Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms”. In: *Frontiers in Neuroscience* 6. DOI: [10.3389/fnins.2012.00070](https://doi.org/10.3389/fnins.2012.00070).
- (Jan. 2015). “Do humans make good decisions?” In: *Trends in Cognitive Sciences* 19.1, pp. 27–34. DOI: [10.1016/j.tics.2014.11.005](https://doi.org/10.1016/j.tics.2014.11.005).
- (2020). “Rationality and Efficiency in Human Decision-Making”. In: *The Cognitive Neurosciences*. Publisher: MIT Press, p. 427.
- Sun, John Z et al. (2012). “A framework for Bayesian optimality of psychophysical laws”. In: *Journal of Mathematical Psychology* 56.6. Publisher: Elsevier, pp. 495–501.
- Sundberg, Kristy A., Jude F. Mitchell, and John H. Reynolds (Mar. 2009). “Spatial Attention Modulates Center-Surround Interactions in Macaque Visual Area V4”. In: *Neuron* 61.6, pp. 952–963. DOI: [10.1016/j.neuron.2009.02.023](https://doi.org/10.1016/j.neuron.2009.02.023).
- Taubert, Jessica, David Alais, and David Burr (Oct. 2016). “Different coding strategies for the perception of stable and changeable facial attributes”. In: *Scientific Reports* 6.1, p. 32239. DOI: [10.1038/srep32239](https://doi.org/10.1038/srep32239).
- Thaler, Richard H and LJ Ganser (2015). “Misbehaving: The making of behavioral economics”. In: Publisher: WW Norton New York.
- Thaler, Richard H and Cass R Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness*.
- Tormala, Zakary L. and Richard E. Petty (Jan. 2007). “Contextual contrast and perceived knowledge: Exploring the implications for persuasion”. In: *Journal of Experimental Social Psychology* 43.1, pp. 17–30. DOI: [10.1016/j.jesp.2005.11.007](https://doi.org/10.1016/j.jesp.2005.11.007).
- Treisman, Michel and Thomas C Williams (1984). “A Theory of Criterion Setting With an Application to Sequential Dependencies”. In: *Psychological review*, p. 44.
- Tremblay, Léon and Wolfram Schultz (Apr. 1999). “Relative reward preference in primate orbitofrontal cortex”. In: *Nature* 398.6729, pp. 704–708. DOI: [10.1038/19525](https://doi.org/10.1038/19525).
- Treue, Stefan and Julio C. Martínez Trujillo (June 1999). “Feature-based attention influences motion processing gain in macaque visual cortex”. In: *Nature* 399.6736, pp. 575–579. DOI: [10.1038/21176](https://doi.org/10.1038/21176).
- Trueblood, J.S. (2012). “Multialternative context effects obtained using an inference task”. In: *Psychonomic Bulletin and Review* 19.5, pp. 962–968. DOI: [10.3758/s13423-012-0288-9](https://doi.org/10.3758/s13423-012-0288-9).
- Trueblood, J.S., S.D. Brown, and A. Heathcote (2014). “The multiattribute linear ballistic accumulator model of context effects in multialternative choice”. In: *Psychological Review* 121.2, pp. 179–205. DOI: [10.1037/a0036137](https://doi.org/10.1037/a0036137).
- (2015). “The fragile nature of contextual preference reversals: Reply to Tsetsos, Chater, and Usher (2015)”. In: *Psychological Review* 122.4, pp. 848–853. DOI: [10.1037/a0039656](https://doi.org/10.1037/a0039656).

References

- Trueblood, J.S. et al. (2013). "Not Just for Consumers: Context Effects Are Fundamental to Decision Making". In: *Psychological Science* 24.6, pp. 901–908. DOI: [10.1177/0956797612464241](https://doi.org/10.1177/0956797612464241).
- Tsetsos, K., M. Usher, and N. Chater (2010). "Preference Reversal in Multiattribute Choice". In: *Psychological Review* 117.4, pp. 1275–1291. DOI: [10.1037/a0020580](https://doi.org/10.1037/a0020580).
