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

Sean Duffy* John Smith†

March 29, 2025

Appendix for Online Publication

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 $3,553$ to $3,751$. 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 $2,181$ to $2,284$.

*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

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 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 length treatments range from $1,093$ to $1,102$.

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 $1,115$ to $1,166$ 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 $A5$: 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 $A6$: 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 $A7$: 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 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 length 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)
CRT sum	–	–	-1.2890^{***}	-1.3116^{***}
			(0.1887)	(0.1881)
Female	–	–	-2.3124^{***}	-2.3138^{***}
			(0.3278)	(0.3270)
Letter dummies	No	Yes	No	Yes
Fixed effects	Yes	Yes	No	No
Demographics	No	No	Yes	Yes
AIC	31,441	31,400	31,758	31,715

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 Left handed 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.

The analysis summarized in Table 1 employed fixed-effects, whereby a unique dummy variable was estimate for each subject. In order to learn the robustness to this assumption, here we conduct random-effects regressions. Specifically, we estimate a compound symmetry covariance matrix clustered by participant. In other words, any two observations involving a participant are assumed to be correlated, but any observations involving different participants are assumed to be 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). We summarize these random-effects regressions in Table *A*9.

Table *A*9: Logistic regressions of the Selected longest line variable

	(1)	(2)
Longest length normalized	-0.0007^{***}	-0.0007^{***}
	(0.0001)	(0.0001)
Number of lines normalized	-0.041^{***}	–
	(0.003)	
Easy treatment dummy	0.368^{***}	0.397^{***}
	(0.014)	(0.016)
Difficult treatment dummy	-0.274^{***}	-0.295^{***}
	(0.011)	(0.011)
Trial	$-0.00028^{\dagger}$	-0.00033^{*}
	(0.00015)	(0.00016)
Letter dummies	*No*	*Yes*
Random effects	*Yes*	*Yes*
Demographics	*No*	*No*
QIC	9935.97	9889.80

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 random-effects correlation estimates. QIC refers to the Quasi-likelihood 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 are not changed. Therefore, the results in Table 1 seem to not be sensitive to the repeated measures specification.

Robustness of the relationship between quality of choice and response times

The reader *should* be concerned about endogeneity in our specifications in Table 5. Response time is an endogenously selected variable and is possibly correlated with the error term. For example, it is possible that our specifications do not sufficiently capture the subject heterogeneity. It is therefore possible that the Selected longest result is being driven by more accurate subjects also tending to have faster response times. It is additionally possible that our specifications do not sufficiently account for the choice set heterogeneity. More difficult choice problems are both less accurate and slower. However, if we are not sufficiently accounting for the heterogeneity of the choice problems, then this effect might be driving the results.

Due to concerns of endogeneity in the results from Table 5, we check whether the Selected longest variable is correlated with the errors in the regressions. 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, 0.88, and 0.87. Further, our qualitative results are not changed when we use the student 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.

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 $A10$.

Table $A10$: Regressions of the log of Response time variable

	(1)	(2)	(3)
Trial	-0.002^{***}	-0.002^{***}	$-$
	(0.0001)	(0.0001)	
Selected longest	-0.015^{***}	-0.041^{***}	-0.017^{***}
	(0.004)	(0.008)	(0.004)
Trial*Selected longest	$-$	0.0005^{***}	$-$
		(0.0001)	
Subject-specific treatment estimates	Yes	Yes	Yes
Subject-specific Trial estimates	No	No	Yes
Letter dummies	No	No	No
Fixed effects	Yes	Yes	Yes
Demographics	No	No	No
AIC	-4087.2	-4087.6	-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. However, we note that the effect decreases across trials.

Further, we note that when we conduct Spearman correlations between the unstandardized residuals and the Select longest variable in specifications $(1) - (3)$, the p-values respectively are 0.630 , 0.579, and 0.628. Further, our qualitative results are not changed when we use the standardized Pearson residuals rather than the unstandardized residuals.

We offer another test of robustness for Table 5. Here we employ specifications without fixed-effects but 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 $A11$.

