An economist and a psychologist form a line: What can imperfect perception of length tell us about stochastic choice?

Duffy, Sean and Smith, John

Rutgers University-Camden

2 April 2020

An economist and a psychologist form a line: What can imperfect perception of length tell us about stochastic choice?*

Sean Duffy† John Smith‡

April 2, 2020

Abstract

Standard choice experiments are hampered by the fact that utility is either unknown or imperfectly measured by experimenters. As a consequence, the inferences available to researchers are limited. By contrast, we design a choice experiment where the objects are valued according to only a single attribute with a continuous measure and we can observe the true preferences of subjects. Subjects have an imperfect perception of the choice objects but can improve the precision of their 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 increases with the length of their choice. Our design allows us to observe the search history, the response times, and make unambiguous conclusions about the optimality of choices. We find a negative relationship between the demanding nature of the choice problems and the likelihood that subjects select the optimal lines. We also find a positive relationship between the demanding nature of the choice problems and the response times. However, we find evidence that suboptimal choices are associated with longer response times than are optimal choices. This result appears to be consistent with Fudenberg, Strack, and Strzalecki (2018). Additionally, our experimental design permits a multinomial discrete choice analysis. Our results suggest that the errors in our data are better described as having a Gumbel distribution rather than a normal distribution. We also observe effects consistent with memory decay and attention. Finally, we find evidence that choices in our experiment exhibit the independence from irrelevant alternatives (IIA) property.

Keywords: judgment, memory, response times, independence from irrelevant alternatives
JEL: C91, D91

*We thank Roberto Barbera, Rahul Bhui, I-Ming Chiu, Carlos Cueva, Jimena González, Hubert János Kiss, Tomás Lejarraga, Marco Mantovani, Andreu Mas-Colell, Richard McLean, Nick Netzer, Ismael Rodriguez-Lara, Tomas Sjöström, Barry Sopher, and Michael Woodford for helpful comments. We thank Jackie Dunn, Steven Gussman, and Lisa Saal for help with the experiment. John Smith thanks Biblioteca de Catalunya. This research was supported by a Rutgers Library OAT award and Rutgers University Research Council Grant #FPRC0000016.

†Rutgers University-Camden, Department of Psychology

‡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; Fax: +1-856-225-6602

1 Introduction

Standard choice experiments are hampered by the fact that utility is either unknown or imperfectly measured by experimenters. As a consequence, the inferences available to researchers are limited. In contrast, we design a choice experiment where objects are valued according to only a single attribute with a continuous measure and we can observe the exact value accruing to the subjects. However, subjects do not always select the optimal choice because they have an imperfect perception of the choice objects.

In our experiment, subjects are given choice sets consisting of lines of various lengths. The subjects are directed to select one of the lines from the choice set. Subjects are paid an amount that increases with the length of their selected line.

While we are able to observe the true objective length of each line, it is well-known that subjects have imperfect perception the objective features of objects (Weber, 1834; Fechner, 1860; Thurstone, 1927a,b). This insight has led researchers to consider that subjects' preferences might be imperfectly perceived by the subjects themselves and this serves as a justification for random choice and random utility models. Since the beginning of this literature, authors have been making explicit references to Weber, Fechner, or Thurstone.[1] Despite this known connection between imperfect perception of objective properties and stochastic choice, to our knowledge, ours is one of only a few experiments that employ the technique of using an objectively measurable object as a proxy for utility.

In our experiment, subjects can only view one line at a time.[2] There are otherwise no restrictions on the nature of their search, as long as the choice occurred within 60 seconds. Subjects have an imperfect perception of the choice objects but can improve the precision of their perception with cognitive effort. Our design allows us to unobtrusively observe the search

[1] See Bradley and Terry (1952), Luce (1959a, 1959b, 1994, 2005), Becker, DeGroot, and Marschak (1963), McFadden (1974, 1976, 1981, 2001), Yellott (1977), Falmagne (1978), Mas-Colell, Whinston, and Green (1995), Ballinger and Wilcox (1997), Loomis, Peterson, Champ, Brown, and Lucero (1998), Butler (2000), Butler and Loomes (2007), Blavatskyy (2008, 2011), Caplin (2012), Lévy-Garboua, Maafi, Masclet, and Terracol (2012), Fudenberg, Iijima, and Strzalecki (2015), Agranov and Ortoleva (2017), Argenziano and Gilboa (2017), Khaw, Li, and Woodford (2017), Alós-Ferrer, Fehr, and Netzer (2018), Caplin, Csaba, Leahy, and Nov (2018), Navarro-Martinez, Loomes, Isoni, Butler, and Alaoui (2018), Cerreia-Vioglio, Dillenberger, Ortoleva, and Riella (2019), Horan, Manzini, and Mariotti (2019), Olschewski, Newell, and Scheibehenne (2019), and Webb (2019).

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

history and the response times of the subjects. To our knowledge, ours is the first example of an experiment where utility is represented by a static, single-attribute physical quantity with an uncountable measure in a (nearly) unrestricted choice setting.

It seems that choice problems with a larger number of lines in the choice set, a choice set of lines of similar lengths, and a choice set involving longer lines are more *demanding* for the subjects. We find a negative relationship between these demanding choice problems and the likelihood that subjects select the optimal lines. We also find a positive relationship between the demanding nature of the choice problems and the response times. However, we find evidence that suboptimal choices are associated with longer response times than are optimal choices.[3] This somewhat counterintuitive result emerges from a model where an agent faces a choice with uncertain utility and there is a constant cost of gathering information about the choice problem. In this setting, which seems to correspond to our experiment, Fudenberg, Strack, and Strzalecki (2018) show that suboptimal decisions will tend to take, on average, a longer time. Additionally, our experimental design permits a multinomial discrete choice analysis. Our results suggest that the errors are better described as having a Gumbel distribution rather than a normal distribution. We also observe effects that are consistent with (possibly endogenous) memory decay and attention. Finally, we find evidence that choices in our dataset are consistent with the independence from irrelevant alternatives (IIA) property.

2 Related literature

2.1 Random utility and random choice

Numerous random utility or random choice experimental and theoretical papers have emerged in an effort to better understand choice.[4] Some authors examine the role of consideration sets

[3]We also note that this result does not appear to be driven by endogeneity.

[4]A partial list of these efforts, not previously mentioned, would include Debreu (1958), 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), Weibull, Mattsson, and Voorneveld (2007), 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), Matějka and McKay (2015), Aguiar, Boccardi, and Dean (2016), Apesteguia, Ballester, and Lu (2017), Dean and Neligh (2017), Ahumada and Ulku (2018), Apesteguia and Ballester (2018), Echenique,

(Masatlioglu, Nakajima, and Ozbay, 2012; Manzini and Mariotti, 2014), private information (Lu, 2016), the preference for randomness (Agranov and Ortoleva, 2017; Cerreia-Vioglio, Dillenberger, Ortoleva, and Riella, 2019), and the preference for flexibility (Ahn and Sarver, 2013) in explaining the apparent randomness in choice data. These factors help us understand choice, but it is our view that imperfect perception of one's preferences is fundamental in every choice setting. In our experiment, there is no plausible preference for randomization, there is no preference for flexibility, there is no private information, there are not multiple attributes that could possibly interact (for instance, as compliments or substitutes), and we can observe the consideration set. Despite the simple and objective nature of our setting, we observe choice that is apparently random.

Further, the experimental study of imperfect perception of one's preferences tends to be hampered by the fact that utility is typically unknown or imperfectly measured. However, because we know the objective value of the imperfectly perceived choice objects, our design yields a unique dataset with which to study random choice.

2.2 Choice involving imperfectly perceived objects

We are not the first authors to study choice in a setting where material outcomes depend on imperfectly perceived objects with objectively measurable properties. For instance, researchers have made payments to subjects as a function of judgments involving the relative quantity of dots (Caplin and Dean, 2015; Dutilh and Rieskamp, 2016), the dominant direction of moving dots (Bhui, 2019a; 2019b), the number of flickering dots (Oud et al., 2016), a dynamic display of dots (Zeigenfuse, Pleskac, and Liu, 2014), the heights of bars of dynamic size (Tsetsos et al., 2016), and the area occupied by objects of various sizes (Polanía, Krajbich, Grueschow, and Ruff, 2014).[5]

To our knowledge, Duffy, Gussman, and Smith (2019) is the only other paper that describes a choice experiment where suboptimal choices are perfectly observable because utility

Saito, and Tserenjigmid (2018), Koida (2018), Kovach and Tserenjigmid (2018), Caplin, Dean, and Leahy (2019), Cattaneo, Ma, Masatlioglu, and Suleymanov (2019), Conte and Hey (2019), and Natenzon (2019).

[5] Gabaix et al. (2006) and Sanjurjo (2015, 2017) also describe experiments where the experimenter knows the choice object with the highest objective value. However, due to computational limitations (rather than imperfect perception in our setting) the subjects often do not make the optimal selection.

is represented by a static, single-attribute physical quantity with an uncountable measure. However, Duffy, Gussman, and Smith (2019) was conducted under time restrictions and the subjects were placed under a differential cognitive load. Therefore, our paper appears to be the only example of an experiment where utility is represented by a static, single-attribute physical quantity with an uncountable measure in a (nearly) unrestricted choice setting.

