Range Effects in Multi-Attribute Choice: An Experiment*

Tommaso Bondi†, Dániel Csaba‡, and Evan Friedman§

December 10, 2019

Abstract

Several behavioral theories suggest that, when choosing between multi-attribute goods, choices are context-dependent. Two theories provide such predictions explicitly in terms of attribute ranges. According to the theory of Focusing (Kőszegi and Szeidl [2012]), attributes with larger ranges receive more attention. On the other hand, Relative thinking (Bushong et al. [2015]) posits that fixed differences look smaller when the range is large. It is as if attributes with larger ranges are over- and under-weighted, respectively. Since the two theories make opposing predictions, it is important to understand which features of the environment affect their relative prevalence. We conduct an experiment designed to test for both of these opposing range effects in different environments. Using choice under risk, we use a two-by-two design defined by high or low stakes and high or low dimensionality (as measured by the number of attributes). In the aggregate, we find evidence of focusing in low-dimensional treatments. Classifying subjects into focusers and relative thinkers, we find that focusers are associated with quicker response times and that types are more stable when the stakes are high.

*We thank seminar audiences at both NYU and Columbia, as well as the SWEET Conference 2018 participants, for insightful comments. We are grateful to Hassan Afrouzi, Ala Avoyan, Andrew Caplin, Alessandra Casella, Mark Dean, Guillaume Fréchette, Alessandro Lizzeri and Andrew Schotter for useful suggestions, and to the Columbia Experimental Laboratory in the Social Sciences for financial support. All errors are our own.

†Department of Economics, New York University, Stern School of Business. Email: tbondi@stern.nyu.edu

‡Department of Economics, New York University. Email: daniel.csaba@nyu.edu

§Department of Economics, Columbia University. Email: ekf2119@columbia.edu

1 Introduction

The majority of our choices involve multiple attributes and non-trivial trade-offs. Is the cheaper car considerably less fuel efficient? Is the better paid job less enjoyable? Standard economic theory assumes that, at least when agents can access the relevant information for all the options, they are able to solve these trade-offs optimally. In particular, determining the relative value of two options only requires computing the difference in hedonic utility they generate, and does not depend on the other available options. A wealth of empirical evidence, both in laboratory and real world settings, has questioned this assumption: agents' preferences over the two options often depends on the framing of the choice problem as well as on the presence of other options, even when these are not chosen. We discuss this literature in greater detail in Section 2.

Let us go back to one of the initial examples: a difference of \$3,000 in two cars' price tags might capture our attention much more than the future monthly \$50 difference in gasoline expenses, thus leading us to buy the cheaper, less fuel-efficient car. On the other hand, imagine that the price difference between the cheapest and most expensive cars in the choice set is \$20,000. In this case, the \$3,000 difference between the two cars we are most carefully considering may seem rather small, leading us to go for the more expensive one.

The example shows a particular type of context-dependence in choice: range effects. These are not new to this paper. In recent influential work, Kőszegi and Szeidl [2012] (henceforth KS) and Bushong et al. [2015] (henceforth BRS) present theoretical models formalizing these two intuitions. Both messages are intuitive: KS argue that some dimensions, by including options that are far apart, naturally stand out to the decision maker, and hence receive disproportionate weight in the decision. For instance, someone deciding between Los Angeles and Chicago might disproportionately base his decision on the weather, and forget about the cost of living, quality of public schools, jobs offered, and so forth, choosing Los Angeles even when not normatively optimal. BRS, on the other hand, offer a very different perspective: whenever the alternatives vastly differ in one dimension, differences of fixed size in that dimension will appear smaller, or, in other terms, they will be under-weighted. For instance, the non-negligible differences in weather between New York and Chicago might appear negligible when the decision maker's choice set includes Los Angeles, and Chicago might end up chosen over New York even when not optimal.

Both theories have gained traction in recent years for their ability to accommodate common choice reversal patterns. We refer to KS and BRS for detailed discussions, and we mention the main findings in Section 3. However, a paradox remains in that the two always make exactly opposite predictions. That is, choice reversals predicted by focusing are never predicted by relative thinking, and vice versa. In this paper, we take advantage of the two papers' unified framework to test for both. We do so using an experimental study to better understand *if*, and *when*, each of these two theories has descriptive power in incentivized, multi-attribute choice.

Our experiment takes place in the lab and involves choice under risk. Since both theories seem empirically plausible, despite offering opposite predictions, we put particular emphasis on designing environments in which one or the other is more likely to appear: in particular, we vary both the stakes and the dimensionality (as measured by the number of attributes) of the lotteries faced by subjects.

Our two-by-two design is motivated by theoretical considerations. Varying stakes is important because it changes the cost-benefit analysis for the agents: when stakes are higher, the benefits from a correct decision increase, but—for a fixed level of dimensionality—costs of thinking remain fixed. Therefore, the disappearance of these phenomena in higher stakes domains might suggest a rational inattention type explanation for them: concentration is costly, and the effort is only worth it when the stakes are high enough. Dimensionality is key—as explicitly mentioned in Bushong et al. [2015]—in understanding *why* these biases might arise. BRS posit that focusing on larger spreads might be justified in environments where alternatives have many attributes, and evaluating each dimension is too costly, or even unfeasible; relative thinking, on the other hand, reflects a widespread perception phenomenon and thus also applies to simpler decisions. Our design allows us to test for this claim explicitly. It also sheds light on how opposing attention and perception patterns can arise as a function of the dimensionality of attributes. In this sense, we speak to the growing debate on (in)attention being either a rational response to costly attentional frictions, or an irrational one reflecting mostly psychological distortions in information processing, as discussed in detail in Handel and Schwartzstein [2018].

A key feature of the experiment is in isolating range effects by controlling for a variety of well known behavioral biases that have received extensive documentation in the literature. In particular, it is easy to see that in two dimensional choices, altering ranges by adding a non-dominating option implies introducing a decoy for one of the existing options. We therefore move away from this scenario and only present subjects with three (or more) dimensional questions, in which we have the ability to avoid decoy effects. This is in contrast with existing studies, in which spurious range effects might be found when in reality decoy effects were driving the results. Second, and equally important, we take the possibility of noisy decision making seriously: subjects might change their answer to the same question even without changing any of the features they are presented with (Agranov and Ortoleva [2017]). Adding a third option might influence subjects in a variety of ways, the majority of which have nothing to do with range effects. We therefore always present subjects with placebo questions in which the added third option does not introduce changes in attribute ranges. We then compare the number of reversals following range changes of different types relative to this case.

The aforementioned considerations are particularly important in light of our findings. We obtain a more nuanced picture than that given by previous studies. At the aggregate level, we find a statistically significant focusing effect in low-dimensional environments, but not in high-dimensional ones. This sharply contrast with the aforementioned conjecture by BRS, who hypothesize a focusing effect can arise as a best response behavior in environments in which evaluating each attribute (or choice) is unfeasible. We find no evidence for relative thinking in aggregate data: the number of reversals that we attribute to relative thinking is not statistically different from that we attribute to noise.

It is important, particularly so in light of the opposite nature of the two phenomena, to also analyze behavior at the individual level. We therefore classify subjects as focusers or relative thinkers whenever they are considerably more likely to show preference reversals predicted by each theory, respectively. We find some evidence that subjects' types are stable across the first and second half of the experiment, suggesting that subject characteristics may be as important as environmental ones to predict the relative likelihood of each bias.

Lastly, what drives, or at least correlates with, this individual heterogeneity? To answer this question, we complement our choice data with detailed demographics measures, as well as response times. We find that focusers indeed answer more quickly, suggesting that the small weight given to some of the information we provide might reflect inattention or impulsivity. Relative thinking appears virtually uncorrelated with response times.

The rest of the paper is organized as follows: Section 2 describes the existing research on the topic, both theoretical and experimental. Section 3 briefly reviews the two models, together with their main implications. Sections 4 and 5 detail our design and results, respectively. We conclude in Section 6.

2 Literature

Broadly speaking, this study relates to the vast literature in economics and psychology dealing with context-dependent choice. For a detailed review of some of the classic papers in this literature, see Tversky and Simonson [1993]; more updated, though less general reviews are contained in Bordalo et al. [2012a], Kőszegi and Szeidl [2012] and Bushong et al. [2015]. Here, we focus on the theories of Focusing and Relative Thinking, as well as the existing evidence related to them (experimental or otherwise).

The conceptually closest theory is Salience Theory [Bordalo et al., 2012a,b, 2013]. Like the two theories we study, Salience Theory introduces a context-dependent, dimension specific decision weight that distorts subjects' computation of utility. Despite the different functional forms, Salience shares the main qualitative predictions of Focusing, in that it posits that some elements of the environment stand out due to larger ranges, and uses this fact to parsimoniously explain several paradoxes in choice under risk. However, Salience also features diminishing sensitivity, which closely resembles Relative Thinking. For this reason, our experiment does not allow us to draw straightforward inference about Salience.

