

Supplemental Material: Descriptions of all tasks

Decision-Making Measures

Temporal discounting was measured with five choice titrators (Green, Fry, & Myerson, 1994). These titrators used both “delay” and “accelerate” frames: The delay frame presented participants with a series of choices between a fixed smaller gift certificate today and varying amounts of a larger gift certificate at a fixed time in the future, thus allowing participants to receive more money at the cost of delayed payment. The accelerate frame instead fixed the larger, future gift certificate and varied the amount of the smaller gift certificate today, thus allowing participants to accelerate payment at the cost of receiving less money. In addition to frame, we also manipulated and counterbalanced the future time (3, 4, or 12 months) and the fixed gift certificates amount (\$50, \$75, \$100, and \$115). The dependent variable for each titrator was each participant’s exponential annual discount rate, δ , as implied by the midpoint between preferring the earlier versus later payments (i.e., the indifference point). Values of δ closer to one (larger) indicated less discounting and thus more patient preferences.

Loss aversion was also measured with five choice titrators (Fehr & Goette, 2007). Each titrator presented participants with a series of choices indicating willingness to play each of a series of binary gambles with a 50% chance of winning some fixed amount (\$6 or \$20) and a 50% chance of losing some varying amount (between \$.50 to \$7 in \$.50 increments or between \$2 to \$24 in \$2 increments). Two of the titrators in the first wave were repeated without change in the fourth wave one year later. We calculated loss aversion coefficients, λ , by dividing the gain amount by the loss amount at the indifferent point (i.e., midpoint between where the participant switches from willing to play the gamble to not willing). Values of λ were reverse coded so that larger values indicated better *loss neutrality*.

Resistance to anchors was measured with six numerical estimation questions (e.g., distance between New York and Cairo), four of which were accompanied by anchors. For each set of three questions, participants saw one question with a high anchor (e.g., 11300 miles), one with a low anchor (e.g., 1400 miles), and one with no anchor. Questions were counterbalanced with anchor condition. We calculated standardized estimates by pooling responses across high, low, and no anchor conditions within question. Thus, the four dependent variables were the z-scores for the two low-anchor questions and reverse-coded z-scores for the two high-anchor questions. Higher values on these measures correspond to less anchoring.

Resistance to framing was measured in four different scenarios that asked participants to choose between a risky option and a sure option and varied the framing of the options as losses or gains (Tversky & Kahneman, 1981). We used the classic Asian disease problem (e.g., number of lives saved or lost), as well as variants involving potential layoffs (e.g., number of factories kept open or shut down), drought (e.g., acres of crops saved or lost) and bankruptcy (e.g., invested money saved or lost). Each participant saw either the gain version or the loss version of each scenario. Gain scenarios offered a choice between a sure but partial gain (e.g., 100% chance that 200 lives will be saved) and a risky option with chances of full gain and no gain (e.g., 33% chance that 600 lives will be saved and 67% chance that 0 lives will be saved). Loss scenarios offered a choice between a sure but partial loss (e.g., 100% chance that 400 lives will be lost) and a risky option with chances of no loss and full loss (e.g., 33% chance that 0 lives will be lost and 67% chance that 600 lives will be lost). Because these framings provide identical choices, Expected Utility maximizers should choose the same option under either framing, but Prospect Theory predicts that people susceptible to the framing will choose the sure option in the gain frame and the risky option in the loss frame. Susceptibility to framing is thus a binary measure of

whether, for each pair of scenarios, choices follow consistent risk preferences regardless of framing (1 if either both risk-averse or both risk-seeking) or not (0).

Financial and debt literacy were measured using a single six-question scale composed of three financial literacy questions (Lusardi & Mitchell, 2006) designed to assess knowledge of fundamental economic concepts and three debt literacy questions (Lusardi & Tufano, 2009) designed to assess knowledge of compound interest and credit card debt (see Table S1).

Table S1. Financial and debt literacy questions.

FL1. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account?

- More than today
- Exactly the same as today
- Less than today
- Do not know

FL2. Do you think that buying a single company stock usually provides a return that is more safe, equally safe, or less safe than the return on a stock mutual fund?

- More safe return than a stock mutual fund
- Equally safe return as a stock mutual fund
- Less safe return than a stock mutual fund
- Do not know

FL3. Is using money in a bank savings account to pay off credit card debt usually a good or a bad idea?

- Good idea
- Bad idea
- Do not know

DL1. Suppose you owe \$1000 on your credit card and the interest rate you are charged is 20% per year compounded annually. If you didn't pay anything off, at this interest rate,

how many years would it take for the amount you owe to double?

