# Range Effects in Economic Choice: The Role of Complexity*

Tommaso Bondi† Dániel Csaba‡ Evan Friedman§ Salvatore Nunnari¶

September 30, 2025

## Abstract

Several behavioral models assume that choice over multi-attribute goods is systematically affected by the ranges of attribute values. Two recurring principles in this literature are *contrast*—whereby attributes with larger ranges attract attention and are therefore overweighted—and *normalization*—whereby attributes with larger ranges are underweighted as fixed differences appear smaller against a larger range. These principles lead to divergent predictions, and yet, both contrast-based and normalization-based models have found strong empirical support, albeit in different contexts and with different experimental designs. The question remains: when does one effect emerge over the other? We experimentally test a unifying explanation: normalization dominates in simple choices, while contrast dominates in complex choices. We conduct an experiment with real-effort tasks in which we manipulate attribute ranges in both simple and complex choices. We find that, indeed, contrast dominates as the number of attributes increases. We also find that contrast emerges with cognitive load induced by time pressure.

**Keywords:** Multi-Attribute Choice; Range Effects; Focusing; Relative Thinking; Salience; Bottom-Up Attention; Context Dependence; Complexity; Experiment
**JEL Classifications:** C91, D91, D12.

*We are grateful to audiences at the 2024 CESifo Area Conference in Behavioral Economics and the 2025 Korean Economic Review International Conference for helpful comments and discussions. This study was approved by Bocconi Research Ethics Committee (Application EA000706). The pre-registrations are available at https://osf.io/5ek4u, https://osf.io/hn9de, and https://osf.io/nq4zy. We acknowledge financial support from the European Research Council (POPULIZATION Grant 852526) and the Agence Nationale de la Recherche (Grant ANR-24-CE28-0028-01).
* Cornell Tech and Johnson School of Management, Cornell University; tbondi@cornell.edu.
† QuantCo; csaba.daniel@gmail.com.
‡ Paris School of Economics, Department of Economics; evan.friedman@psemail.eu.
§ Bocconi University, Department Economics; CEPR; CESifo; salvatore.nunnari@unibocconi.it.

# 1. Introduction

Many decisions we make, both big and small, involve choosing between options with multiple attributes. When buying a house, we consider price and quality. When deciding where to live, we consider the weather, school system, opportunities for work, and many other attributes. These are complex decisions, so we rely on simplified heuristics. In particular, decisions may be systematically affected by the ranges (or spreads) of attribute values.

Consider the decision of where to live. Los Angeles has much better weather than Chicago, but suppose that Chicago is slightly better in most of the many other attributes. It may be that people tend to be happier in Chicago, but many will still choose Los Angeles because weather is salient as the only attribute for which the two locations substantially differ, i.e., for which there is a large range of values. To take another example, suppose one is deciding which of two houses to buy. House A is slightly bigger than house B, but costs 50,000 dollars more. This price difference may seem like a lot when the range of prices across all houses on the market is only 100,000 dollars, but may seem insignificant when the range is 2,000,000 dollars, in which case one is more likely to choose house A.

Both of these hypothetical examples are plausible. However, they suggest opposite heuristics. In the first example, it is as if people put more weight on attributes with large ranges—a principle we refer to as *contrast*, formalized in salience and focusing models (Bordalo, Gennaioli, and Shleifer, 2012, 2013, 2022; Kőszegi and Szeidl, 2013). In the second example, it is as if people put less weight on attributes with large ranges—a principle we refer to as *normalization*, formalized in relative thinking and divisive coding models (Soltani, De Martino, and Camerer, 2012; Webb, Glimcher, and Louie, 2020; Bushong, Rabin, and Schwartzstein, 2021; Landry and Webb, 2021). These models make opposing assumptions and divergent predictions. And yet, both contrast-based and normalization-based theories have found strong empirical support, albeit in very different contexts and on the basis of very different experimental designs (Section 2 provides a discussion). The question remains: when does one effect emerge over the other?

In this paper, we experimentally test the hypothesis—first conjectured in Bushong et al. (2021)—that contrast emerges in complex environments with many attributes, whereas normalization may predominate in simpler ones. The idea is that people naturally use relative values when considering a *given attribute*. This is because fixed differences loom smaller against a larger range. However, people can only attend to a *small number of attributes*. When there are many

attributes, people attend to those with the largest ranges, as these are the most consequential.

Although both principles are embodied in multiple models, we operationalize them in our design by using the two frameworks that provide the sharpest, most minimalistic formulations: focusing (Kőszegi and Szeidl, 2013) for contrast and relative thinking (Bushong et al., 2021) for normalization. In these models, each attribute receives a weight that depends on that attribute's range of utility values within the choice set. In focusing, the weights are *increasing* in the range (contrast), whereas, in relative thinking, the weights are *decreasing* in the range (normalization). Hence, the two models share a common formalism, but are defined by single, opposing axioms. They therefore serve as clean operationalizations of the opposing forces of attention that we seek to identify.[1]

To test for the role of complexity, we document range effects in both low- and high-attribute choice sets (where the available options have 2 and 4 attributes, respectively). Hence, we focus on the number attributes as our measure of complexity, which captures a natural dimension of complexity in this context.[2] We find that, indeed, contrast dominates as the number of attributes increases.

In our real-effort experiment, participants are asked to rank four work contracts—each consisting of a wage and a number of tasks to be completed. In our low-attribute treatments, which closely follow the design in Bushong et al. (2021), the two attributes are "money" and "effort." Two contracts are held fixed across treatments, whereas the other two contracts vary—our instrument for manipulating attribute ranges. We identify range effects through choice reversals between the fixed contracts: one direction for focusing and the other for relative thinking. Our high-attribute treatments replicate this basic design, except with more complex contracts, constructed from the original by including two more attributes—additional tasks of a different kind (that is, requiring different types of effort and denominated in different units).[3] Hence, we identify range effects in both simple and complex environments, while keeping other aspects of the design as similar as possible.

[1] By comparison, salience theory (Bordalo et al., 2012, 2013, 2022) combines two axioms—ordering and diminishing sensitivity—and would predict no treatment effects in our design, since average attribute values are held constant when ranges are manipulated. Nonetheless, its ordering axiom is meant to capture the idea that large attribute differences attract attention, i.e., the essence of contrast in bottom-up attention.

[2] In a similar spirit, Puri (2025) argues, theoretically and experimentally, that the number of outcomes of a lottery captures an important element of its complexity.

[3] The additional attributes are held fixed across high-attribute treatments. Hence, the range manipulations are directly comparable across low- and high-attribute treatments.

We find no evidence for range effects in the simple environment, but we find a large and highly significant contrast effect (via focusing) in the complex one. We also find that the difference-in-difference is large and highly significant—and hence we provide evidence for the emergence of contrast when the environment becomes more complex. Our finding that focusing emerges as the number of attributes increases is natural, confirming the conjecture of Bushong et al. (2021) and helping to unify disparate results in the literature.

The lack of evidence for range effects in our low-attribute treatment deserves more scrutiny. We think of both focusing and relative thinking as reactions to limited cognitive resources. Hence, we conjectured that our low-attribute treatment was so simple that our participants were not sufficiently cognitively constrained for either focusing or relative thinking to emerge as dominant heuristics. This led us to run a new low-attribute treatment in which we increase cognitive load by imposing time pressure.[4] A priori, the predicted effect is ambiguous. The time pressure may lead to focusing if people only have time to consider one attribute carefully, or it may induce relative thinking if people still consider both attributes, but must now rely more heavily on relative thinking for assessing attribute values.

We find that time pressure leads to a significant contrast effect (via focusing), with a significant difference compared to the treatment without time pressure. Hence, contrast may emerge even in very simple environments, at least when other constraints on the decision maker are present. Overall, our results paint a consistent picture in which both complexity and cognitive load lead decision-makers to focus on salient attributes, consistent with the contrast principle.

The paper is organized as follows. Section 2 reviews the literature, Section 3 introduces the experimental design, Section 4 describes the theoretical framework, Section 5 presents the results, Section 6 discusses the broader implications of our findings, and Section 7 concludes.

## 2. Literature Review

Our paper relates to the large and fast growing literature on context effects, and specifically range effects, particularly as studied by Kőszegi and Szeidl (2013) and Bushong et al. (2021). We start by describing the two theories, then review related theoretical and experimental work, and conclude with broader connections to complexity in choice.

[4] Deck et al. (2021) show that time pressure has similar effects on math performance and elicited risk preference as a number of other methods thought to increase cognitive load (e.g., a number memorization task), supporting our interpretation of time pressure as increasing cognitive load.

The principle of *contrast* is captured in focusing theory (Kőszegi and Szeidl, 2013), which posits that people overweight attributes along which options differ most. This implies a tendency to favor alternatives with concentrated advantages, as attention is drawn to those dimensions. For example, when comparing Los Angeles and Chicago, a focuser would overweight weather differences while underweighting other dimensions in which the cities differ less. Closer to our experiment, when comparing job offers, a large spread in pay can draw attention toward salary at the expense of effort requirements.

Focusing theory has received substantial theoretical and experimental attention (Karlan et al. 2016, Dertwinkel-Kalt et al. 2017, Nunnari and Zápal 2025, Taubinsky and Rees-Jones 2018, Woodford 2012, Fallucchi and Kaufmann 2021). Experiments explicitly testing for focusing effects include Andersson et al. (2021) and Dertwinkel-Kalt et al. (2022).

Andersson et al. (2021) ask participants to choose between streams of payments over time, modeling the dates at which payments are received as separate attributes. They find that participants are more likely to choose a given option – say, Option A – when additional options are added that increase the range of payments in the attribute that is Option A's advantage (that is, the date at which Option A offers a larger payment than other options). This is what focusing predicts: by increasing the range associated with A's advantage, participants pay more attention toward it, making A more attractive. They also find that the focusing effect diminishes when payments are presented numerically rather than graphically or when participants are forced to wait before submitting their answers. The latter can be seen as an indirect test of the role of complexity: being forced to reason more, participants overcome their tendency to (over)focus on the high spread attributes and are instead capable of computing the more complex trade-off involving all (or, at least, more) attributes.

