Visual judgments of length in the economics laboratory: Are there brains in stochastic choice?

Duffy, Sean and Gussman, Steven and Smith, John

Rutgers University-Camden

1 March 2021

Visual judgments of length in the economics laboratory: Are there brains in stochastic choice?*

Sean Duffy† Steven Gussman‡ John Smith§

March 1, 2021

Abstract

We design an induced value choice experiment where the objects are valued according to only a single attribute with a continuous measure. Subjects have an imperfect perception of the choice objects but can reduce their imperfect perception with cognitive effort. Subjects are given a choice set involving several lines of various lengths and are told to select one of them. They strive to select the longest line because they are paid an amount that is increasing in the length of their selection. This "idealized" choice experiment produces a dataset that is uniquely suited to study apparently random choice. We also manipulate the available cognitive resources of the subjects by imposing either a high or low cognitive load. We find that both choices and the allocation of effort are affected by the material incentives in the choice problem and the available cognitive resources. We find evidence that optimal choices have shorter deliberation times than suboptimal choices, which is consistent with previous theoretical predictions. The distribution of errors can have significant implications for the specification of stochastic choice models. Specifications where errors have a Gumbel distribution appear to provide a better fit than those with a normal distribution. Despite that the cognitive load manipulation affects both choice and search, it is notable that neither the Gumbel distribution results nor the relationship between optimal choice and deliberation time appear to be affected by the available cognitive resources. This perhaps suggests that these results are general and persistent features of choice.

Keywords: cognitive load, choice overload, memory, search, imperfect perception

JEL: C91, D12

*We thank Carlos Alós-Ferrer, William Appleman, Roberto Barbera, Rahul Bhui, Oriol Carbonell-Nicolau, I-Ming Chiu, David Dickinson, Noha Emara, Andrea Isoni, Tomás Lejarraga, Jacopo Magnani, Marco Mantovani, Andreu Mas-Collel, Richard McLean, Nick Netzer, Tomas Sjöström, Barry Sopher and Michael Woodford for helpful comments. We thank Jackie Dunn, Lisa Saal, and Bob Young for help with the experiment. John Smith thanks Biblioteca de Catalunya. This research was supported by Rutgers University Research Council Grants #18-AA-00143 and #FPRC0000016.

†Rutgers University-Camden, Department of Psychology

‡Rutgers University-Camden, Digital Studies Center

§Corresponding Author; Rutgers University-Camden, Department of Economics, 311 North 5th Street, Camden, New Jersey, USA 08102; Email: smithj@camden.rutgers.edu; Phone: +1-856-225-6319

1 Introduction

We design a choice experiment where objective values are perfectly observable, we can observe important aspects of the decision process, and we vary the cognitive resources available to the subject. This dataset is uniquely suited to answer questions such as: Do available cognitive resources affect choice or deliberation? Are optimal choices faster than sub-optimal choices? Do available cognitive resources affect this relationship? Can we say anything about the distribution of errors? Do available cognitive resources affect the distribution of errors?

Consider a subject making a binary choice between a bag of potato chips and a can of soda. The choice from this set would allow the experimenter to conduct an inference of the preferences of the subject. However, this inference is noisy and it is not straightforward to detect a suboptimal choice.

If preferences are also elicited by a supplementary method (for example, eliciting either willingness to pay or a ranking of the objects) the experimenter could compare the choice with this alternate measure. However, both the choice and the supplementary elicitation are noisy. In the case that preferences are not elicited by a different method, the experimenter would only be able to identify that a suboptimal action was taken when intransitive choices were made. In contrast, we design an experiment where we are able to determine–without noise–whether subjects selected a suboptimal action.

For some time, economists have been conducting induced value experiments, where the experimenter imposes a value on various outcomes.[1] We distinguish our experiment from the majority of these induced value experiments, as subjects in our experiment imperfectly perceive the objective values of the objects and this produces a dataset that is uniquely suited to study stochastic choice.

The objects of choice are lines of various lengths. Subjects attempt to select the longest line because they are paid an amount that is increasing in the length of their selection. While we are able to observe the true objective length of each line, it is well-known that subjects have an imperfect perception of objectively measurable objects (Weber, 1834; Fechner, 1860; Thur-

[1] For example, see Smith (1976).

stone, 1927a,b). In other words, even where objects have objectively measurable properties, perception of them is imperfect.[2,3]

Certain regularities regarding imperfect perception have been known for some time. Perhaps the oldest regularity is that the larger the stimuli, the more difficult it is to detect absolute differences between stimuli (Fechner, 1860). For example, it is often more difficult to determine the heaviest between a 5kg object and a 5.5kg object than it is to determine the heaviest between a 1kg object and a 1.5 kg object. This regularity is sometimes referred to as *Weber's Law*.

Further, the imperfect perception of objective quantities has led researchers to consider that one's preferences might be imperfectly perceived and this has served as a justification for random choice or random utility models. For instance, Bradley and Terry (1952), Luce (1959a,b), Becker, DeGroot, and Marschak (1963), McFadden (1974, 1976, 1981, 2001), Yellott (1977), and Falmagne (1978) each make explicit reference to Weber, Fechner, or Thurstone.[4] However, despite this known connection between imperfect perception of objective properties and stochastic choice, to our knowledge, our paper and Duffy and Smith (2020) are the only examples of incentivized experiments where suboptimal choices are perfectly observable because utility is represented by a static, single-attribute physical quantity with an uncountable measure.[5]

Subjects can only view one line at a time. This design simulates the feature that deliberation about the desirability of an object compared to another object crucially involves the memory of the assessments of the objects. This design also allows us to observe the search history of subjects.

[2] The vast majority of this literature conducts experiments that are not incentivized. Below, we discuss the exceptions.

[3] Researchers have been studying the judgments of the lengths of lines for some time (Münsterberg, 1894; Cattell, 1902).

[4] More recently, these authors have been cited by Machina (1985), Luce (1994, 2005), Mas-Colell, Whinston, and Green (1995), Ballinger and Wilcox (1997), Loomis et al. (1998), Butler (2000), Butler and Loomes (2007), Blavatskyy (2008, 2011), Rieskamp (2008), Caplin (2012), Lévy-Garboua et al. (2012), Fudenberg, Iijima, and Strzalecki (2015), Caplin (2016), Agranov and Ortoleva (2017), Argenziano and Gilboa (2017), Khaw, Li, and Woodford (2017), Navarro-Martinez et al. (2018), Cerreia-Vioglio et al. (2019), Horan, Manzini, and Mariotti (2019), Olschewski, Newell, and Scheibehenne (2019), Alós-Ferrer and Garagnani (2020), Caplin et al. (2020), and Alós-Ferrer, Fehr, and Netzer (2021).

[5] This design also suggests that there will not be an undetected relationship between one of several attributes from a previous choice and one of several attributes of a subsequent choice.

Subjects make their choice when under a cognitive load. This experimental manipulation is designed to affect the available cognitive resources of subjects, so that the relationship between cognition and behavior can be observed.[6] Some choices are made when required to remember a 6-digit number (high cognitive load) and others when required to remember a 1-digit number (low cognitive load). We have observations about the searches and the choices of subjects in both cognitive load treatments.

We find that the choices are affected by both the features of the choice set and the available cognitive resources, as manipulated by cognitive load. Specifically, we find that subjects in the high load treatment make inferior line selections. We also find that the quality of the selection decreases in the number of lines in the choice set, in the similarity of the lengths of the lines in the choice set, and in the lengths of the lines in the choice set.

We likewise find that the searches are affected by both the incentives of the choice problem and the available cognitive resources. In particular, we find that subjects in the high load treatment conduct worse searches in that they spend less time deliberating than do subjects in the low load treatment. We also find that the time deliberating is increasing in the similarity of the lengths of the lines in the choice set.

We find that a measure of deliberation time is negatively related to selecting the longest line in the choice set. A prediction of this result emerges from a setting where an agent faces a choice between options with uncertain utility and there is a cost of gathering information about the choice problem. Fudenberg, Strack, and Strzalecki (2018) show that, in this setting, which seems to correspond to our experiment, suboptimal decisions will tend to have longer deliberation times than optimal decisions.

McFadden (1974) and Yellot (1977) show that the distribution of error terms has significant implications for the appropriate stochastic choice model specification. Our design permits a multinomial discrete choice analysis on choice among single-attribute objects with an objective value, where we can examine the distribution of the errors. Our analysis suggests that the errors are better described as having a Gumbel distribution rather than a normal distribution.

We find evidence of choice overload in our setting, where the choice set is small and the

[6] For instance, see Duffy and Smith (2014) and Deck and Jahedi (2015).

objects are simple. Finally, we observe the effects of limited cognition, consistent with memory decay and attention.

We hope that our results on stochastic choice can help inform stochastic choice models. We note that neither the Gumbel distribution results nor the relationship between optimal choice and deliberation time appear to be affected by the available cognitive resources. This perhaps suggests that these results are general and persistent features of choice. We also hope that the technique of inducing an objective value based on an imperfectly perceived object can help investigations of stochastic choice, and presently unforeseen applications, in future studies.

2 Related literature

2.1 Random utility and stochastic choice

In order to make sense of the apparent randomness in choice data, researchers have advanced random utility and stochastic choice models. The classic efforts include Bradley and Terry (1952), Debreu (1958), Luce (1959a,b), and Becker, DeGroot, and Marschak (1963). Numerous other random utility or stochastic choice experimental and theoretical papers have emerged in an effort to better understand choice.[7,8] The conceptualization that utility is random has also lead to significant advances in econometrics (McFadden, 1974, 1976, 1981, 2001).

Some of the recent choice literature has focused on consideration set effects, whereby the decision maker does not consider the entire set of objects and this is not necessarily observable to the experimenter.[9] However, with our experimental design, we can observe the

[7]A partial list of these efforts, not previously mentioned, would include Tversky (1969), Loomes, Starmer, and Sugden (1989), Sopher and Gigliotti (1993), Loomes and Sugden (1995), Sopher and Narramore (2000), Gul and Pesendorfer (2006), Rubinstein and Salant (2006), Tyson (2008), Caplin, Dean, and Martin (2011), Conte, Hey, and Moffatt (2011), Reutskaja, Nagel, Camerer, and Rangel (2011), Wilcox (2011), Gul, Natenzon, and Pesendorfer (2014), Loomes and Pogrebna (2014), Woodford (2014), Caplin and Dean (2015), Caplin and Martin (2015), Cubitt, Navarro-Martinez, and Starmer (2015), Lu (2016), Apesteguia, Ballester, and Lu (2017), Dean and Neligh (2017), Ahumada and Ulku (2018), Apesteguia and Ballester (2018), Echenique, Saito, and Tserenjigmid (2018), Koida (2018), Conte and Hey (2019), Natenzon (2019), and Kovach and Tserenjigmid (2021),

[8]For a partial list from the psychology literature, see Regenwetter, Dana and Davis-Stober (2011), Regenwetter, Dana, Davis-Stober, and Guo (2011), Regenwetter and Davis-Stober (2012), Birnbaum and Schmidt (2008, 2011), and Birnbaum (2011).

[9]For instance, see Masatlioglu, Nakajima, and Ozbay (2012), Manzini and Mariotti (2014), Aguiar, Boc-

consideration set and the objective lengths of the lines. We find that the longest viewed line is not selected in many trials and this selection is affected by available cognitive resources. Our analysis therefore suggests that, while there are possibly also consideration set effects, imperfect perception about one's preferences is a key component to understanding stochastic choice.

Because economic agents cannot process all available and relevant information, the allocation of attention has been used to improve our understanding of a number of choice settings.[10] In our setting, we find that the provision of effort and attention are affected both by the incentives in the choice set and by the available cognitive resources.

Matějka and McKay (2015) offer a rational inattention foundation for discrete choice models. Agents optimally allocate costly attention in order to better understand the true state of nature.[11] Specifically, the agents can reduce the Shannon entropy associated with the choice setting by incurring costs associated with attention. The authors show that this implies a random choice specification, similar to Luce (1959a). In our experiment, there is a similar process as subjects devote cognitive effort in order to select the longest line in the choice set.

2.2 Incentivized, induced value experiments with imperfectly perceived objects

We are not the first authors to study behavior in a setting where material outcomes depend on choice involving imperfectly perceived objects. Researchers have conducted incentivized choice experiments when the judgments are based on the relative quantity of static dots (Caplin and Dean, 2015; Dutilh and Rieskamp, 2016), the quantity of dynamic dots (Zeigenfuse, Pleskac, and Liu, 2014), the dominant direction of moving dots (Bhui, 2019a,b), and the number of flickering dots (Oud et al., 2016). These papers are different from ours in many respects, perhaps most notably because the imperfect perception in these settings could (in principle) be eliminated by carefully counting the discrete and finite measures. In this sense, counting could be considered an objective measure of value. Because the line pixels in our experiment

cardi, and Dean (2016), Caplin, Dean, and Leahy (2019), and Cattaneo et al. (2020).

[10] For example, see Caplin (2016) for an overview.

[11] Also see Weibull, Mattsson, and Voorneveld (2007).

cannot be individually identified, they cannot be counted, and subjects must rely on their subjective perception in order to make their choice.

To our knowledge, there are only three instances of papers that study choice where material outcomes depend on choice involving imperfectly perceived objects with an uncountable measure. However, each differs from our setting. Tsetsos et al. (2016) study choice that involves judgements of the heights of bars. Such a measure is uncountable, however the size of the bars within each trial is dynamic: the subjects are charged with estimating the distribution within a trial. By contrast, the size of each line in our setting is static within each trial.

Polanía et al. (2014) examine choice in a setting where outcomes are based on the area occupied by the image of various objects. Area is also an uncountable measure. However, the images have different shapes and so the objects vary according to several meaningful attributes.

As we do here, Duffy and Smith (2020) conduct an experiment where subjects select among lines and are paid as a function of the lengths of the selected lines. Therefore, to our knowledge, Duffy and Smith (2020) is the only other example of an incentivized choice experiment in a setting where outcomes depend on imperfectly perceived static objects with an uncountable measure that varies only according to a single relevant attribute.

This is an attractive setting to study apparently random choice because values can be completely characterized by a single value and it is therefore straightforward to produce a dataset that can study the distribution of the errors. Further, there will neither be substitutes nor compliments among the attributes, whereby a previous decision will interact with a subsequent decision.