2.3 Response times and choice

Observations of response times have been used to gain insights on choice, beyond those available by simply observing the outcome of the decision.[6] One key insight from this literature is that longer response times tend to be associated with choices or judgments that are closer to indifference.[7] We observe similar results in our setting: choice sets with lines of more similar lengths are associated with longer response times. We also observe longer response times both for choice sets with a larger number of lines and for choice sets with longer lines.

In many choice problems, it is not clear if a suboptimal decision was made. However, in perceptual decision problems, optimal choices can easily be distinguished from suboptimal choices. Researchers find a negative relationship between accuracy of these perceptual choices and response times.[8] In other words, suboptimal choices tend to take a longer time than do optimal actions.[9,10] Fudenberg, Strack, and Strzalecki (2018), demonstrate that this relationship emerges from a model of an agent in a choice problem with unknown utility and a cost of acquiring information about the choice options. This setup appears to be similar to our

[6] See Clithero (2018) and Spiliopoulos and Ortmann (2018) for recent overviews of the response time literature in economics.

[7] For instance, see Henmon (1911), Volkmann (1934), Dashiell (1937), Mosteller and Nogee (1951), Jamieson and Petrusic (1977), Hey (1995), Moffatt (2005), Chen and Fischbacher (2016), Alós-Ferrer, Granić, Kern, and Wagner (2016), Echenique and Saito (2017), Alós-Ferrer and Garagnani (2019), and Konovalov and Krajbich (2019).

[8] For instance, see Henmon (1911), Kellogg (1931), Swensson (1972), and Bhui (2019b).

[9] Swensson (1972), Luce (1986), and Ratcliff and McKoon (2008) note that this relationship tends to hold when the experimenter emphasizes the accuracy of the decision. The authors also note that the opposite relationship between errors and response times holds when the speed of the decision is stressed by the experimenter.

[10] In general settings, caution should be used when examining the relationship between response times and suboptimal actions. For example, Rubinstein (2013) finds shorter response times on responses that are clearly mistakes. By contrast Kiss, Rodriguez-Lara, Rosa-Garcia (2020) find that in decisions where there is a dominant strategy, suboptimal decisions are associated with longer response times. Achtziger and Alós-Ferrer (2014) find that if the optimal decision is in conflict with reinforcement learning then optimal responses tend to have longer response times, but in the event that they are aligned, then optimal responses tend to have shorter response times.

experiment. Consistent with the predictions of Fudenberg, Strack, and Strzalecki, we find evidence that suboptimal decisions are associated with longer responses times. Further, our result does not appear to be endogenous.

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 21.5 inch (54.6 cm) Dell EliteDisplay E221 monitors. E-Prime imposed a resolution of 1024 pixels by 768 pixels. A total of 112 subjects participated in the experiment.

3.2 Line selection task

In each trial, 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 in the center 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 particular line. To view another line, subjects click on its corresponding label. This would make the new line appear and the old line disappear.

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, both 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 (0.73 cm) increments from 160 pixels (7.4 cm) to 304 pixels (14.1 cm). The lines each had a height of 0.36 cm and were the identical shade of grey. Each of the 10 possible longest line lengths appeared in 10 trials and in random order.

There were three difficulty 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 $\{-1, ..., -10\}$. In the *medium* treatment, one line was exactly 11 pixels shorter than the longest and the other differences were drawn from a uniform on $\{-11, ..., -30\}$. In the *easy* treatment, one line was exactly 31 pixels shorter than the longest, and the other differences were drawn from a uniform on $\{-31, ..., -60\}$. Therefore, the shortest possible line was 100 pixels. The difficult, medium, and easy treatments each occurred with probability $\frac{1}{3}$, in random order, and were drawn with replacement. The subjects were not informed of the existence of these treatments.

Adjacent to each letter label was a box indicating that the subject currently *selected* that line. Subjects could change this selection at any time during the allotted 60 seconds. The subjects could view the time remaining, rounded to the nearest second. See Figure 1 for a screenshot[11] and Figure 2 for a characterization of the regions, which were not visible to the subjects.

[11] See https://osf.io/f7gu4 for the full set of screenshots.

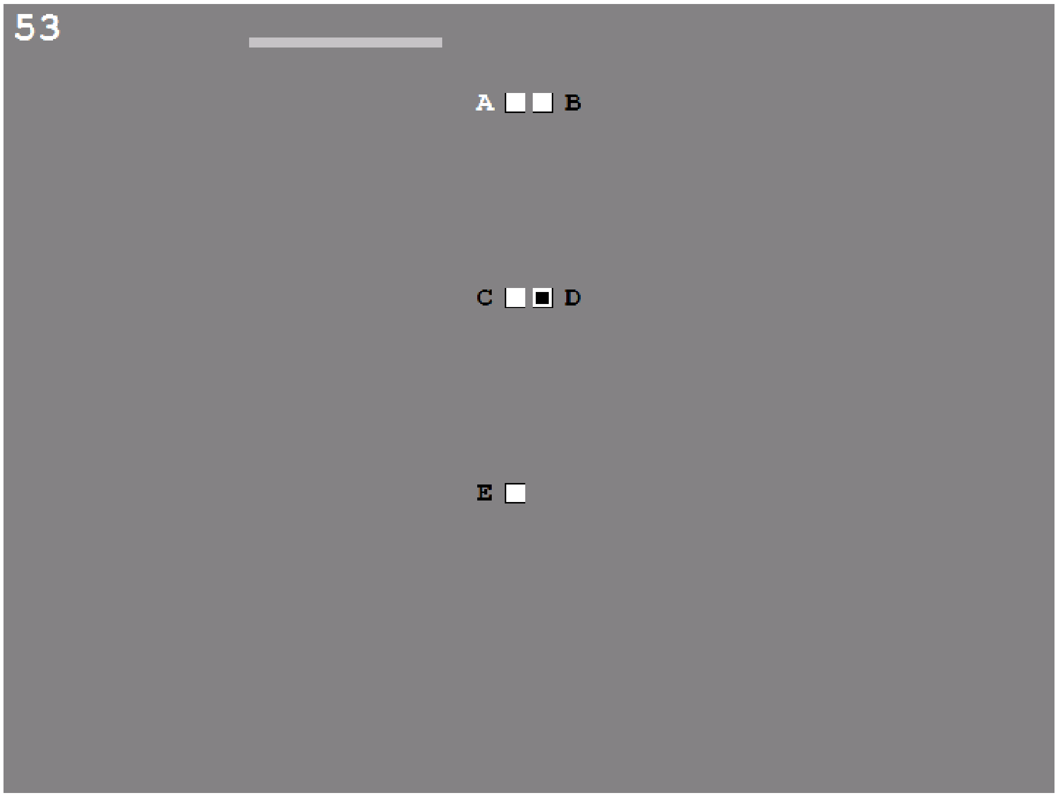

Figure 1: Screenshot from a trial with 5 lines in the choice set, where line A is being viewed, line D is currently selected, and there are 53 seconds remaining.

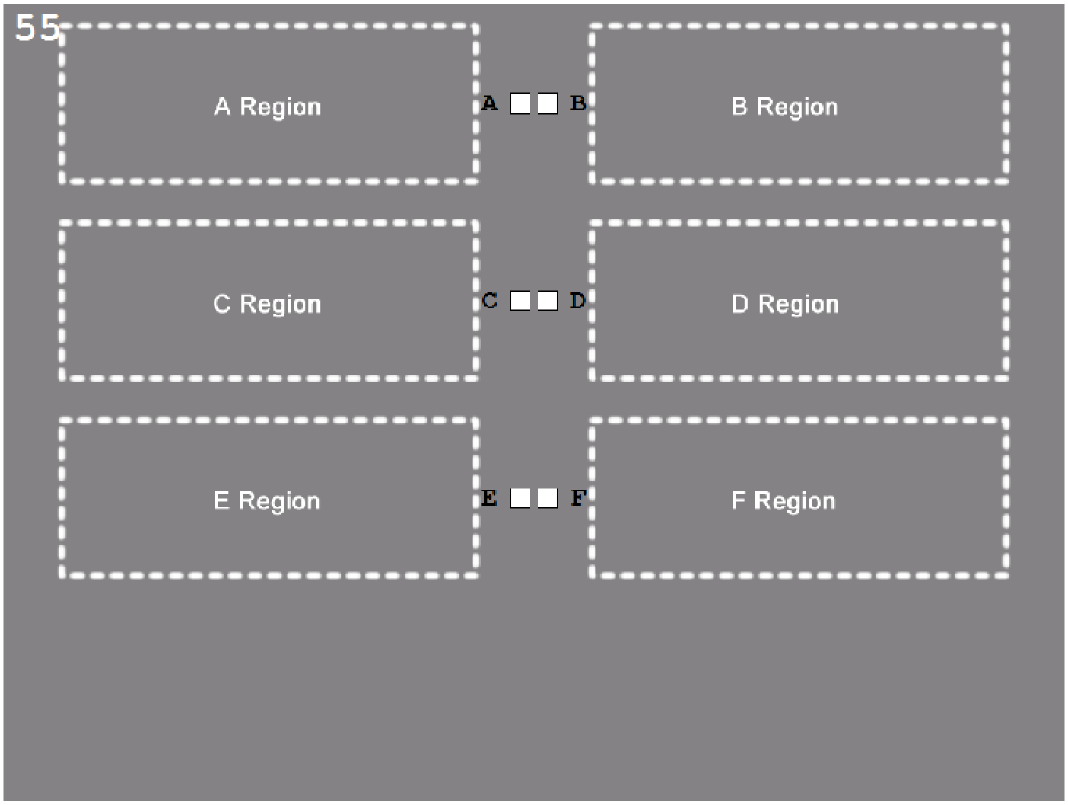

Figure 2: A characterization of the regions, invisible to the subjects, which contain the corresponding lines.