In a related experiment on focusing in intertemporal choice, Dertwinkel-Kalt et al. (2022) provide experimental evidence for "concentration bias," the tendency to overweight advantages that are concentrated in time. As an example, consider the choice between a single 20-dollar payment in one month and four 5-dollar payments in each of the next four weeks. Clearly, neither standard discounting nor present bias could explain a preference for the former over the latter, as the latter option involves the same total compensation, paid earlier on average. In their experiment, the authors show that participants commit to too much overtime work (i.e. that enters utility negatively) when it is dispersed over multiple days in exchange for a bonus that is concentrated in time. Moreover, this concentration (or focusing) bias is quantitatively important, as it increases

participants' willingness to work by 22.4% beyond what standard discounting models can account for.

Salience theory (Bordalo et al., 2012, 2013, 2022) also links attention to contrast, as formalized in the *ordering* axiom, which implies that an attribute of a given object stands out more when its value deviates further from the choice-set average. These models also feature *diminishing sensitivity*, whereby the marginal increase in salience decreases as attribute values deviate even further from the choice-set average. Because in our design the average values of each attribute remain constant as ranges are manipulated, salience theory formally predicts no treatment effect.[5] Nonetheless, our experiment still speaks to the ordering axiom conceptually, as it isolates whether large attribute differences attract attention—the core idea behind contrast in bottom-up attention. Other contrast-based approaches include the sparsity model by Gabaix (2014), in which decision makers selectively attend to the few dimensions that matter most. While not formulated in terms of attribute ranges, sparsity provides a microfoundation for why attention is drawn to dimensions with greater variability, closely aligning with the idea of contrast.

The opposing principle, *normalization*, is embodied in relative thinking (Bushong et al., 2021), which posits that fixed differences loom smaller against a larger range. In our running examples, adding Los Angeles to the choice set makes the Chicago–New York weather difference appear negligible, and the presence of a very high-paying job reduces the perceived importance of any fixed pay gap. Put differently, whereas focusing predicts increased attention (and, thus, decision weight) given to attributes with large ranges, relative thinking predicts the opposite: the very presence of a large range makes fixed differences seem small, and so the high-range attribute is discounted by the decision maker. Relative thinking has also been widely studied (Shah et al. 2015, Hirshman et al. 2018, Dertwinkel-Kalt and Köster 2020, Castillo 2020, Landry and Webb 2021, Strulov-Shlain 2023), including recent experimental tests by Somerville (2022) and Azar and Voslinsky (2024).

Somerville (2022) conducts a laboratory experiment in which the prices of high- and low-quality variants of multiple products are varied. The data provide clear evidence of choice-set dependence consistent with relative thinking: price increases that expand the range of prices in the choice set lead to more purchases. Structural estimates imply economically meaningful effect sizes: the average participant is willing to pay around 17% more when a seemingly irrelevant option

[5]For experimental tests of salience theory's distinctive predictions, see Mormann and Frydman (2018) and Dertwinkel-Kalt and Köster (2020). For an axiomatization of salience theory that emphasizes how the ordering property differentiates salience theory from other models, see Lanzani (2022).

is added to the choice set. Similarly, Azar and Voslinsky (2024) examine a scenario in which participants are offered to do real-effort tasks and are paid piece-rate for every correct answer. However, some participants are paid a higher fixed payment. Consistent with relative thinking, effort is lower when the fixed payment is higher: the piece-rate seems smaller in comparison, resulting in less effort.

Beyond relative thinking, normalization mechanisms feature prominently in the neuroeconomic literature on *divisive normalization*, a canonical neural computation observed across species.[6] Choice experiments by Louie et al. (2013), Khaw et al. (2017), and Glimcher and Tymula (2023) confirm distinct predictions of the model. Landry and Webb (2021) demonstrate that a variant of this model labeled as *pairwise normalization* can reproduce a broad set of context effects including attraction effects. Together, these frameworks provide a rich set of behavioral and theoretical underpinnings for why fixed differences loom smaller in wider ranges.

More broadly, our paper contributes to a growing literature on complexity in economic choice, which emphasizes that cognitive limitations and costly information processing play a central role in shaping beliefs and behavior. Complexity has been shown to matter in diverse canonical settings, including choice under uncertainty (Enke and Graeber, 2023; Enke and Shubatt, 2024; Oprea, 2024b; de Clippel et al., 2024; Puri, 2025), intertemporal choice (Enke et al., Forthcoming), and information processing and acquisition (Ba et al., 2024; Guan et al., 2025).[7] A recurring theme in this literature is that increasing complexity can attenuate sensitivity to fundamentals (Enke et al., 2024) or lead decision makers to rely more heavily on simplifying heuristics (Arrieta and Nielsen, 2024). Our contribution is to connect these insights to the study of range effects. In particular, we show that the relative prevalence of contrast and normalization—two competing principles in theories of attention and perception—depends systematically on the complexity of the environment. In this way, our results highlight that complexity shapes not only the extent of behavioral departures from standard models, but also the form they take, by shifting which heuristic governs choice.

# 3. Experimental Design

**Overall Structure.** Our experiment follows a $2 \times 2$ design, with each treatment defined by (i) which attribute ranges are manipulated and (ii) the number of attributes (i.e., the degree of com-

[6] Bucher and Brandenburger (2022) characterize when divisive normalization is an "efficient code."
[7] See Oprea (2024a) for recent reviews of this literature.

plexity of the decision-making environment). Each participant took part in exactly one treatment, i.e. a between-participant design. The four treatments are: wide money-2 attributes (WM2), wide effort-2 attributes (WE2), wide money-4 attributes (WM4), and wide effort-4 attributes (WE4).

The basic idea behind the experiment is as follows. In each treatment, participants give their preference ranking between 4 work contracts, in a similar setup to that of Bushong et al. (2021). Each contract consists of a certain number of real-effort tasks and an amount of money to be received upon completion of the work. By using 4 contracts, we can hold 2 contracts fixed and vary the other 2 contracts across treatments. This allows us to document how the choice between the 2 fixed contracts is affected by attribute ranges, which are controlled by varying the other 2 contracts.

All experiments were conducted on the Prolific online platform. Based on pre-registered power calculations, we collected the data of approximately 400 participants per treatment, for a total of 1,603 participants. The study was pre-registered on OSF (https://osf.io/hn9de).[8]

**2-Attribute Treatments.** First consider the 2-attribute treatments. Across WM2 and WE2, two contracts—call them $B$ and $C$—are held fixed, while the other two contracts—call them $A$ and $D$—vary. In both treatments, $A$ is attribute-wise dominant and $D$ is attribute-wise dominated, so that $A$ is expected to be ranked first and $D$ is expected to be ranked last. Contract $C$ offers more money than $B$, but also requires more effort than $B$. We are interested in the fraction of participants choosing $B$ over $C$. In the WM2 treatment, options $A$ and $D$ increase the range in the money dimension, and so we refer to them as $A_m$ and $D_m$; in the WE2 treatment, options $A$ and $D$ increase the range in the effort dimension, and so we refer to them as $A_e$ and $D_e$. The exact contracts we use are given in Table 1. In "Details of work contracts" below, we discuss details of the task and the units in the table; in "Calibration and selection of contracts", we discuss how these exact contracts were selected on the basis of a calibration exercise.

**4-Attribute Treatments.** The 4-attribute treatments are similar. Contracts $B^{'}$ and $C^{'}$ are held fixed across WM4 and WE4, whereas $A^{'}$ and $D^{'}$ vary, and we are interested in the fraction of participants choosing $B^{'}$ over $C^{'}$. The key idea behind our design is that the 4 contracts $(A^{'},B^{'},C^{'},D^{'})$ are constructed from $(A,B,C,D)$ by "appending" two additional attributes—another two real-effort tasks. Hence, in WM4, options $A^{'}_{m}$ and $D^{'}_{m}$ increase the range in the money dimension;

[8] We pre-registered the collection of 400 participants per treatment. Because of how recruiting works on Prolific, we ended up with 401 participants for 2 treatments. Results are unchanged if we include only the first 400 participants who completed the study in each treatment.

| | Wide Money | |
|---|---|---|
| | Wage | Task 1 |
| $A_m$ | 4.80 | 9 |
| $B$ | 2.20 | 9 |
| $C$ | 2.70 | 12 |
| $D_m$ | 0.10 | 12 |

| | Wide Effort | |
|---|---|---|
| | Wage | Task 1 |
| $A_e$ | 2.70 | 1 |
| $B$ | 2.20 | 9 |
| $C$ | 2.70 | 12 |
| $D_e$ | 2.20 | 20 |

Table 1: *2-Attribute Contracts*

in WE4, $A_e^{'}$ and $D_e^{'}$ increase the range in the *first* effort dimension. The attribute values in the second and third effort dimensions are held fixed across WM4 and WE4. As before, $A^{'}$ and $D^{'}$ are attribute-wise dominant and dominated, respectively. The exact contracts we use are given in Table 2.

| | Wide Money | | | |
|---|---|---|---|---|
| | Wage | Task 1 | Task 2 | Task 3 |
| $A_m^{'}$ | 4.80 | 9 | 5 | 2 |
| $B^{'}$ | 2.20 | 9 | 6 | 5 |
| $C^{'}$ | 2.70 | 12 | 10 | 3 |
| $D_m^{'}$ | 0.10 | 12 | 11 | 6 |

| | Wide Effort | | | |
|---|---|---|---|---|
| | Wage | Task 1 | Task 2 | Task 3 |
| $A_e^{'}$ | 2.70 | 1 | 5 | 2 |
| $B^{'}$ | 2.20 | 9 | 6 | 5 |
| $C^{'}$ | 2.70 | 12 | 10 | 3 |
| $D_e^{'}$ | 2.20 | 20 | 11 | 6 |

Table 2: *4-Attribute Contracts*

**Identifying Range Effects.** Clearly, in the standard model, i.e. without any set-dependent preferences, an individual's preference ranking between $B$ and $C$ should not depend on $A$ and $D$. However, with range effects, choice may be affected. Notice that $B$'s advantage relative to $C$ is in the effort dimension, whereas $C$'s advantage relative to $B$ is in the money dimension. For an individual participant, a choice reversal from $B$ in WM2 to $C$ in WE2 would indicate relative thinking, and a choice reversal from $C$ in WM2 to $B$ in WE2 would indicate focusing. In our between-participant design, we cannot identify choice reversals on the individual level.[9] However, as we show in Section 4, we can identify *lower bounds* on the shares of participants who are relative thinkers and focusers.

