How well do laboratory-derived estimates of time preference predict real-world behaviors? Comparisons to four benchmarks

Accepted Manuscript, January 2023. Journal of Experimental Psychology: General

Daniel M. Bartels*, Ye Li†, Soaham Bharti*

*University of Chicago Booth School of Business, †University of California Riverside

Author's Note: Daniel M. Bartels is a Professor of Marketing at The University of Chicago Booth School of Business. Ye Li is an Associate Professor of Management at University of California Riverside. Soaham Bharti is a doctoral student at The University of Chicago Booth School of Business. Correspondence concerning this article should be addressed to Daniel Bartels (bartels@uchicago.edu), The University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, Illinois 60637.
Some of the ideas and data appearing in this manuscript have been presented at annual meetings of the Association of Consumer Research (2022), the Society for Judgment and Decision Making (2022), and the Society for the Quantitative Analyses of Behavior (2019). The data, materials, and analysis code for this paper have been made available for review at and will be publicly available once the paper is published: https://researchbox.org/459&PEER_REVIEW_passcode=ZPMQZP.

ACKNOWLEDGMENTS

We thank Stephen Atlas, Eddy Davelaar, Berkeley Dietvorst, Lin Fei, Alex Imas, Eric Johnson, Antonia Krefeld-Schwalb, Nick Reinholtz, Jack Soll, Christoph Ungemach, Oleg Urminsky, and two anonymous reviewers for their generous and helpful comments on the manuscript. We are also grateful to Oystein Daljord, Lorien Elleman, Lin Fei, Trevor Kvaran, Justin Landy, Sanjog Misra, Chris Nosko, Bill Revelle, and Matt Taddy for their suggestions regarding our analyses. We thank Stephen Atlas, Shane Frederick, Todd Gureckis, Alex Imas, Eric Johnson, Ki Kim, Stephan Meier, Rachel Meng, Oded Netzer, Ben Newell, Daniel Read, Nick Reinholtz, Marc Scholten, Dan Shaffer, Olivier Toubia, Christoph Ungemach, Oleg Urminsky, Liad Weiss, and Gal Zauberman for their suggestions regarding Supplemental Material Section A of the manuscript. Finally, we express our gratitude to Steve Atlas, Sachin Banker, Manel Baucells, Daniel Benjamin, Sudeep Bhatia, Andy Brownback, Eugene Caruso, Gretchen Chapman, Helen Colby, Xianchi Dai, Mette Trier-Daamgard, Stefano DellaVigna, J.P. Dube, Keith Ericson, Bernd Figner, Geoff Fisher, Jeremy Fox, Shane Frederick, Dan Goldstein, Christina Gravert, Adam Harris, Lisheng He, Hal Hershfield, Gunter Hitsch, Alex Imas, Pranav Jindal, Bryan Jones, Michael Kuhn, Robyn LeBoeuf, Selin Malkoc, Samuel M. McClure, Joe McGuire, Stephan Meier, Ashwani Monga, Tim Pleskac, Drazen Prelec, Daniel Read, Crystal Reeck, Stian Reimers, Sally Sadoff, Adam Sanborn, Benjamin Scheibehenne, Marc Scholten, Charlie Sprenger, Neil Stewart, Christoph Ungemach, Elke Weber, Haewon Yoon, Gal Zauberman, and six other experts who wished to remain anonymous for their participation in Study 2 in the paper.

ABSTRACT

A large literature implicates time preference (i.e., how much an outcome retains value as it is delayed) as a predictor of a wide range of behaviors, because most behaviors involve sooner and delayed consequences. We aimed to provide the most comprehensive examination to date of how well laboratory-derived estimates of time preference relate self-reports of 36 behaviors, ranging from retirement savings to flossing, in a test-rest design using a large sample (N = 1,308) and two waves of data collection separated by 4.5 months. Time preference is significantly—albeit modestly—associated with about half of the behaviors; this is true even when controlling for 15 other demographic variables and psychologically-relevant scales. There is substantial variance in the strengths of associations that is not easily explained. Time preference's predictive validity falls in the middle of these 16 possible predictors. Finally, we ask time preference researchers (N = 55) to predict the variation in the relationship between time preference and behaviors, and although they are reasonably well-calibrated, these experts tend to overestimate the predictive power of time preference estimates. We discuss implications of invoking time preference as a predictor and/or determinant of behaviors with delayed consequences in light of our findings.