The choice within each trial was the line that was selected when either the subject hit the Enter key or when the allotted 60 seconds elapsed. The earnings on this task were increasing in the length of the choice in that trial. Specifically, if a line x pixels in length was selected then in that trial the subject earned:

$$\$5 * \frac{(x-100)}{(304-100)}.$$

In other words, the payment was $5 multiplied by the fraction of the difference between the selected line and the shortest possible line with the difference between the longest possible line and the shortest possible line. If subjects did not select a line before they hit Enter or before time expired, it was assumed that the selected line had a length of 100, thereby earning 0 in that trial.

There was a 3 second stage between any two line judgment tasks where the subjects were told, "Relax... wait 3 seconds before new trial. DO NOT press RETURN without clicking on a box."

At the start of every trial, the location of the mouse on the screen was randomly determined. Each computer had the mouse speed setting of 6 out of 11.

3.3 Unincentivized practice

The subjects had an unincentivized practice on the line selection task. 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.

3.4 Survey questions

After every line trial was completed, but before the subjects were paid, the subjects were given a set of survey questions, administered via paper. We elicited the gender of the subject, the handedness (right or left) of the subject, and the standard versions of the 3 Cognitive Reflection Test (CRT) questions (Frederick, 2005).[12,13]

3.5 Experimental details

Three line selection trials were randomly selected for payment. Additionally, subjects were 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 $14.50.

There were 112 subjects each completing 100 line selection trials. However, there were 80 trials in which no line was viewed by the subject, 145 trials in which no line was selected by the subject, and 211 trials (1.88%) in which either no line was selected or no line was viewed. We

[12] See https://osf.io/f7gu4 for the version that was given to the subjects.

[13] The subjects were also asked to provide an optional estimate of their grade point average (between 0.0 and 4.0). However, the response rate was sufficiently low (92 of 112 subjects offered some response) and many responses were difficult to interpret. Therefore, we do not include this in our analyses. Additionally, the subjects were asked to provide an optional estimate of their SAT or ACT percentile rank. Again, many responses were incomplete or difficult to interpret, therefore we do not include these responses in the analysis.

exclude these trials from the analysis because we do not know how to interpret these trials. We therefore have 10,989 valid line selection trials. The data is available at https://osf.io/f7gu4.

3.6 Discussion of the design

We do not put any constraints on the nature of the search, beyond the 60 second limit and the restriction that only one line can be viewed at a time. This allows us to measure response times and search histories as unobtrusively as possible.

The design of the interface was motivated by Duffy, Gussman, and Smith (2019), who found a negative relationship between accuracy on a trial and the distance between the line with the longest length and its letter label. However, the options with a larger distance between the line and the label also tended to have a larger time that had elapsed since the line was last viewed at the end of the trial. Therefore, Duffy, Gussman, and Smith were not able to distinguish between the temporal explanation and the distance explanation. However, we improve on this design in that every option has the identical distance between the line and the label.

4 Results

4.1 Optimality of choices

Here we explore the optimality of choices. We define the *Selected longest* variable to be 1 if the choice was the longest line in the choice set, and 0 otherwise. We conduct regressions with the Selected longest variable as dependent variable. As the dependent variable is binary, we employ a logistic specification. Since the Selected longest variable appears 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.[14] 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

[14] See Tables $A1 - A4$ for the summary statistics. for Selected longest.

combinations of letter-number of lines.[15] Due to the repeated nature of the observations, we also offer fixed-effects specifications where we estimate a dummy variable for each subject. We run other specifications that control for the demographics of the subjects: whether the subjects reported being left handed, whether the subjects reported being female, and the sum of their CRT score. We summarize these regressions in Table 1.

Table 1: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Longest line normalized	-0.0044^{***}	-0.0044^{***}	-0.0042^{***}	-0.0042^{***}
	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Number of lines normalized	-0.257^{***}	–	-0.242^{***}	–
	(0.018)		(0.018)	
Easy treatment dummy	2.273^{***}	2.291^{***}	2.205^{***}	2.219^{***}
	(0.098)	(0.099)	(0.097)	(0.097)
Difficult treatment dummy	-1.729^{***}	-1.747^{***}	-1.630^{***}	-1.646^{***}
	(0.055)	(0.056)	(0.053)	(0.053)
Trial	$-0.0017^{\dagger}$	-0.0019^{*}	$-0.0016^{\dagger}$	-0.0017^{*}
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Letter dummies	No	Yes	No	Yes
Fixed effects	Yes	Yes	No	No
Demographics	No	No	Yes	Yes
AIC	9716.7	9655.7	9893.5	9836.1

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 (Akaike, 1974). Each regression has $10,989$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, we find a negative relationship between the quality of the choice and the features of the choice problem that we describe as demanding.[16] We find that the accuracy of the choice decreases when there is a larger number of lines (choice overload effects) and decreases in the difficulty treatments. Additionally, we see that the accuracy decreases in the length of the longest line. This result could be interpreted as suggesting that subjects are

[15] As in Table $A4$. However, in the analysis immediately below we do not explore the effect of the letter label on the quality of the choice. We postpone our discussion of this issue until later in the paper.

[16] In Table $A8$, we examine the robustness of these results. We conduct the analogous tobit regressions with Longest line minus the selected line as the dependent variable. Our results are not changed.

simply worse at judging longer lines than shorter lines. This explanation is consistent with classic psychology (Fechner, 1860). On the other hand, it is possible that subjects expended less effort on trials with longer lines because they knew that they would earn more on these trials. These effort-wealth effects provide another explanation for the negative coefficient estimates for the Longest line variable. We will say more about this matter below. Finally, we find evidence that choices become less accurate across trials.

4.2 Quality of the Searches

The analysis above suggests a relationship between the quality of the choice and features of the choice problem that we describe as demanding. It is not clear whether this relationship is driven by low quality searches in these demanding choice problems. One measure of the quality of the search is the number of lines viewed in the trial, where a higher number would suggest more effort in searching. We define the *View clicks* variable to be the number of total line view clicks in the trial. We conduct an analysis, identical to Table 1, with the exception that the dependent variable is View clicks and the regression is linear, not logistic. Table 2 summarizes this analysis.

Table 2: Regressions of the View clicks variable

	(1)	(2)	(3)	(4)
Longest line normalized	0.002*	0.002*	0.002*	0.002*
	(0.0008)	(0.0008)	(0.0009)	(0.0009)
Number of lines normalized	1.876***	–	1.896***	–
	(0.026)		(0.031)	
Easy treatment dummy	−2.084***	−2.090***	−2.006***	−2.006***
	(0.090)	(0.090)	(0.107)	(0.107)
Difficult treatment dummy	1.338***	1.333***	1.374***	1.376***
	(0.090)	(0.090)	(0.106)	(0.106)
Trial	−0.024***	−0.024***	−0.024***	−0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Letter dummies	*No*	*Yes*	*No*	*Yes*
Fixed effects	*Yes*	*Yes*	*No*	*No*
Demographics	*No*	*No*	*Yes*	*Yes*
AIC	60559.2	60545.9	64415.6	64399.7

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 10,989 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The results summarized in Table 2 seem to suggest that there is a positive relationship between the quality of the search and the demanding nature of the choice problem. The number of view clicks is increasing in the length of the longest line, increasing in the number of lines in the choice set, and increasing in the difficulty treatments. This suggests to us that more, not less, effort was devoted to demanding choice problems. We also note that there is a negative relationship between view clicks and trial, suggesting a decrease in effort across trials.

We admit that the number of lines viewed is one of many measures of the quality of the search. Another such measure is the response time on the trial. We conduct an analysis, identical to that in Table 2, but we employ the log of the response time as the dependent variable.[17] Table 3 summarizes this analysis.[18]

Table 3: Regressions of the log of Response time variable

	(1)	(2)	(3)	(4)
Longest line normalized	0.0003***	0.0003***	0.0003***	0.0003***
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Number of lines normalized	0.083***	–	0.083***	–
	(0.0012)		(0.0014)	
Easy treatment dummy	−0.118***	−0.118***	−0.117***	−0.116***
	(0.004)	(0.004)	(0.005)	(0.005)
Difficult treatment dummy	0.057***	0.057***	0.056***	0.056***
	(0.004)	(0.004)	(0.005)	(0.005)
Trial	−0.002***	−0.002***	−0.002***	−0.002***
	(0.00006)	(0.00006)	(0.00007)	(0.00007)
Letter dummies	*No*	*Yes*	*No*	*Yes*
Fixed effects	*Yes*	*Yes*	*No*	*No*
Demographics	*No*	*No*	*Yes*	*Yes*
AIC	−6089.7	−6066.4	−2591.5	−2555.1

[17] See Tables $A1 - A4$ for the summary statistics.