**Details of Work Contracts.** Task 1, which is used in all four treatments is a "counting task" in which participants must count every appearance of a particular symbol within an image. The same task was used previously in Bushong et al. (2021). Tasks 2 and 3 are only used in the

[9] A within-participant design would expose individuals to multiple ranking tasks featuring some of the same work contracts. This might affect choices in the second ranking task, especially as participants may be affected by the ranges observed in the first ranking task.

4-attribute treatments. Task 2 is a "translation task" in which participants must translate a 7-digit number into a string of letters according to a key, and Task 3 is a "waiting task" in which participants must wait for a certain number of minutes and click a button as it appears at random intervals on their screens. The units displayed in the tables are thus images, numbers, and minutes, respectively. The money offered to participants is denominated in British pounds, the standard currency on the Prolific platform. Screenshots of the experimental interface, showing all three tasks, are in Appendix 7.

**Incentives.** All participants received 2 pounds simply for completing the experiment. In addition, choices were incentivized. After ranking the 4 contracts, with 90% probability, the experiment ended and the participant received only the completion fee. With 10% probability, the participant had to perform additional work for additional pay (in order to receive the completion fee): 2 of the 4 contracts were randomly chosen, and of those two, the participant had to complete the contract that they ranked higher. This procedure ensures that it is incentive compatible for participants to truthfully report their preference rankings.

The average earnings (including both the completion fee and the performance-based bonus for participants selected to complete additional tasks) were 2.32 pounds. The average earnings conditional on being selected to complete additional tasks were 5.34 pounds. The median completion time was 9m33s with an average reward per hour of 12.57 pounds (exceeding the rate suggested by the recruiting platform).

**Procedures.** Participants first read instructions. They were then required to answer comprehension questions. If they did not answer all questions correctly after two attempts, they were not allowed to continue the experiment and did not receive any payment. If they successfully completed the comprehension questions, they continued to Part 1 in which they gained experience with the task (in the case of 2-attribute treatments) or tasks (in the case of 4-attribute treatments). In Part 2, they ranked the 4 work contracts. With 90% probability, the experiment ended and the participant received the completion fee. With 10% probability, the participant proceeded to Part 3: they were given a work contract to complete, selected in the manner described in "Incentives" above. If they completed the work, they received the corresponding additional pay.

Labels $A$, $B$, $C$, $D$, etc. were not shown to participants. Instead, the order of the contracts was randomized at the participant level, and the contracts were labeled Option 1, Option 2, Option 3, and Option 4 (see Figures 6 and 7 of Appendix 7 for screenshots of the experimental interface).

**Calibration and Contracts' Selection.** We used a pre-registered calibration procedure (https://osf.io/5ek4u) to select the contracts $B$ and $C$ that were used in the 2-attribute treatments. The idea was to find contracts $B$ and $C$ such that approximately 50% of participants rank $B$ above $C$, i.e. a population-level calibration.

Our calibration procedure is sequential in nature. In the first step, we asked 300 participants for their preference ranking over the contracts $(A'', B'', C'', D'')$ in Table 3 (using the same incentives and experimental procedures as in our main experiment). Note that $A''$ is strictly attribute-wise dominant and $D''$ is strictly attribute-wise dominated. We found that 48.5% chose contract $B''$ over $C''$, which is between 42% and 58%—our preregistered stopping criterion—and therefore we set $B = B''$ and $C = C''$ for the main experiment. Had the fraction preferring $B''$ over $C''$ been less than 42% or greater than 58%, we would have proceeded to the second step by asking another 300 participants for their preference ranking over a different set of 4 contracts $(A''', B''', C''', D''')$—the exact set depending on which of $B''$ or $C''$ was favored. If the fraction of participants choosing $B'''$ over $C'''$ was between 42% and 58%, we would have selected $B = B'''$ and $C = C'''$ for the main experiment. Otherwise, we would have proceeded in this fashion with another set of 4 contracts, and so on, until the stopping criterion was achieved. The exact sequence of contracts and stopping criterion were preregistered.

With $B$ and $C$ selected based on this calibration, $A_m$ and $D_m$ were chosen in WM2 so as to increase the range of values in the money dimension, while leaving the effort dimension unchanged. Similarly, $A_e$ and $D_e$ were chosen in WE2 to increase the range of values in the effort dimension, while leaving the money dimension unchanged. As in Bushong et al. (2021), $A$ and $D$ are "symmetric" about $B$ and $C$ in the sense that they do not affect the average values of both attributes in the choice set. Furthermore, $A$ and $D$ increase ranges as much as possible while avoiding negative or zero values, so the smallest amount of money across all contracts is 0.10 (contract $D_m$ in WM2) and the smallest number of tasks across all contracts is 1 (contract $A_e$ in WE2).

As discussed, in the 4-attribute treatments, $B'$ and $C'$ are constructed from $B$ and $C$ by appending two additional attributes—another 2 real-effort tasks. The values of these additional attributes were not selected on the basis of a calibration. Rather, these values were chosen so that tasks 2 and 3 were relatively less important quantitatively: for both $B'$ and $C'$, the expected time required for these tasks is fairly small compared to task 1, and the differences across $B'$ and $C'$ in the expected time required for these tasks are also fairly small. Hence, even without having calibrated $B'$ and $C'$, we would expect approximately half of all participants to favor $B'$ over $C'$ in the absence of

range effects.

| | Wage | Task 1 |
|---|---|---|
| $A''$ | 2.80 | 8 |
| $B''$ | 2.20 | 9 |
| $C''$ | 2.70 | 12 |
| $D''$ | 2.10 | 13 |

Table 3: *Contracts Used in Calibration Study*

**Discussion of Design.** The 2-attribute treatments are closely based on the experiments in Bushong et al. (2021). There are four primary differences between our 2-attribute treatments and their design. First, the exact contracts are slightly different (see Table 9 of Appendix 7 for the exact contracts used by Bushong et al. 2021). In particular, we simplified the counting task (task 1) by using a simpler image[10] and used fewer tasks in our work contracts. This was necessary so that the 4-attribute treatments, which involve additional tasks, did not take too long.[11] We also paid in pounds as opposed to dollars.[12] Second, we ran our experiments on Prolific as opposed to MTurk. We made the switch because Prolific, which has grown in use since Bushong et al. (2021) ran their original experiments, is now known to have much higher data quality (Peer et al., 2022). Third, we used a different calibration procedure, which also used many more participants. Fourth, we used a larger sample size (400 per treatment as opposed to 276-294).[13]

An advantage of using real-effort contracts is that they allowed us to finely adjust the number of tasks involved and vary the number of attributes. Further, as opposed to lotteries or intertemporal choices (in which each alternative represents a series of payments or an amount of effort or money at each of several time periods), we feel that using several real-effort tasks makes it less likely that participants will integrate all attributes together into a single index, as participants must evaluate the subjective trade-offs between money and different types of effort. Another aspect

[10] Our image was a matrix containing $8 \times 13 = 104$ symbols, as opposed to $10 \times 15 = 150$ symbols in Bushong et al. (2021).

[11] To avoid differential selection into treatments, all treatments were advertised to participants in the same way, with the same completion fee. Hence, the expected durations of 2- and 4-attribute treatments could not be too different without resulting in either very high or very low expected average payments per minute for one of the treatments.

[12] The wages associated with work contracts are similar in the two studies but we offered a greater completion fee (2 GBP instead of 1 USD). At the same time, there was a period of high inflation between the two studies and, thus, our greater completion fee is partially offset by the lower purchasing power. It is also possible that stakes had a different salience or were perceived differently in the two studies since the minimum and suggested hourly rates on Prolific are larger than the typical payment for completing a task on MTurk.

[13] The sample size of 400 per treatment was selected so that we are powered enough to detect the effects reported in Bushong et al. (2021). The details of the power calculations are given in the pre-analysis plan (https://osf.io/hn9de).

of our design that makes integration across attributes difficult is the fact that different tasks are defined with different units of measure (for example, the waiting task is expressed in minutes to wait while the counting and translation tasks are expressed in number of tasks to complete). Intuitively, it is the evaluation of these trade-offs, we conjecture, that give rise to range effects.

Our design is minimal in the following sense. First, 2 attributes is clearly the minimum to detect range effects. Second, while we could have used 3 attributes instead of 4 for the complex choices, it is important for our design that the complex contracts are constructed by appending additional attributes to the 2-attribute contracts. Had we appended only 1 additional attribute, this would have necessarily given only one of $B'$ or $C'$ an additional advantage, unless of course this additional attribute had the same value across $B^{'}$ and $C^{'}$—a feature we wanted to avoid. Hence, we append 2 additional attributes with task 2 being an advantage for $B^{'}$ and task 3 being an advantage for $C^{'}$.

**Time Pressure.** As discussed in Section 1, following our original experiment manipulating complexity through the number of attributes, we decided to design a second experiment to shed light on the mechanism behind the observed results. To this end, we designed a pair of low-attribute treatments in which we increase cognitive load by imposing time pressure.

The new WE and WM treatments were identical to the original 2-attribute treatments, except that each participant was given an additional bonus of £0.5 if their ranking of contracts was submitted within 30 seconds (and their choice was randomly selected to be implemented).[14] The exact amount of time—30 seconds—was chosen with the idea that it would add cognitive load, while still allowing participants enough time to fully understand the contracts and consider both attributes.[15] The time pressure treatments were separately pre-registered after observing the initial data (https://osf.io/nq4zy).