Similar to our findings, Duffy and Smith find that both choice and searches respond to the features of the choice problem. The authors also find that Gumbel errors better fit the data than normal errors. Additionally, Duffy and Smith find a negative relationship between deliberation time and selecting the longest available line.

However, Duffy and Smith do not manipulate any other aspect of the choice problem. In contrast, here we manipulate the available cognitive resources during the line selection task, and this provides additional clues about the nature of choice and search.

Further, Duffy and Smith (2020) permit up to 60 seconds for the line selection task and allow subjects to click to proceed to the following stage. Here, we allocate 15 seconds for the line selection task and we do not permit subjects to click to proceed to the following stage. Therefore, in our experiment, shorter deliberation times do not yield the material benefit of completing the session faster. Despite these differences, similar to Duffy and Smith (2020), we find that deliberation times decrease across trials and are negatively related to optimal choices.

2.3 Cognitive load manipulation

There is a large literature that employs the cognitive load manipulation in order to affect the available cognitive resources of subjects. Although much of this research appears in the psychology literature, the technique is more frequently appearing in the economics literature,[12] including in strategic settings.[13] Most relevant to our purposes, research finds that subjects in a high cognitive load treatment fail to process available and relevant information (Gilbert, Pelham, and Krull, 1988; Swann et al., 1990) and choices can be less consistent (Franco-Watkins, Rickard, and Pashler, 2010; Olschewski, Rieskamp, and Scheibehenne, 2018). We also note that subjects under a cognitive load tend to perform worse on visual judgment tasks.[14]

To our knowledge, there are only two other examples of papers that employ the cognitive load manipulation in an incentivized choice setting without social considerations or objective risk: Lee, Amir, and Ariely (2009) and Drichoutis and Nayga (2020).

Lee, Amir, and Ariely (2009) study intransitive choices among pair-wise decisions made

[12] For instance, see Benjamin, Brown, and Shapiro (2013), Schulz et al. (2014), Deck and Jahedi (2015), Hauge et al. (2016), and Deck, Jahedi, and Sheremeta (2021). Although the cognitive load manipulation in economic settings can sometimes produce null results or results that are difficult to interpret (Achtziger, Alós-Ferrer, and Ritschel, 2020).

[13] See Milinski and Wedekind (1998), Roch et al. (2000), Cappelletti, Güth, and Ploner (2011), Carpenter, Graham, and Wolf (2013), Duffy and Smith (2014), Allred, Duffy, and Smith (2016), Buckert, Oechssler, and Schwieren (2017), and Duffy, Naddeo, Owens, and Smith (2021).

[14] See Morey and Cowan (2004), Allen, Baddeley, and Hitch (2006), Cocchi et al. (2011), Morey and Bieler (2013), Zokaei, Heider, and Husain (2014), and Allred et al. (2016).

while their subjects are under a cognitive load.[15] Surprisingly, the authors find that subjects under a high cognitive load make fewer intransitive choices than subjects under a low cognitive load. However, these are real world objects that have attributes whose desirability is not observable to the experimenters. Further, the repeated nature of the experiment makes it difficult to determine if the attributes from previous choices affected subsequent choices (either because the attributes are regarded as complements or substitutes). By contrast our subjects make judgments on objects that have a value based on only a single objective attribute.

Drichoutis and Nayga (2020) find that a high cognitive load does not increase internal inconsistency on a GARP budget allocation task. By contrast, we find that the cognitive load manipulation negatively affects choices and searches.

2.4 Deliberation times and choice

There is a long history of measuring response times in order to gain clues on deliberation and researchers tend to find that longer deliberation times are associated with settings where the elements of the choice set are similarly valued.[16] Likewise, we find that trials involving choice sets with lines of similar lengths tend to have longer deliberation times.

We also find that trials where the longest line was selected tend to have shorter deliberation times than trials where the longest line was not selected. This finding also has precedence in the experimental literature.[17] We note that this result emerges from a model of an agent in a choice problem with unknown utility and a cost of acquiring information about the elements of the choice set (Fudenberg, Strack, and Strzalecki, 2018). Consistent with the predictions of Fudenberg, Strack, and Strzalecki, we find evidence that suboptimal decisions are associated with longer deliberation times, although we caution that our results could be driven by endogenous effects.

[15] See Experiment 4.

[16] For instance, see Henmon (1911), Volkmann (1934), Dashiell (1937), Mosteller and Nogee (1951), Hey (1995), Moffatt (2005), Chen and Fischbacher (2016), Alós-Ferrer et al. (2016), and Alós-Ferrer and Garagnani (2020).

[17] For instance, see Henmon (1911), Kellogg (1931), Bhui (2019b), and Duffy and Smith (2020).

2.5 Induced valuation experiments with information processing limitations

Our experiment presents subjects with a decision problem with an objectively optimal solution. However because of their limitations, subjects are not able to attain the optimal solution with certainty.[18] This feature also appears in Gabaix et al. (2006) and Sanjurjo (2015, 2017). There subjects are given a multi-attribute choice problem where each attribute value is represented by a number. Since subjects cannot fully process the available information, despite that there is an objectively optimal solution, the optimal solution is not attained with certainty. Also similar to our setting, subjects must click on the information in order to make it appear. In this way, similar to this multi-attribute literature, we can observe the process of search.[19]

3 Experimental design

3.1 Overview

The experiment was programmed on E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The sessions were performed on standard 23 inch (58.42 cm) Dell Optiplex 9030 AIO monitors. E-Prime imposed a resolution of 1024 pixels by 768 pixels. A total of 92 subjects participated in the experiment.

3.2 Line selection task

In each round, subjects were presented a choice set of lines that ranged in number between 2 and 6. Each of these choice set sizes occurred with probability $\frac{1}{5}$ and were drawn with replacement. Subjects were able to only view one line at a time. The lines were labeled in alphabetic order at the bottom of the screen. Letters A and B always represented the first two options, and consecutive letters were added as needed. Subjects could view a particular line by clicking on the letter label that corresponds to that line. To view another line, subjects click on its corresponding label. This makes the new line appear and the old line disappear.

[18] Also see Caplin, Dean, and Martin (2011) and Geng (2016) for induced valuation experiments where subjects select among options that pay amounts implied by arithmetic operations and subjects do not always select optimally.

[19] Also see Payne, Braunstein, and Carroll (1978) and Payne, Bettman, and Johnson (1993).

Each line appeared within a rectangular region of 400 pixels in the horizontal direction and 150 pixels in the vertical direction. The boundaries of these regions were not visible to the subjects. The lines were randomly offset vertically and horizontally within these regions such that there was a minimum cushion between the line and the edge of the region. This cushion was 20 pixels in the horizontal direction and 10 pixels in the vertical direction. The offsetting was fixed for each line throughout each trial. The regions were non-overlapping and arranged in 2 columns and 3 rows, with the regions for A and B in the top row, the regions for C and D in the middle row, and the regions for E and F in the bottom row.

The length of the lines in any trial were determined by subtracting various amounts from the *longest line*. There were 10 possible longest line lengths, ranging in 16 pixel (0.80 cm) increments, from 160 pixels (8.0 cm) to 304 pixels (15.1 cm). The lines each had a height of 0.38 cm.

There were three line length treatments. In the *difficult* treatment, one line was exactly one pixel shorter than the longest, and the other differences were drawn from a uniform on $\{-1, ..., -11\}$. In the *medium* treatment, one line was exactly 12 pixels shorter than the longest and the other differences were drawn from a uniform on $\{-12, ..., -39\}$. In the *easy* treatment, one line was exactly 40 pixels shorter than the longest, and the other differences were drawn from a uniform on $\{-40, ..., -100\}$. The difficult, medium, and easy treatments each occurred with probability $\frac{1}{3}$, in random, interspersed order, and drawn with replacement. The subjects were not informed of the existence of these treatments. Because the lines in the choice set are generated by subtracting amounts from the longest line, the longest line offers a partial characterization of the distribution of lengths in the choice set.

Below each letter label was a box indicating that the subject currently *selected* that line. Subjects could change this selection during the allotted time.

The line selection task lasted 15 seconds. The subjects could view the time remaining, rounded to the nearest second. The choice within each trial was the line that was selected when the 15 seconds expired. If subjects did not select a line before time expired, it was assumed that the selected line had a length of 0. Regardless of their actions in the stage, subjects would

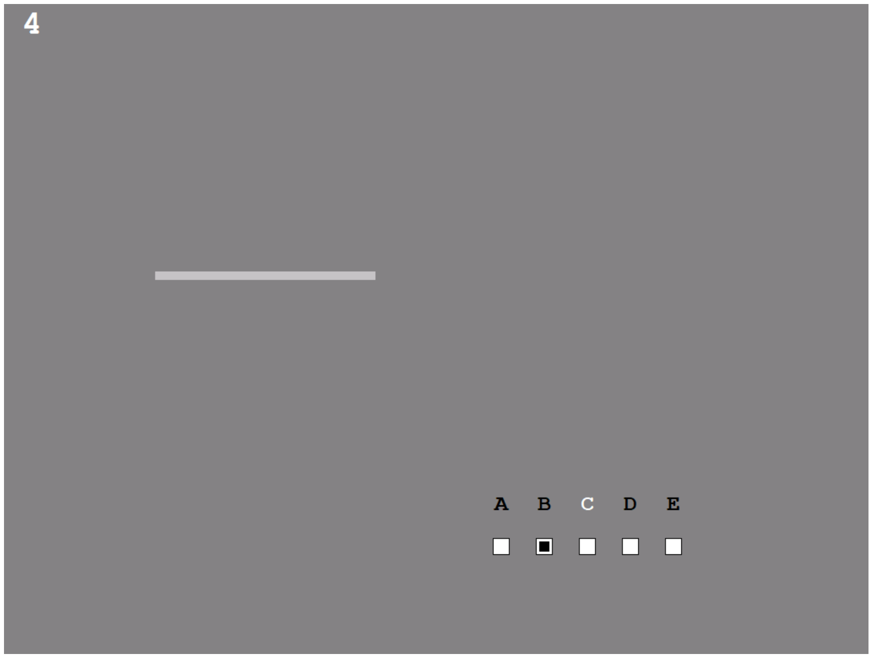

Figure 1: Screenshot from a trial with 5 lines in the choice set, where line C is being viewed, line B is currently selected as the longest, and there are 4 seconds remaining.

only advance to the following screen when the 15 seconds had expired. Therefore, 15 seconds was both the minimum and maximum time in this stage. Subjects were not permitted to use measurement aids and were required to base their choice exclusively on their perception. See Figure 1 for a screenshot[20] and Figure 2 for a characterization of the regions, which were not visible to the subjects.

3.3 Cognitive load treatments

There were 50 trials where subjects were given a 6-digit number to remember, which we refer to as *high load*. There were 50 trials where subjects were given a 1-digit number to remember, which we refer to as *low load*. The cognitive load treatments were interspersed randomly and drawn without replacement. Each of the 10 longest line lengths were presented 5 times in the high load treatment and 5 times in the low load treatment, randomly interspersed and drawn without replacement.

[20] See https://osf.io/srpzh/ for the full set of screenshots.

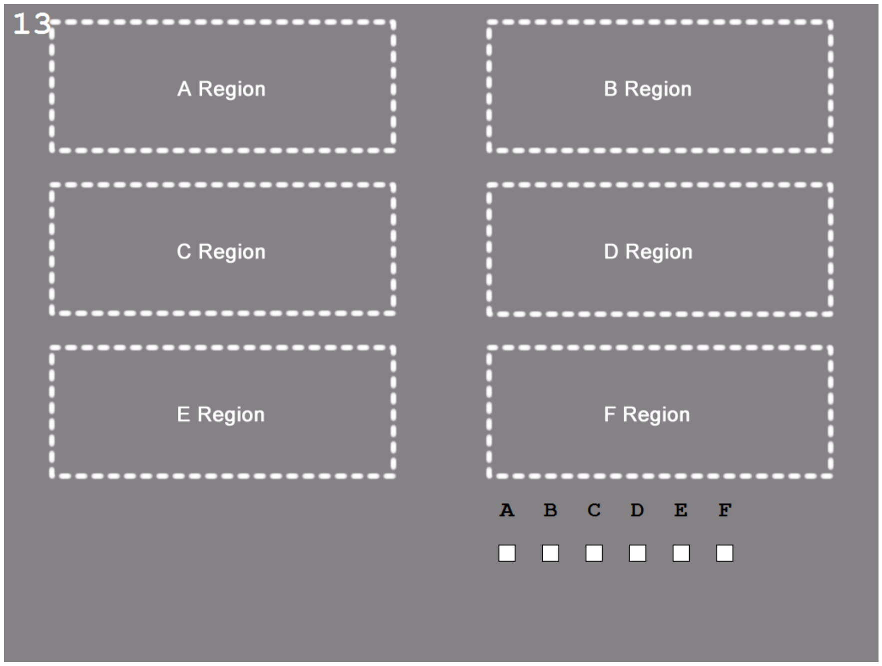

Figure 2: A characterization of the regions, invisible to the subjects, which contain the corresponding lines.

3.4 Timeline of the trials

At the start of each trial, subjects were presented with a number to be memorized. Subjects were given 5 seconds to commit the number to memory[21] and would only proceed to the following stage when the 5 seconds had expired. Subjects then proceeded to the line selection task, which lasted 15 seconds. Finally, subjects were presented with the screen that elicited the memorization number. There was neither a maximum nor minimum time for this stage.

3.5 Unincentivized practice

Prior to the incentivized portion of the experiment, subjects had unincentivized practice remembering both a 1-digit and a 6-digit number. In contrast to the incentivized portion of the experiment, here subjects were told if their responses were correct. If a response did not contain the correct number of digits then subjects were directed to repeat the practice memorization task.

[21]The subjects could not view the time remaining in this stage, as these numbers could interact with the memorization number.

Additionally, subjects had an unincentivized practice on the line selection task. Subjects were given this practice with a choice set of 5 lines in the medium difficulty treatment. If the subjects did not view any lines, did not select a line that they viewed, or did not select any lines, the subjects were informed of this and were directed to repeat the practice line selection task.

3.6 Payment details

Subjects completed 100 line selection tasks and 100 memorization tasks. Those who correctly completed all 100 memorization tasks were paid for 30 randomly determined line selections, those who correctly completed 99 were paid for 29, those who correctly completed 98 were paid for 28, and so on, until subjects who correctly completed 70 or fewer memorization tasks were not paid for any of the line selection tasks. The earnings for the line selection task were paid at a rate of \$1 per 240 pixels (or \$0.4167 per 100 pixels). In addition to these payments, subjects were also paid a \$5 show-up fee. Subjects were paid in cash and amounts were rounded up to the nearest \$0.25. Subjects earned a mean of \$26.00.