- 2 years
- Less than 5 years
- More than 5 but less than 10 years
- More than 10 years
- Do not know

DL2. You owe \$3,000 on your credit card. You pay a minimum payment of \$30 each month. At an Annual Percentage Rate of 12% (or 1% per month), how many years would it take to eliminate your credit card debt if you made no additional new charges?

- Less than 5 years
- Between 5 and 10 years
- Between 10 and 15 years
- Never, you will continue to be in debt
- Do not know

DL3. You purchase an appliance which costs \$1,000. To pay for this appliance, you are given the following two options: a) pay 12 monthly installments of \$100 each, b) borrow at a 20% annual interest rate and pay back \$1,200 a year from now. Which is the more advantageous offer?

- Option (A)
- Option (B)
- They are the same
- Do not know

Cognitive Measures

Fluid intelligence

Our most widely recognized measure of fluid intelligence is the *Raven's Progressive Matrices* task (Raven, 1962), a non-verbal test of inductive and analytic reasoning. The version we used, adapted from Salthouse, Pink and Tucker-Drop (2008) consisted of patterns in the form of 3x3 matrices with one cell missing. Participants had to determine the underlying rules that produce the pattern of rows and columns in the matrix which of the eight choice options correctly completed the pattern. Participants had 10 minutes to answer up to 18 matrices. Each item was presented on its own page and participants could choose to skip a question by selecting “no answer” but could not return to earlier items. Performance on this task was measured by the number of correct responses (0-18) with no penalty for incorrect responses.

Letter Sets (Ekstrom, French, Harman, & Dermen, 1976; Thurstone, 1962) is another measure of inductive and reasoning ability. The version of the task used in this study was also adapted from Salthouse et al. (2008). In this task, participants were presented with five sets of letters (e.g., NOPQ, DEFL, ABCD, HIJK, and UVWX) and they had to find the rule that related four of the five sets by checking the one which did not fit that rule (e.g., DEFL). Participants had 10 minutes to complete up to 15 items. As with the Raven Progressive Matrices, each item was presented on its own page and a “no answer” option allowed participants to skip items. The score for this task was calculated by the number of correct responses with no penalty for incorrect responses.

Number Series (Thurstone, 1962) is yet another measure of inductive and reasoning ability with particular emphasis on quantitative reasoning. The version we have used in this study is a block adaptive test developed by (McArdle & Woodcock, 2009) for the HRS 2010. Each item consisted of a series of numbers (e.g., 23, 26, 30, 35, ___), and participants identified

the number that correctly completed the series. All participants saw the same three items in the first block. The number of items answered correctly determined the difficulty of the three items in a second block. Thus, each participant completed six items in total, and a Rasch score was calculated based on which second block the participant completed and how many answers they got right in each block.

The *Cognitive Reflection Test* (CRT; Frederick, 2005; Frederick, Loewenstein, & O'Donoghue, 2002) consists of three mathematical questions that yield quick, impulsive, but incorrect first responses, which need to be inhibited to arrive at the correct answer. According to Frederick, the “*three items on the CRT are ‘easy’ in the sense that their solution is easily understood when explained, yet reaching the correct answer often requires the suppression of an erroneous answer that springs ‘impulsively’ to mind*” (Frederick, 2005, p. 27). The dependent variable used is the number of correct responses (0-3) with no penalty for incorrect responses.

Numeracy (Lipkus, Samsa, & Rimer, 2001) is the ability to understand probability and mathematical concepts. The numeracy task we used in our study consists of 11 questions that test comprehension and manipulation of proportions, percentages, and probabilities. The dependent variable used is the number of correct responses (0-11) with no penalty for incorrect responses.

Crystallized intelligence

Shipley’s vocabulary (Zachary, 1986) is a synonym vocabulary task that measures vocabulary knowledge. In our version, adapted from CREATE’s Common Core Battery of Measures (Czaja, Charness, Dijkstra, et al., 2006; Czaja, Charness, Fisk, et al., 2006), participants choose from among four words the one most similar in meaning to a target word. Participants had 10 minutes to complete up to 40 items split into two screens of 20 items each. In this task, a visible timer counted down on the upper left hand corner of the computer screen. The dependent measure used for this task is the number of correct responses (0-40) with no penalty

for incorrect responses.

Antonym vocabulary also measures vocabulary knowledge, using items developed by Salthouse (1993). In contrast to the Shipley's synonym vocabulary however, in this task participants choose from among five words the one most nearly opposite in meaning to a target word. Participants had five minutes to complete up to ten items, with a visible timer. Each item was presented on its own page but participants could choose to skip an item by selecting "no answer". Their score was the number of correct responses (0-10) with no penalty for incorrect responses.