[18] Although negative AIC values are somewhat unusual, negative values are possible and it remains the case that lower values indicate a better fit.

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 $10,989$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

It seems that response times are larger for decisions that are more demanding: a larger number of lines in the choice set, the difficulty of the treatment, and the length of the longest line in the trial. Based on the results of Tables 2 and 3, it seems that subjects expend more, not less, effort on demanding trials. Therefore we cannot explain the longest line length results of Table 1 as due to diminished effort.

4.3 Relationship between quality of choice and consideration sets

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 $10,956$ trials where subjects viewed the longest line, there are 3302 observations where the longest line was not selected. However, in none of the 33 trials where subjects did not view the longest line, was the longest line selected. Therefore in our data, 99.0% of the suboptimal choices occurred in trials where subjects 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.

In Table 1, we explored whether subjects optimally select the longest line by conducting regressions with the Selected longest line variable. Another question is whether subjects selected the longest line, among the lines that were viewed. We define the *Selected longest line viewed* variable to be 1 if the longest line among those viewed was selected, and 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 4.

Table 4: Logistic regressions of Selected longest line viewed variable

	(1)	(2)	(3)	(4)
Longest line normalized	-0.0042^{***}	-0.0042^{***}	-0.0040^{***}	-0.0040^{***}
	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Number of lines normalized	-0.249^{***}	–	-0.235^{***}	–
	(0.018)		(0.018)	
Easy treatment dummy	2.304^{***}	2.322^{***}	2.240^{***}	2.253^{***}
	(0.100)	(0.100)	(0.099)	(0.099)
Difficult treatment dummy	-1.719^{***}	-1.737^{***}	-1.626^{***}	-1.643^{***}
	(0.055)	(0.056)	(0.053)	(0.053)
Trial	-0.0014	$-0.0016^{\dagger}$	-0.0014	$-0.0015^{\dagger}$
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Letter dummies	No	Yes	No	Yes
Fixed effects	Yes	Yes	No	No
Demographics	No	No	Yes	Yes
AIC	9712.7	9651.9	9865.6	9808.8

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 $10,989$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The results of Table 4 are similar to those in Table 1. Further, the suboptimal choices that we observe in our experiment do not appear to be driven by consideration set effects. Rather, the suboptimal choices appear to be driven by imperfect perception of the objectively measurable choice objects.

4.4 Relationship between quality of choice and response times

Response times of the trials in which the longest line was selected ($mean = 12.250s$, $SD = 8.426$) are smaller than the response times in trials in which the longest line was not selected ($mean = 15.730s$, $SD = 10.428$), according to a Wilcoxon Two-Sample Test ($Z = 18.915$, $p < 0.001$). This effect is further robust across longest line treatments, number of line treatments, and difficulty treatments.[19]

In order to more carefully investigate this matter, we conduct regressions with the log of Response time as the dependent variable. We employ specifications similar to those in

[19]See Tables $A5 - A7$.

Table 3, however we include Selected longest as an independent variable. Further, for those specifications without fixed-effects, we include an independent variable that is the average of the log of the Response times for that particular subject. We summarize this analysis in Table 5.

Table 5: Regressions of the log of Response time variable

	(1)	(2)	(3)	(4)
Longest line normalized	0.0003***	0.0003***	0.0003***	0.0003***
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Number of lines normalized	0.083***	–	0.082***	–
	(0.0012)		(0.0012)	
Easy treatment dummy	−0.115***	−0.115***	−0.114***	−0.114***
	(0.004)	(0.004)	(0.004)	(0.004)
Difficult treatment dummy	0.051***	0.052***	0.051***	0.052***
	(0.004)	(0.004)	(0.004)	(0.004)
Trial	−0.002***	−0.002***	−0.002***	−0.002***
	(0.00006)	(0.00006)	(0.00006)	(0.00006)
Selected Longest	−0.016***	−0.014**	−0.016***	−0.014***
	(0.004)	(0.004)	(0.004)	(0.004)
Sub's Average log RT	–	–	0.994***	0.992***
			(0.014)	(0.014)
Letter dummies	*No*	*Yes*	*No*	*Yes*
Fixed effects	*Yes*	*Yes*	*No*	*No*
Demographics	*No*	*No*	*Yes*	*Yes*
AIC	−6093.4	−6067.4	−6680.2	−6651.1

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 10,989 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

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.

The reader is likely concerned about endogeneity because the Selected longest variable is possibly correlated with the errors in the regressions. However, when we conduct Spearman correlations between the unstandardized residuals and the Selected longest variable in specifications $(1)-(4)$, the p-values, respectively, are 0.94, 0.92, 0.92, and 0.89. When we conduct Pearson correlations, the qualitative results are not changed. Further, our qualitative results are not changed when we use the student residuals rather than the unstandardized residuals.[20] We interpret this as suggesting that our specifications are not suffering from an endogeneity bias caused by the inclusion of the Selected longest variable. We conclude that the relationship between response times and the optimality of the choice is not driven by a possible subject-level relationship between skill at the task and the speed by which it is accomplished.

4.5 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.[21] 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:

$$U_j = V_j + \epsilon_j,$$

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:

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

Further, the non-stochastic components to the RUMs are not typically observable. Therefore researchers include 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, researchers also es-

[20] In Table $A9$, we offer a robustness check of Table 5. There we estimate a treatment variable coefficient for every subject. Our qualitative results are not changed.

[21] For instance, see McFadden (1974, 1976, 1981, 2001).

timate $\overline{\beta} = (\beta_1, ..., \beta_n)$. In these settings, the estimate of 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:

$$V_j = \beta * Length_j,$$

where β is a scalar.

We also note that there is a number 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, $e^{-e^{-\epsilon}}$. (Confusingly, this is also referred to as the Type I extreme-value distribution, the double exponential distribution, 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.

We run one specification where the stochastic component has the Gumbel distribution and is independently and 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:

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

We refer to this *Conditional Logistic* model as "Logit" and denote it 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 "Probit" and denote it as specification (2).

It is not clear that the linear specification of Length is the most appropriate. There are researchers who argue that there is not a linear relationship between stimuli and the perception of stimuli, but rather they are related by the log function (Fechner, 1860).[22] According to this classic view in psychology, the non-stochastic component of the utility function should be specified as:

$$V_j = \beta * \log(Length_j).$$

With this "Log" specification, we conduct an analysis that assumes that the errors have a Gumbel distribution. We denote this specification as (3). Also within this Log specification, we conduct an analysis that assumes that errors have a normal distribution. We denote this specification as (4).

There are other researchers who argue that the relationship between stimuli and the perception of the stimuli is neither linear nor logarithmic, but is rather described by the power function (Stevens, 1961). Research suggests that the specification of this power function depends on the type of stimulus. Some research suggests that the exponent in the power function, when the stimulus is length, is 1.04 (Teghtsoonian, 1971). According to this view, the non-stochastic component of the utility function should be specified as:

$$V_j = \beta * (Length_j)^{1.04}.$$

With this "Power" specification, we conduct an analysis that assumes that the errors have a Gumbel distribution. We denote this specification as (5). Also within this Power specification, we conduct an analysis that assumes that errors have a normal distribution. We denote this specification as (6).

We report the Akaike Information Criterion (AIC, Akaike, 1974) for each specification, restricted to a particular number of lines treatment. We also report the estimate of β for each model in each setting. These analyses[23] are summarized in Table 6.

[22] This is a version of what is sometimes referred to as Fechner's Law.

[23] Each specification was executed with the MDC (multinomial discrete choice) procedure in SAS. Specifications (1), (3), and (5) were performed with the *clogit* option. Specification (2), (4), and (6) were performed with the *mprobit* option.

Table 6: Comparisons of different multinomial discrete choice models

		Linear		Log		Power	
		Logit	Probit	Logit	Probit	Logit	Probit
		(1)	(2)	(3)	(4)	(5)	(6)
2 Lines	β est.	0.126	0.097	64.193	49.104	0.097	0.075
	AIC	1917	1928	1899	1913	1919	1930
3 Lines	β est.	0.139	0.104	69.951	51.525	0.108	0.080
	AIC	2602	2607	2615	2629	2603	2608
4 Lines	β est.	0.144	0.163	72.918	57.223	0.112	0.112
	AIC	3218	3513	3197	3228	3220	3384
5 Lines	β est.	0.135	0.194	66.324	45.846	0.105	0.074
	AIC	4046	5166	4066	4113	4047	4067
6 Lines	β est.	0.131	0.089	64.954	43.736	0.101	0.069
	AIC	4394	4407	4386	4402	4397	4410

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

Regardless of the specification (Linear, Log, or Power) and regardless of the number of lines treatment, we find that the Logit specification has a lower AIC than the corresponding Probit specification. We interpret these results as suggesting that the models, which assume that the errors have a Gumbel distribution, provide a better fit than the models that assume that errors have a normal distribution.