3.7 Discussion of the design

Considerable attention has been devoted to the design of incentives where subjects make multiple (but similar) decisions. The majority of designs either pay for every decision or pay for a single randomly selected decision.[22] However, our experiment largely falls outside of this discussion because there is an auxiliary task (memorization) and a primary task (line selection). Both of these tasks need to be incentivized but the design requires additional consideration.

The goals of our incentive scheme are as follows: strongly incentivize the memorization task, keep incentives for memorization in each trial independent from incentives for the line selection task in that particular trial, and have equal material incentives in the line selection task for high and low load trials. To strive for these goals, we do not provide feedback on the memorization tasks and we pay a number of randomly selected line selection outcomes that is

[22] For example, see Cox, Sadiraj, and Schmidt (2015) and Charness, Gneezy, and Halladay (2016).

decreasing in the number of incorrect memorization tasks. Only 5 subjects, out of 92, failed to correctly perform at least 70 memorization tasks, suggesting that the incentive scheme was sufficiently calibrated. In addition, as feedback was not given on the memorization task, it is not clear whether subjects realized that they were near or below 70 correct. Finally, while incorrectly answering a specific memorization task decreases overall incentives, this affects both high and low load trials equally and we are primarily interested in the difference between these treatments.

Subjects were given inflexible time constraints. These fixed times were implemented so that subjects were not able to strategically allocate their time in the experiment. For instance, our design prevents subjects in the high cognitive load treatment from spending less time in the line judgment task so that they could proceed quickly to the memorization task.

The boundaries of the regions that contained the lines were not visible to the subjects. Our concern was that any such aid would differentially benefit the judgment of the lengths of extreme (very short or very long) lines. Lines always appeared in the identical location within a trial.

Although the letter labels and their spatial orientations might suggest a natural order in which to view the lines, subjects could view the lines in any order. For example, it was permitted to click back-and-forth between lines and it was permitted to view a particular line on multiple occasions before all the lines were viewed. In summary, there were no restrictions on the nature of the line search beyond the time constraint and the constraint that only one line could be viewed at a time.

4 Results

4.1 Cognitive load

A larger fraction of memorization tasks were correctly completed in the low cognitive load treatment (97.6%, 4490 of 4600) than the high cognitive load treatment (85.8%, 3947 of 4600) according to a two-sample Wilcoxon rank-sum test[23], $Z = 20.53$, $p < 0.001$. See Table A1 for a characterization of the distribution of correct memorization tasks by cognitive load treatments.[24]

Further, 77 of the 92 subjects successfully completed more than 85% of their memorization tasks correctly. This suggests to us that the incentives were sufficient to elicit cognitive effort on these tasks.

Recall that subjects were given 15 seconds to make a decision and were not allowed to click to proceed before this time. It appears that in many trials, subjects stopped viewing lines and making line selections before the 15 seconds had elapsed. This suggests to us that the subjects had concluded their effort for the trial, and would have clicked to proceed, if such were possible. We therefore refer to the time between the start of the trial and the last click in that trial as the *Implicit response time*.

We find that implicit response times are smaller in the high cognitive load treatment ($mean = 9.586s$, $SD = 3.463$) than in the low cognitive load treatment ($mean = 10.081s$, $SD = 3.439$) according to a two-sample Wilcoxon rank-sum test[25] $Z = 6.73$, $p < 0.001$. It is possible to interpret this result as evidence that our cognitive load manipulation affected the decision-making process of subjects (Achtziger, Alós-Ferrer, and Ritschel, 2020).

4.2 Quality of choices

Here we explore the optimality of choices. We define the *Selected longest* variable to be a 1 if the choice was the longest available line, and a 0 otherwise. We conduct regressions with the Selected longest variable as dependent variable. Since the dependent variable is binary, we employ a logistic specification. We include the High load variable, which obtains a 1 in the high load treatment, and a 0 otherwise. Further, since the Selected longest variable appears

[23] The difference is also significant according to a one-sample Wilcoxon signed-rank test $S = 95296.5$, $p < 0.001$.

[24] We report the cognitive load summary statistics of two other repeated and incentivized experiments. Duffy and Smith (2014) found that 592 of 600 (98.67%) 2-digit numbers were correctly remembered and 676 of 840 (80.48%) 7-digit numbers were correctly remembered. Further, Duffy, Naddeo, Owens, and Smith (2021) found that 6362 of 6500 (98.87%) 1-digit numbers were correctly remembered and 5718 of 6500 (87.96%) 6-digit numbers were correctly remembered.

[25] The number of observations of this test is identical to that in the preceding test.

to be affected by the difficulty treatments, the number of lines treatments, the longest line treatments, and the letter that contained the longest line, we include these as independent variables.[26] For the difficulty treatments, we include dummy variables indicating whether the treatment was Easy or whether the treatment was Difficult. To account for the letter label of the longest line, we offer specifications where we estimate a unique dummy variable for each of the 20 combinations of letter-number of lines as in Table $A5$. However, in the analysis immediately below we do not explore the effect of the letter label on the quality of the choice.[27] Due to the repeated nature of the observations, we also offer fixed-effects specifications where we estimate a dummy variable for each subject. We summarize these regressions in Table 1.

Table 1: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	-0.0270^{**}	-0.0299^{**}	-0.0268^{**}	-0.0291^{**}
	(0.0093)	(0.0101)	(0.0092)	(0.0100)
Longest line normalized	-0.0006^{***}	-0.0006^{***}	-0.0006^{***}	-0.0006^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.054^{***}	–	-0.054^{***}	–
	(0.003)		(0.003)	
Easy treatment dummy	0.355^{***}	0.389^{***}	0.367^{***}	0.407^{***}
	(0.014)	(0.016)	(0.014)	(0.016)
Difficult treatment dummy	-0.285^{***}	-0.311^{***}	-0.286^{***}	-0.314^{***}
	(0.011)	(0.012)	(0.011)	(0.012)
Trial	0.00007	0.00005	0.00005	0.00002
	(0.00016)	(0.00017)	(0.00016)	(0.00017)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	8339.6	8182.4	8173.6	8016.6

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, we see that the high load coefficient is negative. This implies that choices are worse in the high cognitive load treatment. We also find that the accuracy of the

[26] See Tables $A2$-$A5$.

[27] We postpone our discussion of this issue until subsections 4.7 and 4.8.

choice decreases when there is a larger number of lines (choice overload effects) and decreases in the difficulty of the decision. Additionally, we see that the accuracy decreases in the length of the longest line. This result could be interpreted as suggesting that subjects are worse at judging longer lines than shorter lines. This explanation is consistent with Weber's law. On the other hand, it is possible that the subjects expended less effort on trials with longer lines because the subjects knew that they would earn more on these trials than on trials with shorter lines. These effort-wealth effects could also explain the negative coefficient estimates for the Longest line variable.

In the appendix, we also report additional analyses that investigate the optimality of choice.[28] Our results are not changed.[29] Together these results imply that the availability of cognitive resources affects the quality of the choice.[30,31]

4.3 Quality of searches

The analysis above suggests that the high cognitive load treatment implied worse choices. Here we explore the effect of the cognitive load on the quality of the searches. We examine the quality of the searches by employing the Implicit response time variable. A larger Implicit response time would seem to be associated with more effort in identifying the optimal line. Table 2 summarizes the analysis, which is analogous to that summarized in Table 1, with the

[28] We restrict the analysis to trials in which the cognitive load task was performed correctly (Table $A6$) and to trials in which at least one line was viewed and one line was selected (Table $A7$). Additionally, we conduct the analogous tobit regressions with the Longest line minus the selected line as dependent variable (Table $A8$). We also include interactions of the difficulty variables with the Longest line variable and with the Number of lines variable (Table $A9$). Finally, we conduct specifications with random-effects rather than fixed-effects (Table $A10$).

[29] Although we note a negative and significant coefficient estimate for Trial in Table $A8$.

[30] The reader might wonder about the relationship between success on the memorization task and optimality in the line selection task. We note a positive relationship between the total number of correct memorization tasks completed by a subject and the total number of instances that the longest line was selected by that subject, according to a Spearman correlation ($r(92) = 0.266$, $p = 0.01$). This suggests a role for individual differences, possibly cognitive ability, attention, motivation, or some other attribute.

[31] A related question is whether subjects were aware that they would not correctly recall the memorization number and this affected their behavior in the line selection task. We designed the incentives to avoid this possibility. In order to avoid trial-specific wealth effects there is no relationship between correctly performing the memorization task in that period with payment for the line selection task in that period. It is still possible that subjects did not understand this matter. This concern prompted us to include Table $A7$, which only includes observations where the memorization task was correct. The results from Table 1 are not changed. We further explore this issue in Table $A11$. The results of Table $A11$ suggest that the effect of the high load treatment in Table 1 is possibly slightly underestimated but possibly slightly overestimated in Table 2 below.

exception that the dependent variable is Implicit response time and the regression is linear, not logistic.

Table 2: Regressions of the Implicit response time variable

	(1)	(2)	(3)	(4)
High load	-0.406^{***}	-0.410^{***}	-0.408^{***}	-0.413^{***}
	(0.0580)	(0.0579)	(0.0510)	(0.0509)
Longest line normalized	0.00239^{***}	0.00234^{***}	0.00239^{***}	0.00235^{***}
	(0.00063)	(0.00063)	(0.00056)	(0.00055)
Number of lines normalized	1.088^{***}	–	1.100^{***}	–
	(0.021)		(0.018)	
Easy treatment dummy	-2.118^{***}	-2.129^{***}	-2.119^{***}	-2.130^{***}
	(0.0712)	(0.0711)	(0.0629)	(0.0628)
Difficult treatment dummy	1.042^{***}	1.030^{***}	0.969^{***}	0.958^{***}
	(0.0711)	(0.0710)	(0.0629)	(0.0628)
Trial	-0.00764^{***}	-0.00750^{***}	-0.00764^{***}	-0.00751^{***}
	(0.00100)	(0.00100)	(0.00101)	(0.00088)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	44974.8	44935.9	42597.4	42565.5

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$.

We find that trials in the high cognitive load treatment have a shorter Implicit response time. This suggests that the cognitive load manipulation is negatively affecting the quality of the searches.[32] We also observe that Implicit response time is increasing in the difficulty of the choice problem. In other words, whereas the penalty for an incorrect line selection is smaller in the difficult treatment, the time devoted to deliberation is increasing in the difficulty of the treatment. These results are consistent with previous experimental findings.[33]

[32] We conduct a similar analysis of the quality of searches by using the number of total line view clicks during the stage as dependent variable (Table $A12$). These *View clicks* results are similar to those summarized in Table 2 with the exceptions below. We find that View clicks are decreasing in the length of the longest line but Implicit response times are increasing in the length of the longest line. Additionally, we find that View clicks are increasing across trials but Implicit response times are decreasing across trials.

[33] See Henmon (1911), Kellogg (1931), Bhui (2019b), and Duffy and Smith (2020).

Further, Implicit response time is increasing in the length of the longest line. According to this measure, more effort–not less–is associated with longer lines. It is possible that subjects are responding to the material incentives, as longer line treatments are more valuable than shorter line treatments. However, it is also possible that this effect was driven by the psychology finding that absolute differences are more difficult to judge for longer than shorter lines (Weber's Law). We further note that the relationship involving the length of the longest line is different when the measure of effort is the number of line views (Table $A12$). Regardless, we are not able to distinguish between the Weber's law explanation and effort-wealth effects explanation for the results in Table 1.

4.4 Relationship between choice and search

We observe both that choices are worse in the high cognitive load treatment and that searches are worse in the high cognitive load treatment. A natural question is whether the worse searches are causing the worse choices. There is a literature that posits that suboptimal choice occurs because subjects do not consider every object in the choice set, but only a subset. Further this consideration set is not typically observable to the experimenter. However, due to our design, we are able to observe whether subjects viewed the longest line.

Among the 9109 trials where subjects viewed the longest line, there are 6354 observations where the longest line was not selected. However, among the 91 trials where subjects did not view the longest line there are 73 observations where the longest line was not selected.[34] Therefore in our data, 98.9% of the suboptimal choices occurred in trials where the subject viewed the longest line. This suggests that the bulk of our suboptimal choices can be explained due to imperfect perception rather than not considering the longest line.[35]

[34] How could the longest line be selected in a trial where the longest line was not viewed? One conjecture is that in these 18 trials, subjects mistakenly thought that the line was viewed and it was perceived to be the longest, when in fact, the line was not actually viewed in that trial. Regardless, we note that this is a very rare occurrence that appears in less than 0.2% of trials.

[35] We conduct an analysis similar to that summarized in Table 1, but the dependent variable measures whether the longest line was selected among the lines viewed in that trial (Table $A13$). The results are similar to those in Table 1.

4.5 Relationship between quality of choice and implicit response times

Recall that, in a theoretical set-up similar to our experimental setting, Fudenberg, Strack, and Strzalecki (2018) find that suboptimal decisions will have longer deliberation times than optimal decisions. Here we test this prediction with our dataset.

Implicit response times of the trials in which the longest line was selected ($mean = 9.094s$, $SD = 3.352$) are smaller than those in trials in which the longest line was not selected ($mean = 11.498s$, $SD = 3.102$), according to a two-sample Wilcoxon rank-sum test ($Z = 31.22$, $p < 0.001$). This effect is robust when restricted to a cognitive load treatment.[36]

In order to more carefully investigate this matter, we conduct regressions with Implicit response time as the dependent variable. We employ specifications similar to those in Table 2, however we include Selected longest as an independent variable. In addition to estimating the standard fixed effects dummies, we estimate an Easy treatment dummy coefficient, a Difficult treatment dummy coefficient, a Number of lines coefficient estimate, and a Longest line coefficient estimate for every subject. Below, we refer to these as the *Subject-specific choice set estimates*. Additionally, we include specifications that estimate a High load dummy coefficient for every subject. We refer to these as the *Subject-specific cognitive load dummies.* We also include a specification where we estimate a Trial coefficient for every subject. We refer to this as *Subject-specific Trial estimates.* Because the cognitive load might affect the subject-specific choice set estimates, for every subject we estimate the choice set coefficients in both the high load treatment and the low load treatment. We refer to this as *Subject-specific cognitive load-choice set interactions.* We summarize this analysis in Table 3.