Closer to Relative Thinking is Range Normalization (Soltani et al. [2012]). The authors investigate mechanisms underlying decoy effects; one key finding is that the strength of the decoy is *increasing* in the distance between the decoy and the original option (that is, decreasing in the quality of the decoy). This suggests a boundedly rational explanation for perception biases, somewhat different from what BRS posit. As previously mentioned, our design purposefully gets rid of decoy effects: we do so exactly because it is not clear how to interpret their relationship with range effects. In this sense, we want to eliminate it as an alternative explanation, limiting our conclusions on the issue studied by Soltani et al. [2012]. Closely related work by Landry and Webb [2019] also attempts to bridge the gap between a neural model of (pairwise) attribute normalization and a range of phenomena observed in the empirical literature, including the asymmetric dominance effect, the compromise effect, the similarity effect, "majority rule" transitivity violations, sub-additive attribute weighting, and comparability biases.

To best of our knowledge, the most closely related work to ours is Somerville [2019].[1]. He also proposes a joint test for focusing and relative thinking, which relies on experimentally varying the prices of high- and low-quality variants of multiple products. He finds "clear evidence of choice-set dependence: price increases that expand the range of prices in the choice set lead to more purchases, and increasing prices while holding price differences fixed induces substitution to higher-quality options". In contrast, we document how these behaviors might arise in different environments, and zoom in on their subject-level correlates.

Moreover, there have been some recent papers trying to test each of the two theories separately in recent years in a variety of domains. Following the chronological order of the two original papers, we start with experimental tests of focusing. Andersson et al. [2016] conduct an experiment in which subjects have to choose between monetary rewards. Crucially, in order to make the problem multidimensional, each sum is to be paid over multiple dates. Subjects are presented with histograms showing, for each date, the different amounts paid by each option. The authors find significant evidence for focusing. Our paper differs from theirs significantly both in our experimental set up and results. Importantly, we present everything in numerical form, to get around the concern that visual perception might play a role. Perhaps as a result of the several differences, our results also considerably differ. We hypothesize visual priming might play an important role in how options are evaluated; this might be relevant to some choice problems, but not others (the same, of course, applies to our design).

Another paper, Dertwinkel-Kalt et al. [2016a], investigates focusing effects in intertemporal choice. The authors find strong evidence that subjects disproportionately prefer future payments when paid in a single large sum compared to several, smaller payments over consecutive (but not later) periods. Fixed costs might play an important role in driving these results: subjects with no focusing bias would still prefer a single, larger if slightly later payment whenever the act of collecting the payment itself is costly (for instance, because they have to walk back to the experimental lab, or keep memory of the dates, and so forth). Part of our experiment could be seen as the risk equivalent of theirs; however, we fail to see such strong preference for single peaked lotteries that only pay well in one

[1] A related but distinct exercise is performed in Hirshman et al. [2018], in which relative thinking is contrasted with mental budgeting.

state.

We also mention two experimental tests of Salience, given its similarity to Focusing. Dertwinkel-Kalt et al. [2016b] and Bordalo et al. [2012a] test distinctive implications of Salience Theory in consumer choice and risk domain respectively, finding strong evidence.

Relative thinking—at least in the theoretical formulation of BRS—has received, to the best of our knowledge, virtually no experimental testing. In particular, the question of whether focusing or relative thinking effects arise in different kinds of incentivized decisions remains open. However, the concept of relative thinking is not new. Two seminal articles introduced the idea of relative thinking to economists: Thaler [1985] and Tversky and Kahneman [1985]. Thaler [1985] conjectures that people exert more effort to save \$5 on a \$25 radio than to save \$5 on a \$500 TV. He introduces what he calls *transaction utility*: fixed percentage savings on baseline prices generate utility that does not (fully) scale with the stakes involved. Similarly, Tversky and Kahneman [1985] report that subjects are willing to drive more to save a fixed amount of money when the baseline price is lower. In similar spirit, failures of fungibility of money—which causes small changes in payoffs to appear much larger than they are—have been documented in the high-stakes, high-experience context of choosing between gasoline and other consumer goods by Hastings and Shapiro [2013].

Ofer Azar conducts multiple experimental tests of different forms of relative thinking. He finds that *i*) relative thinking is present in hypothetical choices (Azar [2011c]), *ii*) but disappears with incentives (Azar [2011a]) and *iii*) relative and not only absolute prices seem to matter: subjects demand greater absolute changes in price to switch from a superior product to the other as both products' baseline prices go up (Azar [2011b])[2]. Lastly, it is worth mentioning that at least some sellers and marketers seem to be aware of consumer relative thinking, and have developed pricing models and frames to exploit it; such strategies are explored in Cunningham [2011].

3 The Two Theories

Following the common notation in Kőszegi and Szeidl [2012] and Bushong et al. [2015], we assume the decision maker is choosing from n options, each of which has k attributes or dimensions (terms we

[2] Saini et al. [2010] offers similar evidence.

use interchangeably). Hedonic utility is separable, and given, for a generic option c^i with attributes $c^i_1, \dots c^i_k$, by $U(c^i) = \sum_{j=1}^k u_j(c^i_j)$. Both theories depart from expected utility in assuming that the decision marker is context-dependent, and thus maximizes the weighted utility,

$$U(c^i) = U(c^i_1, \dots, c^i_k) = \sum_{j=1}^{k} w(R_j) u_j(c^i_j), \quad R_j = \max_{h=1,\dots,n} u_j(c^h_j) - \min_{h=1,\dots,n} u_j(c^h_j)$$

The key object of interest is the **attention weight**, $\mathbf{w}(\mathbf{R})$. According to standard utility, $\mathbf{w}'(\mathbf{R}) = \mathbf{0}$. According to *focusing*, $\mathbf{w}'(\mathbf{R}) > \mathbf{0}$; that is dimensions with greater spread (in utility terms) receive greater attention. *Relative thinking* makes exactly the opposite prediction: $\mathbf{w}'(\mathbf{R}) < \mathbf{0}$: differences of fixed size will look smaller if in dimensions with greater spreads—while still satisfying the intuitive "more is better" condition that $Rw(R)$ is increasing in R.

Clearly, changes in $w(\cdot)$ can potentially affect the chosen option. Given the fact that focusing and relative thinking posit opposite functional forms for $w(R)$, it is reasonable to conjecture that they never predict the same deviation from expected utility. This is indeed the case: denoting by EU, F and RT expected utility, focusing and relative thinking respectively, we have the following:

Lemma 1. *If $c^i \in C^F(c^1, \dots, c^n)$ but $c^i \notin C^{EU}(c^1, \dots, c^n)$, then $c^i \notin C^{RT}(c^1, \dots, c^n)$. Conversely, if $c^i \in C^{RT}(c^1, \dots, c^n)$ but $c^i \notin C^{EU}(c^1, \dots, c^n)$, then $c^i \notin C^F(c^1, \dots, c^n)$.*

This is because if $c^i \in C^F(c^1, \dots, c^n)$ but $c^i \notin C^{EU}(c^1, \dots, c^n)$, this means that c^i has comparatively high value in dimensions with large ranges, and comparatively low value in dimensions with small ranges. Moreover, if all dimensions were equally weighted, c^i would not be chosen (since $c^i \notin C^{EU}(c^1, \dots, c^n)$). This is *a fortiori* true when the weight of dimensions in which c^i does not rank near the top goes up, and that of dimensions in which it does goes down. Hence, the option will not be chosen under relative thinking.

This fact is important for our experiment, since *i*) we can design additional options to nudge subjects towards either bias, and *ii*) when observing reversals, we don't have to worry about the two theories being simultaneously implied.

To get intuition for the two theories, consider a risk neutral, perfectly patient agent choosing between a \$20 payment to be received today, and two \$10 payments to be received in one and two days respectively. Whereas standard utility theory predicts indifference, focusers will choose \$20 today since $w(20) \cdot 20 > 2 \cdot w(10) \cdot 10$ if $w'(\cdot) > 0$: the large difference in today's payoffs stands out more than the two smaller ones in the next two days. Relative thinkers will make the opposite choice.

4 Experimental Design

Overview. In each question, subjects are asked to choose lotteries from a set of two or three lotteries. In describing the questions, we write, for example, $\{A, B, C\}$ to indicate a choice from 3 lotteries labelled A, B, and C.