4.6 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 that is consistent with memory decay.

There appears to be 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.[24] Below, we test whether there is such a relationship. Similar to the analysis summarized in Table 1, the dependent variable is Selected longest and we include the same specifications for the treatment variables and the Trial variable.

[24]See Table $A4$.

We introduce the *Distance from last* variable, which provides a measure of the alphabetic distance between the letter label of the longest line and the last alphabetic 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 0. In the 3 Line treatment, if A is the longest then the variable is 2, if B is the longest then 1, and if C is the longest then 0. We include Distance from last as an independent variable. We also include specifications with the interaction between the Trial variable and the Distance from last variable. For identification reasons, we do not include the Letter dummy variables. We also include specifications with an independent variable of the average of the Selected longest variable for that particular subject. We include these with the demographics and we denote their inclusion with *Demographics+*. We summarize these regressions in Table 7.

Table 7: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Distance from last	-0.151^{***}	-0.149^{***}	-0.138^{***}	-0.137^{***}
	(0.021)	(0.021)	(0.039)	(0.038)
Longest line normalized	-0.005^{***}	-0.004^{***}	-0.005^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Number of lines normalized	-0.183^{***}	-0.178^{***}	-0.183^{***}	-0.178^{***}
	(0.021)	(0.021)	(0.021)	(0.021)
Easy treatment dummy	2.282^{***}	2.258^{***}	2.282^{***}	2.258^{***}
	(0.098)	(0.098)	(0.098)	(0.098)
Difficult treatment dummy	-1.736^{***}	-1.702^{***}	-1.736^{***}	-1.702^{***}
	(0.056)	(0.055)	(0.056)	(0.055)
Trial	-0.0018^{*}	$-0.0017^{\dagger}$	-0.0014	-0.0014
	(0.0009)	(0.0009)	(0.0013)	(0.0013)
Trial*Distance from last	–	–	-0.0003	-0.0002
			(0.0006)	(0.0006)
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*Yes*	*No*	*Yes*	*No*
Demographics+	*No*	*Yes*	*No*	*Yes*
AIC	9668.3	9498.6	9670.1	9500.5

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the subject-specific dummies in the fixed effects regressions or the demographics estimates. AIC refers to the Akaike information criterion. Each regression has 10,989 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In each specification, the Distance from last coefficient estimate is significant and negative. This suggests the presence of memory decay in the subjects.

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.). However, lines viewed in the more distant past are recalled with a lower precision: either the location of the longest line or the length of the longest line. To explore this possibility, we define the variable *Time since longest* to be the time elapsed since subjects viewed the longest line when the trial ended. Table 8 demonstrates the relationship between the Time since longest variable and the letter label of the longest line in that trial.

Table 8: Time since longest line by number of lines and letter label of the longest

	A	B	C	D	E	F
2 Lines	0.828 s	0.485 s	–	–	–	–
3 Lines	1.393 s	1.064 s	0.954 s	–	–	–
4 Lines	2.096 s	2.026 s	1.400 s	1.239 s	–	–
5 Lines	3.012 s	2.629 s	2.403 s	2.214 s	1.662 s	–
6 Lines	3.991 s	3.480 s	2.978 s	2.338 s	2.819 s	2.201 s

Table 8 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 such a relationship is related to the optimality of choice in the trial. To do so, we conduct an analysis similar to Table 7, however we employ the Time since longest variable, rather than the Distance from last variable. We summarize these regressions in Table 9.

Table 9: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Time since longest	-0.844^{***}	-0.794^{***}	-0.666^{***}	-0.623^{***}
	(0.022)	(0.021)	(0.037)	(0.035)
Longest line normalized	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Number of lines normalized	0.077^{**}	0.063^{**}	0.082^{***}	0.069^{**}
	(0.023)	(0.023)	(0.024)	(0.023)
Easy treatment dummy	2.442^{***}	2.432^{***}	2.457^{***}	2.451^{***}
	(0.124)	(0.122)	(0.125)	(0.122)
Difficult treatment dummy	-1.677^{***}	-1.650^{***}	-1.681^{***}	-1.655^{***}
	(0.069)	(0.067)	(0.069)	(0.067)
Trial	-0.005^{***}	-0.005^{***}	0.0006	0.0006
	(0.001)	(0.001)	(0.0015)	(0.0014)
Trial*Time since longest	–	–	-0.004^{***}	-0.004^{***}
			(0.0007)	(0.0007)
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*Yes*	*No*	*Yes*	*No*
Demographics+	*No*	*Yes*	*No*	*Yes*
AIC	6816.0	6741.4	6785.2	6712.0

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the subject-specific dummies in the fixed effects regressions or the demographics estimates. AIC refers to the Akaike information criterion. Each regression has $10,989$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

We find 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. This result is consistent with the possibility that subjects suffer from memory decay. We also note the negative Trial-Time since longest interaction. This suggests that the effects consistent with memory decay are exacerbated across trials. While the other coefficient estimates resemble those from Table 1, it should be noted that the Number of lines estimate is positive. This suggests that, when accounting for the Time since longest variable, there is a positive relationship between the number of lines and selecting the longest line.

However, the results summarized in Table 9 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 our regressions could suffer from an endogeneity bias.

Overall the results from Tables 7 and 9 suggest effects that are consistent with memory decay. It appears that lines viewed in the more distant past are remembered with a lower precision. Duffy, Gussman, and Smith (2019) also found evidence of memory decay, yet their interface had a feature that not every item in the choice set had an equal distance between the line and its letter label. It is possible that the Duffy, Gussman, and Smith design exacerbated a differential memory decay. However, in our interface, every line is equidistant from its letter label and we find differential effect consistent with (possibly endogenous) memory decay.

4.7 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.[25] One measure of attention is the total time spent viewing a line. We perform an analysis, similar to Tables 7 and 9, but with *Time viewing longest* as an independent variable. This variable sums the time, possibly across multiple durations, which the subject viewed the longest line in the choice set. We summarize these regressions in Table 10.

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

Table 10: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Time viewing longest	0.164***	0.134***	0.126***	0.103***
	(0.011)	(0.010)	(0.017)	(0.016)
Longest line normalized	−0.005***	−0.005***	−0.005***	−0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of lines normalized	−0.260***	−0.255***	−0.259***	−0.254***
	(0.019)	(0.019)	(0.019)	(0.019)
Easy treatment dummy	2.412***	2.365***	2.412***	2.363***
	(0.099)	(0.098)	(0.099)	(0.098)
Difficult treatment dummy	−1.778***	−1.739***	−1.781***	−1.742***
	(0.057)	(0.056)	(0.057)	(0.056)
Trial	0.002*	0.0015	−0.0011	−0.0012
	(0.0009)	(0.0009)	(0.0015)	(0.0015)
Trial*Time viewing longest	−	−	0.0009**	0.0008*
			(0.0003)	(0.0003)
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*Yes*	*No*	*Yes*	*No*
Demographics+	*No*	*Yes*	*No*	*Yes*
AIC	9459.3	9329.5	9453.6	9325.9

We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the subject-specific dummies in the fixed effects regressions or the demographics estimates. AIC refers to the Akaike information criterion. Each regression has 10,989 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. A similar result is reported by Krajbich and Rangel (2011). We interpret this as consistent with attention being related to choice. We also find that the interaction between Trial and Time viewing longest is positive. This provides some evidence that the attention effects become stronger across trials.

Similar to the analysis summarized in Table 9, we interpret these results with caution due to the possibility of endogeneity introduced by including the Time viewing longest variable. Similar to the analysis from Table 9, the Spearman correlations between the Time viewing longest variable and the residuals (both unstandardized and Pearson standardized) are each significant at 0.001. These results suggest that (possibly endogenous) attention is related to choice.

4.8 Independence from irrelevant alternatives

Our dataset provides a unique opportunity to test the independence from irrelevant alternatives (IIA) property. Recall that our choice sets always involve a *longest* line and another line that is a specific amount shorter than this longest line. The difference between these lines is 1 pixel in the difficult treatment, 11 pixels in the medium treatment, and 31 pixels in the easy treatment. Choice sets with more than 2 lines are constructed by including lines that have lengths less than or equal to the shorter of these lines. We now refer to lines in the difficult treatment 1 pixel shorter than the longest, lines in the medium treatment 11 pixels shorter than the longest, or lines in the easy treatment 31 pixels shorter than the longest to be the *second longest* line in the choice set.

To test IIA in our setting, we can observe if the ratio of the probability that the longest line is selected and the probability that the second longest line is selected varies with the size of the choice set.[26] Because not every trial has a unique second longest line, we only consider observations with a unique second longest line. Further, since we are interested in the relative selection of the longest and the second longest line, we consider only trials in which either the longest or the second longest lines were selected. This allows us to interpret the Selected longest variable as the relative choice between the longest and the second longest lines. Due to the non-monotonic relationship between the size of the choice set and the choices, we estimate a unique dummy variable for each choice set size. In order to investigate the effect of the size of the choice set on choice, we report the Wald statistic for the Number of lines dummy variables. We also report the associated p-value.

The other elements of the analysis are similar to that presented above. We summarize these regressions in Table 11.