[36] Implicit response times of high load trials in which the longest line was selected ($mean = 8.851s$, $SD = 3.298$) are smaller than those in trials in which the longest line was not selected ($mean = 11.214s$, $SD = 3.259$), according to a two-sample Wilcoxon rank-sum test ($Z = 21.82$, $p < 0.001$). Likewise, Implicit response times of low load trials in which the longest line was selected ($mean = 9.335s$, $SD = 3.388$) are smaller than those in trials in which the longest line was not selected ($mean = 11.789s$, $SD = 2.906$), according to a two-sample Wilcoxon rank-sum test ($Z = 22.43$, $p < 0.001$).

Table 3: Regressions of the Implicit response time variable

	(1)	(2)	(3)	(4)
Trial	-0.0079^{***}	-0.0080^{***}	–	–
	(0.0009)	(0.0009)		
High load	-0.420^{***}	–	–	–
	(0.050)			
Selected longest	-0.684^{***}	-0.682^{***}	-0.664^{***}	-0.648^{***}
	(0.066)	(0.066)	(0.066)	(0.067)
Subject-specific				
choice set estimates	*Yes*	*Yes*	*Yes*	*Yes*
cognitive load dummies	*No*	*Yes*	*Yes*	*Yes*
Trial estimates	*No*	*No*	*Yes*	*Yes*
cognitive load-choice set interactions	*No*	*No*	*No*	*Yes*
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
AIC	42318.7	42108.1	42549.3	42572.4

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts or the subject-specific estimates. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$.

In every specification, the Selected Longest variable is negative and significant. We interpret this as suggesting that, even when controlling for the fixed characteristics of the subjects, suboptimal choices tend to take longer than optimal choices.[37] We also note that this relationship is robust across our specifications that account for the cognitive load treatments.

Due to the possible endogeneity present in the analysis, we conduct Spearman correlations between the unstandardized residuals and the Selected longest variable. In specifications (1) – (4), the p-values, respectively, are 0.029, 0.028, 0.035, and 0.040.[38] We conclude that, as predicted by Fudenberg, Strack, and Strzalecki, there is a negative relationship between implicit response times and the optimality of choices. However, we acknowledge that this is possibly driven by endogeneity.

[37] In Table $A14$, we offer a robustness check of Table 3. Our qualitative results are not changed.

[38] When we restrict the analysis to the 9001 observations with a line view and a line selection, the p-values, respectively, are 0.066, 0.070, 0.078, and 0.086.

4.6 Multinomial discrete choice analysis and the nature of the stochastic utility

An assumption in multinomial discrete choice analysis is that choice is stochastic because of an unobserved stochastic component in the utility function.[39] A common specification in these random utility models (RUM) is that there is a non-stochastic component of the utility function and an additive stochastic component. For example, option j would have utility:

$$U_j = V_j + \epsilon_j \ , \tag{1}$$

where V_j is the non-stochastic component and ϵ_j is the random component. RUMs typically assume that agents select the item with the largest realized utility. Specifically, a choice of i from the set $K = \{1, ..., k\}$ arises when:

$$V_i + \epsilon_i \geq V_j + \epsilon_j \text{ for every } j \in K. \tag{2}$$

Further, the non-stochastic components to the RUMs are not typically observable. Therefore the researcher includes a set of observable features possibly relevant to the choice j, $\overline{x_j} = (x_{j1}, ..., x_{jn})$. In order to account for the effect of each of these factors, the analyst also estimates $\overline{\beta} = (\beta_1, ..., \beta_n)$. In these settings, the non-stochastic component is $V_j = \overline{\beta} * \overline{x_j}$. However, in our setting, the length of the line is the only relevant attribute. Therefore the non-stochastic component of option j simplifies to:

$$V_j = \beta * Length_j, \tag{3}$$

where β is a scalar.

We also note that there can be a range of specifications of the stochastic component. For instance, ϵ_j might be assumed to be normally distributed. On the other hand, the stochastic component might also be assumed to have the Gumbel distribution, $e^{-e^{-\epsilon}}$. (Regrettably, this is also referred to as the Type I extreme-value distribution, the double exponential distribution,

[39] See McFadden (1974, 1976, 1981, 2001).

and the log-Weibull distribution.) In our experiment, we can perfectly observe the objective lengths of the lines and the choices made by the subjects. We can therefore run specifications that employ either of these assumptions of the error distribution and observe which provides a better fit of the data, given the objective lengths of the lines in the choice set.

We run one specification where the stochastic component has the Gumbel distribution and is identically distributed for every option. As McFadden (1974) and Yellot (1977) show, this structure implies the Luce (1959a) stochastic choice model, whereby the probability that option j is selected from set K is:

$$P(j) = \frac{e^{\beta * Length_j}}{\sum_{k \in K} e^{\beta * Length_k}}. \tag{4}$$

We denote this *Conditional Logistic* model as specification (1).

We also run a specification where the stochastic component is assumed to have a normal distribution and is independently and identically distributed for every option. Yellot (1977) shows that this corresponds to Case V of Thurstone (1927a). We refer to this *Multinomial Probit* model as "Multi Probit 1" and denote it as specification (2).

Further, we run a specification where the stochastic component is assumed to be Gumbel but the options are not identically distributed. Specifically, each option has a stochastic component distributed $e^{-e^{-\frac{\epsilon}{\theta_i}}}$ where θ_i varies by the option. This specification is the Heteroschedastic Extreme-Value (HEV) model, introduced by Bhat (1995). For identification purposes, the final two options are assumed to be identically distributed with the unit scale: $\theta_k = \theta_{k-1} = 1$. We denote the HEV model as specification (3).

Finally, we run a specification where the stochastic component is assumed to be normally but non-identically distributed. This Multinomial Probit specification assumes that the standard deviations of the options can be different but that they are also independently distributed. Note that similar to the HEV model, for identification purposes, we assume that the standard deviation of the final two choices are identical. We refer to this Multinomial Probit model as "Multi Probit 2" and denote it as specification (4).

Note that we exclude observations where subjects did not specify a choice before time expired.[40] We report the Akaike Information Criterion (AIC, Akaike, 1974) and the Bayesian Information Criterion (BIC, Schwarz, 1978) for each model, restricted to a particular number of lines treatment. We also report the estimate of β for each model in each setting. These analyses are summarized in Table 4.[41,42]

Table 4: Comparisons of different multinomial discrete choice models

		Cond Logit	Multi Probit 1	HEV	Multi Probit 2	Trials
		(1)	(2)	(3)	(4)	
2 Lines	β est.	0.131	0.098	–	–	1785
	AIC	1417	1432			
	BIC	1422	1437			
3 Lines	β est.	0.128	0.086	0.118	0.067	1871
	AIC	2088	2140	2078	2145	
	BIC	2094	2146	2089	2156	
4 Lines	β est.	0.115	0.076	0.121	0.084	1826
	AIC	2718	2801	2709	2820	
	BIC	2723	2807	2726	2837	
5 Lines	β est.	0.110	0.108	0.113	0.116	1780
	AIC	3181	3383	3186	3282	
	BIC	3186	3389	3208	3304	
6 Lines	β est.	0.094	0.062	0.070	0.046	1780
	AIC	3775	3808	3613	3684	
	BIC	3780	3813	3641	3711	

We provide the estimates of β, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the various models restricted to treatments with identical numbers of lines. Each of the estimates for β are significantly different from 0 with $p < 0.001$.

For both AIC and BIC, every value for the Conditional Logit model (1) is lower than that for the analogous Multinomial Probit 1 model (2). Additionally for both measures, every value for the HEV model (3) is lower than that for the analogous Multinomial Probit 2 model (4). We interpret these results as suggesting that the models that assume that errors have

[40] Therefore, the numbers of trials as reported in Table $A3$ are different than those reported below.

[41] Each specification was executed with the MDC (multinomial discrete choice) procedure in SAS. Specification (1) was performed with the *clogit* option. Specification (2) was performed with the *mprobit* option. Specification (3) was performed with the *hev* option and the Hardy integration method. Specification (4) was performed with the *mprobit* option.

[42] Note that for the case of 2 Lines, the Conditional Logistic regression is identical to the HEV specification, and the Multinomial Probit 1 is identical to the Multinomial Probit 2 specification. Therefore, we do not report specifications (3) and (4) for the 2 Lines treatment.

a Gumbel distribution provide a better fit than comparable models that assume that errors have a normal distribution. These results are not changed when we restrict to either trials in the high load treatment or trials in the low load treatment.[43] These results are largely unchanged when we restrict $\beta = 0.1$.[44] The results are also unchanged when we estimate different Length coefficients for high and low load, and we consider specifications where the letter label affects choice.[45] We interpret these results as evidence that the assumption that the errors have a Gumbel distribution is a better fit than the assumption that the errors have a normal distribution. Further, as we do not find evidence that these results are affected by the cognitive load treatment, it is possible Gumbel errors are general and persistent features of choice.

4.7 Memory decay and choice

Reutskaja, Nagel, Camerer, and Rangel (2011) report that the quality of choices tend to be diminishing in number of items viewed between the last item viewed and the best item viewed. Here we examine whether our subjects exhibit similar behavior consistent with memory decay.

Table $A5$ demonstrates the relationship between the quality of choice and the letter label of the longest line. This suggests a relationship between the quality of the choice and number of letters alphabetically between the letter label of the longest line and the last letter label in the choice set. Below, we test whether there is such a relationship.

We introduce the variable *Distance from last*, which provides a measure of the alphabetic distance between the letter label of the longest line and the last letter label in the choice set. For instance, in the 2-Line treatment, if line A is the longest then the variable is 1 and if line B is the longest then it is 0. In the 3-Line treatment, if A is the longest then the variable is 2, if B is the longest then it is 1, and if C is the longest then 0. We include Distance from the last as an independent variable. We also include specifications with the interaction between the High load dummy and the Distance from last variable. For identification reasons, we do not include the Letter dummy variables. We summarize these regressions in Table 5.

[43] See Tables $A15$ and $A16$.
[44] See Table $A17$.
[45] See Table $A18$.

Table 5: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	-0.0278^{**}	-0.0276^{**}	-0.0275^{**}	-0.0273^{**}
	(0.0093)	(0.0093)	(0.0092)	(0.0092)
Distance from last	-0.042^{***}	-0.042^{***}	-0.042^{***}	-0.042^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
High load * Distance from last	–	-0.0065	–	-0.0072
		(0.0067)		(0.0066)
Longest line normalized	-0.0005^{***}	-0.0005^{***}	-0.0005^{***}	-0.0005^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.033^{***}	-0.033^{***}	-0.033^{***}	-0.033^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Easy treatment dummy	0.359^{***}	0.359^{***}	0.372^{***}	0.372^{***}
	(0.014)	(0.014)	(0.014)	(0.014)
Difficult treatment dummy	-0.285^{***}	-0.285^{***}	-0.286^{***}	-0.286^{***}
	(0.011)	(0.011)	(0.011)	(0.011)
Trial	0.00006	0.00006	0.00004	0.00004
	(0.00016)	(0.00016)	(0.00016)	(0.00016)
Letter dummies	No	No	No	No
Fixed effects	No	No	Yes	Yes
AIC	8221.9	8223.7	8049.4	8051.2

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, we observe a negative relationship between Distance from last and the quality of the choice. This is consistent with the hypothesis suggested by Table $A5$. We also note that there is not a significant interaction between the cognitive load and the Distance from last variable. Additionally, each of the other coefficient estimates are relatively unchanged from the analysis summarized in Table 1.

One explanation for the negative coefficient estimates for the Distance from last variable is that subjects view the lines in alphabetical order (A then B then C etc.). Even though there were no restrictions on the order of viewing the lines, such a search might reduce the need to recall which non-optimal lines have been previously viewed. Alphabetic search implies that lines to the "left" have been viewed and lines to the "right" have not. Alphabetic search and the condition that lines viewed in the more distant past are recalled with a lower precision offers an explanation for the apparent relationship between the letter label and line selection accuracy.

To explore this explanation, we define the variable *Time since longest* to be the time elapsed since the subject viewed the longest line when the trial ended. Table 6 demonstrates the relationship between the Time since longest variable and the letter label of the longest line.

Table 6: Time since longest line by number of lines and letter label of the longest

	A	B	C	D	E	F
2 Lines	2.491 s	1.452 s	–	–	–	–
3 Lines	2.801 s	3.464 s	1.347 s	–	–	–
4 Lines	3.150 s	3.335 s	3.232 s	1.810 s	–	–
5 Lines	3.404 s	3.472 s	3.664 s	3.125 s	2.461 s	–
6 Lines	4.117 s	3.986 s	3.627 s	3.270 s	3.211 s	1.800 s

Table 6 suggests that there is a negative relationship between the Time since longest variable and the number of letter labels alphabetically between the longest line and the last letter label in the choice set. Here we test whether there is such a relationship. To do so, we conduct an analysis similar to Table 5, however we employ the Time since longest variable rather than the Distance from last variable. We summarize these regressions in Table 7. We interpret these results with caution due to the possibility of endogeneity by including the Time since longest variable.

Table 7: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	-0.0136	-0.0144	-0.0130	-0.0136
	(0.0089)	(0.0090)	(0.0088)	(0.0089)
Time since longest	-0.042^{***}	-0.042^{***}	-0.044^{***}	-0.044^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
High load * Time since longest	–	-0.0013	–	-0.0016
		(0.0026)		(0.0027)
Longest line normalized	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.040^{***}	-0.040^{***}	-0.041^{***}	-0.041^{***}
	(0.0033)	(0.0033)	(0.0033)	(0.0033)
Easy treatment dummy	0.378^{***}	0.378^{***}	0.377^{***}	0.376^{***}
	(0.013)	(0.013)	(0.013)	(0.013)
Difficult treatment dummy	-0.236^{***}	-0.236^{***}	-0.235^{***}	-0.235^{***}
	(0.011)	(0.011)	(0.011)	(0.011)
Trial	0.00015	0.00015	0.00012	0.00012
	(0.00015)	(0.00015)	(0.00015)	(0.00015)
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*No*	*No*	*Yes*	*Yes*
AIC	6703.6	6705.2	6582.0	6583.7

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, there is a negative relationship between the time elapsed since the longest line was viewed at the end of the trial and the quality of the choice. On the other hand, we do not find evidence that this relationship is affected by the cognitive load. Further, we observe qualitatively similar results to those found above, with the exception of the estimate for the High load variable. In none of the specifications is the estimate significant at 0.05.