The experiment consists of two blocks. Each is similar in structure, and differs only in specific details of the lottery choices, so we focus our discussion on Block 1. Each question in Block 1 falls into one of four treatments of a 2×2 design. Each treatment is defined by whether the lottery choices are low or high *stakes* (s/S) and low or high *dimension* (d/D). Within each treatment, there are 4 *main* questions: $\{A, B\}$, presenting only two "baseline" lotteries, $\{A, B, R_A\}$, which expands the range in a single dimension in which A is better than B, $\{A, B, R_B\}$, which does the same for B, and $\{A, B, C\}$, which introduces a *range neutral* third option C. In addition, there is a 5th question, $\{A_f, B_f, C_f\}$, which we call the *FOSD* question (for "first-order stochastic dominance") for reasons that will become clear. This gives 5 questions within each treatment, and thus 20 questions within the block.

The following lists all of the Block 1 lotteries organized into treatments.

Block 1: main

Low Stakes (s), Low Dimension (d)

	35%	45%	20%
A	10	7	4
B	5	11	4
R_A	1	8	6
R_B	8	2	7
C	6	8	5

Low Stakes (s), High Dimension (D)

	25%	35%	10%	15%	15%
A	10	7	4	5	3
B	5	11	4	2	6
R_A	1	8	6	3	4
R_B	8	2	7	3	5
C	6	8	5	3	4

High Stakes (S), Low Dimension (d)

	35%	45%	20%
A	20	14	8
B	10	22	8
R_A	2	16	12
R_B	16	4	14
C	12	16	10

High Stakes (S), High Dimension (D)

	25%	35%	10%	15%	15%
A	20	14	8	10	6
B	10	22	8	4	12
R_A	2	16	12	6	8
R_B	16	4	14	6	10
C	12	16	10	6	8

Block 1: FOSD

Low Stakes (s), Low Dimension (d)

	35%	45%	20%
A_f	9	2	5
B_f	5	9	2
C_f	2	4	9

Low Stakes (s), High Dimension (D)

	25%	35%	10%	15%	15%
A_f	9	2	9	4	5
B_f	5	9	2	9	2
C_f	2	4	4	2	9

High Stakes (S), Low Dimension (d)

	35%	45%	20%
A_f	18	4	10
B_f	10	18	4
C_f	4	8	18

High Stakes (S), High Dimension (D)

	25%	35%	10%	15%	15%
A_f	18	4	18	8	10
B_f	10	18	4	18	4
C_f	4	8	8	4	18

Block 2 (displayed in detail in the Appendix) has the same basic structure as Block 1, so there are 40 questions in all.

Lotteries and range effects. We explain the features of the lotteries within each treatment and the rationale for how the questions are constructed to identify range effects.

Lotteries A and B are constructed to be similarly attractive to subjects. B has a slightly higher expected value than A but is also arguably riskier. The choice $\{A, B\}$ is meant to determine each subject's "baseline" choice. When R_A is introduced, i.e. subjects are given the choice $\{A, B, R_A\}$, the range of prizes in the first state in which A is better than B is increased. Similarly, in the choice $\{A, B, R_B\}$, the range of the prizes in the second state in which B is better than A is increased. For a subject who chooses A at baseline, i.e. $C(\{A, B\}) = A$, there are two types of "choice reversal". $C(\{A, B, R_B\}) = B$ indicates *focusing*, and $C(\{A, B, R_A\}) = B$ indicates *relative thinking*. Similarly, if $C(\{A, B\}) = B$, $C(\{A, B, R_A\}) = A$ indicates focusing and $C(\{A, B, R_B\}) = A$ indicates relative thinking. All other patterns in choice are inconclusive with respect to these theories.

Since the dimensionality of the choice problem increases with the addition of a third option, we expect to observe some reversals with the introduction of any third option due to decision noise. Hence, we also include the choice problem $\{A, B, C\}$ in which C is range-neutral. If, in the aggregate there are more choice reversals indicating range effects than choice reversals with the introduction of C, we interpret this as evidence of range effects. $\{A, B, C\}$ is thus a control to estimate the magnitude of switches due to decision noise.

Since all choices of R_A, R_B, or C are inconclusive, these lotteries are meant to be unattractive relative to both A and B. While none of these lotteries is dominated by A or B, they have comparatively low expected values.

In addition to the main choices, we also include an FOSD question $\{A_f, B_f, C_f\}$ within each treatment. B_f first-order stochastically dominates the others, and hence is the objective best choice for any expected utility maximizer. The choices are constructed to not be obvious, in that B_f does not dominate the others state-by-state. There is one FOSD question within each treatment, which mirrors the features of the treatment in that it can be either high or low stakes and high or low dimension. The FOSD questions are included to see if subjects are paying attention and to see if the prevalence of mistakes depends on the stakes or dimension. They also break up the appearance of the main questions, reducing the probability subjects will simply repeat their past choices.

Treatments and blocks. Taking as reference the low stakes-low dimension treatment of Block 1, the other treatments in Block 1 are constructed from this one. The low stakes-high dimension treatment "mixes" in two additional states. That is, each low stakes-high dimension lottery is approximately the reduction of a compound lottery in which the corresponding low dimension lottery is chosen with some probability and another two state lottery is chosen with complementary probability. The lotteries in the high stakes treatments–both low and high dimension–are simply the corresponding low stakes lotteries in which all the prizes are multiplied by 2.

Block 2 is constructed from Block 1 by *translating* all of the lotteries. The low stakes lotteries in Block 2 are the same as the corresponding ones in Block 1, except $1 is added to all the prizes. Similarly, the high stakes lotteries in Block 2 are the same as the corresponding ones in Block 1, except $2 is added to all the prizes. Translation is a range-neutral operation, so we think of the corresponding choice problems in Block 1 and Block 2 as being essentially the same. The translation is merely to obfuscate this fact for subjects.

Randomization and presentation to subjects. All questions appear in random order within the block, so subjects will see questions interspersed from different treatments. The lotteries themselves appear on a grid, with both the rows and columns in random order. This is to control for whatever tendency subjects may have to always choose the first lottery, say, and also to obfuscate the fact that subjects see the same lottery several times throughout the experiment. For half of the sessions, what we call Block 2 appears first.

In all cases, as in the screenshot below, lotteries presented to subjects have neutral labels: A and B if there are two lotteries and A, B, and C if there are three lotteries, regardless of which underlying lotteries they refer to.

Please select your preferred lottery among the presented options.

	20%	**35%**	**45%**
A	8	20	14
B	14	16	4
C	8	10	22

A **B** **C**

Incentives and Questionnaire. Everything so far described takes place in Section 1. There is also a short questionnaire in Section 2 in which we (1) elicit risk attitudes via a multiple price list (as in Holt and Laury [2002]), (2) ask subjects to make some expected value type calculations to gauge ability and willingness to make such calculations, and (3) collect basic demographic data.

Subjects are paid a \$5 show-up fee, and in addition, are paid one randomly selected question from Section 1 and a randomly selected row from the multiple price list in Section 2. Total earnings range from \$6.10 to \$32.85[3], and the average is \$18.

4.1 Identification of Range Effects

Our task is to aggregate the data into measures of focusing and relative thinking. Following our previous discussion, the focusing measure estimates the probability of choice reversals that indicate focusing *minus* the probability of choice reversals due to the introduction of the range-neutral option, and similarly for relative thinking. As will become clear, the data can be consistent with one of the two effects, both, or neither.

Our identification strategy is best explained through example, so we consider the data from the *sd* treatment in Table 1 (see the Appendix for analogous tables for the other treatments). This gives the choice frequencies from the 4 main questions, which is further broken into blocks and "baseline pick"–the choice from $\{A, B\}$. Since choosing the third option (R_A, R_B, or C) is not counted toward a range effect or a control switch, we first drop all subjects within a treatment who made such a choice. This is necessary because otherwise there could be more opportunities to display a range effect than a control switch or vice versa, which would invalidate the measure[4].

sd	Block 1				Block 2			
Baseline pick→	A		B		A		B	
	A	B	A	B	A	B	A	B
Baseline: $\{A, B\}$	49	0	0	42	47	0	0	44
Range-A: $\{A, B, R_A\}$	31	18	14	28	29	18	15	29
Range-B: $\{A, B, R_B\}$	28	21	15	27	25	22	10	34
Control: $\{A, B, C\}$	24	25	13	29	33	14	11	33

Table 1: *Choice frequencies from sd.* This table reports the choice frequencies from the *sd* treatment, broken into blocks and "baseline pick', after dropping subjects who chose the third option (R_A, R_B, or C) anywhere in the treatment.

[3] All of the prizes in our lotteries are multiples of \$1. The 0.10 and 0.85 come from Holt and Laury [2002].

[4] At one extreme, if every subject chose C, then there would be no control switches and effectively no control group.