[26]See Table $A10$ for the summary statistics involving the longest line and the second longest line choices.

Table 11: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Longest line normalized	–	-0.004^{***}	-0.004^{***}	-0.004^{***}
		(0.0007)	(0.0007)	(0.0007)
Easy treatment dummy	2.208^{***}	2.224^{***}	2.224^{***}	2.259^{***}
	(0.125)	(0.125)	(0.125)	(0.126)
Difficult treatment dummy	-1.388^{***}	-1.386^{***}	-1.387^{***}	-1.437^{***}
	(0.065)	(0.065)	(0.065)	(0.067)
Trial	–	–	-0.001	-0.001
			(0.001)	(0.001)
Number of lines Wald statistic	4.77	4.32	4.31	4.84
p-value of Wald statistic	0.31	0.36	0.37	0.30
Letter dummies	*No*	*No*	*No*	*No*
Fixed effects	*No*	*No*	*No*	*No*
Demographics+	*No*	*No*	*No*	*Yes*
AIC	6634.3	6606.9	6607.5	6429.7

We only consider trials with a unique second longest line and those in which either the longest line or the second longest line was selected. We provide the coefficient estimates and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Number of lines dummy estimates, or the demographics estimates. We report the Wald statistic and its corresponding p-value for the Number of lines dummy estimates. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has $8,628$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In every specification, the size of the choice set is not significantly related to the relative choice of the longest and the second longest lines. We interpret this as consistent with the IIA property. The other results in Table 11, are largely consistent with the analyses summarized above. The relative choice of the longest and second longest lines are decreasing in the difficulty treatments and decreasing in the length of the longest line. However, we note that the Trial coefficient estimate is not significant.

The results in Table 11 seem to suggest that the well-documented violations of IIA stem from details in a specific choice setting (for instance, a particular profile of multiple attributes) rather than being a general feature across all choice settings.

5 Conclusion

We conduct an idealized choice experiment where the choice set is populated with objects that are valued according to only a single, static attribute with a continuous measure and we can observe the true preferences of subjects. Subjects have an imperfect perception of the choice objects but can improve the precision of their perception with cognitive effort. Subjects are given a choice set involving several lines of various lengths and are told to select one of them. Subjects are paid an amount that increases with the length of their choice and they therefore strive to select the longest line. This design allows us to make unambiguous conclusions about the optimality of choices. We find a negative relationship between the optimality of choice and the number of lines in the choice set, the lengths of the lines in the choice set, and the similarity of the lengths of the lines in the choice set. We note that this apparent random choice emerges from a setting without a preference for randomization, without a preference for flexibility, without private information, without multiple attributes that could possibly interact, and is not the result of consideration set effects.

Our design allows us to observe the search history and the response times. We also find a positive relationship between response times and the number of lines in the choice set, the lengths of the lines in the choice set, and the similarity of the lengths of the lines in the choice set. However, we find evidence that suboptimal choices are associated with longer response times than are optimal choices. This result is consistent with the predictions of Fudenberg, Strack, and Strzalecki (2018), who study a model where an agent faces a choice with uncertain utility and there is a constant cost of gathering information about the choice problem. The authors show that suboptimal decisions will tend to take, on average, a longer time. We note that their theoretical model seems to closely resemble our experimental setup. We also note that our statistical results do not appear to suffer from an endogeneity bias.

Since we know the objective value of each object in the choice set, our experiment produces choice data that can investigate the statistical distribution of the errors. When we conduct a multinomial discrete choice analysis, we find that the errors are better described as having a Gumbel distribution rather than a normal distribution. We also observe effects that are consistent with (possibly endogenous) memory decay and attention.

Finally, we find evidence that our dataset is consistent with the independence from irrelevant alternatives (IIA) property. Given the general nature of our choice setting, we interpret this as suggesting that IIA could be a general feature of choice, and that violations of IIA only occur in specific choice settings, such as those with certain profiles of multiple attributes.

We admit that our results related to memory decay and attention possibly suffer from an endogeneity bias. In the future we are interested in studying memory decay effects by manipulating the time between the last line viewed and the response. We are also interested in studying attention effects, by manipulating the time that the subject is permitted to view the lines. These exogenous interventions should facilitate the investigation of memory decay and attention.

We are also interested to learn whether the effects that we find, also appear in settings with other objectively measurable quantities. For instance, paying subjects as a function of tones, color shade, temporal duration, or weight. We are further interested to learn the affect of various other payment schemes. We leave these questions for future research.

References

Achtziger, Anja and Alós-Ferrer, Carlos (2014): "Fast or rational? A response-times study of Bayesian updating," *Management Science*, 60(4), 923–938.

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

Aguiar, Victor H., Boccardi, Maria Jose, and Dean, Mark (2016): "Satisficing and stochastic choice," *Journal of Economic Theory*, 166, 445–482.

Ahn, David S. and Sarver, Todd (2013): "Preference for flexibility and random choice," *Econometrica*, 81(1), 341–361.

Ahumada, Alonso and Ulku, Levent (2018): "Luce rule with limited consideration," *Mathematical Social Sciences*, 93, 52–56.

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

Alós-Ferrer, Carlos, Fehr, Ernst, and Netzer, Nick (2018): "Time will tell: recovering preferences when choices are noisy," Working paper, University of Zurich.

Alós-Ferrer, Carlos and Garagnani, Michele (2019): "Strength of preference and decision making under risk," Working paper, University of Zurich.

Alós-Ferrer, Carlos, Granić, Đura-Georg, Kern, Johannes, and Wagner, Alexander K. (2016): "Preference reversals: Time and again," *Journal of Risk and Uncertainty*, 52(1), 65–97.

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

Apesteguia, Jose, Ballester, Miguel A., and Lu, Jay (2017): "Single-Crossing Random Utility Models," *Econometrica*, 85(2), 661–674.

Argenziano, Rossella and Gilboa, Itzhak (2017): "Psychophysical foundations of the Cobb–Douglas utility function," *Economics Letters*, 157, 21–23.

Armel, K. Carrie, Beaumel, Aurelie, and Rangel, Antonio (2008): "Biasing simple choices by manipulating relative visual attention," *Judgment and Decision Making*, 3(5), 396–403.

Armel, K. Carrie and Rangel, Antonio (2008): "The Impact of Computation Time and Experience on Decision Values," *American Economic Review*, 98(2), 163–168.

Ballinger, T. Parker, and Wilcox, Nathaniel T. (1997): "Decisions, Error and Heterogeneity," *Economic Journal*, 107(443), 1090–1105.

Becker, Gordon M., DeGroot, Morris H., and Marschak, Jacob (1963): "Stochastic models of choice behavior," *Systems Research and Behavioral Science*, 8(1), 41–55.

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

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

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

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

Bradley, Ralph Allan and Terry, Milton E. (1952): "Rank analysis of incomplete block designs: I. The method of paired comparisons," *Biometrika*, 39(3-4), 324–345.

Butler, David 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, David and Loomes, Graham (1988): "Decision difficulty and imprecise preferences," *Acta Psychologica*, 68(1-3), 183–196.

Butler, David J. and Loomes, Graham C. (2007): "Imprecision as an account of the preference reversal phenomenon," *American Economic Review*, 97(1), 277–297.

Caplin, Andrew (2012): "Choice Sets as Percepts," In *Neuroscience of Preference and Choice: Cognitive and Neural Mechanisms*, Dolan, Raymond and Sharot, Tali (Eds.), Academic Press, Waltham MA, 295–304.

Caplin, Andrew, Csaba, Daniel, Leahy, John, and Nov, Oded (2018): "Rational Inattention, Competitive Supply, and Psychometrics," Working paper, New York University.

Caplin, Andrew and Dean, Mark (2015): "Revealed preference, rational inattention, and costly information acquisition," *American Economic Review*, 105(7), 2183–2203.

Caplin, Andrew, Dean, Mark, and Leahy, John (2019): "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," Working Paper, Columbia University.

Caplin, Andrew, Dean, Mark, and Martin, Daniel (2011): "Search and satisficing," *American Economic Review*, 101(7), 2899–2922.

Caplin, Andrew and Martin, Daniel (2015): "A testable theory of imperfect perception," *Economic Journal*, 125(582), 184–202.

Cattaneo, Matias D., Ma, Xinwei, Masatlioglu, Yusufcan, and Suleymanov, Elchin (2019): "A Random Attention Model," *Journal of Political Economy*, forthcoming.

Cerreia-Vioglio, Simone, Dillenberger, David, Ortoleva, Pietro, and Riella, Gil (2019): "Deliberately stochastic," *American Economic Review*, 109(7), 2425–2445.

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

Clithero, John A. (2018): "Response times in economics: Looking through the lens of sequential sampling models," *Journal of Economic Psychology*, 69, 61–86.

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

Conte, Anna, Hey, John D., and Moffatt, Peter G. (2011): "Mixture models of choice under risk," *Journal of Econometrics*, 162(1), 79–88.

Cubitt, Robin P., Navarro-Martinez, Daniel, and Starmer, Chris (2015): "On preference imprecision," *Journal of Risk and Uncertainty*, 50(1), 1–34.