# 4. Theoretical Framework

The formal models of focusing and relative thinking use a common framework. Anticipating the experiment, suppose there are two possible choice sets with 4 contracts each: $X_m$ =

[14] A timer counting down from 30 seconds was displayed on the screen. The ranking could still be submitted after 30 seconds, and failing to submit within 30 seconds had no impact on the probability choices were implemented. The bonus of £0.5 for submitting within 30 seconds is low relative to the payments for completing contracts $B$ and $C$ to dissuade participants from submitting their rankings before considering the contracts carefully.

[15] We chose 30 seconds after observing that the median response time in the original 2-attribute treatments was 35 seconds. Pooling across both time pressure treatments, 66% of participants submitted rankings within 30 seconds.

$\{A_m, B, C, D_m\}$ and $X_e = \{A_e, B, C, D_e\}$. Using $X \in \{X_m, X_e\}$ to denote an arbitrary choice set, the utility to individual participant $i$ from contract $x \in X$ with $K$ attributes is given by

$$u^i(x) = \sum_{k=1}^{K} u_k^i(x_k),$$

where $x_k \in \mathbb{R}$ is contract $x$'s $k$th attribute and $u_k^i(\cdot) \in \mathbb{R}$ is the utility over attribute $k$. Note that utility is assumed to be separable over attributes for simplicity.

However, choice is not determined by that which maximizes utility, but that which maximizes *weighted utility*

$$u^i(x;X) = \sum_{k=1}^{K} w^i(\Delta_k^i) u_k^i(x_k),$$

where $\Delta_k^i$ is the range of utility values in attribute $k$, defined as $\Delta_k^i := \max_{x \in X}(u_k^i(x_k)) - \min_{y \in X}(u_k^i(y_k))$, and $w^i : [0, \infty) \to (0, \infty)$ is a weighting function.

In focusing, the weighting function is assumed to be increasing in the range, while in relative thinking, it is assumed to be decreasing. Hence, the models make opposing assumptions to capture the principles of contrast and normalization, respectively. To be able to identify range effects from data, we make the following assumption, which rules out indifference and imposes that the weighting function is monotonic in range (for a fixed number of attribtues).

**Assumption 1.** For all $i$, (1) $u^i(B;X) \neq u^i(C;X)$ for $X \in \{X_m, X_e\}$ and (2) either

(a) $w^i(\cdot)$ is strictly decreasing, i.e. $i$ is a *relative thinker*,

(b) $w^i(\cdot)$ is strictly increasing, i.e. $i$ is a *focuser*, or

(c) $w^i(\cdot)$ is constant, i.e. $i$ is a *utility maximizer*.

Let $x \succ_X^i y$ denote that individual $i$ ranks $x$ over $y$ in choice set $X \in \{X_m, X_e\}$. The following proposition shows how to identify range effects when the same individual chooses from both $X_m$ and $X_e$, i.e. a within-participant design.

**Proposition 1.** *(1) If $B \succ_{X_m}^i C$ and $C \succ_{X_e}^i B$, then $i$ is a relative thinker. (2) If $C \succ_{X_m}^i B$ and $B \succ_{X_e}^i C$, then $i$ is a focuser. (3) If $\succ_{X_e}^i = \succ_{X_m}^i$, then $i$'s type is ambiguous: they may be a relative thinker, focuser, or utility maximizer.*

Parts (1) and (2) of the proposition show that choice reversals indicate range effects. Part (3) makes clear that individuals who make the same choice in $X_m$ and $X_e$ can be any of the three

types. Clearly, they may be standard utility maximizers, but they may also be relative thinkers or focusers for whom the range effects are insufficient to affect their choices. Note that this is not the same as saying that they have weak range effects, because these individuals may strongly favor $B$ over $C$, say, in the absence of range effects in the sense that $u^i(B) \gg u^i(C)$. In such case, even very strong range effects may be insufficient to affect choice.

In our experiment, however, we use a between-participant design (as motivated in "Identifying Range Effects" above): no individual $i$ chooses from both $X_m$ and $X_e$. What does our experiment allow us to identify about the distribution of range effects in the population?

Let there be some joint distribution over $(u^i, w^i)$ in the population. Let $P^{xy}$ be the induced probability of drawing an individual $i$ with $x \succ^i_{X_m} y$ and $y \succ^i_{X_e} x$, and let $P^x(X)$ be the probability of drawing an individual $i$ with $x \succ^i_X y$ in choice set $X \in \{X_m, X_e\}$. Clearly, *at least* a share $P^{BC}$ of individuals are relative thinkers, and *at least* a share $P^{CB}$ are focusers. The remaining share $P^0 := P^{BB} + P^{CC}$ is comprised of individuals for whom the range manipulation is insufficient to change their choice: they may be either relative thinkers, focusers, or utility maximizers. Notice that because we do not observe individuals choosing from both menus, $P^{xy}$ is not observable, but $P^x(X)$ is.

Clearly, $P^{BC} + P^{CB} + P^{BB} + P^{CC} = 1$, $P^B(X_m) = P^{BC} + P^{BB}$, and $P^B(X_e) = P^{CB} + P^{BB}$. Therefore, we have that $P^B(X_m) - P^B(X_e) = P^{BC} - P^{CB}$. Thus, $P^B(X_m) - P^B(X_e)$ gives the share of individuals who are definitely relative thinkers minus the share of individuals who are definitely focusers. We thus identify *lower bounds* on the shares of individuals who are focusers or relative thinkers by taking the difference across wide money and wide effort in the fraction of participants choosing $B$ over $C$:

**Proposition 2.** *(1) The share of relative thinkers is at least* $P^F := max\{P^B(X_m) - P^B(X_e), 0\}$. *(2) The share of focusers is at least* $P^R := max\{P^B(X_e) - P^B(X_m), 0\}$

We are also interested in how range effects change across 2- and 4-attribute choices. For this, we focus on the *difference-in-difference* statistic $(P^B_4(X'_m) - P^B_4(X'_e)) - (P^B_2(X_m) - P^B_2(X_e))$, where the subscript indicates the number of attributes. Assuming that the unobserved fraction of participants who make consistent choices is the same across 2- and 4-attribute choices, i.e. $P^0_2 = P^0_4$, this statistic is proportional to $(P^{BC}_4 - P^{BC}_2)$—the increase in the share of participants with choice reversals in the focusing direction.[16]

[16] Simple algebra gives that $(P^B_4(X'_m) - P^B_4(X'_e)) - (P^B_2(X_m) - P^B_2(X_e)) = 2(P^{BC}_4 - P^{BC}_2) - (P^0_4 - P^0_2)$.

# 5. Results

## 5.1. Complexity Manipulation

As discussed, we identify range effects in the population from the percentage of individuals ranking $B$ over $C$ (in 2-attribute treatments) and $B'$ over $C'$ (in 4-attribute treatments). We summarize the results in Table 4.

| | | | Wide Effort | Wide Money | Difference |
|---|---|---|---|---|---|
| All participants | Count | 2 att: $B \succ C$ | 193/401 | 189/400 | – |
| | | 4 att: $B' \succ C'$ | 174/401 | 122/400 | – |
| | Percent | 2 att: $B \succ C$ | 48.1% | 47.3% | 0.9% |
| | | 4 att: $B' \succ C'$ | 43.4% | 30.5% | 12.9% |
| Participants ranking A first/D last | Count | 2 att: $B \succ C$ | 157/329 | 170/364 | – |
| | | 4 att: $B' \succ C'$ | 141/318 | 99/354 | – |
| | Percent | 2 att: $B \succ C$ | 47.7% | 46.7% | 1.0% |
| | | 4 att: $B' \succ C'$ | 44.3% | 28.0% | 16.4% |

Table 4: *Summary of Results: 2- and 4- attribute treatments*

In 2-attribute treatments, we find that 48.1% of participants prefer $B$ to $C$ in wide effort, and 47.3% prefer $B$ to $C$ in wide money. The small difference, of 0.9%, is not significant ($p = 0.80$ for difference in proportions), but is in the focusing direction. Hence, our results differ from those of Bushong et al. (2021), who found—despite many similarities in their design—a relative thinking effect of approximately 10%.[17] We discuss this further in Section 6.3.

In 4-attribute treatments, we find that 43.4% of participants prefer $B'$ to $C'$ in wide effort, and 30.5% prefer $B'$ to $C'$ in wide money. Hence, there is a large focusing effect of 12.9%, which is highly significant ($p < 0.001$ for difference in proportions).

We also find that the difference-in-difference, i.e. 12.9% - 0.9% = 12.0%, is highly significant ($p = 0.01$ based on a $t$-test). Hence, the tendency to focus increases with the number of dimensions. This confirms the main hypothesis that focusing emerges as the complexity of choices increases.

The results are even stronger if we only consider participants who respect attribute-wise dominance: among participants who rank $A$ first and $D$ last, we estimate that at least 16.4% of participants are focusers in 4-attribute treatments, whereas the estimate in 2-attribute treatments is

[17] As we show in Appendix Table 10, our estimate of 0.9% is statistically different from the 10% figure reported in Bushong et al. (2021) ($p = 0.068$ based on a $t$-test).

unchanged at 1.0%. Presumably, that the effects get stronger after dropping the participants who fail to rank $A$ first and $D$ last reflects that these participants are noisier, and so their inclusion attenuates effects toward zero. In Appendix 7, we show that all the results of this section are robust to this and various other specifications.

Finally, we note that our interpretation of more attributes as implying greater complexity is validated by response times: the mean response time when ranking 2-attribute work contracts is 44.3 seconds while the mean response time when ranking 4-attribute work contracts is 49.6 seconds. This difference is highly significant based on a two-sided t-test ($p = 0.005$) and the distributions of response times in 2-attribute and 4-attribute treatments are statistically different according to non-parametric tests (the $p$-value of a Kolmogorov–Smirnov test and the $p$-value of a Wilcoxon-Mann–Whitney test are both $< 0.0001$).

## 5.2. Cognitive Load Manipulation

The evidence that contrast (operationalized through focusing) emerges as the number of attributes increases is consistent with the conjecture of Bushong et al. (2021) and with the view that attention is shaped by the complexity of the environment. By contrast, the absence of range effects in our low-attribute treatments calls for explanation. We interpret both contrast and normalization as heuristics triggered by limits in cognitive resources. In very simple environments, such as our two-attribute choices, participants may not be sufficiently constrained for either mechanism to be triggered, leading to a pattern that is indistinguishable from standard utility maximization.