First, let's calculate excess focusing from Block 1. We have 49 subjects choosing A and 42 choosing B from the baseline set $\{A, B\}$. R_A increases the range in which A is better than B while R_B increases the range where B is better than A. Hence, from the group that chose A at baseline, the 21 subjects who chose B from $\{A, B, R_B\}$ exhibit focusing, while from the group that chose B at baseline, the 14 subjects who chose A from $\{A, B, R_A\}$ exhibit focusing. The numbers of subjects who exhibit similar switching behavior in the control set $\{A, B, C\}$ from their baseline choices are 25 and 13 respectively. The excess focusing propensities for those who picked A and B at baseline respectively are:

$$\underbrace{\left(\frac{21}{49} - \frac{25}{49}\right)}_{\text{baseline } A \text{, Block 1}}, \qquad \underbrace{\left(\frac{14}{42} - \frac{13}{42}\right)}_{\text{baseline } B \text{, Block 1}}.$$

Similarly, we calculate the measures for Block 2:

$$\underbrace{\left(\frac{22}{47} - \frac{14}{47}\right)}_{\text{baseline } A \text{, Block 2}}, \qquad \underbrace{\left(\frac{15}{44} - \frac{11}{44}\right)}_{\text{baseline } B \text{, Block 2}}.$$

In order to get an overall measure, we take a weighted average with weights based on the number of subjects within each group:

$$\frac{49}{49+42+47+44}\underbrace{\left(\frac{21}{49} - \frac{25}{49}\right)}_{\text{baseline } A \text{, Block 1}} + \frac{42}{49+42+47+44}\underbrace{\left(\frac{14}{42} - \frac{13}{42}\right)}_{\text{baseline } B \text{, Block 2}} +$$
$$\frac{47}{49+42+47+44}\underbrace{\left(\frac{22}{47} - \frac{14}{47}\right)}_{\text{baseline } A \text{, Block 2}} + \frac{44}{49+42+47+44}\underbrace{\left(\frac{15}{44} - \frac{11}{44}\right)}_{\text{baseline } B \text{, Block 2}},$$

which simplifies to

$$F_{excess} = \underbrace{\left(\frac{21+14+22+15}{182}\right)}_{\text{focusing}} - \underbrace{\left(\frac{25+13+14+11}{182}\right)}_{\text{control switching}} = 4.9\%,$$

where 182 is simply two times the sample size.

General methodology. In general, our overall excess focusing measure for any given treatment will be calculated just as the above. To this end, we use the notation $X^Y_{c,b}$ to denote the number of choices from set c in Block b among subjects who chose Y in the baseline[5]. For example, $B^A_{\{A,B,C\},1}$ indicates the number of choices of B from $\{A, B, C\}$ in Block 1 among the subjects who chose A at baseline. Using N to denote the number of subjects, the excess focusing measure is defined as:

$$\begin{aligned} F_{excess} =& F_{gross} - F_{control} \\ =& \left(\frac{B^A_{\{A,B,R_B\},1} + A^B_{\{A,B,R_A\},1} + B^A_{\{A,B,R_B\},2} + A^B_{\{A,B,R_A\},2}}{2N} \right) \\ & - \left(\frac{B^A_{\{A,B,C\},1} + A^B_{\{A,B,C\},1} + B^A_{\{A,B,C\},2} + A^B_{\{A,B,C\},2}}{2N} \right), \end{aligned} \tag{1}$$

which is decomposed into a *gross* and *control* component. Analogously, the excess relative thinking measure is defined as

$$\begin{aligned} RT_{excess} =& RT_{gross} - RT_{control} \\ =& \left(\frac{B^A_{\{A,B,R_A\},1} + A^B_{\{A,B,R_B\},1} + B^A_{\{A,B,R_A\},2} + A^B_{\{A,B,R_B\},2}}{2N} \right) \\ & - \left(\frac{B^A_{\{A,B,C\},1} + A^B_{\{A,B,C\},1} + B^A_{\{A,B,C\},2} + A^B_{\{A,B,C\},2}}{2N} \right). \end{aligned} \tag{2}$$

Note that it so happens to be that the control component of both F_{excess} and RT_{excess} are the same (i.e. $F_{control} = RT_{control}$).

5 Results

We summarize our findings from 117 subjects.

[5] The notation resembles that for maps: X^Y is a switch from Y to X, similar to how $f \in X^Y$ is a mapping from space Y to space X.

5.1 Response Times and Decision Accuracy

We begin our analysis of the data by exploring response times and decision accuracy, as proxied by performance on the FOSD questions. This is to get a sense of the degree to which subjects are paying attention and if they are responsive to differences across treatments. This is especially important for interpreting our findings in relation to those obtained by other researchers.

In the left panel of Figure 1, we plot a histogram of all response times, i.e. across all subjects and questions regardless of treatment. In the right panel, we plot a histogram of average response times for each subject. It is clear that there is considerable heterogeneity, but it is noteworthy that while response times on some individual questions are very low, the average is 22 seconds, and more than 90% of subjects take at least 10 seconds on average.

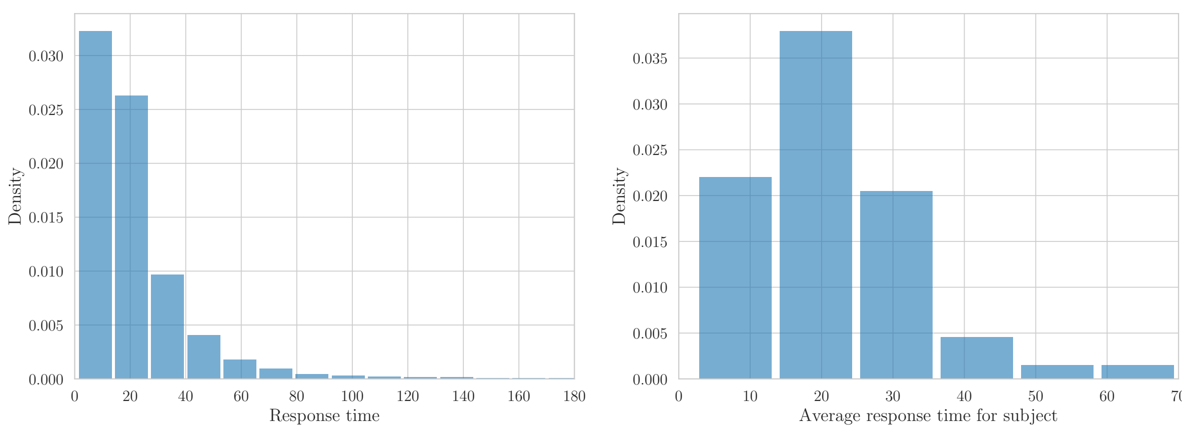

Figure 1: *Distribution of response times.* The left panel plots a histogram of all response times across all subjects and questions. The right panel plots the histogram of subjects' average response times across questions.

In Figure 2, we plot the average response time by question rank, and break it into different treatments. We also plot averages pooling across the low dimension treatments (sd and Sd), which we label d for "low dimension". Similarly, we use D, s, and S for high dimension, low stakes, and high stakes, respectively. There is a clear downward trend in response times for all treatments.