Table $A11$: Regressions of the log of Response time variable

	(1)	(2)	(3)	(4)
Longest length normalized	0.0003^{***}	0.0003^{***}	0.0003^{***}	0.0003^{***}
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Number of lines normalized	0.082^{***}	–	0.082^{***}	–
	(0.0012)		(0.0012)	
Easy treatment dummy	-0.114^{***}	-0.114^{***}	-0.114^{***}	-0.114^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Difficult treatment dummy	0.051^{***}	0.052^{***}	0.051^{***}	0.051^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Trial	-0.002^{***}	-0.002^{***}	-0.002^{***}	-0.002^{***}
	(0.00006)	(0.00006)	(0.0001)	(0.0001)
Selected longest	-0.016^{***}	-0.014^{***}	-0.040^{***}	-0.040^{***}
	(0.004)	(0.004)	(0.008)	(0.008)
Trial*Selected longest	–	–	0.0005^{***}	0.0005^{***}
			(0.0001)	(0.0001)
Sub's Average log RT	0.994^{***}	0.992^{***}	0.993^{***}	0.991^{***}
	(0.014)	(0.014)	(0.014)	(0.014)
Letter dummies	No	Yes	No	Yes
Fixed effects	No	No	No	No
Demographics	Yes	Yes	Yes	Yes
AIC	-6680.2	-6651.1	-6677.2	-6647.8

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 $p < 0.1$.

Again, we note the negative and significant Selected longest coefficient estimates. Again, we note the positive and significant interaction between Trial and Selected longest.

When we conduct Spearman correlations between the unstandardized residuals and the Select longest variable in specifications $(1) - (4)$, the p-values respectively are 0.92, 0.89, 0.85, and 0.84. Further, our qualitative results are not changed when we use the standardized Pearson residuals rather than the unstandardized residuals.

We interpret these results as suggesting that our specifications are not suffering from an endogeneity bias caused by the inclusion of the Selected longest variable.

Robustness of the multinomial discrete choice models

In order to further explore the robustness of our results, we conduct an analysis similar to Table 6. However, because sensitivity to line length might be affected by the letter label of the line we account for this in the analysis below. In other words, the non-stochastic component to the utility for line j: $V_j = \beta_j * f(Length_j)$, where $f(.)$ has a linear, log, or power specification. These analyses are summarized in Table $A12$.

Table $A12$: Comparisons of different multinomial discrete choice models

	Linear		Log		Power	
	Logit	Probit	Logit	Probit	Logit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
2 Lines	1919	1930	1900	1914	1921	1932
3 Lines	2594	2599	2609	2623	2595	2599
4 Lines	3184	3225	3166	6036	3187	3364
5 Lines	4021	5694	4039	7135	4022	5900
6 Lines	4310	4321	4304	4321	4312	4324

We provide the Akaike Information Criterion (AIC) for the various models restricted to treatments with identical numbers of lines. We do not provide the coefficient estimates.

Regardless of the specification (Linear, Log, or Power) or the number of lines in the choice set, the AIC for the logit specification is smaller than that for the probit specification.

In order to further account for possible heterogeneity, we include specifications that estimate a unique dummy for every letter label in the choice set. In these specifications, the non-stochastic component to the utility for line j is: $V_j = \beta * f(Length_j) + LetterLabelDummy_j$. This analysis is summarized in Table $A13$.

Table $A13$: Comparisons of different multinomial discrete choice models

	Linear		Log		Power	
	Logit	Probit	Logit	Probit	Logit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
2 Lines	1921	1932	1902	1916	1922	1933
3 Lines	2597	2603	2611	2625	2598	2603
4 Lines	3188	3397	3169	3230	3190	3353
5 Lines	4020	5176	4040	5875	4022	4077
6 Lines	4314	4327	4306	4324	4317	4330

We provide the Akaike Information Criterion (AIC) for the various models restricted to treatments with identical numbers of lines. We do not provide the letter label dummy estimates.

Again, regardless of the specification (Linear, Log, or Power) or the number of lines in the choice set, the AIC for the logit specification is smaller than that for the probit specification.