We note that the results summarized in Table 7 should be viewed with caution. The Spearman correlations between the Time since longest variable and the residuals (both unstandardized and Pearson standardized) are each significant at 0.001. This suggest that endogeneity is present in our regressions.

It seems that choices are worse when the longest line is more alphabetically distant from the last letter label in the choice set and the longer the time since the longest line was viewed.

Taken together, our results are consistent with memory decay: lines viewed in the more distant past are remembered with lower precision. We note that while these results are consistent with limited cognition, we also note that we do not find a relationship between the memory decay effects and the cognitive load manipulation. Finally, we note that these effects might be exacerbated by the relatively smaller spatial distance between F region and the F box compared with the larger spatial distance between the A region and the A box.

4.8 Attention and choice

Testing for evidence consistent with memory decay is not the only such investigation on the effects of limited cognitive resources. Here we investigate the role of attention in choice.

Research finds that the time that a subject spends viewing (or fixated on) an object in a choice setting is associated with a higher likelihood of selecting the object.[46] Additionally, the visual psychology literature also finds that spatial resolution of abstract objects and visual information processing are enhanced by attention.[47]

One measure of attention is the total time spent viewing a line. In Table 8, we summarize the *Time viewing* variable by the number of lines treatment and the letter label.

Table 8: Time viewing by number of lines and letter label

	A	B	C	D	E	F
2 Lines	6.338 s	6.909 s	–	–	–	–
3 Lines	4.356 s	3.675 s	5.195 s	–	–	–
4 Lines	3.238 s	2.966 s	2.953 s	4.104 s	–	–
5 Lines	2.733 s	2.443 s	2.367 s	2.454 s	3.262 s	–
6 Lines	2.263 s	2.080 s	1.993 s	2.005 s	1.975 s	2.938 s

In Table 9, we report the Time viewing variable but restricted to the letter label of the longest line.

[46] See Armel, Beaumel, and Rangel (2008), Armel and Rangel (2008), Krajbich, Armel, and Rangel (2010), and Krajbich and Rangel (2011).

[47] For instance, see Yeshurun and Carrasco (1998), Carrasco and McElree (2001), Carrasco, Williams, and Yeshurun (2002), and Liu, Abrams, and Carrasco (2009).

Table 9: Total time viewing longest by number of lines and letter label of the longest

	A	B	C	D	E	F
2 Lines	8.410 s	9.028 s	–	–	–	–
3 Lines	7.020 s	6.010 s	7.805 s	–	–	–
4 Lines	5.622 s	5.252 s	5.074 s	6.351 s	–	–
5 Lines	5.047 s	4.374 s	4.351 s	4.170 s	5.040 s	–
6 Lines	3.992 s	3.772 s	3.600 s	3.778 s	3.806 s	4.994 s

Table 9 suggests that when the longest line is closer to the end of the alphabetic sequence, subjects spend more time viewing the longest line. Table $A5$ suggests that when the longest line is closer to the end of the sequence, there is a higher likelihood that the longest line was selected. Here we test whether there is a relationship between the time viewing the longest line and the likelihood that the longest line is selected. We perform an analysis similar to Tables 5 and 7 but with Time viewing longest as an independent variable. We summarize these regressions in Table 10. Similar to the analysis summarized in Table 7, we interpret these results with caution due to the possibility of endogeneity by including the Time viewing longest variable.

Table 10: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	-0.0270^{**}	-0.0255^{**}	-0.0264^{**}	-0.0244^{**}
	(0.0092)	(0.0094)	(0.0091)	(0.0093)
Time viewing longest	0.020^{***}	0.020^{***}	0.019^{***}	0.018^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
High load * Time viewing longest	–	0.0039	–	$0.0047^{\dagger}$
		(0.0028)		(0.0028)
Longest line normalized	-0.0005^{***}	-0.0005	-0.0005^{***}	-0.0005^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.031^{***}	-0.031^{***}	-0.033^{***}	-0.033^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Easy treatment dummy	0.340^{***}	0.341^{***}	0.353^{***}	0.354^{***}
	(0.014)	(0.014)	(0.014)	(0.014)
Difficult treatment dummy	-0.274^{***}	-0.274^{***}	-0.276^{***}	-0.276^{***}
	(0.011)	(0.011)	(0.011)	(0.011)
Trial	-0.000015	-0.000013	-0.00003	-0.00003
	(0.00016)	(0.00016)	(0.00016)	(0.00016)
Letter dummies	No	No	No	No
Fixed effects	No	No	Yes	Yes
AIC	8185.3	8186.7	8060.5	8061.3

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, the quality of the choice is increasing in the time viewing the longest line.[48] A similar result is reported by Krajbich and Rangel (2011). Also in every specification, we do not find a significant interaction between the cognitive load and the time viewing the longest line. The qualitative results involving the other variables are similar to those found above.

We note that the results summarized in Table 10 should be viewed with caution. The Spearman correlations between the Time since longest variable and the residuals (both unstandardized and Pearson standardized) are each significant at 0.001. This suggests that endogeneity is present in these regressions.

Our results suggest that (endogenous) attention is related to choice. However, we do not find that the cognitive load manipulation affects this relationship.

5 Conclusion

We observe behavior in an "idealized" choice setting where we know the true preferences of the subjects, but subjects have an imperfect perception of the objects in the choice set. Subjects can reduce their imperfect perception through cognitive effort. The objects of choice are lines of various lengths and subjects are paid an amount increasing in the length of the selected line. This feature allows us to make unambiguous conclusions about the optimality of choices. Subjects also make their choices in different cognitive load treatments, which are designed to manipulate their available cognitive resources. We are also able to observe aspects of the search, such as the number of lines viewed and the deliberation time.

Are there brains in stochastic choice? Our results suggest a qualified "yes." In our choice setting, we found that differences in available cognitive resources, as manipulated by cognitive

[48] We find similar results if we measure attention with the number of view clicks on the longest line or whether subjects viewed the longest line 2 or more times.

load, implied differences in both choice and search. Specifically, choices and searches are worse in the high cognitive load treatment.

We further find that choices and searches respond to features of the choice set. For example, the likelihood of optimal choice is decreasing and the deliberation times are increasing in the similarity of the lengths of the lines in the choice set. This suggests to us that the allocation of cognitive effort responds to the details of choice set.

We also find evidence that suboptimal choices are associated with longer deliberation times than are optimal choices. This is consistent with the implications of a model (Fudenberg et al., 2018) that is similar to our experimental setting. However, we admit that we cannot rule out the possibility that our results on this are endogenous.

Additionally, we find evidence of choice overload in our setting, where the choice set is small and the objects are simple. We also observe limited cognition effects, consistent with memory decay and attention. However, we note that these effects, which are consistent with memory decay and attention, are not affected by the cognitive load manipulation.

Many random utility models posit that there is a non-stochastic component and an additive stochastic component, which is also referred to as an error term. The distribution of the error term has significant implications for the specifications of stochastic choice models.[49] An additional advantage of our design is that we are well-positioned to test the nature of these errors. We run specifications that assume normally distributed errors and analogous specifications that assume errors have a Gumbel distribution. We find that the Gumbel specifications provide a better fit. We interpret this as suggesting that the assumption of Gumbel errors is more accurate than the assumption of normal errors. We also note that this result does not appear to depend on the available cognitive resources and this suggests that Gumbel errors could be a general and persistent feature of choice.

We admit that there is much work to be done on the topic. We are interested to learn the insights gleaned from eye-tracking and neuroeconomics techniques in our setting.[50] We are also interested in whether our results on Gumbel errors extend to other stimuli with uncountable

[49] See McFadden (1974) and Yellot (1977).
[50] For instance, see Summerfield and Tsetsos (2012).

measures, for example brightness, loudness, etc. This research would help shed light on the generality of the specifications of stochastic choice models.

Whereas our design entailed objects valued according to only a single attribute, we hope that future designs will study behavior in settings where the objects are valued based on multiple attributes (Gabaix et al., 2006; Sanjurjo, 2015, 2017). Specifically, we are interested to learn if classic multiple attribute effects, such as the decoy effect, can be replicated in this setting and if the attributes interact as compliments or substitutes.[51]

Finally, in our design, subjects were forced to select only a single object from the choice set. We are interested to study behavior if subjects are not constrained to select only one, and are able to select more than one object. Such a multiple selection could be interpreted as indifference if the received object was randomly selected among the chosen objects. We leave these and other interesting questions to future research.

References

Achtziger, A. Alós-Ferrer, C., & Ritschel, A. (2020). Cognitive load in economic decisions. Working paper, University of Zurich.

Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy, 125*(1), 40–68.

Aguiar, V.H., Boccardi, M. J., & Dean, M. (2016). Satisficing and stochastic choice. *Journal of Economic Theory, 166*, 445–482.

Ahumada, A., & Ulku, L. (2018). Luce rule with limited consideration. *Mathematical Social Sciences, 93*, 52–56.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

[51] See Tsetsos, Chater, and Usher (2012) for a different design in a multi-attribute experiment.

Allen, R.J., Baddeley, A.D., & Hitch, G.J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General, 135*(2), 298–313.

Allred, S., Crawford, L.E., Duffy, S., & Smith, J. (2016). Working memory and spatial judgments: Cognitive load increases the central tendency bias. *Psychonomic Bulletin and Review, 23*(6), 1825–1831.

Allred, S., Duffy, S., & Smith, J. (2016). Cognitive Load and Strategic Sophistication. *Journal of Economic Behavior and Organization, 125*, 162–178.

Alós-Ferrer, C., Fehr, E., & Netzer, N. (2021). Time will tell: recovering preferences when choices are noisy. *Journal of Political Economy*, forthcoming.

Alós-Ferrer, C., & Garagnani, M. (2020). Strength of preference and decision making under risk. Working paper, University of Zurich.

Alós-Ferrer, C., & Garagnani, M. (2021). Choice consistency and strength of preference. *Economics Letters, 198*, 109672

Alós-Ferrer, C., Granić, Đ.G., Kern, J., & Wagner, A.K. (2016). Preference reversals: Time and again. *Journal of Risk and Uncertainty, 52*(1), 65–97.

Apesteguia, J., & Ballester, M.A. (2018). Monotone stochastic choice models: The case of risk and time preferences. *Journal of Political Economy, 126*(1), 74–106.

Apesteguia, J., Ballester, M.A., & Lu, J. (2017). Single-Crossing Random Utility Models. *Econometrica, 85*(2), 661–674.

Argenziano, R., & Gilboa, I. (2017). Psychophysical foundations of the Cobb–Douglas utility function. *Economics Letters, 157*, 21–23.

Armel, K.C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making, 3*(5), 396–403.

Armel, K.C., & Rangel, A. (2008). The Impact of Computation Time and Experience on Decision Values. *American Economic Review, 98*(2), 163–168.

Ballinger, T.P., & Wilcox, N.T. (1997). Decisions, Error and Heterogeneity. *Economic Journal, 107*(443), 1090–1105.

Becker, G.M., DeGroot, M.H., & Marschak, J. (1963). Stochastic models of choice behavior. *Systems Research and Behavioral Science, 8*(1), 41–55.

Benjamin, D.J., Brown, S.A., & Shapiro, J.M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association, 11*(6), 1231–1255.

Bhat, C.R. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B: Methodological, 29*(6), 471–483.

Bhui, R. (2019a). A statistical test for the optimality of deliberative time allocation. *Psychonomic Bulletin and Review, 26*(3), 855–867.

Bhui, R. (2019b). Testing Optimal Timing in Value-Linked Decision Making. *Computational Brain and Behavior, 2*(2), 85–94.

Birnbaum, M.H. (2011). Testing mixture models of transitive preference: Comment on Regenwetter, Dana, and Davis-Stober (2011). *Psychological Review, 118*(4), 675–683.

Birnbaum, M.H., & Schmidt, U. (2008). An experimental investigation of violations of transitivity in choice under uncertainty. *Journal of Risk and Uncertainty, 37*(1), 77–91.

Birnbaum, M.H., & Schmidt, U. (2010). Testing transitivity in choice under risk. *Theory and Decision, 69*(4), 599–614.

Blavatskyy, P.R. (2008). Stochastic utility theorem. *Journal of Mathematical Economics, 44*, 1049–1056.

Blavatskyy, P.R. (2011). Probabilistic risk aversion with an arbitrary outcome set. *Economics Letters, 112*(1), 34–37.

Bradley, R.A., & Terry, M.E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika, 39*(3-4), 324–345.

Buckert, M., Oechssler, J., & Schwieren, C. (2017). Imitation under stress. *Journal of Economic Behavior and Organization, 139*, 252–266.

Butler, D.J. (2000). Do non-expected utility choice patterns spring from hazy preferences? An experimental study of choice 'errors'. *Journal of Economic Behavior and Organization, 41*(3), 277–297.

Butler, D., & Loomes, G. (1988). Decision difficulty and imprecise preferences. *Acta Psychologica, 68*(1-3), 183–196.

Butler, D.J., & Loomes, G.C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review, 97*(1), 277–297.

Caplin, A. (2012). Choice Sets as Percepts. In *Neuroscience of Preference and Choice: Cognitive and Neural Mechanisms*, Dolan, R., & Sharot, T. (Eds.), Academic Press, Waltham MA, 295–304.

Caplin, A. (2016). Measuring and modeling attention. *Annual Review of Economics, 8*, 379–403.

Caplin, A., Csaba, D., Leahy, J., & Nov, O. (2020). Rational Inattention, Competitive Supply, and Psychometrics. *Quarterly Journal of Economics, 135*(3), 1681–1724.

Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review, 105*(7), 2183–2203.

Caplin, A., Dean, M., & Leahy, J. (2019). Rational Inattention, Optimal Consideration Sets, and Stochastic Choice. *Review of Economic Studies, 86*(3), 1061–1094.

Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review, 101*(7), 2899–2922.

Caplin, A., & Martin, D. (2015). A testable theory of imperfect perception. *Economic Journal, 125*(582), 184–202.

Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology, 32*(6), 940–950.

Carpenter, J., Graham, M., & Wolf, J. (2013). Cognitive Ability and Strategic Sophistication. *Games and Economic Behavior, 80*(1), 115–130.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences, 98*(9), 5363–5367.

