

# Supplemental Appendix to Who is “Behavioral”? Cognitive Ability and Anomalous Preferences

Daniel J. Benjamin\*  
*Harvard University*

Sebastian A. Brown  
*Harvard University*

Jesse M. Shapiro  
*University of Chicago and NBER*

May 5, 2006

## Abstract

In this appendix, we summarize the findings of a pilot study conducted with Harvard undergraduates in February and May of 2004.

## 1 Laboratory Procedures for the Harvard Pilot

### 1.1 Participants

Participants were 60 undergraduates (virtually all Harvard students), recruited through on-campus posters and e-mail solicitations. We promised students \$5 for participating in a 45-minute experiment, with the possibility to earn more “depending on your responses [in the experiment].” We allowed only non-economics majors to participate because we were concerned that economics students would be familiar with our preference elicitation procedures. In total, we held six sessions, on February 8, 21 and 22 and May 14, 15 and 16, 2004.

### 1.2 Procedure

An experimenter guided participants through a questionnaire in unison by reading instructions aloud. The questionnaire contained four preference-elicitation sections: small-stakes risk preferences, short-term time preferences, fairness preferences, and idiosyncratic consumption preferences. The order of the sections differed across sessions; this order had no effect on the results.

At the end of the questionnaire, after all the preference-elicitation questions, we asked participants for their major, year in school, gender, and highest Math and Verbal SAT I scores. 57 out of 60 participants had taken the SAT. As a check on the accuracy of the self-reported SAT scores, we included a section of SAT-like math questions near the end of the questionnaire (after participants had responded to all the preference-elicitation sections). The correlation between performance on the SAT-like math questions and self-reported Math SAT score was 0.33, which is significantly different from 0 ( $p = 0.012$ ).

We paid all participants \$5 in cash for their participation immediately at the completion of the session. We also paid participants (by check) for their choices so that our preference-elicitation

---

\*E-mail: benjamin@fas.harvard.edu, sebastian.brown@post.harvard.edu, jmshapiro@uchicago.edu.

procedures were incentive-compatible. As described below, for some sections we paid participants immediately after the experiment. For other sections, we mailed checks to participants within a week of their participation.

*Small-Stakes Risk Preference.* Each question offered the choice between \$0.50 for sure and a gamble that gave a 50% chance of winning \$0 and a 50% chance of winning  $X$ , where  $X$  took the values \$0.95, \$1.05, \$1.15, \$1.25, and \$1.35 in ascending order. We gave participants an example question in the instructions and the opportunity to ask questions. Participants had 15 seconds to answer the five questions.

Participants knew from the instructions that after they made a selection for each question, we would roll a die five times to determine their payment for this section. We asked a participant to roll the die to maximize our credibility. We paid participants for this section in the check that we mailed within a week of the participant's participation.

*Short-Term Time Preference.* After an example question and an opportunity to ask the experimenter about the instructions, participants chose between \$5.00 today and  $X$  a week from now, where  $X$  was \$5.05, \$5.35, \$5.55, \$5.75, \$5.95, and \$6.15 in ascending order. Among the 57 participants for whom SAT scores are available, 6 did not answer all of the time preference questions (5 of whom were subjected to cognitive load, described below). These participants have been omitted from our analysis.

After giving participants 15 seconds to answer the six questions, we asked a participant to roll the die that would select the question to be implemented. The instructions explained that participants would receive a check to pay them for this section. That check would be post-dated by a week if the participant had chosen (B) for the relevant question. We gave participants their check immediately after the session.

*Fairness Preference.* To implement an anonymous Dictator Game, we informed participants that they had been randomly assigned to another participant in the same session, but that no one would ever find out who had been assigned to whom. To rule out simple reciprocity concerns, the experimenter made it clear that the other person had, in turn, been assigned to a different participant. As a result, each participant would affect the payoff of another participant but would not be affected by that other participant.

Participants were told they had been given \$1.00, and they had the opportunity to give away \$0.00, \$0.25, \$0.50, \$0.75, or \$1.00 to another participant. Participants were given 15 seconds to make a decision.