We note that the Logit and Probit analyses summarized in Tables 6, $A12$, and $A13$, did not account for subject-specific heterogeneity, and this possibly affects our results. In order to address this possibility, we offer an analysis, where we estimate a subject-specific Letter label dummy. In these specifications, the non-stochastic component to the utility for line j and subject i is: $V_{i,j} = \beta * f(Length_j) + LetterLabelDummy_{i,j}$. In Table $A14$ we summarize this analysis.

Table $A14$: Comparisons of different multinomial discrete choice models

	Linear		Log		Power	
	Logit	Probit	Logit	Probit	Logit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
2 Lines	2060	2073	2038	2058	2062	2075
3 Lines	2836	2842	2856	2870	2837	2843
4 Lines	3498	4235	3475	4385	3502	4122
5 Lines	4345	5538	4361	5757	4347	5411
6 Lines	4735	4755	4729	4756	4738	4758

We provide the Akaike Information Criterion (AIC) for the various models restricted to treatments with identical numbers of lines. We do not provide the letter label dummy estimates.

In each specification summarized in Table $A14$, the logit version has a lower AIC than the probit version. The results of Tables $A12$, $A13$, and $A14$ suggest that the results in Table 6 are robust to the specification.

More on the IIA analysis

Table $A15$ 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. Keep in mind that these values are averaged across the length of the longest line, subject heterogeneity, letter labels, etc.

Table *A*15: Longest and Second longest choices by difficulty treatment and number of lines in the choice set

	Easy treatment				
	2	3	4	5	6
Longest	800	697	671	615	537
Second longest	23	11	9	12	22
Fraction	0.972	0.984	0.968	0.981	0.961
	Medium treatment				
	2	3	4	5	6
Longest	575	528	467	444	396
Second longest	121	119	83	94	92
Fraction	0.826	0.816	0.849	0.825	0.811
	Difficult treatment				
	2	3	4	5	6
Longest	409	286	234	202	124
Second longest	356	239	193	164	105
Fraction	0.535	0.545	0.548	0.552	0.541

Only trials with a unique second longest line are included. The upper panel summarizes the averaged data from the easy treatment, the middle panel summarizes the averaged data from the medium treatment, and the lower panel summarizes the averaged data from the difficult treatment. Easier treatments and treatments with smaller choice sets are more likely to have a unique second longest line. There are a total of $8,628$ observations.

We conduct an analysis, as that summarized in Table 8, however, we assume that *Number of lines* is a continuous, not categorical variable. These regressions are summarized in Table *A*16.

Table $A16$: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)	(5)
Longest length normalized	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}	–
	(0.00007)	(0.00007)	(0.00007)	(0.00007)	
Number of lines normalized	-0.0005	-0.0005	-0.0021	0.0235	0.0812^{***}
	(0.0022)	(0.0022)	(0.0022)	(0.0378)	(0.0198)
Easy treatment dummy	0.223^{***}	0.222^{***}	0.216^{***}	0.213^{***}	–
	(0.009)	(0.009)	(0.009)	(0.008)	
Difficult treatment dummy	-0.138^{***}	-0.138^{***}	-0.138^{***}	-0.136^{***}	–
	(0.008)	(0.008)	(0.008)	(0.008)	
Trial	–	-0.0001	-0.0001	-0.0001	–
		(0.0001)	(0.0001)	(0.0001)	
Letter dummies	No	No	No	Yes	No
Fixed effects	No	No	Yes	Yes	No
Demographics+	No	No	No	No	No
AIC	6605.2	6605.8	6582.0	6567.2	8387.9

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 marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the subject-specific estimates, or the letter label 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 $p < 0.1$.

We note that none of the Number of lines estimates are significant, and the qualitative results from Table 8 are unchanged.

We now investigate whether choices involving the longest three lines in our data are consistent with IIA. For the bookkeeping reason described above, here we exclude trials where there is not a unique second or unique third longest line. Table $A17$ lists the number of instances where the longest line, the second longest line, or the third longest line was selected. These observations are grouped by difficulty treatment and the number of lines in the choice set.