The *Information* task (WAIS-III) (Wechsler, 1997) also adapted from CREATE (Czaja, Charness, Dijkstra, et al., 2006; Czaja, Charness, Fisk, et al., 2006) consisted of questions that measure general factual knowledge about events, objects, places, and people. Online administration required participants to read the questions on the computer screen and type their responses rather than having the questions read to them by a tester and answering verbally. We used the acceptable responses for each item as listed in the CREATE manual (Czaja, Charness, Dijkstra, et al., 2006) to determine whether an answer is correct. Because we could not prompt participants to give further details for a question, we could only code answers as correct (1) or incorrect (0) (instead of the original 0, 1, 2 scoring). Participants answered 28 questions without time restriction. The dependent measure used for this task is the number of correct responses (0-28) with no penalty for incorrect responses.

Inhibitory Control

The *Spatial 1-back* (Spatial Speed Match¹) task belongs to a family of N-back tasks that measure the capacity to update and actively manipulate working memory contents. The version

¹ All inhibitory control tasks were adapted from tasks developed by Lumos Labs (Lumosity.com). The name of each task in parentheses is the name for the task on the Lumosity website.

of the task we employed in our study tests visuo-spatial information processing. Participants were presented with three circles, two white and one blue, which formed a triangle on the screen. In each trial, they saw the same triangle but the blue circle could appear in one of the other corners of the triangle. Therefore, participants had to evaluate if the location of the blue circle was had changed from the trial before (1-back). As the accuracy level on this task was very high for young and old people, the dependent measure we used is participants' average reaction time.

The *Stroop* (Color Match) task measures cognitive flexibility and response inhibition capacity (Stroop, 1935). In the version used in this study, each trial displayed two words on the screen. The word on the left appeared in black font and indicated the name of a color (i.e., blue, green, red, etc.). The word on the right was also the name of a color and appeared in either the same color font as the semantic value of the color word on the left of the screen (e.g., congruent trials) or in a different color font (e.g., incongruent trials). Participants judged as quickly as possible if the semantic value of the word on the left corresponded to the font color of the word on the right. Participants had two minutes to finish as many trials as possible. The dependent measure was the difference in reaction times between incongruent and congruent trials, coded such that higher values correspond to better inhibition capacity.

The *Flanker* (Lost in Migration) task measures focus and resistance to interference. In this task, participants see a central item in a certain orientation flanked by distracting items in either the same or different orientations. In the version used in this study, participants saw five birds flying in a "V" shape. The task was to report the direction the middle bird (target) was flying by pressing one of the four arrow keys. In congruent trials, the other four birds were all flying in the same direction as the target and, in incongruent trials, they were all flying in a different direction. Participants had 45 seconds to finish as many trials as possible. The dependent measure was the difference in reaction times between incongruent and congruent

trials, coded such that higher values correspond to better inhibition capacity.

REFERENCES

- Czaja, S. J., Charness, N., Dijkstra, K., Fisk, A. D., Rogers, W. A., & Sharit, J. (2006). CREATE common core battery of measures: Technical report No. CREATE-2006-01. Atlanta, GA: Center for Research and Education on Aging and Technology Enhancement (CREATE).
- Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21*(2), 333-352.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Fehr, E., & Goette, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. *American Economic Review, 97*(1), 298-317.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature, 40*(2), 351-401.
- Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science*(5), 33-36.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making, 21*, 37-44.
- Lusardi, A., & Mitchell, O. S. (2006). *Financial Literacy and Planning: Implications for Retirement Wellbeing*. Pension Research Council. Wharton School, University of Pennsylvania.
- Lusardi, A., & Tufano, P. (2009). *Debt Literacy, Financial Experiences, and Overindebtedness*. Retrieved from http://www.dartmouth.edu/~alusardi/Papers/Lusardi_Tufano.pdf
- McArdle, J. J., & Woodcock, R. W. (2009). *Adaptive testing with selected Woodcock cognitive tests*. Longitudinal Research Institute. Charlottesville, VA.
- Raven, J. (1962). *Advanced Progressive Matrices, Set II*. London: H. K. Lewis.
- Salthouse, T. A. (1993). Speed mediation of adult age differences in cognition *Developmental Psychology, 29*, 722-738.
- Salthouse, T. A., Pink, J. E., & Tucker-Drop, E. M. (2008). Contextual analysis of fluid intelligence. *Intelligence, 36*, 464-486.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662.
- Thurstone, T. G. (1962). *Primary mental ability for Grades 9-12 (Rev. ed.)*. Chicago, IL: Science Research Associates.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale III (3rd ed.)*. San Antonio, TX: The Psychological Corporation.
- Zachary, R. A. (1986). *Shipley Institute of Living Scale: Revised manual*. Los Angeles: Western Psychological Services.