*Idiosyncratic Preference.* As a placebo section, we asked participants five binary preference questions. We asked them whether they preferred chocolate or coffee ice cream, red or silver-colored cars, cats or dogs as pets, Pepsi or Coke to drink, and Butterfinger or Kit Kat as a candy to eat. We informed participants that they would actually receive their preferred candy as part of their payment.

*Cognitive Load Manipulation.* We subjected a randomly chosen half of the participants to "cognitive load" during each section of the questionnaire. During each of the preference-elicitation sections of the questionnaire and during the SAT-like math questions, participants heard a CD of piano notes while they filled out the section. In each section, half the participants were required to remember the number of times they heard a specific sequence of musical tones. The sequence to be remembered varied across questionnaires. To incentivize participants in the cognitive load condition to pay attention to the tones, we made payment for that section contingent on correct recall of the number of repetitions of the sequence of tones.

## 2 Evidence on Cognitive Ability and Cognitive Load

### 2.1 Relationship between Cognitive Ability and Preference Anomalies

Our measure of cognitive ability is the participant’s Math SAT score. Participants’ Math SAT scores were all between 600 and 800, with a mean of about 750 and a median of 760. Almost a quarter of our respondents reported a perfect score of 800. The nationwide average Math SAT score was about 500 during the time period in which our participants took the exam (College Board, 2001). Since the distribution of Math SAT scores is highly non-normal, we conduct our analysis by comparing participants with below-median Math and Verbal SATs to participants with above-median scores.<sup>1</sup>

Table 1 shows the results of a probit analysis of the relationship between SAT scores and measured preferences. Column (1) shows that participants with above-median Math SAT scores have a 24 percentage point greater chance of behaving risk-neutrally, which is economically large and statistically significant at the 10% level. In column (2) when we include measures of both Math and Verbal SAT score, the estimated effect of an above-median Math SAT score increases and becomes significant at the 5% level, whereas we find a statistically insignificant negative effect of having an above-median Verbal SAT score. Our estimates in columns (3) and (4) of the effect of SAT scores show positive (though not quite statistically significant) effects of mathematical ability on patience and a small and insignificant negative effect of verbal ability.<sup>2</sup>

As columns (5) and (6) show, we find no evidence of a relationship between Math SAT score and “selfishness” (defined as keeping 100% of the dollar available for splitting). (We also find no relationship between cognitive ability and our incentivized idiosyncratic preference—choice of Kit Kat vs. Butterfinger candy bar.)<sup>3</sup> The point estimate indicates that more cognitively able individuals are slightly less likely to behave selfishly, but this estimate is statistically indistinguishable from zero.

### 2.2 Effects of Cognitive Load on Choices

Column (1) of table 2 shows that cognitive load reduced the number of correct answers on our six-question battery of SAT-like math questions by about .7 on average, and a Mann-Whitney test rejects the null of no effect at the 10% level ( $p = .080$ ).<sup>4</sup> Columns (2)-(4) show that in this sample, cognitive load did not have a statistically significant effect on measured risk-neutrality, expressed time preference, or Dictator Game behavior.<sup>5</sup>

---

<sup>1</sup>When we instead use specifications that are linear in SAT score, the estimated coefficients are comparable in sign and magnitude, but less statistically precise and in general not statistically distinguishable from zero.

<sup>2</sup>Interestingly, in a follow-up data collection two months after the final round of experiments, we found a positive (but statistically significant) relationship between Math SAT scores and the number of days it took the participant to cash the check she received in the discounting task. Evidently, higher-scoring participants were somewhat more willing to delay cashing the check, as well as to have the check dated later.

<sup>3</sup>There is no statistically significant relationship between Math SAT score and choices in three of our four other, hypothetical idiosyncratic preferences. The sole exception is that individuals with higher Math SAT scores are statistically more likely to prefer cats over dogs as pets. Neither of us interprets this to mean that cats are better than dogs.

<sup>4</sup>Additional tests (not shown) confirm that the cognitive load manipulation was successfully randomized—we find no evidence of a correlation between cognitive load and SAT score, gender, or year in school.