Keywords: time preference, temporal discounting, intertemporal choice, time discounting, delay discounting.

Most decisions entail tradeoffs between sooner outcomes and later outcomes—from the mundane (i.e., whether to order dessert) to the consequential (i.e., how much to save for retirement). Accordingly, most choice theories either assume or estimate people's time preferences as captured by a temporal discount factor or rate—that is, how much value an outcome retains or loses as it is delayed—to describe or predict these behaviors. A large literature articulates how estimates of time preference can help us understand and predict a variety of behaviors involving delayed consequences (for reviews, see Ericson & Laibson, 2019; Frederick et al., 2002; Read & Read, 2004). Hundreds of papers have explored the association between measures of time preference and behaviors, many of which we note below (Urminsky & Zauberman, 2015).

In this paper, we build on these previous explorations and aim to provide a thorough examination of how well laboratory-derived measures of time preference relate to a wide range of behaviors—by assessing more behaviors, with more covariates, using a test-retest design, and comparing to more benchmarks. We measured time preference in an online survey and, following prior work (e.g., Bradford et al., 2017; Chabris et al., 2008; Reimers et al., 2009), examined the correlations between time preference and people's self-reports of a variety of behaviors, as well as other demographic and psychologically-relevant variables. As a secondary goal, we also assessed how well experts could predict the relationships between time preference and behaviors.

Associations of time preference with behaviors

One major motivation for research on time preference is that it is a potential predictor of a wide range of behaviors. Papers have found significant associations between time preference estimates and behaviors involving delayed consequences, particularly when explaining differences between substance-using and control populations (MacKillop et al., 2011; Reynolds, 2006). Studies with the general population have found that estimates of time preference significantly correlate with smoking (Bradford, 2010; Reynolds et al., 2004; Sutter et al., 2013), alcohol use (Bradford et al., 2017; Sutter et al., 2013; Vuchinich & Simpson, 1998), credit card debt (Bradford et al., 2017), mortgage choice (Atlas et al., 2017), credit scores (Li et al., 2015; Meier & Sprenger, 2012), savings (Angeletos et al., 2001; Bradford et al., 2017; Sutter et al., 2013), educational attainment (Duckworth & Seligman, 2005; Falk et al., 2018; Reed & Martens, 2011), gambling (Dixon et al., 2003; Petry, 2001), exercise (Bradford, 2010; Bradford et al., 2017), and more (Bradford et al., 2017; Chabris et al., 2008; Reimers et al., 2009).

Most behaviors have delayed consequences, but *ex ante*, it is not obvious which behaviors we should expect to correlate most strongly with estimates time preference or why. With the previous significant correlations noted, researchers have also found *non*significant correlations between measures of time preference and behaviors that one might expect to be related to time preference, including dental check-up frequency (Farrell & Fuchs, 1982), dietary behaviors flossing, gambling, percentage of income saved, late credit card payments, wealth (Chabris et al., 2008), efficient energy use, and using sunscreen (Bradford et al., 2017). Similarly, evidence linking time preference to obesity and BMI is mixed (Barlow et al., 2016).

Studies assessing how well time preference predicts behaviors often do not control for other relevant and/or potentially confounding variables. Most studies that control for covariates typically incorporate demographics and a few other variables directly relevant to the target behaviors, as opposed to psychological factors frequently associated with a range of behaviors. A few notable exceptions include studies that controlled for risk and ambiguity aversion (Bradford, 2010; Bradford et al., 2017; Sutter et al., 2013) or cognitive ability (Chabris et al., 2008; Li et al., 2013, 2015).

Overview of studies

We intend to provide a more comprehensive approach to examining the relationship between time preference and behavior on several dimensions. First, we incorporate a wider range of target behaviors with delayed outcomes, adding behaviors based on their variation across a set of 23 differentiating characteristics as assessed in a separate norming study (see Supplemental Material Section A). Second, we collect a large and widely-varied set of additional variables to control for and compare to, including various personality scales and predictors relevant to financial decision-making. Third, with some exceptions (Li et al., 2013, 2015), studies of the correlation between time preference and behaviors do not account for attenuation due to imperfect measurement reliability.[1] We address this by collecting two waves of data, separated

[1] *Attenuation* occurs when correlations are underestimated due to measurement error in one or both underlying variables (Spearman, 1904). Correlations can be *disattenuated* by adjusting for measurement reliability: disattenuated $r_{x,y}=r_{x,y}/\mathrm{sqrt}(r_{x,x}\times r_{y,y})$.

by 4.5 months, to assess the test-retest reliability of time preference, the 36 behaviors, and the 15 covariates. Fourth, we try to understand the considerable heterogeneity that we observe in the relationship between time preference and these 36 behaviors. Fifth, we elicit expert forecasts of these correlations to examine whether the widely-varying magnitudes of correlation that we observe are, in fact, predictable.