To probe this interpretation, we designed an additional low-attribute treatment that raises cognitive demands by imposing time pressure. Theoretical predictions in this case are ambiguous. If time pressure limits deliberation to a single dimension, we would expect contrast to dominate. If, instead, participants still consider both attributes but must rely on quicker, relative assessments, normalization may prevail. This treatment therefore provides a clean test of how cognitive load shapes the balance between the two principles.

Table 5 presents the results of the time pressure (TP) treatments alongside the results from the original 2-attribute treatments without time pressure (NTP).

In the time pressure treatments, we find that 59.1% of participants prefer $B$ to $C$ in wide effort, and 49.6% prefer $B$ to $C$ in wide money, implying a large a statistically significant focusing effect

| | | | Wide Effort | Wide Money | Difference |
|---|---|---|---|---|---|
| All participants | Count | NTP: $B \succ C$ | 193/401 | 189/400 | – |
| | | TP: $B \succ C$ | 237/401 | 199/401 | – |
| | Percent | NTP: $B \succ C$ | 48.1% | 47.3% | 0.9% |
| | | TP: $B \succ C$ | 59.1% | 49.6% | 9.5% |
| Participants ranking A first/D last | Count | NTP: $B \succ C$ | 157/329 | 170/364 | – |
| | | TP: $B \succ C$ | 199/312 | 165/336 | – |
| | Percent | NTP: $B \succ C$ | 47.7% | 46.7% | 1.0% |
| | | TP: $B \succ C$ | 63.8% | 49.1% | 14.7% |

Table 5: *Summary of Results: With and without time pressure (2 attributes)*

of 9.5% ($p = 0.01$ for difference in proportions). Moreover, we also find that this is statistically different from the (null) effect found in the original treatments: the difference-in-difference estimate of 9.5% - 0.9% = 8.6% is significant ($p = 0.08$ based on a $t$-test). Hence, we find that focusing emerges with the cognitive load induced by time pressure.

The results are even stronger if we only consider participants who respect attribute-wise dominance: among participants who rank $A$ first and $D$ last, the size of the focusing effect becomes 14.7% ($p < 0.001$) and the difference-in-difference, capturing the effect of the time pressure, becomes 14.7% - 1.0% = 13.7% ($p = 0.01$). In Appendix 7, we show that all the results of this section are robust to this and various other specifications.

# 6. Discussion

## 6.1. Toward a Unified Model

We have found that focusing emerges under increased complexity and cognitive load. Future work should develop microfounded models that predict such effects endogenously. Short of this goal, we provide a reduced-form framework that can incorporate both focusing and relative thinking, and can account for our treatment effects. We then discuss possible microfoundations.

We suppose that participants decide which attributes to focus on, and then, conditional on attending to a given attribute, may use the relative thinking heuristic to assess value. This suggests that the *overall weighting function* $w : [0,\infty) \to (0,\infty)$ may take a multiplicative form $w = f \cdot r$, where $f : [0,\infty) \to (0,\infty)$ gives the *focusing weights*, and $r : [0,\infty) \to (0,\infty)$ gives the *relative thinking weights*.[18] By assumption, $f$ is (weakly) increasing, and $r$ is (weakly) decreasing. We

[18]This multiplicative formulation follows Online Appendix D of Bushong et al. (2021), who propose a specific para-

suppose that the propensity toward focusing and relative thinking may depend on the number of attributes $K$ and a parameter $\theta \in \mathbb{R}$, representing cognitive load. Hence, the overall weight placed on an attribute with range $\Delta$ is given by $w(\Delta;K,\theta) = f(\Delta;K,\theta)r(\Delta;K,\theta)$.

We begin by formally defining an order on the space of weighting functions that allows us to say whether there is an increased tendency toward focusing or relative thinking.

**Definition 1.** For any two functions $h, g : [0,\infty) \to (0,\infty)$, we write $h \succ g$ to mean that $h(\delta + \epsilon)/h(\delta) > g(\delta+\epsilon)/g(\delta)$ for all $\delta \in [0,\infty)$ and $\epsilon \in (0,\infty)$, and we write $h \sim g$ to mean that $h(\delta) = \alpha \cdot g(\delta)$ for all $\delta \in [0,\infty)$ and some constant $\alpha > 0$. Going from focusing weight $f$ to $f'$, if $f' \succsim f$, we say there is an *an increased tendency to focus*. Similarly, going from relative thinking weight $r$ to $r'$, if $r \succsim r'$, there is an *increased tendency toward relative thinking*.

The order $\succsim$ is a partial order in general, but for simplicity, we make the following completeness assumption to aid identification.

**Assumption 2.** Focusing and relative thinking weights are always ordered as we vary $\theta$ and $K$. That is, for all $h \in \{f, r\}$, $K$, $K'$, $\theta$, and $\theta'$ (1) $h(\cdot;K,\theta) \succsim h(\cdot;K,\theta')$ or $h(\cdot;K,\theta') \succsim h(\cdot;K,\theta)$, and (2) $h(\cdot;K,\theta) \succsim h(\cdot;K',\theta)$ or $h(\cdot;K',\theta) \succsim h(\cdot;K,\theta)$.

We now interpret our experimental results through the lens of this framework.

By increasing the number of attributes from $K = 2$ to $K = 4$, our results suggest that $w(\cdot;4,\theta) \succ w(\cdot;2,\theta)$. A priori, this only rules out that there is simultaneously an increased tendency toward relative thinking ($r(\cdot;2,\theta) \succsim r(\cdot;4,\theta)$) *and* a reduced tendency toward focusing ($f(\cdot;2,\theta) \succsim f(\cdot;4,\theta)$). However, if we think of focusing as the process of determining which attributes to attend to and relative thinking as the process of assessing attribute values *conditional on* attending to a given attribute, relative thinking may not depend on $K$ at all, i.e. $r(\cdot;K,\theta) \sim r(\cdot;K',\theta)$ for all $K$ and $K'$. Under this assumption, we may conclude that there is an increased tendency to focus: $f(\cdot;4,\theta) \succ f(\cdot;2,\theta)$.

By increasing time pressure, we increase cognitive load from $\theta$ to $\theta' > \theta$ and find that $w(\cdot;2,\theta') \succ w(\cdot;2,\theta)$. As before, this does not pin down the effects of the treatment on either $f$ or $r$. In general, we think that increasing cognitive load has the tendency to exacerbate bias. This is because we imagine agents exerting costly effort to overcome biases, and the effect of cognitive load is to increase such costs. In our case, we imagine the agent as exerting costly effort to

metric form of focusing weights to illustrate how focusing can dominate with many attributes.

reduce the focusing and/or relative thinking biases as part of some joint optimization, and so the effects of cognitive load will be sensitive to the nature of the cost function. However, if we assume that increasing cognitive load weakly increases both biases ($r(\cdot;2,\theta) \gtrsim r(\cdot;2,\theta')$ and $f(\cdot;2,\theta') \gtrsim f(\cdot;2,\theta)$), then we conclude that, once again, there is an increased tendency to focus (that dominates any increase in relative thinking): $f(\cdot;2,\theta') \succ f(\cdot;2,\theta)$.

## 6.2. Relationship with (Rational and Behavioral) Inattention

Our results are broadly consistent with models of *inattention* whereby the agent pays more attention to, and thus overweights, certain attributes as a reaction to limited cognitive resources. We argue that the results can be explained by both fast and intuitive "System 1" thinking, as well as slow and deliberate "System 2" thinking (Kahneman, 2013), with both systems potentially playing an important role.

With time pressure, it is as if the agent pays more attention to attributes with larger ranges. To explain this, we note that if an attribute's range proxies for the prior variance of attribute values and inspecting a given attribute is costly, then it is rational (System 2) to pay more attention to those attributes with the highest range as these are the most consequential ex-ante.[19] In the context of such a model, increasing time pressure is similar to increasing the cost of inspecting attributes. As a result, the ex-ante less important attributes—those with smaller ranges—will be inspected less, getting less weight. Reinforcing this rational theory, the attributes with the highest ranges are likely the most salient and instinctively attention-grabbing, and therefore favored by System 1.[20] In such case, time pressure leads to more weight on the high-range attributes because there is simply less time to pass to System 2.

By increasing the number of attributes, it is as if the agent pays more attention to attributes with a larger range. A rational (System 2) explanation is that, with more attributes, it is costly to determine which attributes have the largest range in the first place. This may exhaust resources that can be put toward inspecting the ex-ante less important (low range) attributes. An automatic (System 1) explanation is that, as the number of attributes increases, the ones with the largest

[19] This feature recalls models of *rational inattention* which typically predict that more attention is paid to more volatile (i.e., with greater prior uncertainty) variables. In their review, Maćkowiak et al. (2023) refer to this as Feature F6. This feature is also broadly consistent with dynamic models of search over multi-attribute goods (e.g. Sanjurjo (2017) and Safonov (2024)).

[20] Salience theory (Bordalo et al., 2012, 2013, 2022) posits that different objects have potentially different salient attributes, with salience increasing in the distance between attribute value and a reference point. In a similar vein, we imagine salience applied on the level of attribute.

ranges simply become more salient, drawing more undirected attention.

While we have not found evidence for relative thinking, the framework of Section 6.1 makes clear that both focusing and relative thinking may co-exist: relative thinking is a heuristic for assessing attribute value after focusing on a given attribute. In the next subsection, we discuss when relative thinking may be more empirically relevant.

## 6.3. Relationship with Bushong, Rabin and Schwarzstein (2021)

Our 2-attribute treatments (without time pressure) are a conceptual replication of Bushong et al. (2021), who found a significant relative thinking effect of 10.8-14.4% when comparing behavior in their wide money and wide effort treatments,[21] whereas we found a null effect in our baseline treatments and a significant focusing effect of 9.5-14.7% with time pressure. In all cases, our effects are statistically different from those reported in Bushong et al. (2021). What accounts for these differences?