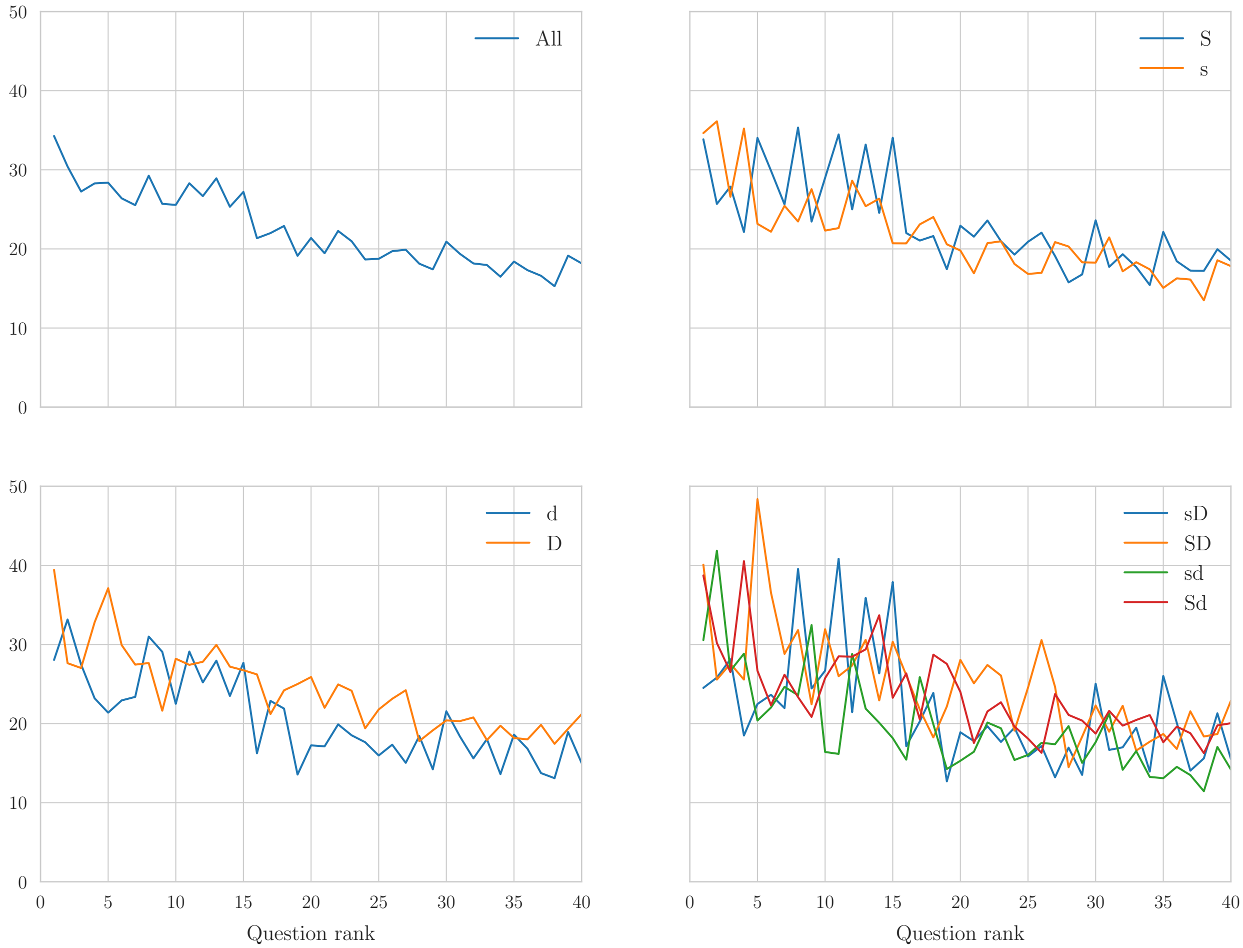

Figure 2: *Trends in response times.* Each panel plots the average of all subjects' response times as a function of question rank (1 through 40) for different treatments.

In Figure 3, we plot the average response times broken down by block and treatment. An all-else-equal increase in dimension has a large positive effect on response times, and an all-else-equal increase in stakes also has a sizeable positive effect on response times. Since there is substantial heterogeneity of response times across subjects, one might be concerned that the pattern is driven by subjects with very high response times. We thus plot the same figure after normalizing response times by dividing each subject's response times by his average response time (averaged across all questions), and the pattern is unchanged.

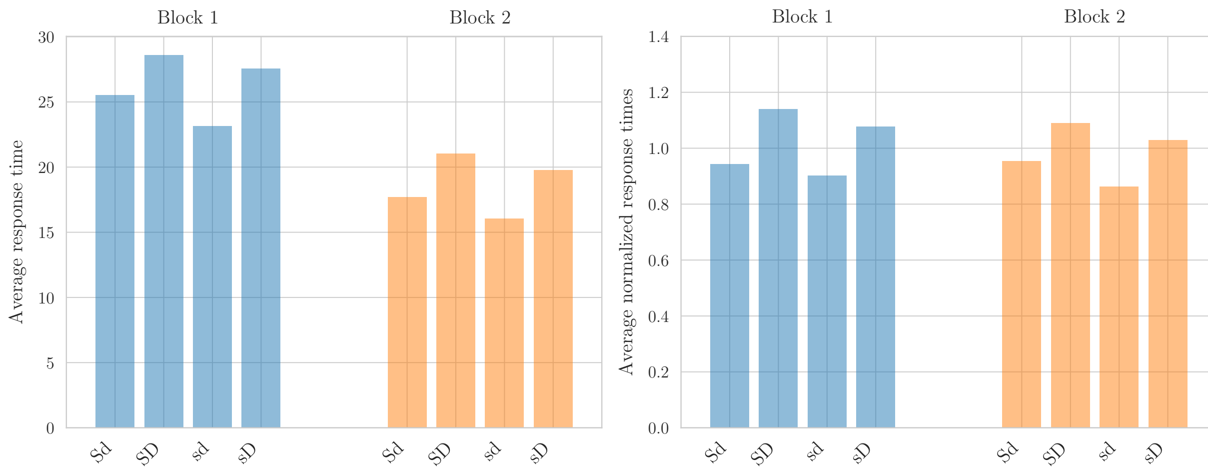

Figure 3: *Response times by treatment and block.* The left panel plots the average response times for each treatment-block. The right panel is similar, but prior to averaging across subjects, each subject's response times are divided by his average response time across all questions.

Next, we explore decision accuracy as proxied by the number of FOSD questions answered correctly. Recall that there are 8 such questions throughout the experiment–one within each treatment-block. Each has three lotteries, one of which first-order stochastically dominates the others. Hence, subjects who uniformly randomize would be expected to answer fewer than 3 correctly. The histogram in Figure 4 shows that nearly all subjects answer 4 or more correctly, the large majority answering at least 7 correctly, and nearly one third answering all 8 correctly.

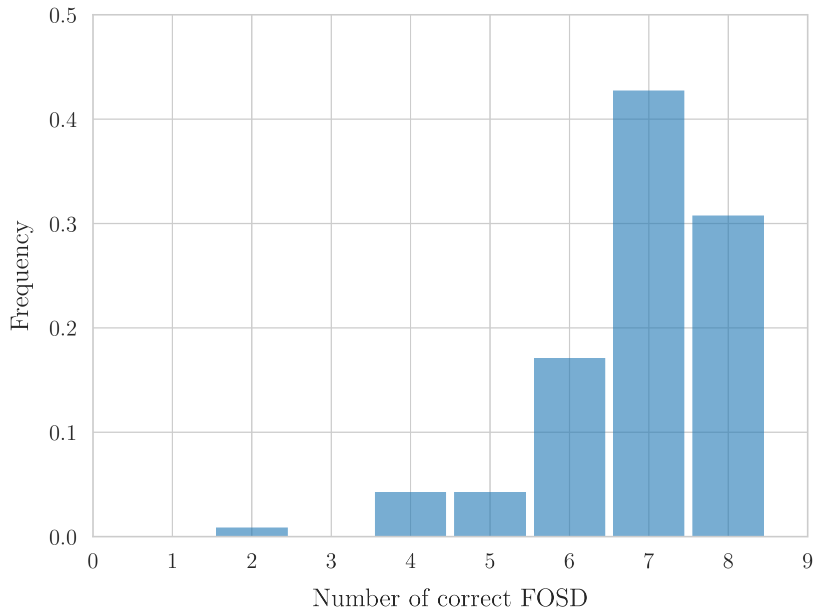

Figure 4: *Distribution of decision accuracies.* The figure plots a histogram of subjects' decision accuracies as measured by the number of FOSD questions answered correctly, which ranges from 0 to 8.

Figure 5 breaks down the previous figure into the four treatments. Within each treatment, subjects can answer up to 2 FOSD questions correctly. An all-else-equal increase in dimension leads to many more mistakes, especially when the stakes are low. An all-else-equal increase in stakes leads to many fewer mistakes, especially when the dimension is high. Recalling that increases in dimension and stakes both increase response times, this pattern makes perfect sense from an optimizing perspective if response time is a measure of effort, and effort itself is costly. An all-else-equal increase in dimensionality would increase effort at optimum but still lead to more mistakes. An all-else-equal increase in stakes would lead to more effort and thus fewer mistakes since the dimensionality remains fixed.