Carrasco, M., Williams, P.E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision, 2*(6), 467–479.

Cattaneo, M.D., Ma, X., Masatlioglu, Y., & Suleymanov, E. (2020). A Random Attention Model. *Journal of Political Economy, 128*(7), 2796–2836.

Cattell, J.M. (1902). The time of perception as a measure of differences in intensity. *Philosophische Studien, 19*, 63–68.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., & Riella, G. (2019). Deliberately stochastic. *American Economic Review, 109*(7), 2425–2445.

Charness, G., Gneezy, U., & Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior and Organization, 131*, 141–150.

Chen, F., & Fischbacher, U. (2016). Response time and click position: Cheap indicators of preferences. *Journal of the Economic Science Association, 2*(2), 109–126.

Cocchi, L., Toepel, U., De Lucia, M., Martuzzi, R., Wood, S.J., Carter, O., & Murray, M.M. (2011). Working memory load improves early stages of independent visual processing. *Neuropsychologia, 49*(1), 92–102.

Conte, A., & Hey, J.D. (2019). Rehabilitating the Random Utility Model. A comment on Apesteguia and Ballester (2018). Working paper, Univerity of York.

Conte, A., Hey, J.D., & Moffatt, P.G. (2011). Mixture models of choice under risk. *Journal of Econometrics, 162*(1), 79–88.

Cox, J.C., Sadiraj, V., & Schmidt, U. (2015). Paradoxes and mechanisms for choice under risk. *Experimental Economics, 18*(2), 215–250.

Cubitt, R.P., Navarro-Martinez, D., & Starmer, C. (2015). On preference imprecision. *Journal of Risk and Uncertainty, 50*(1), 1–34.

Dashiell, J.F. (1937). Affective value-distances as a determinant of esthetic judgment-times. *American Journal of Psychology, 50*(1/4), 57–67.

Dean, M., & Neligh, N. (2017). Experimental Tests of Rational Inattention. Working Paper, Columbia University.

Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica, 26*(3), 440–444.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review, 78*, 97–119.

Deck, C., Jahedi, S., & Sheremeta, R. (2021). On the Consistency of Cognitive Load. *European Economic Review*, forthcoming.

Drichoutis, A.C., & Nayga, R. (2020). Economic rationality under cognitive load. *Economic Journal, 130*(632), 2382–2409.

Duffy, S., Naddeo, J.J., Owens, D., & Smith, J. (2021). Cognitive load and mixed strategies: On brains and minimax. Working paper, Rutgers University-Camden.

Duffy, S., & Smith, J. (2014). Cognitive Load in the Multi-player Prisoner's Dilemma Game: Are There Brains in Games? *Journal of Behavioral and Experimental Economics, 51*, 47–56.

Duffy, S., & Smith, J. (2020). An economist and a psychologist form a line: What can imperfect perception of length tell us about stochastic choice? Working paper, Rutgers University-Camden.

Dutilh, G., & Rieskamp, J. (2016). Comparing perceptual and preferential decision making. *Psychonomic Bulletin and Review, 23*(3), 723–737.

Echenique, F., Saito, K., & Tserenjigmid, G. (2018). The perception-adjusted Luce model. *Mathematical Social Sciences, 93*, 67–76.

Falmagne, J.C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology, 18*(1), 52–72.

Fechner, G.T. (1860). *Elemente der Psychophysik*. (*Elements of psychophysics*, translated 1966. Holt, Rinehart, and Winston, New York.)

Franco-Watkins, A.M., Rickard, T.C., & Pashler, H. (2010). Taxing executive processes does not necessarily increase impulsive decision making. *Experimental Psychology, 57*, 193–201.

Fudenberg, D., Iijima, R., & Strzalecki, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica, 83*(6), 2371–2409.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review, 108*(12), 3651–3684.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96*(4), 1043–1068.

Geng, S. (2016). Decision time, consideration time, and status quo bias. *Economic Inquiry, 54*(1), 433–449.

Gilbert, D.T., Pelham, B.W., & Krull, D.S. (1988). On Cognitive Busyness: When Person Perceivers Meet Persons Perceived. *Journal of Personality and Social Psychology, 54*(5), 733–740.

Gul, F., Natenzon, P., & Pesendorfer, W. (2014). Random choice as behavioral optimization. *Econometrica, 82*(5), 1873–1912.

Gul, F., & Pesendorfer, W. (2006). Random expected utility. *Econometrica, 74*(1), 121–146.

Hauge, K.E., Brekke, K.A., Johansson, L.O., Johansson-Stenman, O., & Svedsäter, H. (2016). Keeping others in our mind or in our heart? Distribution games under cognitive load. *Experimental Economics, 19*(3), 562–576.

Henmon, V.A.C. (1911). The relation of the time of a judgment to its accuracy. *Psychological Review, 18*(3), 186–201.

Hey, J.D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review, 39*(3-4), 633–640.

Horan, S., Manzini, P., & Mariotti, M. (2019). When Is Coarseness Not a Curse? Comparative Statics of the Coarse Random Utility Model. Working paper, University of Sussex.

Kellogg, W.N. (1931). The time of judgment in psychometric measures. *American Journal of Psychology, 43*(1), 65–86.

Khaw, M.W., Li, Z., & Woodford, M. (2017). Risk aversion as a perceptual bias. Working paper, National Bureau of Economic Research.

Koida, N. (2018). Anticipated stochastic choice. *Economic Theory, 65*(3), 545–574.

Kovach, M., & Tserenjigmid, G. (2021). The focal Luce model. *American Economic Journal: Microeconomics*, forthcoming.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences, 108*(33), 13852–13857.

Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research, 36*(2), 173–187.

Lévy-Garboua, L., Maafi, H., Masclet, D., & Terracol, A. (2012). Risk aversion and framing effects. *Experimental Economics, 15*(1), 128–144.

Liu, T., Abrams, J., & Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science, 20*(3), 354–362.

Lleras, J.S., Masatlioglu, Y., Nakajima, D., & Ozbay, E.Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory, 170*, 70–85.

Loomes, G., & Pogrebna, G. (2014). Measuring individual risk attitudes when preferences are imprecise. *Economic Journal, 124*(576), 569–593.

Loomes, G., Starmer, C., & Sugden, R. (1989). Preference reversal: information-processing effect or rational non-transitive choice? *Economic Journal, 99*(395), 140–151.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review, 39*(3-4), 641–648.

Loomis, J., Peterson, G., Champ, P., Brown, T., & Lucero, B. (1998). Paired comparison estimates of willingness to accept versus contingent valuation estimates of willingness to pay. *Journal of Economic Behavior and Organization, 35*(4), 501–515.

Lu, J. (2016). Random choice and private information. *Econometrica, 84*(6), 1983–2027.

Luce, R.D. (1959a). *Individual choice behavior: A theoretical analysis*. Wiley: New York.

Luce, R.D. (1959b). On the possible psychophysical laws. *Psychological Review, 66*(2), 81–95.

Luce, R.D. (1994). Thurstone and Sensory Scaling: Then and Now. *Psychological Review, 101*(2), 271–277.

Luce, R.D. (2005). Measurement analogies: Comparisons of behavioral and physical measures. *Psychometrika, 70*(2), 227–251.

Machina, M.J. (1985). Stochastic choice functions generated from deterministic preferences over lotteries. *Economic Journal, 95*(379), 575–594.

Manzini, P., & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica, 82*(3), 1153–1176.

Masatlioglu, Y., Nakajima, D., & Ozbay, E.Y. (2012). Revealed Attention. *American Economic Review, 102*(5), 2183–2205.

Mas-Colell, A., Whinston, M.D., & Green, J.R. (1995). *Microeconomic Theory*, New York: Oxford University Press.

Matějka, F., & McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review, 105*(1), 272–298.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*, Zarembka, P. (Ed.), New York, Academic Press, 105–142.

McFadden, D. (1976). Quantal choice analaysis: A survey. *Annals of Economic and Social Measurement, 5*(4), 363–390.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*, Manski, C., & McFadden, D. (Eds.), Cambridge, MA, MIT Press, 198–272.

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

Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the Prisoner's Dilemma. *Proceedings of the National Academy of Sciences, 95*(23), 13755–13758.

Moffatt, P.G. (2005). Stochastic choice and the allocation of cognitive effort. *Experimental Economics, 8*(4), 369–388.

Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy, 59*(5), 371–404.

Morey, C.C., & Bieler, M. (2013). Visual short-term memory always requires general attention. *Psychonomic Bulletin and Review, 20*(1), 163–170.

Morey, C.C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin and Review, 11*(2), 296–301.

Münsterberg, H. (1894). Studies from the Harvard Psychological Laboratory: (I). *Psychological Review, 1*(1), 34–60.

Natenzon, P. (2019). Random choice and learning. *Journal of Political Economy, 127*(1), 419–457.

Navarro-Martinez, D., Loomes, G., Isoni, A., Butler, D., & Alaoui, L. (2018). Boundedly rational expected utility theory. *Journal of Risk and Uncertainty, 57*(3), 199–223.

Olschewski, S., Newell, B., & Scheibehenne, B. (2019). How Basic Cognition Influences Experience-Based Economic Valuation. Working paper University of Basel.

Olschewski, S., Rieskamp, J., & Scheibehenne, B. (2018). Taxing cognitive capacities reduces choice consistency rather than preference: A model-based test. *Journal of Experimental Psychology: General, 147*(4), 462–484.

Oud, B., Krajbich, I., Miller, K., Cheong, J.H., Botvinick, M., & Fehr, E. (2016). Irrational time allocation in decision-making. *Proceedings of the Royal Society B: Biological Sciences, 283*(1822), 20151439.

Payne, J.W., Bettman, J.R., & Johnson, E.J. (1993). *The Adaptive Decision Maker.* Cambridge University Press.

Payne, J.W., Braunstein, M.L., & Carroll, J.S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Performance, 22*(1), 17–44.

Polanía, R., Krajbich, I., Grueschow, M., & Ruff, C.C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3), 709–720.

Psychology Software Tools, Inc. [E-Prime 2.0] (2012). Retrieved from http://www.pstnet.com.

Regenwetter, M., Dana, J., & Davis-Stober, C.P. (2011). Transitivity of preferences. *Psychological Review, 118*(1), 42–56.

Regenwetter, M., Dana, J., Davis-Stober, C.P., & Guo, Y. (2011). Parsimonious testing of transitive or intransitive preferences: Reply to Birnbaum (2011). *Psychological Review, 118*(4), 684–688.

Regenwetter, M., & Davis-Stober, C.P. (2012). Behavioral variability of choices versus structural inconsistency of preferences. *Psychological Review, 119*(2), 408–416.

Reutskaja, E., Nagel, R., Camerer, C.F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review, 101*(2), 900–926.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34* (6), 1446–1465.

Roch, S.G., Lane, J.A.S., Samuelson, C.D., Allison, S.T., & Dent, J.L. (2000). Cognitive Load and the Equality Heuristic: A Two-Stage Model of Resource Overconsumption in Small Groups. *Organizational Behavior and Human Decision Processes, 83*(2), 185–212.

Rubinstein, A., & Salant, Y. (2006). A model of choice from lists. *Theoretical Economics, 1*(1), 3–17.

Sanjurjo, A. (2015). Search, memory, and choice error: an experiment. *PloS ONE, 10*(6), e0126508.

Sanjurjo, A. (2017). Search with multiple attributes: Theory and empirics. *Games and Economic Behavior, 104*, 535–562.

Schulz, J.F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology, 41*, 77–87.

Schwarz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464.

Smith, V.L. (1976). Experimental economics: Induced value theory. *American Economic Review, 66*(2), 274–279.

Sopher, B., & Gigliotti, G. (1993). Intransitive cycles: Rational Choice or random error? An answer based on estimation of error rates with experimental data. *Theory and Decision, 35*(3), 311–336.

Sopher, B., & Narramore, J.M. (2000). Stochastic choice and consistency in decision making under risk: An experimental study. *Theory and Decision, 48*(4), 323–350.

Summerfield, C., & Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Frontiers in Neuroscience, 6*, 70.

Swann, W.B., Hixon, G, Stein-Seroussi, A., & Gilbert, D.T. (1990). The Fleeting Gleam of Praise: Cognitive Processes Underlying Behavioral Reactions to Self-Relevant Feedback. *Journal of Personality and Social Psychology, 59* (1), 17–26.

Thurstone, L.L. (1927a). A law of comparative judgment. *Psychological Review, 34*(4), 273–286.

Thurstone, L.L. (1927b). Psychophysical Analysis. *American Journal of Psychology, 38*(3), 368–389.

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences, 109*(24), 9659–9664.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences, 113*(11), 3102–3107.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*(1), 31–48.

Tyson, C.J. (2008). Cognitive constraints, contraction consistency, and the satisficing criterion. *Journal of Economic Theory, 138*(1), 51–70.

Volkmann, J. (1934). The relation of the time of judgment to the certainty of judgment. *Psychological Bulletin, 31*(9), 672–673.

Weber, E. (1834). *De Tactu.* (*The Sense of Touch*, translated 1978. Academic Press, New York.)

Weibull, J.W., Mattsson, L.G., & Voorneveld, M. (2007). Better may be worse: Some monotonicity results and paradoxes in discrete choice under uncertainty. *Theory and Decision, 63*(2), 121–151.

Wilcox, N.T. (2011). ‘Stochastically more risk averse:’ A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics, 162*(1), 89–104.

Woodford, M. (2014). Stochastic choice: An optimizing neuroeconomic model. *American Economic Review, 104*(5), 495–500.

Yellott, J.I. (1977). The relationship between Luce’s Choice Axiom, Thurstone’s Theory of Comparative Judgment, and the double exponential distribution. *Journal of Mathematical Psychology, 15*(2), 109–144.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*(6706), 72–75.

Zeigenfuse, M.D., Pleskac, T.J., & Liu, T. (2014). Rapid decisions from experience. *Cognition, 131*(2), 181–194.

Zokaei, N., Heider, M., & Husain, M. (2014). Attention is required for maintenance of feature binding in visual working memory. *Quarterly Journal of Experimental Psychology, 67*(6), 1191–1213.

Appendix for Online Publication

Cognitive load summary statistics

As each of the 92 subjects attempt 50 high load memorization tasks and 50 low load memorization tasks, Table $A1$ presents a characterization of the subject-level distribution of the number of correct memorization tasks by cognitive load treatment and the number pooled across treatments.