Table $A17$: Longest, second longest, and third longest choices by difficulty treatment and number of lines

	Easy treatment				
	2	3	4	5	6
Longest	800	697	654	586	497
Second longest	23	11	9	12	20
Third longest	–	2	6	11	7
Fraction 1 : 2	0.972	0.984	0.986	0.980	0.961
Fraction 2 : 3	–	0.846	0.600	0.545	0.741
Fraction 1 : 3	–	0.997	0.991	0.982	0.986
	Medium treatment				
	2	3	4	5	6
Longest	575	528	451	416	358
Second longest	121	119	81	89	90
Third longest	–	26	46	40	30
Fraction 1 : 2	0.826	0.816	0.848	0.837	0.799
Fraction 2 : 3	–	0.821	0.638	0.690	0.750
Fraction 1 : 3	–	0.953	0.907	0.912	0.923
	Difficult treatment				
	2	3	4	5	6
Longest	409	286	216	168	99
Second longest	356	239	159	121	84
Third longest	–	139	110	90	74
Fraction 1 : 2	0.534	0.545	0.576	0.581	0.541
Fraction 2 : 3	–	0.632	0.591	0.573	0.531
Fraction 1 : 3	–	0.673	0.663	0.651	0.572

Only trials with unique second and third longest lines 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 and third longest lines. There are a total of $8,852$ observations, with $8,274$ from the longest and second longest, and 578 from the third longest.

As in Table $A15$, these values are averaged across the length of the longer lines, subject heterogeneity, letter labels, etc. We use the data summarized in Table $A17$ to investigate whether these fractions ($1:2$, $2:3$, and $1:3$) are consistent with IIA

There are $8,852$ observations in Table $A17$. Obviously, there are no third longest observations in choice sets with only two lines. There are $2,284$ observations in choice sets with two lines and there are $6,568$ observations in choice sets with three or more lines. We want to investigate whether the fractions change as a function of the size of the choice set. Each

observation from a choice set with three or more lines (whether the longest, second longest, or third longest) appears in two separate fraction calculations. On the other hand, each observation from a choice set with two lines appears in only a single fraction calculation. Hence, each of the $6,568$ observations in the choice sets with three or more lines appears twice in our dataset, and each of the $2,284$ observations in the choice sets with two lines appear only once. Our dataset therefore has $15,420$ observations.

To compare across fractions, we describe the dependent variable as *Selected longer*. Clearly the fraction values depend on which fraction is calculated ($1:2$, $2:3$, or $1:3$). For this reason, we include *Fraction dummies* to account for these differences.

An additional complication is that, unlike the data summarized in Table $A15$, the difference between any pairs of lines in the fractions are not completely characterized by the difficulty treatment. Specifically, the differences between the lengths of the lines are only characterized by the difficulty treatment in the Fraction $1:2$ values. The differences between the lengths of the lines in the Fraction $2:3$ and Fraction $1:3$ are not completely characterized by the difficulty treatment. It is for this reason that we include a *Line difference* variable, which is the longer line minus the shorter line. Due to the redundance introduced by the Line difference variable, we exclude the difficulty treatment dummies. Our analysis is otherwise similar to specifications (1)-(4) in Table 8. We summarize this analysis in Table $A18$.

Table $A18$: Logistic regressions of the Selected longer line variable

	(1)	(2)	(3)	(4)
Longer length normalized	-0.00022^{***}	-0.00022^{***}	-0.00022^{***}	-0.00022^{***}
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Line difference	0.0094^{***}	0.0094^{***}	0.0091^{***}	0.0089^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Trial	$-$	-0.00006	-0.00006	-0.00007
		(0.00006)	(0.00006)	(0.00006)
Number of lines Wald statistic	1.74	1.71	1.57	0.85
p-value of Wald statistic	0.78	0.79	0.81	0.93
Fraction dummies	Yes	Yes	Yes	Yes
Letter dummies	No	No	No	Yes
Fixed effects	No	No	Yes	Yes
Demographics+	No	No	No	No
AIC	$11,046.0$	$11,047.1$	$10,939.1$	$10,864.3$

We only consider trials with unique second and third longest lines. 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 Number of lines dummy estimates, the subject-specific estimates, the letter label estimates, or the fraction estimates. We report the Wald statistic and its corresponding p-value for the Number of lines dummy estimates, which has 4 degrees of freedom. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has $15,420$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In none of the specifications is the Number of lines variable significantly related to the choice probabilities. In order to test the robustness of this result, we perform an analysis similar to that summarized in Table $A18$, but we assume that the Number of lines variable is continuous, not categorical. This analysis is summarized in Table $A19$.