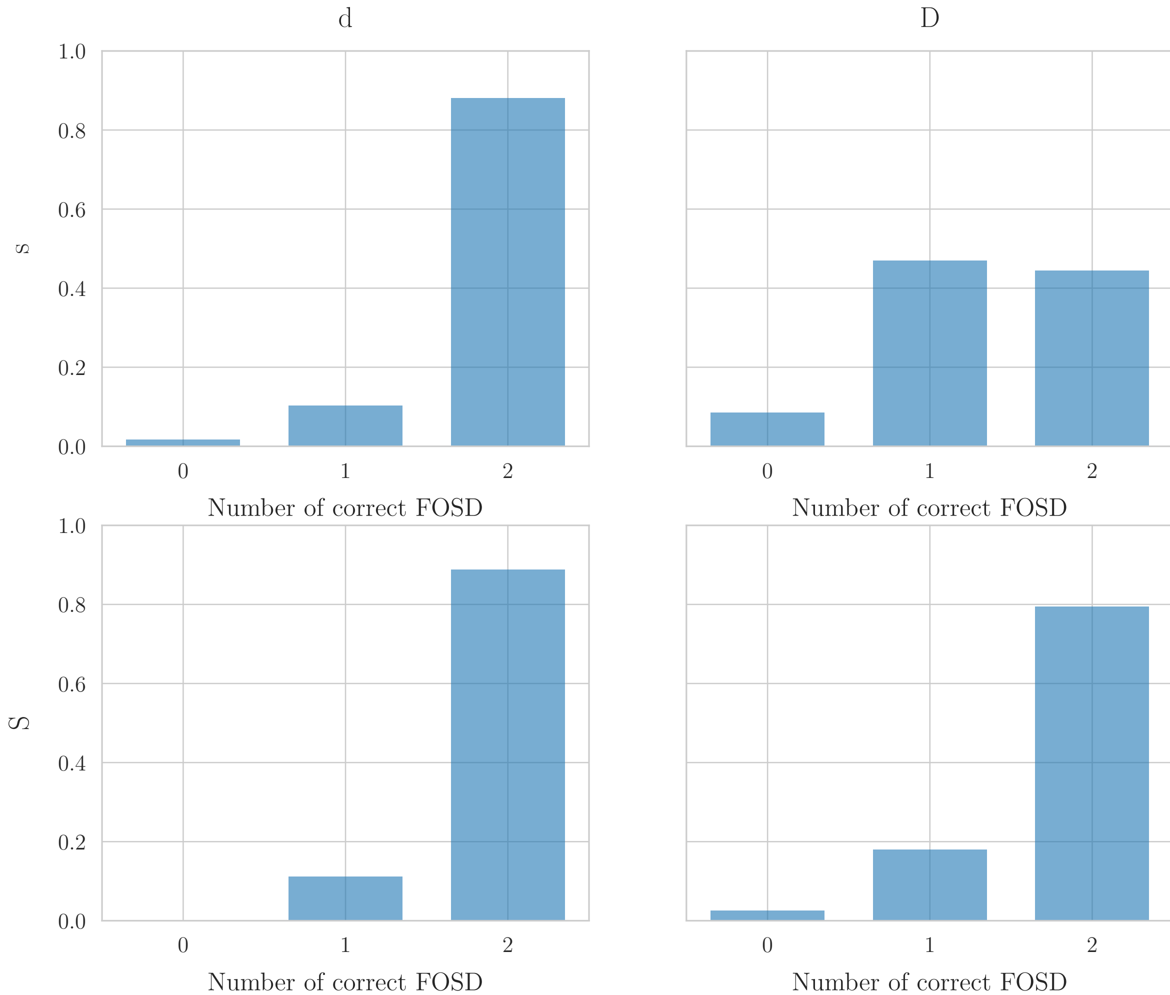

Figure 5: *Decision accuracy by treatment.* Each panel is a histogram of subjects' decision accuracies as measured by the number of FOSD questions answered correctly within each treatment, which ranges from 0 to 2.

Figure 6 shows that the proportion of FOSD questions answered correctly significantly drops in the second block. We have already established that response times go down steadily throughout the experiment. This could be due to some combination of learning (less effort required for accuracy) and boredom (higher cost of effort). Plausibly, if the learning effect were large, the proportion of FOSD questions answered correctly might increase, but we see just the opposite, which suggests boredom is playing a role.

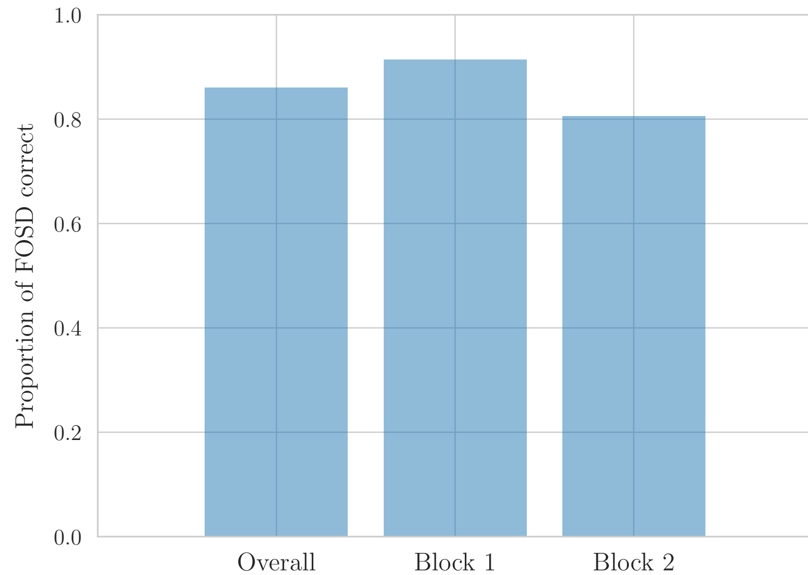

Figure 6: *Decision accuracy by block.* We plot the proportion of FOSD questions answered correctly within each block and throughout experiment as a whole.

5.2 Range Effects

We have established that our treatments and the duration of the experiment itself have had the desired effect of leading to variation in response times and accuracy on FOSD questions. We now turn to documenting range effects and how they interact with these variations.

To estimate range effects, we re-express the formulas from Section 4.1 in terms of linear regressions. This is convenient for inference and allows for the inclusion of controls.

Within each block-treatment, subject i chooses either A or B at baseline, and then makes up to three types of reversals: focusing (F), relative thinking (RT), and/or control (C). We simply regress an indicator for choice reversals on indicators for the type of reversal. In the basic regression equation (3), $i \in \{1, ..., 117\}$ indexes subjects, $q \in \{F, RT, C\}$ indexes the reversal (or question) type, $b \in \{1, 2\}$ indexes the block, and $t \in \{sd, sD, Sd, SD\}$ indexes the treatment. Indicator $\mathbf{1}\{E\}$ equals 1 under event E and 0 otherwise. In some regressions, we include X_{ibt}, a vector of controls aggregated across question types within each subject-block-treatment (and hence there is no q subscript).

$$\mathbf{1}\{reversal\}_{iqbt} = \beta_0 + \beta_1 \mathbf{1}\{q = F\}_{iqbt} + \beta_2 \mathbf{1}\{q = RT\}_{iqbt} + \beta_3 X_{ibt} + \epsilon_{iqbt} \tag{3}$$

Without controls, it is easy to see that regression estimates $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_0$, satisfy $\hat{\beta}_1 = F_{excess}$, $\hat{\beta}_2 = RT_{excess}$, and $\hat{\beta}_0 = F_{control} = RT_{control}$, and hence the question of whether or not there are range effects is reduced to interpreting standard regression output. Table 2 reports the regression results for each of the four treatments (columns 1-4). Columns 5-8 are for low dimension (d), high dimension (D), low stakes (s), and high stakes (S). Following the discussion in Section 4.1, we first dropped subjects within each treatment who ever switched to a third option (R_A, R_B, or C).

	(1) Sd	(2) SD	(3) sd	(4) sD	(5) d	(6) D	(7) s	(8) S
β_1	0.033	-0.044	0.049	0.005	0.042*	-0.018	0.026	-0.009
	(0.220)	(0.836)	(0.114)	(0.454)	(0.085)	(0.724)	(0.178)	(0.618)
β_2	0.020	-0.106	-0.011	0.034	0.003	-0.031	0.013	-0.048
	(0.333)	(0.992)	(0.603)	(0.195)	(0.461)	(0.844)	(0.330)	(0.940)
β_0	0.280***	0.394***	0.346***	0.340***	0.316***	0.365***	0.343***	0.342***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	450	540	546	618	996	1,158	1,164	990
R^2	0.001	0.008	0.003	0.001	0.002	0.001	0.000	0.002

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The p-values, shown in parentheses, are based on one-sided tests of the null hypothesis that $\beta_j \leq 0$. Standard errors clustered at the subject level.

Table 2: *Range Effects*. Each column reports results from regression (3) run on data from a different treatment or set of treatments pooled together.

Figure 7 plots the estimates for each treatment and corresponds to columns 1-4 of the table. Each panel corresponds to a different treatment, with estimates $F_{gross} = \hat{\beta}_1 + \hat{\beta}_0$ in blue and $RT_{gross} = \hat{\beta}_2 + \hat{\beta}_0$ in orange. The black line corresponds to $\hat{\beta}_0 = F_{control} = RT_{control}$, and hence the distance between colored bars and the black line give the treatment effects: $F_{excess} = \hat{\beta}_1$ and $RT_{excess} = \hat{\beta}_2$.

Overall, the evidence for range effects is relatively weak, though the focusing effect is marginally significant in each low dimension treatment separately and significant in the low dimensional treatments pooled together.