Table $A1$: Distribution of subjects by number of correct memorization tasks

	Restricted to cognitive load treatments							
Number correct	$46-50$	$41-45$	$36-40$	$31-35$	$26-30$	$21-25$	< 21	Total
High load	50	17	11	5	4	3	2	92
Low load	88	4	0	0	0	0	0	92

	Pooled across cognitive load treatments							
Number correct	$96-100$	$91-95$	$86-90$	$81-85$	$76-80$	$71-75$	< 71	Total
Pooled	40	24	13	4	5	1	5	92

The upper panel characterizes the subject-level distribution of the number of correct memorization tasks by cognitive load treatment. The lower panel characterizes the subject-level distribution of the correct memorization tasks across both cognitive load treatments.

Selected longest summary statistics

Table $A2$ characterizes the Selected longest variable in the cognitive load and difficulty treatments.

Table $A2$: Selected longest variable by difficulty treatment

	Easy	Medium	Difficult	Pooled
High load	94.6%	73.1%	37.0%	68.9%
	1497 of 1582	1124 of 1538	548 of 1480	3169 of 4600
Low load	96.8%	76.3%	38.5%	69.6%
	1440 of 1487	1140 of 1495	623 of 1618	3203 of 4600
Pooled	95.7%	74.6%	37.8%	69.3%
	2937 of 3069	2264 of 3033	1171 of 3089	6372 of 9200

It appears to be the case that the difficulty treatments were successful in that the longest line is more likely to be selected in the easy treatment. Table $A3$ characterizes the variable by cognitive load and number of lines treatments.

Table $A3$: Selected longest variable by number of lines treatment

	2 Lines	3 Lines	4 Lines	5 Lines	6 Lines
High load	79.0%	74.0%	71.1%	62.3%	57.9%
	710 of 899	690 of 932	674 of 948	580 of 931	515 of 890
Low load	78.0%	75.0%	68.0%	66.4%	61.1%
	700 of 899	720 of 960	613 of 902	588 of 886	582 of 953
Pooled	78.4%	74.5%	69.6%	64.3%	59.5%
	1410 of 1798	1410 of 1892	1287 of 1850	1168 of 1817	1097 of 1843

It also appears that the probability that the longest line is selected is decreasing in the number of available lines. This appears to be suggestive of choice overload, even from a choice set of only a few simple objects of choice. Table $A4$ characterizes the variable in the cognitive load and longest line length treatments.

Table $A4$: Selected longest variable by longest line length treatment

	160	176	192	208	224	240	256	272	288	304
High load	71.1%	72.0%	69.1%	70.7%	70.4%	70.4%	66.7%	71.5%	64.4%	62.6%
Low load	71.7%	73.9%	75.0%	69.8%	69.4%	68.5%	66.3%	68.0%	67.6%	66.1%
Pooled	71.4%	72.9%	72.1%	70.2%	69.9%	69.5%	66.5%	69.8%	66.0%	64.3%

The Pooled values each have 920 observations. The values restricted to a cognitive load treatment each have 460 observations.

This suggests that the quality of choices decreases in the length of the longest line. In Table $A5$ we characterize the variable according to the number of lines and the letter label of the longest line.

Table $A5$: Selected longest variable by number of lines and letter label of the longest

	A	B	C	D	E	F
2 Lines	77.0%	79.9%	–	–	–	–
	705 of 916	705 of 882				
3 Lines	72.5%	72.5%	78.7%	–	–	–
	470 of 648	457 of 630	483 of 614			
4 Lines	64.8%	62.0%	71.6%	79.3%	–	–
	289 of 446	279 of 450	351 of 490	368 of 464		
5 Lines	64.1%	58.0%	62.8%	70.8%	66.0%	–
	236 of 368	215 of 371	219 of 349	250 of 353	248 of 376	
6 Lines	50.8%	52.8%	50.0%	60.2%	64.5%	78.7%
	167 of 329	161 of 305	144 of 288	197 of 327	180 of 279	248 of 315

There appear to be differences in accuracy conditional on the letter label of the longest line. Tables $A2 - A5$ suggest the relevant variables that should be included in the analysis of the Selected longest line variable.

More on the quality of choices

In order to investigate the optimality of choices, in Table 1 we summarized logistic regressions of the Selected longest variable. Here we conduct the analogous analysis, but restricted to the 8437 trials in which the cognitive load task was performed correctly. We summarize these regressions in Table $A6$.

Table $A6$: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	-0.0201^{*}	-0.0233^{*}	-0.0217^{*}	-0.0244^{*}
	(0.0094)	(0.0102)	(0.0093)	(0.0102)
Longest line normalized	-0.0005^{***}	-0.0006^{***}	-0.0005^{***}	-0.0006^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.052^{***}	–	-0.051^{***}	–
	(0.003)		(0.003)	
Easy treatment dummy	0.354^{***}	0.389^{***}	0.369^{***}	0.410^{***}
	(0.014)	(0.016)	(0.014)	(0.016)
Difficult treatment dummy	-0.279^{***}	-0.306^{***}	-0.278^{***}	-0.307^{***}
	(0.011)	(0.013)	(0.012)	(0.013)
Trial	0.00008	0.00006	0.00004	0.00002
	(0.00016)	(0.00018)	(0.00016)	(0.00018)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	7496.6	7358.3	7339.6	7201.3

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 8437 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The qualitative results are not changed from those in Table 1. Likewise, we conduct the analysis in Table 1 but restricted to the 9001 trials in which a line was viewed and a line was selected. We summarize these regressions in Table *A*7.

Table *A*7: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
High load	−0.0213*	−0.0237*	−0.0215***	−0.0232*
	(0.0087)	(0.0095)	(0.0086)	(0.0094)
Longest line normalized	−0.0005***	−0.0005***	−0.0005***	−0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	−0.048***	–	−0.047***	–
	(0.003)		(0.003)	
Easy treatment dummy	0.360***	0.397***	0.368***	0.410***
	(0.013)	(0.015)	(0.013)	(0.015)
Difficult treatment dummy	−0.266***	−0.293***	−0.265***	−0.294***
	(0.011)	(0.012)	(0.011)	(0.012)
Trial	−0.00019	−0.00023	−0.00017	−0.00022
	(0.00015)	(0.00017)	(0.00015)	(0.00016)
Letter dummies	*No*	*Yes*	*No*	*Yes*
Fixed effects	*No*	*No*	*Yes*	*Yes*
AIC	7873.4	7702.5	7761.9	7594.0

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9001 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Again, the qualitative results are not changed from those in Table 1. In order to further explore the robustness the results in Table 1, we perform the analogous exercise but we analyze the *Longest minus selected variable*, defined to be the length of the longest line minus the length of the selected line. As this variable is bounded below by 0 we perform tobit regressions. The regressions are otherwise identical to those in Table 1. We summarize these tobit regressions in Table $A8$.

Table $A8$: Tobit regressions of Longest minus selected variable

	(1)	(2)	(3)	(4)
High load	6.795***	7.038***	6.689***	6.948***
	(1.823)	(1.826)	(1.775)	(1.777)
Longest line normalized	0.133***	0.132***	0.131***	0.131***
	(0.020)	(0.020)	(0.019)	(0.019)
Number of lines normalized	9.934***	–	9.843***	–
	(0.661)		(0.645)	
Easy treatment dummy	−53.638***	−53.784***	−56.172***	−56.316***
	(2.955)	(2.963)	(2.974)	(2.982)
Difficult treatment dummy	34.928***	34.795***	34.286***	34.180***
	(2.082)	(2.085)	(2.033)	(2.037)
Trial	−0.172***	−0.171***	−0.173***	−0.171***
	(0.031)	(0.031)	(0.031)	(0.031)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	35693	35646	35415	35370

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$.

Similar to Table 1, the accuracy of the choice decreases when there is a larger number of lines, decreases in the length of the longest line, and decreases in the difficulty of the decision. Further, in every specification, we see that the high load coefficient is positive. This implies that choices are worse in the high cognitive load treatment. The negative and significant trial coefficient perhaps suggests that very poor line selections become less frequent across trials.

Here we include an analysis similar to that summarized in Table 1, but we include interactions with the difficulty treatment variables. Specifically, we include the difficulty interactions involving the Longest line normalized variable and the Number of lines normalized. This is summarized in Table $A9$.

Table $A9$: Logistic regressions of the Selected longest line variable

	(1)	(2)
High load	-0.0264^{**}	-0.0261^{**}
	(0.0092)	(0.0091)
Easy treatment dummy	0.2025^{***}	0.2026^{***}
	(0.0089)	(0.0088)
Difficult treatment dummy	-0.3874^{***}	-0.4013^{***}
	(0.0120)	(0.0124)
Number of lines normalized	-0.0604^{***}	-0.0615^{***}
	(0.0056)	(0.0056)
Easy*Number of lines norm	0.0465^{***}	0.0498^{***}
	(0.0060)	(0.0060)
Difficult*Number of lines norm	-0.0078	-0.0088
	(0.0085)	(0.0086)
Longest line normalized	-0.00109^{***}	-0.00112^{***}
	(0.00017)	(0.00017)
Easy*Longest line norm	0.00097^{***}	0.00099^{***}
	(0.00018)	(0.00018)
Difficult*Longest line norm	0.00089^{***}	0.00092^{***}
	(0.00026)	(0.00026)
Trial	0.00007	0.00004
	(0.00016)	(0.00016)
Letter dummies	No	No
Fixed effects	No	Yes
AIC	8328.7	8162.3

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The results are largely unchanged from Table 1, but we mention the possibly interesting interaction effects. We find positive and significant interactions with the Longest line variable. This suggests that the diminishing optimality with line length is strongest in the medium difficulty treatment. Also, the interaction of the Number of lines and the Easy variable is positive and significant. This suggests that the diminishing optimality in the number of lines is weaker in the easy treatment than the medium treatment. However, we note that the interaction of the Number of lines and the Difficult variable is not significant.

To further explore the robustness of the results in Table 1, we conduct random-effects regressions. Whereas specifications (3) and (4) in Table 1 employed fixed-effects, here we conduct a similar analysis but with random-effects. Specifically, we estimate an exchangeable covariance matrix, clustered by subject. In other words, we assume a unique relationship between any two observations involving a particular subject. However, we assume that observations involving two different subjects are statistically independent. The regressions are estimated using Generalized Estimating Equations (GEE). Since GEE is not a likelihood based method, Akaike Information Criterion is not available. Therefore, we provide the Quasilikelihood information criterion (QIC). This is summarized in Table $A10$.

Table $A10$: Logistic regressions of the Selected longest line variable

	(1)	(2)
High load	-0.0269^{*}	-0.0287^{*}
	(0.0105)	(0.0114)
Longest line normalized	-0.0005^{***}	-0.0006^{***}
	(0.0001)	(0.0001)
Number of lines normalized	-0.054^{***}	–
	(0.004)	
Easy treatment dummy	0.362^{***}	0.395^{***}
	(0.019)	(0.023)
Difficult treatment dummy	-0.287^{***}	-0.305^{***}
	(0.012)	(0.012)
Trial	0.00007	0.00005
	(0.00019)	(0.00020)
Letter dummies	No	Yes
Random effects	Yes	Yes
QIC	8357.2	8209.1

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the correlation estimates. QIC refers to the Quasi-likelihood information criterion. Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The results are largely unchanged from Table 1 and our results do not appear to be sensitive to the specification of the repeated measures.

Memorization task and line selection task

Here we explore the relationship between success on the memorization task in a trial and outcomes in the line selection task in that trial. This effort is driven by the concern that subjects might have known that the memorization task would be incorrect in that trial and this could have affected outcomes in the line selection task. While incentives were designed to prevent this from occurring, we look for evidence of such outcomes.

In order to measure success on the memorization task, we define the *Cognitive load correct* variable to be a 1 if the cognitive load task in that trial was correct, and a 0 otherwise. We conduct regressions analogous to those in Tables 1 and 2, with the following differences. We restrict attention to a cognitive load treatment, we include Cognitive load correct as an independent variable, and we do not include the High load dummy. We therefore conduct 16 regressions and we report the relevant estimates of the effect of the Cognitive load correct variable. The upper panel is analogous to Table 1, with Selected longest as dependent variable and we report the marginal effects of the Cognitive load correct variable. The lower panel is analogous to Table 2, with Implicit response time as dependent variable and we report the coefficient estimates of the Cognitive load correct variable.

Table $A11$: Cognitive load correct estimates restricted to cognitive load treatments

	DV: Selected longest, Marginal effects			
	(1)	(2)	(3)	(4)
Cognitive load correct, restricted to HL	0.0656^{***}	0.0660^{**}	$0.0374^{\dagger}$	0.0362
	(0.0189)	(0.0206)	(0.0208)	(0.0229)
Cognitive load correct, restricted to LL	0.1103^{**}	0.1182^{**}	0.0999^{*}	0.1080^{*}
	(0.0397)	(0.0424)	(0.0389)	(0.0423)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
	DV: Implicit response time, Coefficient estimates			
	(1)	(2)	(3)	(4)
Cognitive load correct, restricted to HL	-0.0877	-0.0799	-0.0813	-0.0675
	(0.1186)	(0.1184)	(0.1154)	(0.1154)
Cognitive load correct, restricted to LL	-0.6250^{*}	-0.5899^{*}	-0.5000^{*}	$-0.4622^{\dagger}$
	(0.2666)	(0.2667)	(0.2375)	(0.2374)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes

Each cell presents the estimates related to the Cognitive load correct variable in a regression restricted to a cognitive load treatment. In the upper panel, we provide the marginal effects evaluated at the sample means and the standard errors in parentheses. The dependent variable is Selected longest. In the lower panel, we provide the coefficient estimates and the standard errors in parentheses. The dependent variable is Implicit response time. Each regression has 4600 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Restricted either to high or low load, regressions without fixed-effects (specifications (1) and (2)) show that trials in which the memorization was correct are more likely to have the longest line selected than trials in which the memorization was not correct. This result could be driven by the individual differences that are behind the within-subject relationship between the total number of correct memorization tasks and the total number of optimal line selections. However, regressions with fixed-effects (specifications (3) and (4)) only show this relationship when restricted to low load.