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

Dean, Mark and Neligh, Nathaniel (2017): "Experimental Tests of Rational Inattention," Working Paper, Columbia University.

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

Duffy, Sean, Gussman, Steven, and Smith, John (2019): "Judgments of length in the economics laboratory: Are there brains in choice?" Working paper, Rutgers University-Camden.

Dutilh, Gilles and Rieskamp, Jörg (2016): "Comparing perceptual and preferential decision making," *Psychonomic Bulletin and Review*, 23(3), 723–737.

Echenique, Federico, Saito, Kota, and Tserenjigmid, Gerelt (2018): "The perception-adjusted Luce model," *Mathematical Social Sciences*, 93, 67–76.

Echenique, Federico and Saito, Kota (2017): "Response time and utility," *Journal of Economic Behavior and Organization*, 139, 49–59.

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

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

Frederick, Shane (2005): "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 19(4), 25–42.

Fudenberg, Drew, Iijima, Ryota, and Strzalecki, Tomasz (2015): "Stochastic choice and revealed perturbed utility," *Econometrica*, 83(6), 2371–2409.

Fudenberg, Drew, Strack, Philipp, and Strzalecki, Tomasz (2018): "Speed, accuracy, and the optimal timing of choices," *American Economic Review*, 108(12), 3651–3684.

Gabaix, Xavier, Laibson, David, Moloche, Guillermo, and Weinberg, Stephen (2006): "Costly information acquisition: Experimental analysis of a boundedly rational model," *American Economic Review*, 96(4), 1043–1068.

Gul, Faruk, Natenzon, Paulo, and Pesendorfer, Wolfgang (2014): "Random choice as behavioral optimization," *Econometrica*, 82(5), 1873–1912.

Gul, Faruk and Pesendorfer, Wolfgang (2006): "Random expected utility," *Econometrica*, 74(1), 121–146.

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

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

Horan, Sean, Manzini, Paola, and Mariotti, Marco (2019): "When Is Coarseness Not a Curse? The Comparative Statics of Coarse Perception in Choice," Working paper, University of Montreal.

Jamieson, Donald G. and Petrusic, William M. (1977): "Preference and the time to choose," *Organizational Behavior and Human Performance*, 19(1), 56–67.

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

Khaw, Mel Win, Li, Ziang, and Woodford, Michael (2017): "Risk aversion as a perceptual bias," Working paper, National Bureau of Economic Research.

Kiss, Hubert Janos, Rodriguez-Lara, Ismael, and Rosa-Garcia, Alfonso (2020): "Does response time predict withdrawal decisions? Lessons from a bank-run experiment," *Review of Behavioral Finance*, forthcoming.

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

Konovalov, Arkady and Krajbich, Ian (2019): "Revealed strength of preference: Inference from response times," *Judgment and Decision Making*, 14(4), 381–394.

Kovach, Matthew and Tserenjigmid, Gerelt (2018): "The focal Luce model," Working paper Virginia Tech.

Krajbich, Ian, Armel, Carrie, and Rangel, Antonio (2010): "Visual fixations and the computation and comparison of value in simple choice," *Nature Neuroscience*, 13(10), 1292–1298.

Krajbich, Ian and Rangel, Antonio (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.

Lévy-Garboua, Louis, Maafi, Hela, Masclet, David, and Terracol, Antoine (2012): "Risk aversion and framing effects," *Experimental Economics*, 15(1), 128–144.

Loomes, Graham and Pogrebna, Ganna (2014): "Measuring individual risk attitudes when preferences are imprecise," *Economic Journal*, 124(576), 569–593.

Loomes, Graham, Starmer, Chris, and Sugden, Robert (1989): "Preference reversal: information-processing effect or rational non-transitive choice?" *Economic Journal*, 99(395), 140–151.

Loomes, Graham and Sugden, Robert (1995): "Incorporating a stochastic element into decision theories," *European Economic Review*, 39(3-4), 641–648.

Loomis, John, Peterson, George, Champ, Patricia, Brown, Thomas, and Lucero, Beatrice (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, Jay (2016): "Random choice and private information," *Econometrica*, 84(6), 1983–2027.

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

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

Luce, R. Duncan (1986): *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.

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

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

Manzini, Paola, and Mariotti, Marco (2014): "Stochastic choice and consideration sets," *Econometrica*, 82(3), 1153–1176.

Masatlioglu, Yusufcan, Nakajima, Daisuke, and Ozbay, Erkut Y. (2012): "Revealed Attention," *American Economic Review*, 102(5), 2183–2205.

Mas-Colell, Andreu, Whinston, Michael D., and Green, Jerry R. (1995): *Microeconomic Theory*, New York: Oxford University Press.

Matějka, Filip and McKay, Alisdair (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 105(1), 272–98.

McFadden, Daniel (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, Zarembka, Paul (Ed.), New York, Academic Press, 105–142.

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

McFadden, Daniel (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, Manski, Charles and McFadden, Daniel (Eds.), Cambridge, MA, MIT Press, 198–272.

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

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

Mosteller, Frederick and Nogee, Philip (1951): "An experimental measurement of utility," *Journal of Political Economy*, 59(5), 371–404.

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

Navarro-Martinez, Daniel, Loomes, Graham, Isoni, Andrea, Butler, David, and Alaoui, Larbi (2018): "Boundedly rational expected utility theory," *Journal of Risk and Uncertainty*, 57(3), 199–223.

Olschewski, Sebastian, Newell, Ben, and Scheibehenne, Benjamin (2019): "How Basic Cognition Influences Experience-Based Economic Valuation," Working paper, University of Basel.

Oud, Bastiaan, Krajbich, Ian, Miller, Kevin, Cheong, Jin Hyun, Botvinick, Matthew, and Fehr, Ernst (2016): "Irrational time allocation in decision-making," *Proceedings of the Royal Society B: Biological Sciences*, 283(1822), 20151439.

Payne, John W., Bettman, James R., and Johnson, Eric J. (1993): *The Adaptive Decision Maker*. Cambridge University Press.

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

Polanía, Rafael, Krajbich, Ian, Grueschow, Marcus, and Ruff, Christian 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.

Ratcliff, Roger and McKoon, Gail (2008): "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks," *Neural Computation*, 20(4), 873–922.

Reutskaja, Elena, Nagel, Rosemarie, Camerer, Colin F., and Rangel, Antonio (2011): "Search dynamics in consumer choice under time pressure: An eye-tracking study," *American Economic Review*, 101(2), 900–926.

Rubinstein, Ariel (2013): "Response time and decision making: An experimental study," *Judgment and Decision Making*, 8(5), 540–551.

Rubinstein, Ariel and Salant, Yuval (2006): "A model of choice from lists," *Theoretical Economics*, 1(1), 3–17.

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

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

Schulz, Jonathan F., Fischbacher, Urs, Thöni, Christian, and Utikal, Verena (2014): "Affect and fairness: Dictator games under cognitive load," *Journal of Economic Psychology*, 41, 77–87.

Sopher, Barry and Gigliotti, Gary (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, Barry and Narramore, J. Mattison (2000): "Stochastic choice and consistency in decision making under risk: An experimental study," *Theory and Decision*, 48(4), 323–350.

Spiliopoulos, Leonidas and Ortmann, Andreas (2018): "The BCD of response time analysis in experimental economics," *Experimental Economics*, 21(2), 383–433.

Stevens, S. S. (1961): "To honor Fechner and repeal his law," *Science*, 133(3446), 80–86.

Swensson, Richard G. (1972): "The elusive tradeoff: Speed vs accuracy in visual discrimination tasks," *Perception and Psychophysics*, 12(1), 16–32.

Teghtsoonian, Robert (1971): "On the exponents in Stevens' Law and the constant in Ekman's Law," *Psychological Review*, 78(1), 71–80.

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, Konstantinos, Moran, Rani, Moreland, James, Chater, Nick, Usher, Marius, and Summerfield, Christopher (2016): "Economic irrationality is optimal during noisy decision making," *Proceedings of the National Academy of Sciences*, 113(11), 3102–3107.

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

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

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

Webb, Ryan (2019): "The (neural) dynamics of stochastic choice," *Management Science*, 65(1), 230–255.

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

Weibull, Jörgen W., Mattsson, Lars-Göran, and Voorneveld, Mark (2007): "Better may be worse: Some monotonicity results and paradoxes in discrete choice under uncertainty," *Theory and Decision*, 63(2), 121–151.

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

Yellott, John 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.

Zeigenfuse, Matthew D., Pleskac, Timothy J., and Liu, Taosheng (2014): "Rapid decisions from experience," *Cognition*, 131(2), 181–194.

Appendix

Summary statistics of the Selected longest variable and response times

Table $A1$ characterizes the Selected longest variable and response times in the difficulty treatments.

Table $A1$: Selected longest variable and response times

Easy	Medium	Difficult	Pooled
96.36%	74.84%	38.50%	69.65%
$9.984s$	$13.834s$	$16.070s$	$13.306s$

The observations within the difficulty treatments range from 3553 to 3751. There are $10,989$ observations in the pooled analysis.

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 $A2$ characterizes the Selected longest variable and response time in the number of lines treatments.