We aim to assess how well time preference predicts behavior, and we find that the answer to this question depends on the basis for comparison. We used four benchmarks, comparison: (i) *to zero*: how often does time preference significantly correlate with each of these 36 behaviors? (ii) *across behaviors*: which variables or domains are more correlated with time preference? (iii) *to other predictors*: how well does time preference do in predicting behavior relative to demographics and other psychologically relevant individual differences, and (iv) *to expert forecasts*: can time preference researchers predict the magnitude of time preference's association with these 36 behaviors? We believe that this benchmark is especially important since expert intuitions about these relationships directly influence which behaviors they study to gather empirical and theoretical support for the predictive validity of time preference.

Broadly, we found that although most correlations between time preference and the 36 behaviors are small to moderate (Cohen, 1988), time preference significantly predicted over half of the behaviors, even when controlling for 15 demographic and psychological covariates. Also, correlations between time preference and behaviors are greater when we consider aggregated (multi-behavior) indices of financial, health, and prudential behavior, consistent with prior work (Chabris et al., 2008). At the same time, our data also suggest that aggregating by domain may not be justifiable due to considerable heterogeneity within domain. We found large differences in these correlations across behaviors that are not well-predicted by domain nor explained by ratings on 23 differentiating characteristics that formed the basis for sampling these 36 behaviors (see Supplemental Material Section A).

In addition, while time preference consistently predicts behavior better than some covariates, it is in the middle of the pack of 16 predictors. We observed this result even though we specifically selected these behaviors because they have delayed consequences. Finally, we found mixed evidence on the predictability of the size of these correlations. Experts' average forecasts of these relationships were positively correlated with the actual correlations in Study 1. However, even though the average expert prediction for the correlation between time preference and the 36 behaviors in our study was modest (r = 0.11), experts still tended to overestimate the associations, particularly for behaviors that were not correlated with time preference.

Study 1 examines the first three benchmarks (i.e., comparison to zero, across behaviors, and across predictors). In addition, Study 1 accounts for the issue of test-retest reliability for all the variables measured by incorporating data collected over two waves separated by about 4.5 months. Given the heterogeneity in predictive validity of time preference across the behaviors, we wanted to assess whether such heterogeneity was explainable or expected by experts. In addition, comparison to expert predictions potentially offers a fourth benchmark for these correlations. Study 2 therefore examines how well forecasts of the correlation between time preference and each of 36 behaviors aligned with the heterogeneity in observed correlations across behaviors. We believe that this benchmark is especially important because expert intuitions about these relationships influence which behaviors are studied, which research questions are asked, and how researchers ultimately assess theoretical support for explanatory models involving time preference.

STUDY 1

In Study 1, we conducted a large, two-wave study measuring people's time preference, 36 self-reported behaviors, and various other psychologically-relevant scales. Our intention was to run a large-scale examination of how well time preference predicts these behaviors while controlling for measurement error.

Methods

Transparency and openness. We report all data exclusions (if any) and all measures in the study. All data are publicly available (see author's note). Data were analyzed using R, version 3.6.2 (R Core Team, 2019) and the following packages: *tidyverse* version 1.3.0 (Wickham et al., 2019); *dplyr* version 1.0.7 (Wickham et al., 2021); *afex* version 1.0-1 (Singmann et al., 2021); *reghelper* version 0.3.6 (Hughes, 2020); *Hmisc* version 4.6-0 (Harrell, 2019). This study's design and analyses were not pre-registered. The protocols of Studies 1 and 2 were approved by the institutional review board.