- Tsetsos, Konstantinos, N. Chater, and M. Usher (2012). "Salience driven value integration explains decision biases and preference reversal". In: *Proceedings of the National Academy of Sciences of the United States of America* 109.24, pp. 9659–9664. DOI: [10.1073/pnas.1119569109](https://doi.org/10.1073/pnas.1119569109).
- Tsetsos, Konstantinos et al. (Mar. 15, 2016). "Economic irrationality is optimal during noisy decision making". In: *Proceedings of the National Academy of Sciences* 113.11, pp. 3102–3107. DOI: [10.1073/pnas.1519157113](https://doi.org/10.1073/pnas.1519157113).
- Turner, B.M. et al. (2018). "Competing theories of multialternative, multiattribute preferential choice". In: *Psychological Review* 125.3, pp. 329–362. DOI: [10.1037/rev0000089](https://doi.org/10.1037/rev0000089).
- Tversky, A. (1972a). "Elimination by aspects: A theory of choice". In: *Psychological Review* 79.4, pp. 281–299. DOI: [10.1037/h0032955](https://doi.org/10.1037/h0032955).
- Tversky, Amos (1972b). "Choice by elimination". In: *Journal of mathematical psychology* 9.4. Publisher: Elsevier, pp. 341–367.
- Tversky, Amos and Daniel Kahneman (1974). "Judgment under uncertainty: Heuristics and biases". In: *science* 185.4157. Publisher: American association for the advancement of science, pp. 1124–1131.
- Usher, M. and J.L. McClelland (2004). "Loss aversion and inhibition in dynamical models of multialternative choice". In: *Psychological Review* 111.3, pp. 757–769. DOI: [10.1037/0033-295X.111.3.757](https://doi.org/10.1037/0033-295X.111.3.757).
- Van den Berg, Ronald et al. (2017). "Recent is more: A negative time-order effect in nonsymbolic numerical judgment." In: *Journal of Experimental Psychology: Human Perception and Performance* 43.6. Publisher: American Psychological Association, p. 1084.
- Van Der Maaten, L. and G. Hinton (2008). "Visualizing data using t-SNE". In: *Journal of Machine Learning Research* 9, pp. 2579–2625.
- Van Kleef, Ellen and Hans CM van Trijp (2018). "Methodological challenges of research in nudging". In: *Methods in Consumer Research, Volume 1*. Elsevier, pp. 329–349.
- Van Bergen, Ruben S and Janneke FM Jehee (2019). "Probabilistic representation in human visual cortex reflects uncertainty in serial decisions". In: *Journal of Neuroscience* 39.41. Publisher: Soc Neuroscience, pp. 8164–8176.
- Von Neumann, John and Oskar Morgenstern (1944). *Theory of games and economic behavior*. Princeton university press.
- Wang, Xiao-Jing (Dec. 2002). "Probabilistic Decision Making by Slow Reverberation in Cortical Circuits". In: *Neuron* 36.5, pp. 955–968. DOI: [10.1016/S0896-6273\(02\)01092-9](https://doi.org/10.1016/S0896-6273(02)01092-9).
- Wark, Barry, Adrienne Fairhall, and Fred Rieke (2009). "Timescales of inference in visual adaptation". In: *Neuron* 61.5. Publisher: Elsevier, pp. 750–761.
- Webb, R, P. W. Glimcher, and K. Louie (2021). "The normalization of consumer valuations: context-dependent preferences from neurobiological constraints". In: *Manag. Sci.*
- Weber, Max and Louis Schneider (1975). "Marginal utility theory and" The Fundamental Law of Psychophysics". In: *Social Science Quarterly*. Publisher: JSTOR, pp. 21–36.

References

- Webster, M. A. (May 20, 2011). “Adaptation and visual coding”. In: *Journal of Vision* 11.5, pp. 3–3. DOI: [10.1167/11.5.3](https://doi.org/10.1167/11.5.3).
- (Nov. 24, 2015). “Visual Adaptation”. In: *Annual Review of Vision Science* 1.1, pp. 547–567. DOI: [10.1146/annurev-vision-082114-035509](https://doi.org/10.1146/annurev-vision-082114-035509).