Table $A2$: Selected longest variable and response times

2 Lines	3 Lines	4 Lines	5 Lines	6 Lines
78.11%	72.67%	69.21%	64.85%	62.91%
$8.469s$	$11.282s$	$13.356s$	$15.704s$	$18.042s$

The observations within the number of lines treatments range from 2181 to 2284.

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. This also suggests that subjects take more time on trials where the choice set is larger. Table $A3$ characterizes the Selected longest variable and response times in the longest line length treatments.

Table $A3$: Selected longest variable and response times

160	176	192	208	224	240	256	272	288	304
76.4%	71.9%	72.1%	71.3%	67.9%	68.4%	69.4%	66.7%	68.6%	63.7%
$12.73s$	$13.18s$	$12.36s$	$13.22s$	$12.76s$	$13.33s$	$13.09s$	$13.63s$	$13.90s$	$14.88s$

The observations within the longest line length treatments range from 1093 to 1102.

There appears to be a relationship between the length of the longest line in a trial and the likelihood that the longest line was selected in that trial. Further, is also appears to be a relationship between response time and the length of the longest line. This suggest to us a positive relationship between the effort spent and the length of the longest line in that trial. In Table $A4$ we characterize the Selected longest variable and the response times according to the number of lines and the letter label of the longest line.

Table $A4$: Selected longest variable and response times by number of lines and letter label of the longest

	A	B	C	D	E	F
2 Lines	77.84%	78.39%	–	–	–	–
	8.546s	8.389s				
3 Lines	72.39%	72.56%	73.08%	–	–	–
	11.272s	11.411s	11.165s			
4 Lines	66.48%	61.90%	75.41%	72.48%	–	–
	13.509s	13.930s	13.015s	13.020s		
5 Lines	59.86%	60.47%	63.82%	70.27%	69.54%	–
	16.087s	16.502s	15.010s	15.974s	15.011s	
6 Lines	54.97%	56.89%	62.43%	73.24%	61.58%	67.79%
	18.461s	17.957s	17.556s	17.712s	19.059s	17.543s

The observations range from 1115 to 1166 within the 2 line treatment, from 707 to 757 in the 3 line treatment, from 525 to 585 in the 4 line treatment, from 430 to 456 in the 5 line treatment, from 341 to 370 in the 6 line treatment.

There appears to be a relationship between both the Selected longest variable and response times with the letter label of the longest line.

Summary statistics for the quality of choice and response times

Response times of the trials in which the longest line was selected is smaller than the response times in trials in which the longest line was not selected is robust across the longest line treatments, the number of line treatments, and the difficulty treatments.

Table *A*5: Response times

	Easy	Medium	Difficult
Longest	9.936	13.498	15.645
Not	11.272	14.833	16.337
Z-stat	2.61	3.50	2.91
p-value	0.009	< 0.001	0.004

The mean response times in seconds within each difficulty treatment, conditional on whether the longest line was selected or not. The results of a Wilcoxon Two-Sample Test (Z-statistic and p-value) are reported.

Table *A*6: Response times

	2 Lines	3 Lines	4 Lines	5 Lines	6 Lines
Longest	8.063	10.722	12.468	14.742	16.735
Not	9.920	12.772	15.353	17.478	20.258
Z-stat	7.832	6.431	7.452	6.224	6.838
p-value	every test < 0.001				

The mean response times in seconds within each number of lines treatment, conditional on whether the longest line was selected or not. The results of a Wilcoxon Two-Sample Test (Z-statistic and p-value) are reported.

Table *A*7: Response times

	160	176	192	208	224	240	256	272	288	304
Longest	11.73	12.07	11.52	12.53	11.89	12.11	12.30	12.55	12.52	13.48
Not	15.95	16.03	14.54	14.92	14.60	15.96	14.89	15.80	16.90	17.35
Z-stat	5.58	8.16	5.74	4.23	5.63	6.12	5.81	5.40	6.52	5.70
p-value	every test <0.001									

The mean response times in seconds within each longest line treatment, conditional on whether the longest line was selected or not. The results of a Wilcoxon Two-Sample Test (Z-statistic and p-value) are reported.

In every treatment, we see that suboptimal choices in the line selection task are associated with a larger response time than are optimal choices.

Robustness of the optimality of choices

In order to investigate the optimality of choices, in Table 1 we summarized logistic regressions of the Selected longest variable. Here we perform the analogous exercise but we analyze the Longest minus selected variable, defined to be the length of the longest line in the trial menus the length of the line selected in that trial. 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 regressions in Table $A8$.

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

	(1)	(2)	(3)	(4)
Longest line normalized	0.027^{***}	0.027^{***}	0.027^{***}	0.027^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
Number of lines normalized	1.621^{***}	–	1.606^{***}	–
	(0.107)		(0.110)	
Easy treatment dummy	-11.571^{***}	-11.520^{***}	-11.617^{***}	-11.591^{***}
	(0.481)	(0.480)	(0.490)	(0.489)
Difficult treatment dummy	3.437^{***}	3.395^{***}	3.407^{***}	3.354^{***}
	(0.341)	(0.340)	(0.350)	(0.349)
Trial	0.015^{**}	0.015^{**}	0.014^{**}	0.015^{***}
	(0.005)	(0.005)	(0.005)	(0.005)
Letter dummies	No	Yes	No	Yes
Fixed effects	Yes	Yes	No	No
Demographics	No	No	Yes	Yes
AIC	31441	31400	31758	31715

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 (Akaike, 1974). Each regression has $10,989$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Similar to Table 1, the accuracy of choice decreases when there is a larger number of lines in the choice set, decreases in the length of the longest line, and decreases in the difficult treatments. We also observe a decrease in accuracy across trial. It seems as if the results presented in Table 1 are robust to the specification of the optimality of the choose.

Robustness of the relationship between quality of choice and response times

In order to test the robustness of Table 5, we conduct a similar analysis, but we perform an estimate of the treatment variables for every subject. 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 treatment estimates*. We also employ a specification where we estimate the Trial coefficient for every subject. We refer to this as *Subject-specific Trial estimates*. We summarize this analysis in Table $A9$.

Table $A9$: Regressions of the log of Response time variable

	(1)	(2)
Trial	-0.002^{***}	–
	(0.0001)	
Select Longest	-0.015^{***}	-0.017^{***}
	(0.004)	(0.004)
Subject-specific treatment estimates	*Yes*	*Yes*
Subject-specific Trial estimates	*No*	*Yes*
Letter dummies	*No*	*No*
Fixed effects	*Yes*	*Yes*
Demographics	*No*	*No*
AIC	−4087.2	−3558.4

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 10,989 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In the regressions where we estimate a unique treatment coefficient for every subject, we see that the Select longest variable remains significant. This suggests a negative relationship between the optimal choice in the line selection task and Response time. Further, we note that when we conduct Spearman correlations between the unstandardized residuals and the Select longest variable in specifications (1) and (2), the p-values respectively are 0.630 and 0.628. When we conduct Pearson correlations, the qualitative results are not changed. Further, our qualitative results are not changed when we use the standardized Pearson residuals rather than the unstandardized residuals. We interpret this as suggesting that our specifications are not suffering from an endogeneity bias caused by the inclusion of the Selected longest variable. The results summarized in Table $A9$ provide additional evidence that the relationship between response times and the optimality of the choice is not driven by a possible negative subject-level relationship between skill at the task and the speed by which it is accomplished.

Summary statistics of the IIA analysis

Table $A10$ lists the number of instances where the longest line and the second longest line was selected in trials with a unique second longest line, by difficulty treatment and the number of lines in the choice set.

Table $A10$: Longest and Second longest choices by difficulty treatment and number of lines in the choice set

	Easy treatment				
	2	3	4	5	6
Longest	801	703	671	615	538
Second longest	23	11	12	14	24
Ratio	34.83	63.91	55.92	43.93	22.42
	Medium treatment				
	2	3	4	5	6
Longest	575	528	467	444	397
Second longest	121	120	83	94	92
Ratio	4.75	4.40	5.70	4.72	4.32
	Difficult treatment				
	2	3	4	5	6
Longest	414	287	235	203	125
Second longest	356	240	193	164	108
Ratio	1.16	1.20	1.22	1.24	1.16

Only trials with a unique second longest line are included. The upper panel summarizes data from the easy treatment, the middle panel summarizes data from the medium treatment, and the lower panel summarizes data from the difficult treatment. Easier treatments and treatments with smaller choice sets are more likely to have a unique second longest line.

In the easy treatment, there is a relationship between the longest and second longest choices across sizes of the choice set according to a χ^2 test ($\chi^2(4) = 12.11$, $p = 0.02$) and a

Kruskal-Wallis test ($H(4) = 12.77$, $p = 0.01$). However, the in the medium treatment, this relationship is not significant according to a χ^2 test ($\chi^2(4) = 3.27$, $p = 0.51$) or a Kruskal-Wallis test ($H(4) = 3.19$, $p = 0.53$). Likewise, in the difficult treatment, this relationship is not significant according to a χ^2 test ($\chi^2(4) = 0.35$, $p = 0.99$) or a Kruskal-Wallis test ($H(4) = 0.38$, $p = 0.98$).