There are differences between our baseline 2-attribute treatments and the experiment of Bushong et al. (2021), as motivated and described in Section 3. We view differences in the value of money (pounds versus dollars, inflation considerations, etc.) and other minor procedural differences as unlikely explanations for the different findings. We consider two other possibilities in turn.

First, unlike for the money attributes, the task attributes differ more substantially across the experiments. Relative to Bushong et al. (2021), we simplified the tasks and reduced their number (as motivated in Section 3). Their task-range manipulation was also larger in absolute terms, though very similar in relative terms.[22] Hence, in Bushong et al. (2021), the task attribute was relatively more important—and its range manipulation stronger—which may have been necessary to evoke relative thinking. Table 9 of Appendix 7 compares our contracts side-by-side with those of Bushong et al. (2021).

Second, whereas Bushong et al. (2021) ran their experiments on Mturk, we ran ours on Prolific. It has been shown that, in other contexts, MTurk samples tend to be much noisier than those on Prolific (Peer et al., 2022), suggesting the possibility that their participants were naturally more cognitively constrained than ours. As we conjecture that cognitive constraints are necessary for relative thinking (as well as focusing), we delve deeper. Bushong et al. (2021) only recruited people who participated in at least 100 prior tasks on MTurk and had an approval rating of 95%,

[21] That is, 10.8% in the full sample and 14.4% among participants who ranked $A$ first and $D$ last.
[22] The ranges were 3 (WM) and 19 (WE) in our treatments, compared to 4 (WM) and 28 (WE) in Bushong et al. (2021).

so it is unclear a priori whether their participants were actually more cognitively constrained. However, their participants submitted their rankings *much* faster than ours, after 19 seconds on median, as opposed to 35 seconds in our case.[23] This suggests the possibility that their participants were more cognitively constrained, perhaps due to a greater opportunity cost of time, causing endogenous time pressure.

If what is driving the difference between our results and those of Bushong et al. (2021) is that our participants are not sufficiently cognitively constrained, one might conjecture that our time pressure treatment would push participants toward relative thinking. However, we find just the opposite (Section 5.2). This suggests that cognitive constraints cannot easily reconcile the two sets of findings.

Of course, cognitive constraints are one of many dimensions of heterogeneity. We embrace the possibility that there may simply be a heterogeneity of "types" in the population, and for whatever reason, the sample of Bushong et al. (2021) involves more individuals who are inclined toward relative thinking.

Hence, while we can only speculate about what is driving the differences between our 2-attribute results and those in Bushong et al. (2021), we conclude that whether or not this paradigm evokes relative thinking is sensitive to the details of the design and/or subject pool. On the other hand, our finding that increased complexity and cognitive load push people toward focusing is likely to be more robust. This is because even individuals who are naturally inclined toward relative thinking may act as focusers if they cannot fully consider each attribute carefully.

# 7. Conclusion

Almost every choice we make is between options with many attributes. Because such choices are complex, people are guided by simplified heuristics. Two recurring principles in the literature are *normalization* and *contrast*. Models based on the normalization heuristic posit that people ascertain value by thinking in relative terms. In this case, fixed differences appear less significant when the range of values is larger, and so people underweight attributes with larger ranges. Models based on the contrast heuristic, on the other hand, posit that people attend to attributes that stand out relative to the context: attributes with larger ranges are particularly salient and thus receive more weight. These heuristics make opposite assumptions and divergent predictions.

[23] Despite the quicker response times, Bushong et al. (2021) found a nearly identical fraction of participants ranking $A$ first and $D$ last (88% as opposed to our 87%).

Both have found found empirical support, albeit in very different contexts. What is driving the differences?

In our experiment, we test the natural hypothesis that normalization dominates in simple choices, while contrast dominates in complex ones. We find that, indeed, contrast (operationalized in our design by focusing à la Kőszegi and Szeidl 2013) emerges as complexity—as measured by the number of attributes—increases. We also show that cognitive load, induced by time pressure, leads to contrast effects in otherwise simple environments.

Our results suggest that, when choices are complex, people attend to those features that are the most consequential, which in our context are the attributes with the largest ranges. This may, in part, be due to undirected attention—that such attributes naturally catch the eye. However, if it were not costly to process attributes, one would expect all attributes to be considered before choices are made, in which case all may be equally weighted. Hence, we think our results are best explained by a form of rational (costly) attention.

Our findings imply that existing models of range effects, while organizing behavior well, do not tell the whole story. Any model of range-dependent attribute-weighting with a monotonic weighting function that does not depend on the number of attributes would be unable to explain our data. To do so, one would have to model the relationship between complexity and attribute-weighting. More broadly, our results suggest that complexity is an important factor shaping biases and attention. Future work should consider other aspects of complexity, including the number of alternatives and the entire distribution—not simply the range—of attribute values.

# References

**Andersson, Ola, Jim Ingebretsen Carlson, and Erik Wengström.** 2021. "Differences Attract: An Experimental Study of Focusing in Economic Choice." *The Economic Journal* 131 (639): 2671–2692.

**Arrieta, Gonzalo, and Kirby Nielsen.** 2024. "Procedural Decision-Making in the Face of Complexity."

**Azar, Ofer H, and Alisa Voslinsky.** 2024. "Examining Relative Thinking in Mixed Compensation Schemes: A Replication Study." *Journal of Economic Behavior & Organization* 218 568–578.

**Ba, Cuimin, J Aislinn Bohren, and Alex Imas.** 2024. "Over- and Underreaction to Information."

**Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer.** 2012. "Salience Theory of Choice Under Risk." *The Quarterly Journal of Economics* 127 (3): 1243–1285.

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

**Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer.** 2022. "Salience." *Annual Review of Economics* 14 (1): 521–544.

**Bucher, Stefan, and Adam Brandenburger.** 2022. "Divisive Normalization Is an Efficient Code for Multivariate Pareto-Distributed Environments." *Proceedings of the National Academy of Sciences* 119 (40): e2120581119.

**Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein.** 2021. "A Model of Relative Thinking." *The Review of Economic Studies* 88 (1): 162–191.

**Castillo, Geoffrey.** 2020. "The Attraction Effect and Its Explanations." *Games and Economic Behavior* 119 123–147.

**de Clippel, Geoffroy, Paola Moscariello, Pietro Ortoleva, and Kareen Rozen.** 2024. "Caution in the Face of Complexity."

**Deck, Cary, Salar Jahedi, and Roman Sheremeta.** 2021. "On the Consistency of Cognitive Load." *European Economic Review.*

**Dertwinkel-Kalt, Markus, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang.** 2022. "Concentration Bias in Intertemporal Choice." *The Review of Economic Studies* 89 (3): 1314–1334.

**Dertwinkel-Kalt, Markus, Katrin Köhler, Mirjam RJ Lange, and Tobias Wenzel.** 2017. "Demand Shifts Due to Salience Effects: Experimental Evidence." *Journal of the European Economic Association* 15 (3): 626–653.

**Dertwinkel-Kalt, Markus, and Mats Köster.** 2020. "Salience and Skewness Preferences." *Journal of the European Economic Association* 18 (5): 2057–2107.

**Enke, Benjamin, and Thomas Graeber.** 2023. "Cognitive Uncertainty." *The Quarterly Journal of Economics* 138 (4): 2021–2067.

**Enke, Benjamin, Thomas Graeber, and Ryan Oprea.** Forthcoming. "Complexity and Time." *Journal of the European Economic Association.*

**Enke, Benjamin, Thomas Graeber, Ryan Oprea, and Jeffrey Yang.** 2024. "Behavioral Attenuation."

**Enke, Benjamin, and Cassidy Shubatt.** 2024. "Quantifying Lottery Choice Complexity."

**Fallucchi, Francesco, and Marc Kaufmann.** 2021. "Narrow Bracketing in Work Choices."

**Gabaix, Xavier.** 2014. "A Sparsity-Based Model of Bounded Rationality." *The Quarterly Journal of Economics*.

**Glimcher, Paul W, and Agnieszka A Tymula.** 2023. "Expected Subjective Value Theory (ESVT): A Representation of Decision Under Risk and Certainty." *Journal of Economic Behavior & Organization* 207 110–128.

**Guan, Menglong, Ryan Oprea, and Sevgi Yuksel.** 2025. "Beyond Instrumental Value: How Complexity Shapes Information Demand."

**Hirshman, Samuel, Devin Pope, and Jihong Song.** 2018. "Mental Budgeting versus Relative Thinking." In *AEA Papers and Proceedings*, Volume 108. 148–152, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

**Kahneman, Daniel.** 2013. *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

**Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman.** 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62 (12): 3393–3411.

**Khaw, Mel W, Paul W Glimcher, and Kenway Louie.** 2017. "Normalized Value Coding Explains Dynamic Adaptation in the Human Valuation Process." *Proceedings of the National Academy of Sciences* 114 (48): 12696–12701.

**Kőszegi, Botond, and Adam Szeidl.** 2013. "A Model of Focusing in Economic Choice." *The Quarterly Journal of Economics* 128 (1): 53–104.

**Landry, Peter, and Ryan Webb.** 2021. "Pairwise Normalization: A Neuroeconomic Theory of Multi-Attribute Choice." *Journal of Economic Theory* 193 105221.

**Lanzani, Giacomo.** 2022. "Correlation Made Simple: Applications to Salience and Regret Theory." *The Quarterly Journal of Economics* 137 (2): 959–987.

**Louie, Kenway, Mel W Khaw, and Paul W Glimcher.** 2013. "Normalization Is a General Neural Mechanism for Context-Dependent Decision Making." *Proceedings of the National Academy of Sciences* 110 (15): 6139–6144.

**Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt.** 2023. "Rational Inattention: A Review." *Journal of Economic Literature* 61 (1): 226–273.

**Mormann, Milica, and Cary Frydman.** 2018. "The Role of Salience and Attention in Choice under Risk: An Experimental Investigation."

**Nunnari, Salvatore, and Jan Zápal.** 2025. "A Model of Focusing in Political Choice." *Journal of Politics* 87 (4): 1465–1481.