Participants. We recruited 1576 U.S. participants in total from Amazon Mechanical Turk (N = 774) and from a market research firm, PureProfile (N = 802), to complete two waves of the study in December 2013 and April 2014 (see Table 1 for descriptive statistics).[2] Of these 1576 participants, 83% finished both waves for a total of 1308 complete two-wave data points (1308/1576; MTurk = 604/774; commercial panel = 704/802). To test for selective attrition, we ran a logistic regression predicting attrition using the wave 1 observations of the 16 predictor variables in our study (see Supplemental Material Table S1). After a Bonferroni adjustment, only panel source, parent's average education level, and numeracy-CRT predicted participation in wave 2, so results pertaining to these variables should be interpreted with caution. For robustness, we replicate our main analyses using only wave 1 data in Supplemental Material Table S2.

Table 1. Study 1 participant summary statistics

Variable	Mean/%	Median	Std. Dev.	Min	Max
Age	40.92	38	14.64	18	86
Gender	41.51% male	-	-	-	-
Source	46.18% MTurk	-	-	-	-
Pretax Household Income (in USD)	26.76% earning $50,000 or above	$30,000 to $39,999	-	Less than $10,000	More than $250,000
Education	41.97% with at least Bachelor's Degree or Equivalent	Associate's Degree or Equivalent	-	Some high school or less	Doctoral degree (e.g., Ph.D.)

Procedure. Participants completed both waves of the survey online after completing a separate intake survey measuring demographics (age, gender, income, education, native/first language, and zip code). Participants completed the first wave between December 14th, 2013 to January 3rd, 2014 (average date of December 18th) and the second wave between April 22nd and May 12th, 2014 (average date of April 25th), about 4.5 months later (*M* = 128.90 days, *SD* = 3.21, range = 110 to 147 days).

[2] We collected data from two pools in case there were data quality issues with either pool. Notably, some concerns over MTurk data quality have arisen in recent years (e.g., Chmielewski & Kucker, 2019). Our data are of high quality by recent standards: 93.37% of our participants passed the attention checks in both waves of this study.

Table 2. Study 1 12-item battery of intertemporal choices

Smaller-Sooner option	Larger-Later option
$504 now	$524 in one month
$600 now	$611 in one month
$777 now	$791 in one month
$1,064 now	$1,153 in one month
$816 now	$860 in three months
$457 now	$551 in six months
$816 now	$5,440 in one year
$213 now	$281 in two years
$816 in six months	$860 in nine months
$816 in six months	$1,028 in one year
$400 in six months	$440 in one and a half years
$840 in six months	$10,125 in two and a half years
$791 today	$777 in one month
$621 in six months	$670 in six months

Note: The last two (gray-font) items are attention-check foils that were included in wave 2 but not evaluated as part of the measure. The order of all 12/14 items was randomized by participant.

Time preference measure. We developed a 12-item battery of intertemporal choice questions posing smaller-sooner versus larger-later tradeoffs (see Table 2). The measure was developed over the course of several pretests that swapped in different alternatives to test different combinations of smaller-sooner and larger later options, using ideas from item response theory. Our goal was to distinguish people who are more patient from those who are less patient, rather than to estimate a specific discount rate, discount factor, or estimate of present bias. (Full details about development of the battery are available in Supplemental Material Section D.) We also recognize that measures like this one (and almost all others using smaller, sooner vs. larger, later monetary tasks) are affected by more than just patience, including many economic considerations, like how much money a person has now and will have in the future relative to current and future needs, their declining marginal utility for money, trust that the later payment will occur and myriad other uncertainties regarding the future, like guesses about inflation, etc. We estimated participants' time preference by simply counting the number of larger-later choice options chosen among the 12 items. Figure 1 presents a histogram of the number of larger-later options chosen. Participants chose an average of 5.25 later options (*SD* = 3.01, *median* = 4.5).