- Wegener, Duane T. et al. (Jan. 2001). “Implications of Attitude Change Theories for Numerical Anchoring: Anchor Plausibility and the Limits of Anchor Effectiveness”. In: *Journal of Experimental Social Psychology* 37.1, pp. 62–69. DOI: [10.1006/jesp.2000.1431](https://doi.org/10.1006/jesp.2000.1431).
- Wei, Xue-Xin and Alan A. Stocker (Sept. 19, 2017). “Lawful relation between perceptual bias and discriminability”. In: *Proceedings of the National Academy of Sciences* 114.38, pp. 10244–10249. DOI: [10.1073/pnas.1619153114](https://doi.org/10.1073/pnas.1619153114).
- Welbourne, Lauren E., Antony B. Morland, and Alex R. Wade (Aug. 2015). “Human colour perception changes between seasons”. In: *Current Biology* 25.15, R646–R647. DOI: [10.1016/j.cub.2015.06.030](https://doi.org/10.1016/j.cub.2015.06.030).
- Wiebel, Christiane B, Guillermo Aguilar, and Marianne Maertens (2017). “Maximum likelihood difference scales represent perceptual magnitudes and predict appearance matches”. In: *Journal of vision* 17.4. Publisher: The Association for Research in Vision and Ophthalmology, pp. 1–1.
- Woodford, Michael (2020). “Modeling Imprecision in Perception, Valuation, and Choice”. In: p. 23.
- Wu, Chunhua and Koray Cosguner (2020). “Profiting from the decoy effect: A case study of an online diamond retailer”. In: *Marketing Science* 39.5. Publisher: INFORMS, pp. 974–995.
- Wyart, V., N. E. Myers, and C. Summerfield (Feb. 25, 2015). “Neural Mechanisms of Human Perceptual Choice Under Focused and Divided Attention”. In: *Journal of Neuroscience* 35.8, pp. 3485–3498. DOI: [10.1523/JNEUROSCI.3276-14.2015](https://doi.org/10.1523/JNEUROSCI.3276-14.2015).
- Wyart, V., A. C. Nobre, and C. Summerfield (Feb. 28, 2012). “Dissociable prior influences of signal probability and relevance on visual contrast sensitivity”. In: *Proc Natl Acad Sci U S A* 109.9, pp. 3593–8. DOI: [10.1073/pnas.1120118109](https://doi.org/10.1073/pnas.1120118109).
- Xiao, Kaida et al. (Jan. 2011). “Visual gamma correction for LCD displays”. In: *Displays* 32.1, pp. 17–23. DOI: [10.1016/j.displa.2010.09.003](https://doi.org/10.1016/j.displa.2010.09.003).
- Xing, J. and D. J. Heeger (Mar. 2001). “Measurement and modeling of center-surround suppression and enhancement”. In: *Vision Research* 41.5, pp. 571–583. DOI: [10.1016/S0042-6989\(00\)00270-4](https://doi.org/10.1016/S0042-6989(00)00270-4).
- Yang, S. and M. Lynn (2014). “More evidence challenging the robustness and usefulness of the attraction effect”. In: *Journal of Marketing Research* 51.4, pp. 508–513. DOI: [10.1509/jmr.14.0020](https://doi.org/10.1509/jmr.14.0020).
- Yuille, Alan and Daniel Kersten (July 2006). “Vision as Bayesian inference: analysis by synthesis?” In: *Trends in Cognitive Sciences* 10.7, pp. 301–308. DOI: [10.1016/j.tics.2006.05.002](https://doi.org/10.1016/j.tics.2006.05.002).
- Yung, Amanda et al. (2015). “Methods to test visual attention online”. In: *JoVE (Journal of Visualized Experiments)* 96, e52470.
- Zaneva, Mirela and Tsvetomira Dumbalska (2020). “Green nudges: Applying behavioural economics to the fight against climate change”. In: *PsyPAG Quarterly goes electronic*, p. 27.
- Zimmermann, Eckart and Gereon R Fink (2016). “Numerosity perception after size adaptation”. In: *Scientific Reports* 6.1. Publisher: Nature Publishing Group, pp. 1–7.