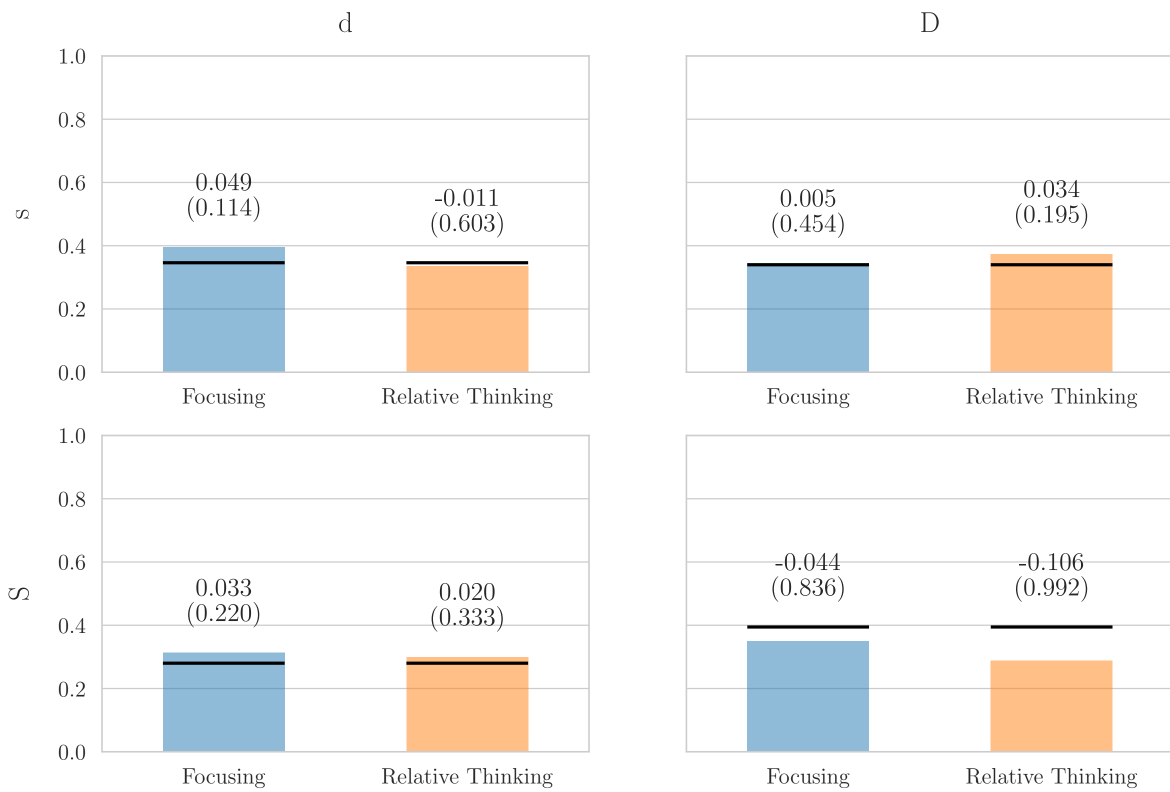

Figure 7: *Range Effects*. This figure plots estimates from Table 2. Each panel corresponds to one of columns 1-4.

5.3 Subject Types and the Determinants of Range Effects

Much is hidden in the aggregate analysis. For instance, is it the case that individual subjects display both focusing and relative thinking? Or is it that there are well-defined types within the population? And if so, are these types stable across treatments and time? Can these types be predicted based on other observables?

Within each treatment and block, we type subjects based on their behavior, classifying them into standard types who satisfy the independence of irrelevant alternatives (IIA), focusers (FOC), and relative thinkers (REL). These are defined by having no choice reversals, focusing choice reversals, or relative thinking choice reversals, respectively. All subjects who do not fit neatly into one of the other categories is classified as a noisy type (NOISY). Since we hypothesize that treatments and the length of the experiment will have an effect on subject types, we allow each block-treatment-subject to be a separate type. Hence, each subject will be typed 8 times based on the responses to the 4 questions within each treatment-block. This is summarized in the Table 3 .

	$\{A,B\}$	$\{A,B,R_A\}$	$\{A,B,R_B\}$	$\{A,B,C\}$
IIA	A	$\neq B$	$\neq B$	$\neq B$
	B	$\neq A$	$\neq A$	$\neq A$
FOC	A	A	B	$\neq B$
	B	B	A	$\neq A$
REL	A	B	A	$\neq B$
	B	A	B	$\neq A$
NOISY	–	–	–	–

Table 3: *Typing procedure.* This table summarizes how types are determined based on a subject's responses within each block-treatment.

Table 4 gives the distribution of types within each treatment, as well as pooled across all treatments. Once subjects are typed, we determine if the type distribution can be explained by various benchmark models. The first model we consider, BM1, is merely illustrative. In this model, subjects choose A and B with 50% probability from each of $\{A,B\}$, $\{A,B,R_A\}$, and $\{A,B,R_B\}$ (choice of R_A or R_B is very rare) and choose C from $\{A,B,C\}$ with probability 0.095, the empirical frequency of choice C across the whole experiment. This model predicts that IIA, FOC, and REL types all have the same frequency, and clearly over-estimates the fraction of noisy types. Our next model, BM2, is more realistic, allowing for another degree of heterogeneity. This model assumes that a fraction χ of subjects are consistent in all of their choices (all of whom will be classified as IIA) and a fraction $1-\chi$ of subjects who are the essentially random types from BM1. We estimate χ from the data so that this model predicts exactly the empirical frequency of IIA types. Compared to this model, we see that in the low-dimensional treatments *sd* and *Sd*, there are significantly more FOC and REL types than predicted by BM2.

	Data					Benchmark	
	sd	sD	Sd	SD	pool	BM1	BM2
IIA	40.2	41.0	40.6	42.3	**41.0**	12.5	**41.0**
FOC	*11.5*	8.1	*11.5*	7.3	9.6	12.5	8.1
REL	*10.3*	9.4	*10.7*	6.8	9.3	12.5	8.1
NOISY	38.0	41.5	37.2	43.6	40.1	62.5	42.8

Table 4: *Type distribution.* This table gives the empirical distribution of types within each treatment and overall, as well as that implied by two different benchmark models.

In the table below, we regress our main variables of interest–response times and the number of correctly answered FOSD questions–on indicators of subject types. Coefficient estimates are simply the averages of these quantities by type, though in columns 3 and 4, we control for treatment and block. Recall that the level of types is block-treatment-subject, and hence the number of FOSD questions answered correctly can be 0 or 1. The omitted category corresponds to NOISY types, and standard errors are clustered by subject. Unsurprisingly, the IIA types are significantly better at the FOSD questions than the NOISY types, while FOC and REL are not. More interestingly, IIA and FOC types spend considerably less time answering questions than REL or NOISY types. That FOC types respond more quickly is particularly interesting. It is unsurprising that if forced to respond quickly, subjects would tend to focus, but it is an important finding that subjects who *choose* to respond quickly also tend to focus. For these subjects, it may be that information acquisition costs are higher.

	(1) Resp. Time	(2) #FOSD correct	(3) Resp. Time	(4) #FOSD correct
IIA	-2.222	0.044**	-1.627	0.044**
	(0.164)	(0.031)	(0.286)	(0.032)
FOC	-2.403	-0.021	-1.359	-0.027
	(0.216)	(0.603)	(0.460)	(0.502)
REL	0.105	-0.048	0.061	-0.055
	(0.941)	(0.284)	(0.965)	(0.217)
Block 2			-7.426***	-0.021
			(0.000)	(0.282)
sD			-2.012**	0.066**
			(0.024)	(0.011)
SD			1.155*	0.002
			(0.052)	(0.932)
sd			-4.042***	0.053*
			(0.000)	(0.060)
Constant	23.565***	0.899***	28.162***	0.880***
	(0.000)	(0.000)	(0.000)	(0.000)
Obs.	936	936	936	936
R^2	0.003	0.011	0.047	0.023

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Standard errors clustered at the subject level. The p-values are shown in parentheses.

Table 5: *Correlates of types.* In regression form, this table gives the average response times and average number of FOSD questions answered correctly by type.

Finally, to determine if types are stable within subject-treatment across blocks, we run Fisher exact tests for each treatment. The null hypothesis is that each subject's type is drawn i.i.d. across blocks, and it is rejected if, on average, a subject's type in Block 1 is predictive of his type in Block 2. Table 6 presents p-values from the test within each of the four treatments. We see that types are considerably more stable across the high-stakes treatments, with the null soundly rejected within the SD treatment.

	sd	sD	Sd	SD
p-value	0.71	0.74	**0.39**	0.04

Table 6: *Stability of types.* This table reports the p-values from Fisher exact tests of whether types are drawn independently across blocks within each treatment.

6 Discussion and Conclusion

In this paper, we experimentally test for focusing and relative thinking in choice under risk. A few interesting and unexpected results emerged. First, we do not confirm previous experimental findings in that each of the two biases is much less prevalent in our data. We offer multiple explanations for this discrepancy, beginning with our different, and we argue more careful, experimental design.