**Oprea, Ryan.** 2024a. "Complexity and its Measurement." *Handbook of Experimental Methods in the Social Sciences.*

**Oprea, Ryan.** 2024b. "Decisions under Risk Are Decisions under Complexity." *American Economic Review* 114 (12): 3789–3811.

**Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer.** 2022. "Data Quality of Platforms and Panels for Online Behavioral Research." *Behavior Research Methods* 54 (4): 1643–1662.

**Puri, Indira.** 2025. "Simplicity and Risk." *The Journal of Finance* 80 (2): 1029–1080.

**Safonov, Evgenii.** 2024. "Slow and Easy: a Theory of Browsing." *Working paper.*

**Sanjurjo, Adam.** 2017. "Search With Multiple Attributes: Theory and Empirics." *Games and Economic Behavior* 104 535–562.

**Shah, Anuj K, Eldar Shafir, and Sendhil Mullainathan.** 2015. "Scarcity Frames Value." *Psychological Science* 26 (4): 402–412.

**Soltani, Alireza, Benedetto De Martino, and Colin Camerer.** 2012. "A Range-Normalization Model of Context-Dependent Choice: A New Model and Evidence." *PLoS Computational Biology* 8 (7): e1002607.

**Somerville, Jason.** 2022. "Range-Dependent Attribute Weighting in Consumer Choice: An Experimental Test." *Econometrica* 90 (2): 799–830.

**Strulov-Shlain, Avner.** 2023. "More Than a Penny's Worth: Left-Digit Bias and Firm Pricing." *The Review of Economic Studies* 90 (5): 2612–2645.

**Taubinsky, Dmitry, and Alex Rees-Jones.** 2018. "Attention Variation and Welfare: Theory and Evidence From a Tax Salience Experiment." *The Review of Economic Studies* 85 (4): 2462–2496.

**Webb, Ryan, Paul W Glimcher, and Kenway Louie.** 2020. "Divisive Normalization Does Influence Decisions With Multiple Alternatives." *Nature Human Behaviour* 4 (11): 1118–1120.

**Woodford, Michael.** 2012. "Inattentive Valuation and Reference-Dependent Choice."

# Appendix

## Additional Tables

| | 2 Attributes | | | 4 Attributes | | | Time Pressure | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Wide Effort | 0.009 | 0.010 | 0.035 | 0.129*** | 0.164*** | 0.558*** | 0.095** | 0.147*** | 0.383*** |
| | (0.035) | (0.038) | (0.141) | (0.034) | (0.037) | (0.148) | (0.035) | (0.039) | (0.142) |
| Constant | 0.473*** | 0.467*** | -0.110 | 0.305*** | 0.280*** | -0.824*** | 0.496*** | 0.491*** | -0.015 |
| | (0.025) | (0.026) | (0.100) | (0.024) | (0.025) | (0.109) | (0.025) | (0.027) | (0.100) |
| Observations | 801 | 693 | 801 | 801 | 672 | 801 | 802 | 648 | 802 |
| Model | LPM | LPM | Logit | LMP | LMP | Logit | LPM | LPM | Logit |
| Sample | All | A First/D Last | All | All | A First/D Last | All | All | A First/D Last | All |

Table 6: *Robustness of Range Effects.* For each of 2-attribute, 4-attribute, and (2-attribute) time pressure treatments, we estimate range effects as the difference in the fraction of participants preferring $B$ to $C$ in wide effort versus wide money treatments using three different specifications: linear probability model (LPM), LPM after dropping participants who fail to rank $A$ first and $D$ last, and logistic regression. A positive sign on the wide effort coefficient indicates focusing.

| | (1) | (2) | (3) |
|---|---|---|---|
| Wide Effort | 0.009 | 0.010 | 0.035 |
| | (0.035) | (0.038) | (0.142) |
| 4 Attributes | -0.167*** | -0.187*** | -0.713*** |
| | (0.034) | (0.035) | (0.148) |
| Wide Effort × 4 Attributes | 0.120** | 0.154** | 0.522** |
| | (0.049) | (0.053) | (0.205) |
| Constant | 0.472*** | 0.467*** | -0.110 |
| | (0.025) | (0.026) | (0.100) |
| Observations | 1602 | 1365 | 1602 |
| Model | LPM | LPM | Logit |
| Sample | All | A First/D Last | All |

Table 7: *Robustness of Difference-in-Differences Analysis: 2 versus 4 Attributes.* We regress an indicator of ranking $B$ over $C$ on indicators for wide effort, 4 attributes, and the interaction between wide effort and 4 attributes. The coefficient on the interaction term gives the difference-in-difference estimate, with a positive sign indicating an increase in focusing with 4 attributes. Each column is based on one of three different specifications: linear probability model (LPM), LPM after dropping participants who fail to rank $A$ first and $D$ last, and logistic regression.

| | (1) | (2) | (3) |
|---|---|---|---|
| Wide Effort | 0.009 | 0.010 | 0.035 |
| | (0.035) | (0.038) | (0.142) |
| Time Pressure | 0.024 | 0.024 | 0.095 |
| | (0.035) | (0.038) | (0.141) |
| Wide Effort × Time Pressure | 0.086* | 0.137** | 0.348* |
| | (0.050) | (0.054) | (0.201) |
| Constant | 0.473*** | 0.467*** | -0.110 |
| | (0.025) | (0.026) | (0.100) |
| Observations | 1603 | 1341 | 1603 |
| Model | LPM | LPM | Logit |
| Sample | All | A First/D Last | All |

Table 8: *Robustness of Difference-in-Differences Analysis: (2-Attribute) Time Pressure versus (2-Attribute) No Time Pressure.* We regress an indicator of ranking $B$ over $C$ on indicators for wide effort, time pressure, and the interaction between wide effort and time pressure. The coefficient on the interaction term gives the difference-in-difference estimate, with a positive sign indicating an increase in focusing with time pressure. Each column is based on one of three different specifications: linear probability model (LPM), LPM after dropping participants who fail to rank $A$ first and $D$ last, and logistic regression.

Bushong et. al. (2021)

| | Wide Money | | | Wide Effort | |
|---|---|---|---|---|---|
| | money | task 1 | | money | task 1 |
| $A_m$ | 4.50 | 14 | $A_e$ | 2.80 | 2 |
| $B$ | 2.20 | 14 | $B$ | 2.20 | 14 |
| $C$ | 2.80 | 18 | $C$ | 2.80 | 18 |
| $D_m$ | 0.50 | 18 | $D_e$ | 2.20 | 30 |

Current paper

| | Wide Money | | | Wide Effort | |
|---|---|---|---|---|---|
| | money | task 1 | | money | task 1 |
| $A_m$ | 4.80 | 9 | $A_e$ | 2.70 | 1 |
| $B$ | 2.20 | 9 | $B$ | 2.20 | 9 |
| $C$ | 2.70 | 12 | $C$ | 2.70 | 12 |
| $D_m$ | 0.10 | 12 | $D_e$ | 2.20 | 20 |

Table 9: *2-Attribute Contracts in Bushong et al. (2021)*

| | (1) |
|---|---|
| Wide Effort | 0.009 |
| | (0.035) |
| BRS | $0.172^{***}$ |
| | (0.039) |
| Wide Effort × BRS | $-0.116^{**}$ |
| | (0.054) |
| Constant | $0.473^{***}$ |
| | (0.025) |
| Observations | 1371 |
| Model | LPM |
| Sample | All |

Table 10: *Comparison between Our 2-Attribute (No Time Pressure) Treatments and Data from Bushong et al. (2021) (BRS).* We regress an indicator of ranking $B$ over $C$ on indicators for wide effort, whether the study is BRS, and the interaction between wide effort and BRS. The coefficient on the interaction term gives the difference in range effects between BRS and the current paper. The negative sign indicates that there is more relative thinking in BRS.

# Experimental Interface

## Instructions

This study has 3 parts and will earn you a payment of £2.00:

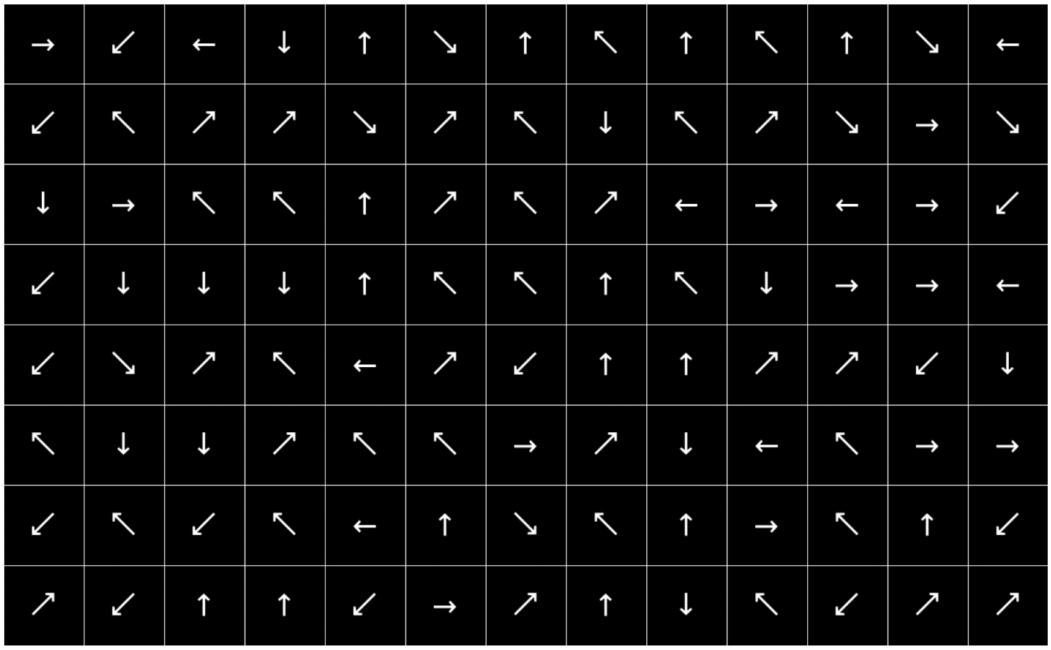
- In Part 1, you will complete work that takes around 3 minutes.
- In Part 2, we will ask about your preferences for doing additional work for a bonus payment.
- In Part 3, you may have to complete additional work for a bonus payment (depending on answers in Part 2).