<sup>5</sup>In earlier pilot tests with Harvard students, we tried two other cognitive load procedures. We required participants to remember seven-digit numbers (as in Shiv and Fedorikhin, 1999, and Hinson, Jameson, and Whitney, 2002), and we also tried playing musical tones at a slower pace than in the actual experiment. Neither of these manipulations reliably influenced the preferences we measured.

## References

- [1] College Board. *College-bound Seniors: A Profile of SAT Program Test Takers*. College Entrance Examination Board, New York, NY, 2001.
- [2] J.M. Hinson, T.J. Jameson, and P. Whitney. Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29:298–306, 2003.
- [3] Baba Shiv and Alexander Fedorikhin. Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26(3):278–292, December 1999.

**Table 1** : *Preferences and Cognitive Ability: Harvard Undergraduates*

Dependent variable	Risk neutral (dummy)		Patient (dummy)		Selfish (dummy)	
	(1)	(2)	(3)	(4)	(5)	(6)
Math SAT $\geq$ median	0.2446 (0.1222)	0.2951 (0.1317)	0.2549 (0.1522)	0.2615 (0.1569)	-0.1635 (0.1421)	-0.1295 (0.1495)
Verbal SAT $\geq$ median		-0.1333 (0.1335)		-0.0252 (0.1499)		-0.1244 (0.1468)
Male	0.1512 (0.1395)	0.1279 (0.1407)	-0.2136 (0.1582)	-0.2158 (0.1586)	-0.3548 (0.1378)	-0.3781 (0.1387)
Year in school (1-4)	0.0710 (0.0715)	0.0771 (0.0720)	0.0817 (0.0847)	0.0812 (0.0848)	-0.1493 (0.0827)	-0.1583 (0.0846)
Mean of dependent variable	0.2982	0.2982	0.5294	0.5294	0.4912	0.4912
Pseudo- $R^2$	0.0809	0.0952	0.0635	0.0639	0.1090	0.1180
$N$	57	57	51	51	57	57

Notes: Standard errors in parentheses. Data are from laboratory study of Harvard undergraduates. Risk neutral indicates that participant made six decisions of the following form in a way consistent with expected-value maximization:

Please circle either Choice A or Choice B.

(A) You get \$0.50 for sure.

(B) If the die comes up 1, 2, or 3, you get  $X$ . If the die comes up 4, 5, or 6, you get nothing.

Patient indicates that participant made six decisions of the following form in a way consistent with the maximization of undiscounted wealth:

Please circle either Choice A or Choice B.

(A) You get \$5.00 right now.

(B) You get  $X$  a week from now.

Selfish indicates that participant kept 100% of sum available for transfer to other participant in Dictator Game.

Data on SAT scores based on participants' self-reports. Median refers to sample median. Demographics include dummies for gender and year in school.

**Table 2** : *Preferences and Cognitive Resources: Harvard Undergraduates*

	Math score (0-6) (1)	Risk neutral (dummy) (2)	Patient (dummy) (3)	Selfish (dummy) (4)
Mean for condition:				
No cognitive load	3.10	0.2424	0.5000	0.3704
Cognitive load	2.40	0.3704	0.5185	0.5455
Test	Mann-Whitney	Fisher exact	Fisher exact	Fisher exact
p-value	0.080	0.397	1.000	0.203
<i>N</i>	60	60	53	60

Notes: Standard errors in parentheses. Data are from laboratory study of Harvard undergraduates. Risk neutral indicates that participant made six decisions of the following form in a way consistent with expected-value maximization:

Please circle either Choice A or Choice B.

(A) You get \$0.50 for sure.

(B) If the die comes up 1, 2, or 3, you get  $X$ . If the die comes up 4, 5, or 6, you get nothing.

Patient indicates that participants made six decisions of the following form in a way consistent with the maximization of undiscounted wealth:

Please circle either Choice A or Choice B.

(A) You get \$5.00 right now.

(B) You get  $X$  a week from now.

Selfish indicates that participant kept 100% of sum available for transfer to other participant in Dictator Game.

Data on SAT scores based on participants' self-reports. Median refers to sample median. Demographics include dummies for gender and year in school.