In particular, focusing appears slightly more prevalent than relative thinking in our data, particularly in simpler environments. This is contrary to the hypothesis advanced by BRS, who conjectured that focusing might represent a useful heuristic when dealing with high-dimensional choice problems. Response times, however, paint a more coherent picture in that focusers tend to answer more quickly than other subjects. We cannot, however, establish a causal relationship beyond reasonable doubt. A clean way to test for this claim would be to artificially impose strict time constraints to subjects, and see whether this leads them, for instance, to focus on the largest payoffs, consistent with focusing.

The lack of relative thinking in our study was at first puzzling to us. We conjecture that relative thinking is heavily dependent on the problem framing. In the case of add-on pricing, or trading off time and financial convenience, relative thinking is likely to prevail. In the more abstract choice problem depicted in our experiment, and in absence of any form of visual priming, it looks like, if anything, subjects—especially those choosing more quickly—were just focusing on stand out payoffs.

While heavily weighting the largest numbers in a table is predicted by focusing (since large numbers usually come with large ranges), the same is not true for relative thinking.

It is important to stress that, although we often referred to both theories as biases, an evaluation of their welfare effects is tricky, something explicitly mentioned in Bushong et al. [2015] and discussed in a more general framework by Handel and Schwartzstein [2018]. Throughout the study, we employed choice reversals for identification exactly to avoid having to estimate utility. In principle, subjects might be aware of choosing this way, and find it satisfactory despite the inconsistencies generated. For example, saving \$20 for our dinner most likely *feels* better than saving \$200 when buying a house. On the other hand, a bonus of \$3000 on any given day might *feel* better than receiving an extra \$1 a day for ten years. A valuable direction for future research would be to find out creative ways to distinguish boundedly rational and richer psychological utility theories from simple perception driven errors.

Another interesting, and more general, direction for future research would be to further our understanding of how dimensionality modifies attention and perception patterns. What kind of violations of utility theory are more likely in different environments? Do lessons we have learned in scenarios in which subjects' choice sets consisted of three options generalise to others in which they are presented with twenty? To what extent can we trust the lessons we have learned about consumer choice two decades ago to hold up to modern scenarios in which agents are presented with a much larger number of options, and information is displayed so differently? This problem seems, to the best of our knowledge, to have been substantially understudied in the economics literature. With this experiment, we are just scraping the surface, but it goes without saying that this research agenda extends much further than range effects.

Appendices

Additional Design Details

Block 2

The following lists all of the Block 2 lotteries organized into treatments.

Block 2: main

Low Dimension (d)

Low Stakes (s)		35%	45%	20%
	A	11	8	5
	B	6	12	5
	R_A	2	9	7
	R_B	9	3	8
	C	7	9	6

High Dimension (D)

Low Stakes (s)		25%	35%	10%	15%	15%
	A	11	8	5	6	4
	B	6	12	5	3	7
	R_A	2	9	7	4	5
	R_B	9	3	8	4	6
	C	7	9	6	4	5

Low Dimension (d)

High Stakes (S)		35%	45%	20%
	A	22	16	10
	B	12	24	10
	R_A	4	18	14
	R_B	18	6	16
	C	14	18	12

High Dimension (D)

High Stakes (S)		25%	35%	10%	15%	15%
	A	22	16	10	12	8
	B	12	24	10	6	14
	R_A	4	18	14	8	10
	R_B	18	6	16	8	12
	C	14	18	12	8	10

Block 2: FOSD

Low Dimension (d)

Low Stakes (s)		35%	45%	20%
	A_f	10	3	6
	B_f	6	10	3
	C_f	3	5	10

High Dimension (D)

Low Stakes (s)		25%	35%	10%	15%	15%
	A_f	10	3	10	5	6
	B_f	6	10	3	10	3
	C_f	3	5	5	3	10

High Stakes (S)		35%	45%	20%
	A_f	20	6	12
	B_f	12	20	6
	C_f	6	10	20

High Stakes (S)		25%	35%	10%	15%	15%
	A_f	20	6	20	10	12
	B_f	12	20	6	20	6
	C_f	6	10	10	6	20

Choice Frequencies

Sd	Block 1				Block 2			
Baseline pick→	A		B		A		B	
	A	B	A	B	A	B	A	B
Baseline: $\{A, B\}$	27	0	0	48	34	0	0	41
Range-A: $\{A, B, R_A\}$	16	11	13	35	23	11	13	28
Range-B: $\{A, B, R_B\}$	15	12	14	34	25	9	9	32
Control: $\{A, B, C\}$	17	10	13	35	25	9	10	31

SD	Block 1				Block 2			
Baseline pick→	A		B		A		B	
	A	B	A	B	A	B	A	B
Baseline: $\{A, B\}$	39	0	0	51	49	0	0	41
Range-A: $\{A, B, R_A\}$	21	18	15	36	30	19	11	30
Range-B: $\{A, B, R_B\}$	20	19	9	42	31	18	6	35
Control: $\{A, B, C\}$	17	22	12	39	29	20	17	24

sd	Block 1				Block 2			
Baseline pick→	A		B		A		B	
	A	B	A	B	A	B	A	B
Baseline: $\{A, B\}$	49	0	0	42	47	0	0	44
Range-A: $\{A, B, R_A\}$	31	18	14	28	29	18	15	29
Range-B: $\{A, B, R_B\}$	28	21	15	27	25	22	10	34
Control: $\{A, B, C\}$	24	25	13	29	33	14	11	33

sD	Block 1				Block 2			
Baseline pick→	A		B		A		B	
	A	B	A	B	A	B	A	B
Baseline: $\{A, B\}$	53	0	0	50	65	0	0	38
Range-A: $\{A, B, R_A\}$	29	24	20	30	46	19	10	28
Range-B: $\{A, B, R_B\}$	35	18	23	27	42	23	11	27
Control: $\{A, B, C\}$	35	18	23	27	44	21	8	30

References

Marina Agranov and Pietro Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68, 2017.

Ola Andersson, Jim Ingebretsen Carlson, Erik Wengström, et al. Differences attract: An experimental study of focusing in economic choice. Technical report, 2016.

Ofer H Azar. Does relative thinking exist in real-world situations? a field experiment with bagels and cream cheese. *Economic Inquiry*, 49(2):564–572, 2011a.

Ofer H Azar. Do people think about absolute or relative price differences when choosing between substitute goods? *Journal of Economic Psychology*, 32(3):450–457, 2011b.

Ofer H Azar. Relative thinking in consumer choice between differentiated goods and services and its implications for business strategy. *Judgment and Decision Making*, 6(2):176, 2011c.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3):1243–1285, 2012a.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience in experimental tests of the endowment effect. *The American Economic Review*, 102(3):47–52, 2012b.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, 2013.

Benjamin Bushong, Matthew Rabin, and Joshua Schwartzstein. A model of relative thinking. *Working Paper, Harvard University, Cambridge, MA*, 2015.

Tom Cunningham. Relative thinking and markups. Technical report, Unpublished Manuscript, Harvard University, 2011.

Markus Dertwinkel-Kalt, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang. Concentration bias in intertemporal choice. 2016a.

Markus Dertwinkel-Kalt, Katrin Köhler, Mirjam RJ Lange, and Tobias Wenzel. Demand shifts due to salience effects: Experimental evidence. Technical report, DICE Discussion Paper, 2016b.

Benjamin Handel and Joshua Schwartzstein. Frictions or mental gaps: What's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives*, 2018.

Justine S Hastings and Jesse M Shapiro. Fungibility and consumer choice: Evidence from commodity price shocks. *The quarterly journal of economics*, page qjt018, 2013.

Samel Hirshman, Devin Pope, and Jihong Song. Mental budgeting versus relative thinking. 108: 148–52, 2018.

Charles A Holt and Susan K Laury. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

Botond Kőszegi and Adam Szeidl. A model of focusing in economic choice. *The Quarterly journal of economics*, page qjs049, 2012.

Peter Landry and Ryan Webb. Pairwise normalization: A neuroeconomic theory of multi-attribute choice. *Working paper*, 2019.

Ritesh Saini, Raghunath Singh Rao, and Ashwani Monga. Is that deal worth my time? the interactive effect of relative and referent thinking on willingness to seek a bargain. *Journal of Marketing*, 74 (1):34–48, 2010.

Alireza Soltani, Benedetto De Martino, and Colin Camerer. A range-normalization model of context-dependent choice: a new model and evidence. *PLoS computational biology*, 8(7):e1002607, 2012.

Jason Somerville. Choice-set-dependent preferences in consumer choice: An experimental test. *Working Paper*, 2019.

Richard Thaler. Mental accounting and consumer choice. *Marketing Science*, 4(3):199–214, 1985.

Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. In *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, pages 107–129. Springer, 1985.

Amos Tversky and Itamar Simonson. Context-dependent preferences. *Management science*, 39(10): 1179–1189, 1993.