You must complete all parts to earn any pay for this study.

---

### Counting Task

In Parts 1 and 3, your job is to count elements in an image. In each page, you will see an image like this:



You will then be asked to count how many times a given element appears in the image.

For example, you might be asked to count how many → there are in the image.

Both the symbols in the image and the symbol you are asked to count will change in each page, so pay close attention.

You must type the correct answer in order to advance to the next page.

If you type an incorrect answer, you will have to wait 10 seconds and then correctly count symbols in a new image before advancing to the next page.

---

In Part 1, you will complete 3 images of the counting task.

When you click to advance to the next page, you will begin Part 1.

Figure 1: Screen with General and Part 1 Instructions, 2-Attribute Treatments

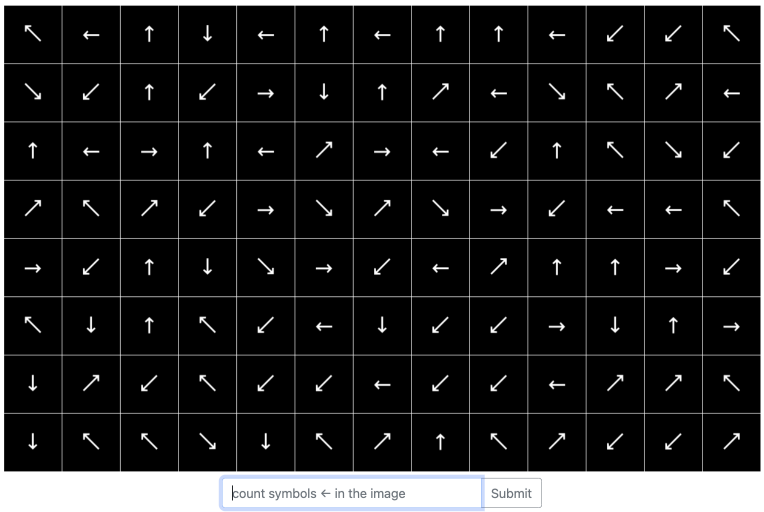



Figure 2: Screen with Sample Counting Task (Task 1)

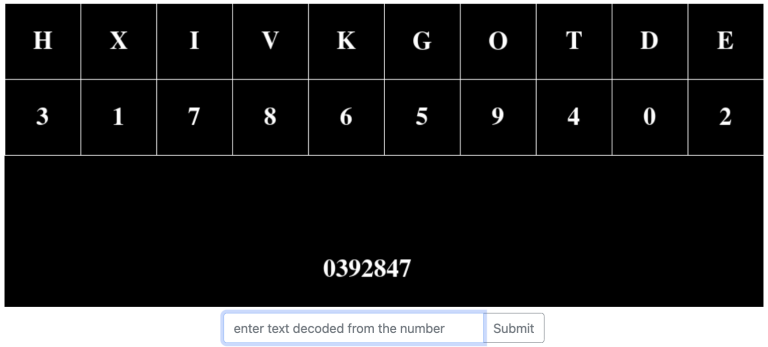



Figure 3: Screen with Sample Translation Task (Task 2)

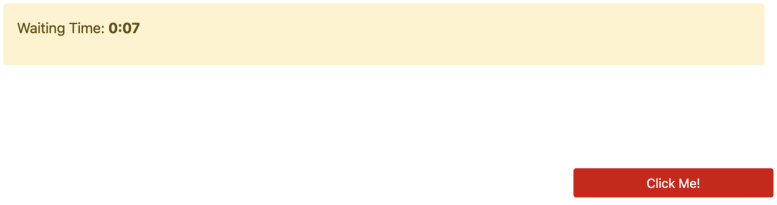



Figure 4: Screen with Sample Waiting Task (Task 3)

## Instructions for Part 2

We will now ask you one question about your willingness to complete additional tasks to increase your earnings.

The task will not change from your experience in Part 1, except that you will see different images.

In order for us to best understand your willingness to complete additional tasks for additional pay, we will show you four options and ask you to rank them. Each option consists of some amount of counting tasks that you must complete and some fixed payment (in pounds). You must rank them by using the mouse and dragging the options into your preferred order (where 1 is your most preferred option and 4 is your least preferred option).

In order to make sure you take the ranking seriously, we will use a particular system.

Out of every ten participants in this study, the computer will randomly choose one of them.

- If you are one of the selected participants, you continue to Part 3:
  - At the beginning of Part 3, the computer will choose two of the available options at random.
  - We will then implement your ranking between these two options: you will complete the number of tasks in the option you assigned a higher ranking to (and receive the corresponding additional payment).
- If you are not one of the selected participants, you do not continue to Part 3 and the study is over.

This may seem complicated, but there is a simple way to maximize the chance you receive your most-preferred outcome: simply answer truthfully, that is, rank the options according to your actual preferences.

---

### Example

This example is for illustrative purposes only.

Suppose you were asked to rank the following options:

- Kiwi
- Watermelon
- Pear
- Strawberry

If you entered a ranking of 1) Strawberry, 2) Kiwi, 3) Watermelon, and 4) Pear and the computer randomly chose (Pear, Kiwi), then you would get a Kiwi. This is because you ranked Kiwi (2nd) above Pear (4th) and, thus, you stated that you prefer a Kiwi to a Pear.

---

### Comprehension Quiz

We will now ask you one question to see if you have understood the instructions. You can attempt twice. If you answer this question incorrectly both times, the study will end and you will not be entitled to any payment.

*Suppose you were asked to rank Apple, Banana, Orange, and Pineapple. If you entered a ranking of 1) Banana, 2) Apple, 3) Pineapple, 4) Orange and the computer randomly chose (Pineapple, Apple), which would you get?*

- ○ *Banana*
- ○ *Orange*
- ○ *Apple*
- ○ *Pineapple*

Figure 5: Screen with Part 2 Instructions and Comprehension Quiz, 2-Attribute Treatments

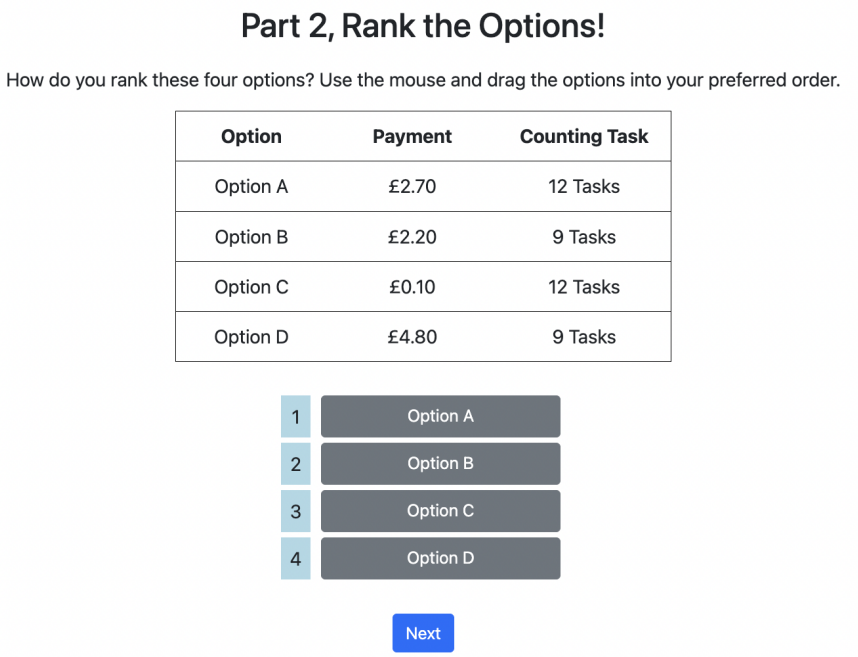



Figure 6: Decision Screen (Ranking Work Contracts), 2 Attributes

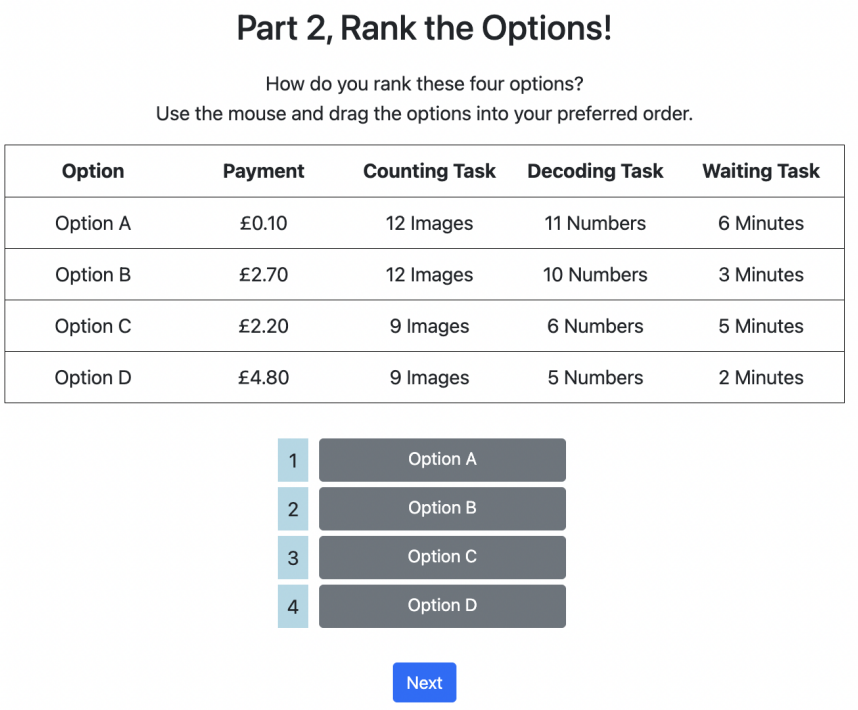



Figure 7: Decision Screen (Ranking Work Contracts), 4 Attributes