Restricted to the high load treatment, we do not find evidence that the Implicit response time is related to the Cognitive load correct variable. This suggests to us that the level of effort in high load trials did not depend on whether the memorization task was correct. On the other hand, in the low load treatment trials, we find evidence that Implicit response time was larger in incorrect trials than in correct trials. Perhaps in these rare incorrect low load trials, subjects invested more effort in the line selection task.

Overall, we find some evidence that the accuracy in the memorization task is related to the behavior in the line selection task in that trial. However, this evidence is confined to the rare instances of incorrect memorization tasks in the low load treatment. Low load trials with an incorrect memorization task are associated with a lower likelihood of selecting the optimal line and a longer deliberation time than low load trials with a correct memorization task. Combined with their relatively rare occurrences, the results of Table $A11$ suggest that the effect of the high load treatment in Table 1 is possibly slightly underestimated but possibly slightly overestimated in Table 2.

More on the quality of searches

In Table 2 we investigated the quality of searches by conducting an analysis with Implicit response time as the dependent variable. Whereas the time deliberating is one measure of effort in the line selection task, it is not the only such measure. Another measure of the effort expended is the number of line view clicks. It would seem that a larger such number would be associated with more effort in identifying the optimal line. We define the *View clicks* variable as the number of total line view clicks during the search stage. We conduct an analysis identical to Table 2, with the exception that the dependent variable is View clicks. Table $A12$ summarizes this analysis.

Table $A12$: Regressions of the View clicks variable

	(1)	(2)	(3)	(4)
High load	-0.339^{***}	-0.346^{***}	-0.340^{***}	-0.348^{***}
	(0.049)	(0.049)	(0.040)	(0.0401)
Longest line normalized	-0.0018^{***}	-0.0018^{***}	-0.0018^{***}	-0.0018^{***}
	(0.0005)	(0.0005)	(0.0004)	(0.0004)
Number of lines normalized	1.082^{***}	–	1.083^{***}	–
	(0.017)		(0.0143)	
Easy treatment dummy	-1.458^{***}	-1.469^{***}	-1.420^{***}	-1.430^{***}
	(0.060)	(0.060)	(0.050)	(0.049)
Difficult treatment dummy	0.654^{***}	0.639^{***}	0.655^{***}	0.643^{***}
	(0.060)	(0.060)	(0.050)	(0.049)
Trial	0.0034^{***}	0.0035^{***}	0.0034^{***}	0.0035^{***}
	(0.0009)	(0.0008)	(0.0007)	(0.0007)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	41891.0	41810.8	38307.7	38209.2

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$.

Here we observe fewer View clicks in the high load treatment. This suggests that the cognitive load manipulation is negatively affecting the quality of the searches. We also observe that View clicks is increasing in the number of available lines. Further, we observe that View clicks is decreasing in the length of the longest line. This suggests that subjects expended less effort in the searches involving longer lines. Additionally, we observe more View clicks in the Difficult treatment and fewer in the Easy treatment. We note that Reutskaja et al. (2011), Krajbich, Armel, and Rangel (2010), and Krajbich and Rangel (2011) find similar results.

More on the relationship between choice and search

In Table 1, we explored whether subjects optimally selects the longest line by conducting regressions with the Selected longest line variable. Another question to ask is whether subjects selected the longest line, among the lines that were viewed. We define the *Selected longest line viewed* variable as a 1 if the longest line among those viewed was selected, and a 0 otherwise. We conduct an analysis similar to Table 1, but rather than using the Selected longest line variable, we employ the Selected longest line viewed variable. We summarize these regressions in Table $A13$.

Table $A13$: Logistic regressions of Selected longest line viewed variable

	(1)	(2)	(3)	(4)
High load	-0.0241^{**}	-0.0268^{**}	-0.0239^{**}	-0.0262^{**}
	(0.0092)	(0.010)	(0.0091)	(0.010)
Longest line normalized	-0.0005^{***}	-0.0005^{***}	-0.0005^{***}	-0.0005^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	-0.051^{***}	–	-0.052^{***}	–
	(0.003)		(0.003)	
Easy treatment dummy	0.360^{***}	0.395^{***}	0.367^{***}	0.407^{***}
	(0.014)	(0.015)	(0.014)	(0.015)
Difficult treatment dummy	-0.281^{***}	-0.307^{***}	-0.284^{***}	-0.312^{***}
	(0.011)	(0.012)	(0.011)	(0.012)
Trial	0.00019	0.00018	0.00018	0.00017
	(0.00016)	(0.00017)	(0.00016)	(0.00017)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes
AIC	8305.4	8134.4	8176.8	8005.0

We provide the marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, or the subject-specific dummies in the fixed effects regressions. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has 9200 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Even when we restrict attention to the set of viewed lines, we still find evidence that subjects in the high load treatment make worse choices. Therefore, consideration set effects cannot fully explain the relationship between cognitive load and the Selected longest variable, as summarized in Table 1. Additionally, we note a negative relationship between the quality of choices among the lines that were viewed and the length of the longest line. Finally, we note the negative relationship between selecting the longest line viewed and the number of lines in the choice set.[52]

Robustness of the relationship between quality of choice and response times

In order to test the robustness of Table 3, we conduct a similar analysis here. We perform an analysis, similar to that summarized in Table 2 but we include Selected longest as an independent variable. Further, for the specifications without fixed-effects, we include an independent variable that is the average of the Implicit response time for that particular subject. We summarize this analysis in Table $A14$.

Table $A14$: Regressions of Implicit response time variable

	(1)	(2)	(3)	(4)
High load	-0.424^{***}	-0.429^{***}	-0.424^{***}	-0.429^{***}
	(0.051)	(0.051)	(0.051)	(0.051)
Longest line normalized	0.0021^{***}	0.0020^{***}	0.0021^{***}	0.0021^{***}
	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Number of lines normalized	1.068^{***}	–	1.057^{***}	–
	(0.018)		(0.018)	
Easy treatment dummy	-1.968^{***}	-1.985^{***}	-1.960^{***}	-1.977^{***}
	(0.064)	(0.064)	(0.064)	(0.064)
Difficult treatment dummy	0.712^{***}	0.714^{***}	0.715^{***}	0.717^{***}
	(0.067)	(0.067)	(0.067)	(0.067)
Trial	-0.0076^{***}	-0.0075^{***}	-0.0076^{***}	-0.0075^{***}
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Selected Longest	-0.705^{***}	-0.673^{***}	-0.668^{***}	-0.639^{***}
	(0.067)	(0.067)	(0.065)	(0.066)
Sub's Average RT	–	–	0.971^{***}	0.969^{***}
			(0.019)	(0.018)
Letter dummies	No	Yes	No	Yes
Fixed effects	Yes	Yes	No	No
AIC	42489.7	42469.2	42480.4	42459.7

[52] Reutskaja et al. (2011) find a similar relationship in their data.

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Letter dummies, the subject-specific dummies in the fixed effects regressions or the demographics estimates. AIC refers to the Akaike information criterion. Each regression has 9200 observations. *** denotes $p < 0.001$.

In every specification, the Selected Longest variable is negative and significant. We interpret this as suggesting that, even when controlling for the fixed characteristics of the subjects, suboptimal choices tend to take longer than optimal choices. Due to the possible endogeneity present in the analysis, we conduct Spearman correlations between the unstandardized residuals and the Selected longest variable. In specifications $(1)-(4)$, the p-values, respectively, are 0.019, 0.019, 0.015, and 0.015.[53] Similar to Table 3, we find a negative relationship between the Selected longest variable and Implicit response time. Although we note that the significant correlations suggest that this result might be driven by endogeneity.

Robustness of the multinomial discrete choice models

In Table 4, we conducted comparisons of different multinomial discrete choice models. In Tables $A15$ and $A16$, we conduct the identical analysis but restricted to either the high cognitive load treatment or the low cognitive load treatment.

[53] When we restrict the analysis to the 9001 observations with a line view and a line selection, the p-values, respectively, are 0.049, 0.048, 0.051, and 0.051.

Table $A15$: Comparisons of different multinomial discrete choice models for high load

		Cond Logit	Multi Probit 1	HEV	Multi Probit 2	Trials
		(1)	(2)	(3)	(4)	
2 Lines	β est.	0.134	0.103	–	–	890
	AIC	689	693			
	BIC	694	698			
3 Lines	β est.	0.134	0.093	0.125	0.084	917
	AIC	983	999	981	996	
	BIC	988	1004	991	1006	
4 Lines	β est.	0.110	0.113	0.109	0.125	933
	AIC	1387	1607	1388	1543	
	BIC	1392	1612	1402	1557	
5 Lines	β est.	0.097	0.128	0.106	0.158	912
	AIC	1680	1998	1678	1756	
	BIC	1685	2003	1698	1775	
6 Lines	β est.	0.089	0.059	0.066	0.044	856
	AIC	1867	1879	1792	1824	
	BIC	1871	1883	1816	1848	

We provide the estimates of β, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the various models restricted to treatments with identical numbers of lines. Each of the estimates for β are significantly different from 0 with $p < 0.001$.

In every model, there is evidence that logistic errors provide a better fit than normal errors in high cognitive load trials. We perform the analysis for trials in the low cognitive load treatment.

Table $A16$: Comparisons of different multinomial discrete choice models for low load

		Cond Logit (1)	Multi Probit 1 (2)	HEV (3)	Multi Probit 2 (4)	Trials
2 Lines	β est.	0.128	0.094	–	–	895
	AIC	730	740			
	BIC	734	745			
3 Lines	β est.	0.123	0.081	0.113	0.074	954
	AIC	1106	1139	1100	1139	
	BIC	1111	1144	1109	1149	
4 Lines	β est.	0.122	0.118	0.119	0.126	893
	AIC	1331	1422	1334	1399	
	BIC	1336	1427	1349	1414	
5 Lines	β est.	0.131	0.129	0.122	0.140	868
	AIC	1486	1613	1488	1583	
	BIC	1490	1618	1507	1602	
6 Lines	β est.	0.100	0.066	0.075	0.048	924
	AIC	1907	1928	1825	1862	
	BIC	1911	1933	1849	1886	

We provide the estimates of β, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the various models restricted to treatments with identical numbers of lines. Each of the estimates for β are significantly different from 0 with $p < 0.001$.

Again, the evidence in our low cognitive load trials suggest that the assumption of Gumbel errors provides a better fit than the assumption of normal errors.

We note that the estimates of β vary among the models summarized in Tables 4, $A15$, and $A16$, which is possibly affecting our results. In order to address this possibility, we offer an analysis, identical to that summarized in Table 4, however we add an additional restriction that $\beta = 0.1$. This analysis is summarized in Table $A17$.

Table $A17$: Comparisons of different restricted multinomial discrete choice models

		Cond Logit (1)	Multi Probit 1 (2)	HEV (3)	Multi Probit 2 (4)	Trials
2 Lines	AIC	1435	1430	–	–	1785
	BIC	1435	1430			
3 Lines	AIC	2116	2154	2087	2154	1871
	BIC	2116	2154	2093	2160	
4 Lines	AIC	2729	2903	2722	2810	1826
	BIC	2729	2903	2733	2821	
5 Lines	AIC	3186	3317	3190	3241	1780
	BIC	3186	3317	3207	3257	
6 Lines	AIC	3776	4153	3691	4097	1780
	BIC	3776	4153	3713	4119	

We provide the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the various models restricted to treatments with identical numbers of lines. We have restricted $\beta = 0.1$ in each specification.

Similar to the analysis summarized in Table 4, with the exception of the 2 Lines treatment, both the AIC and BIC are lower for the specifications with Gumbel errors than for normal errors. In 17 of 18 comparisons, the AIC of the Gumbel error specification is lower than that for the normal error specification. Likewise, in 17 of 18 comparisons, the BIC of the Gumbel error specification is lower than that for the normal error specification.

In order to further explore the robustness of our results, we conduct an analysis similar to specifications (1) and (2) in Table 4, however we estimate different Length coefficients for the high load and the low load. We denote the coefficient estimates for the high and low load trials, respectively, as β_{HL} and β_{LL}. In order to further account for possible heterogeneity, we include specifications that estimate the effects of the letter label on choice. In specifications (3) and (4) we account for the letter label by including a Letter label dummy. In these specifications, the non-stochastic component to the utility for line j is: $V_j = \beta_i * Length_j + LetterLabelDummy_j$ for $i \in \{HL, LL\}$. This analysis is summarized in Table $A18$.

Table $A18$: Comparisons of different multinomial discrete choice models

		Not including letter label		Including letter label		
		Cond Logit	Multi Probit	Cond Logit	Multi Probit	Trials
		(1)	(2)	(3)	(4)	
2 Lines	β_{HL} est.	0.1335	0.1031	0.1337	0.1031	1785
	β_{LL} est.	0.1280	0.0937	0.1277	0.0934	
	AIC	1419	1433	1422	1436	
	BIC	1430	1444	1444	1458	
3 Lines	β_{HL} est.	0.1338	0.0932	0.1367	0.0947	1871
	β_{LL} est.	0.1231	0.0810	0.1244	0.0819	
	AIC	2089	2139	2068	2119	
	BIC	2100	2150	2096	2147	
4 Lines	β_{HL} est.	0.1100	0.0979	0.1128	0.0978	1826
	β_{LL} est.	0.1220	0.1181	0.1239	0.1204	
	AIC	2718	2944	2643	2893	
	BIC	2729	2955	2676	2926	
5 Lines	β_{HL} est.	0.0970	0.1338	0.0992	0.1338	1780
	β_{LL} est.	0.1309	0.1209	0.1351	0.1209	
	AIC	3166	3645	3086	3655	
	BIC	3177	3656	3124	3693	
6 Lines	β_{HL} est.	0.0889	0.0589	0.0952	0.0624	1780
	β_{LL} est.	0.1003	0.0661	0.1070	0.0699	
	AIC	3773	3806	3461	3505	
	BIC	3784	3817	3505	3548	

We provide the estimates of β, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the various models restricted to treatments with identical numbers of lines. Each of the estimates for β_{HL} and β_{LL} are significantly different from 0 with $p < 0.001$.

Similar to the results found above, in every comparison of either AIC and BIC, the value in the Conditional Logit model (1) is lower than that for the analogous Multinomial Probit model (2). Also for both measures, the Conditional Logit model (3) values are lower than those in the analogous Multinomial Probit model (4). We again interpret these results as suggesting that the models that assume that errors have a Gumbel distribution provide a better fit than comparable models that assume that errors have a normal distribution.