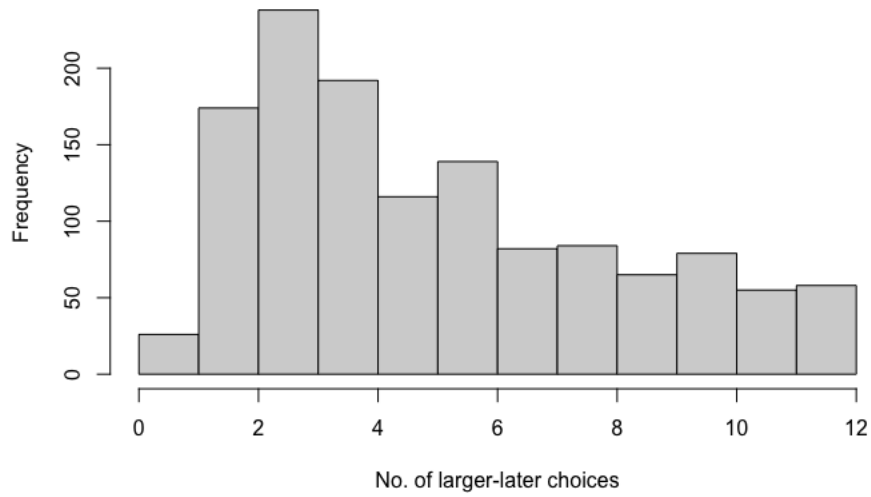


***Figure 1*. Distribution of responses on the 12-item time preference measure.**

Of course, no measure of time preference uses every possible monetary value and delay, and it is possible that the relationship between estimates of time preference and behaviors could be affected by these parameters.[3] For this reason and others, we also collected a second, prominent measure of time preference as a benchmark, DEEP Time (Toubia et al., 2013), an adaptive measure that estimates parameters for people's long term discount factor (δ) and present bias parameter (β) as stipulated by the quasi-hyperbolic discounting model (Laibson, 1997). This measure uses smaller amounts of money and shorter delays than our 12-item measure. We chose to focus on the 12-item measure because (i) it makes no parametric assumptions, (ii) is easier to implement and compute for researchers, (iii) because we did not have hypotheses about how these behaviors would be differentially predicted by β versus δ, and (iv) because the number of larger, later choices on our 12-item measure had higher test-retest reliability than did parameter estimates from the DEEP method ($r_{\text{12-item}} = .70$ vs. $r_\beta = .63$ and $r_\delta = .63$, $z\text{s} = 3.37$, $p < .001$). Note that the 12-item measure correlates highly with the DEEP Time parameter estimates for δ ($r = 0.76$) but less well with β ($r = 0.37$).

Behaviors. We asked participants to self-report the extent to which they do 36 behaviors (or about outcomes associated with those behaviors). We report response scales for each behavior in Supplemental Material Section C and descriptive statistics in Supplemental Material

[3] We thank an anonymous reviewer for raising this point.

Table S3. In addition, an early goal of the project was to sample behaviors that would theoretically be highly differentiated across the 23 characteristics that a separate set of participants would rate for each behavior (see Supplemental Material Section A). This theoretically-driven sampling procedure, which includes suggestions from several time preference researchers and some of the papers cited earlier, resulted in introducing measures for some additional behaviors. See Supplemental Material Section C for the wordings, scaling, and sources for all measures.

We scored behaviors so that larger values indicate what participants in a separate norming study judged to be more far-sighted, prudent, or responsible behavior (see Supplemental Material Section E). This required reverse-scaling responses for missing credit payments, accumulating a lot of credit card debt, accumulating a lot of educational loan debt, smoking, gaining excess weight, using recreational drugs, drinking coffee, overeating, drinking alcoholic beverages, getting tattoos, gambling or buying lottery tickets, leaving dirty dishes overnight, and driving recklessly. In the tables and figures in this paper, the word "NOT" precedes these behaviors to indicate reverse scaling. In addition, for behaviors with highly skewed distributions of responses, we log-transformed the responses for both waves and averaged these log-transformed scores for: accumulating a lot of credit card debt, having kids when older, getting tattoos, earning a large income, keeping physically active, and actively exercising.

Other variables. We included 15 covariates to serve as our third benchmark. For demographics, we measured age, gender, parents' education levels, and a sample (MTurk vs. market research panel) variable to control for other unmeasured differences. Own education level and income were 2 of the 36 behavioral dependent variables. We averaged parents' education levels into a single measure for ease of reporting. We measured five dimensions of personality using the 44-item Big Five Inventory (John & Srivastava, 1999), and assessed six other scales that have been related to temporal discounting and financial decision making: (i) impulsiveness, as measured by the Barratt Impulsiveness Scale (Patton et al., 1996); (ii) financial literacy (Fernandes et al., 2014); (iii) a combined 8-item (Li et al., 2013) questionnaire measuring Numeracy (Lipkus et al., 2001) and Cognitive Reflection (Frederick, 2005), which were correlated 0.87 and thus collapsed into a single Numeracy/CRT scale; (iv) tightwad-spendthrift tendency—whether people typically spend more/less than they would like to (Rick et al., 2008); (v) planning propensity—people's propensity to plan out the use of their money over the next few days and next 1-2 months (Lynch et al., 2010); and (vi) risk preference, as assessed by the single-item Eckel-Grossman measure (2002). Controlling for the 15 covariates also allows us to assess the unique association time preference has with behavior.