Table $A19$: Logistic regressions of the Selected longer line variable

	(1)	(2)	(3)	(4)
Longer length normalized	-0.00022^{***}	-0.00022^{***}	-0.00022^{***}	-0.00022^{***}
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Number of lines normalized	0.0003	0.0003	-0.0010	-0.013
	(0.0014)	(0.0014)	(0.0013)	(0.020)
Line difference	0.0093^{***}	0.0093^{***}	0.0091^{***}	0.0089^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Trial	–	-0.00006	-0.00006	-0.00007
		(0.00006)	(0.00006)	(0.00006)
Fraction dummies	Yes	Yes	Yes	Yes
Letter dummies	No	No	No	Yes
Fixed effects	No	No	Yes	Yes
Demographics+	No	No	No	No
AIC	$11,041.6$	$11,042.7$	$10,934.1$	$10,864.3$

We only consider trials with unique second and third longest lines. 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 subject-specific estimates, the letter label estimates, or the fraction estimates. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has $15,420$ observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Again, in none of the specifications is the Number of lines variable significantly related to the choice probabilities.

It is possible that linear specifications in Table 8 might not adequately capture the joint or interaction effects of the line lengths. We therefore present an analysis that estimates a dummy variable for each of the 30 combinations of longest lengths and difficulty treatments. We refer to these as *Length-difficulty dummies*. The analysis is otherwise identical to that specifications (1)-(4) in Table 8. We summarize this analysis in Table $A20$.

Table $A20$: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Trial	–	−0.0001	−0.0001	−0.0001
		(0.0001)	(0.0001)	(0.0001)
Number of lines Wald statistic	4.59	4.58	5.81	5.36
p-value of Wald statistic	0.33	0.33	0.21	0.25
Length-difficulty dummies	Yes	Yes	Yes	Yes
Letter dummies	No	No	No	Yes
Fixed effects	No	No	Yes	Yes
Demographics+	No	No	No	No
AIC	6627.3	6627.7	6598.8	6580.9

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 marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the Number of lines dummy estimates, the subject-specific estimates, the letter label estimates, or the Letter-difficulty estimates. We report the Wald statistic and its corresponding p-value for the Number of lines dummy estimates, which has 4 degrees of freedom. AIC refers to the Akaike information criterion (Akaike, 1974). Each regression has $8,628$ observations from 112 subjects with an average of 77.0 observations per subject. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Similar to the results summarized in Table 8, the size of the choice set is not significantly related to the relative choice of longest and second longest lines, which is consistent with IIA.

Similarly, the linear specifications in Table $A16$ might not sufficiently capture the the joint or interaction effects of the line lengths. We conduct an analysis that employs the Length-difficulty dummies, but the linear Number of lines specification as in Table $A16$. We summarize this analysis in Table $A21$.

Table $A21$: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
Number of lines normalized	−0.0005	−0.0006	−0.0021	0.0240
	(0.0022)	(0.0022)	(0.0021)	(0.0370)
Trial	–	−0.0001	−0.0001	−0.0001
		(0.0001)	(0.0001)	(0.0001)
Length-difficulty dummies	Yes	Yes	Yes	Yes
Letter dummies	No	No	No	Yes
Fixed effects	No	No	Yes	Yes
Demographics+	No	No	No	No
AIC	6625.8	6626.2	6597.6	6580.9

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 marginal effects evaluated at the sample means and the standard errors in parentheses. We do not provide the estimates of the intercepts, the subject-specific estimates, the letter label estimates, or the Letter-difficulty 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 $p < 0.1$.

Similar to the results summarized in Table $A16$, these results are consistent with IIA.

More on Memory Decay

The results summarized in Table 11 should be viewed with caution because of the possible endogenous nature of Time since longest as an independent variable. For example, subjects who are worse at the search (and therefore tend to have a larger Time since longest variable) also happen to be worse at the task. Or, Time since longest might simply reflect the likelihood that the subject mistakenly considers another line to be longer than the actual longest line.

To check the possibility of such endogeneity, we examine whether the Time since longest variable is correlated with the error term. 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.

In order to test the robustness of the results from Table 11, we employ a different specification of possible memory decay. We define the *View clicks since longest* variable to be the number of view clicks since the longest line was viewed, at the end of the trial. Note that

the variable has a value of 0 if the trial ended with the longest line visible. This variable is reminiscent of the *Number fixations since best last seen* variable employed by Reutskaja et al. (2011). Their results are summarized in Figure $3B$. We conduct an analysis similar to that summarized in Table 11, but with the View clicks since longest variable. This analysis is summarized in Table $A22$.

Table $A22$: Logistic regressions of the Selected longest line variable

	(1)	(2)	(3)	(4)
View clicks since longest	-0.253^{***}	-0.237^{***}	-0.259^{***}	-0.244^{***}
	(0.009)	(0.008)	(0.014)	(0.013)
Longest length normalized	-0.0009^{***}	-0.0008^{***}	-0.0009^{***}	-0.0008^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of lines normalized	0.029^{***}	0.027^{***}	0.029^{***}	0.027^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Easy treatment dummy	0.453^{***}	0.442^{***}	0.453^{***}	0.442^{***}
	(0.017)	(0.016)	(0.017)	(0.016)
Difficult treatment dummy	-0.259^{***}	-0.253^{***}	-0.259^{***}	-0.253^{***}
	(0.013)	(0.012)	(0.013)	(0.012)
Trial	-0.00006	-0.00005	-0.0002	-0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Trial*View clicks since longest	–	–	0.0001	0.0001
			(0.0002)	(0.0002)
Letter dummies	No	No	No	No
Fixed effects	Yes	No	Yes	No
Demographics+	No	Yes	No	Yes
AIC	6282.4	6174.1	6284.1	6175.6

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 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 that the View clicks since longest estimates are negative and significant in each of our specifications. This suggests a negative relationship between the quality of the choice and the number of lines viewed since the longest line.

Also similar to the results summarized in Table 11, we need to be concerned with possible endogeneity. We conduct Spearman correlations between the View clicks since longest vari-

able and the residuals (both unstandardized and Pearson standardized). Each correlation is significant at 0.001. This suggests that the relationship, similar to that summarized in Table 11, could suffer from an endogeneity bias.

More on attention

In an effort to learn if the effects found in Table 12 are endogenous, we conduct Spearman correlations between the Time viewing longest variable and the residuals (both unstandardized and Pearson standardized). We find that every correlation is significant at 0.001. We therefore cannot rule out that these results are driven by endogeneity.

Product rule

Although our analysis suggests that errors are better described as Gumbel rather than normal and our observations are consistent with IIA, we note that such behavior also implies the Product Rule. Perhaps the most prominent statement of the Product Rule is Theorem 2 in Luce (1959a) on page 16. If we denote $P(a, b)$ as the probability that a is selected from a menu of $\{a, b\}$ then we can write the Product Rule as:

$$P(a,b) * P(b,c) * P(c,a) = P(b,a) * P(c,b) * P(a,c),$$

where $a > b > c$. We refer to the set $\{a, b, c\}$ as a *triple*. Unfortunately, our dataset does not contain observations of choices from $\{a, b\}$, choices from $\{b, c\}$, and choices from $\{a, c\}$, which are needed to test the Product Rule.

However, Duffy and Smith (2025a) describe a similar line length judgment task with triples. Specifically, there are 117 subjects making two binary choices on every component of 56 triples. See Duffy and Smith (2025a) Table $A12$ for the percent correct on these binary choices and Table $A13$ for a description of the 56 triples. This dataset of the triples (NoCLUAL-FromPaymentLines-TriplesCalculation-SAS.xlsx) can also be found at https://osf.io/f7gu4/.

We test whether the left-hand side of the Product Rule expression is significantly different than the right-hand side. They are not significantly different according to a t-test ($t = 1.24$,

$p = 0.22$) and a non-parametric Signed-rank test ($S = 185$, $p = 0.13$). We therefore find evidence that the choices in Duffy and Smith (2025a) are consistent with the Product Rule.