Results

Benchmark One: Comparison of predictive validity of time preference against zero

Figure 2 and Table 3 show the degree of association between time preference and each of our variables using five different methods. All methods used data from the participants who completed both waves of data collection, and the association between time preference and each of our variables are depicted graphically in Supplemental Material Figure S1. The first column of Table 3 reports the Pearson correlations between our measure of time preference (higher scores = more patience/less discounting) and the 36 self-reported behaviors, with each variable averaged across two waves of data collection.

In column two, we report the range of correlations we obtained from the four combinations of two waves of data (i.e., wave 1/2 time preferences × wave 1/2 behaviors). In the same column, we also report, in parentheses, the number of times the relationship between time preference and the behaviors was statistically significant (at the $p < .05$ level).

The third column reports the corresponding disattenuated correlations (Spearman, 1904). This procedure is a correction for correlations that accounts for the imperfect test-retest reliability of our measures.[4] To do this, we used the reliabilities (see Table 4) for each variable and divided the correlation coefficient by the product of the square root of the product of the reliabilities of the relevant variables.

The fourth column reports the standardized coefficients (i.e., betas) for time preference from 36 separate regressions predicting each behavior as a function of time preference while

[4] We also calculated ORIV (Obviously Related Instrumental Variables) correlations and multiple regression coefficients, which is an econometric method for adjusting for attenuation (Gillen et al., 2019). See Supplemental Material Table S4 for these results. One limitation of such diattenuation-based approaches is they are unable to assess the relative influence of preference fluctuations versus measurement noise. Our analyses reveal consistent results when performed over data from the first wave only (see Table S2), suggesting that this limitation of the method may not be a pressing concern in our data.

controlling for the 15 covariates (i.e., the unique associations between time preference and behavior). Column five presents the range of standardized coefficients, analogous to column two.

Comparing across the five columns, whether or not a relationship reaches statistical significance is pretty consistent across the various ways we analyze the data. As depicted in Figure 2, the size of the various coefficients is similar across these different specifications. As another robustness check, we also present correlations and standardized coefficients for analyses using rank-ordered data in Supplemental Material Table S2.

The correlation with time preference was significant at $p < .05$ for 20 of the 36 behaviors ($p < .01$ for 17 behaviors; $p < .001$ for 12 behaviors). Similarly, in regressions with all 15 covariates, time preference emerged as a positive and significant predictor at $p < .05$ for 18 of the 36 behaviors ($p < .01$ for 8 behaviors; $p < .001$ for 5 behaviors). Thus, even when controlling for all other covariates, time preference was a significant predictor for over half of the behaviors. However, most of the correlation coefficients between time preference and the behaviors were moderate or small in size. The 25th, 50th, and 75th percentiles of absolute correlation coefficients were .03, .07, and .12 (.05, .10, and .16 for the corresponding absolute disattenuated correlations; .02, .06, and .07 for the median absolute standard betas).[5]

So, by the first benchmark, time preference predicts many behaviors significantly, but maybe not as strongly as some might expect it to. We also found that the inference one reaches about the relative size of each correlation does not vary much across different analyses and is relatively robust to issues related to measurement error.[6]

---

[5] We thank an anonymous reviewer for suggesting an additional regression specification controlling for demographic (age, gender, and income). Supplementary Materials Table B3 presents these analyses. While some associations were weaker when controlling for demographics (e.g., getting routine physicals, accumulating wealth, education level, having kids when older), none of these differences were significant.

[6] Supplementary Materials Table B4 presents the association between these behaviors and the DEEP measures of time preference. Overall, the results are highly consistent between the 12-item measure and the δ parameter from DEEP. In particular, both the measures significantly predict a largely overlapping set of behaviors. On the other hand, results with the β parameter were less consistent. This is reasonable, since the 12-item measure was designed as a measure to discriminate among people with generally different time preferences (which correspond better to δ), whereas the purpose of the β parameter in quasi-hyperbolic models is to characterize a discontinuity—a change in time preferences for different (i.e., now vs. not-now) time periods.

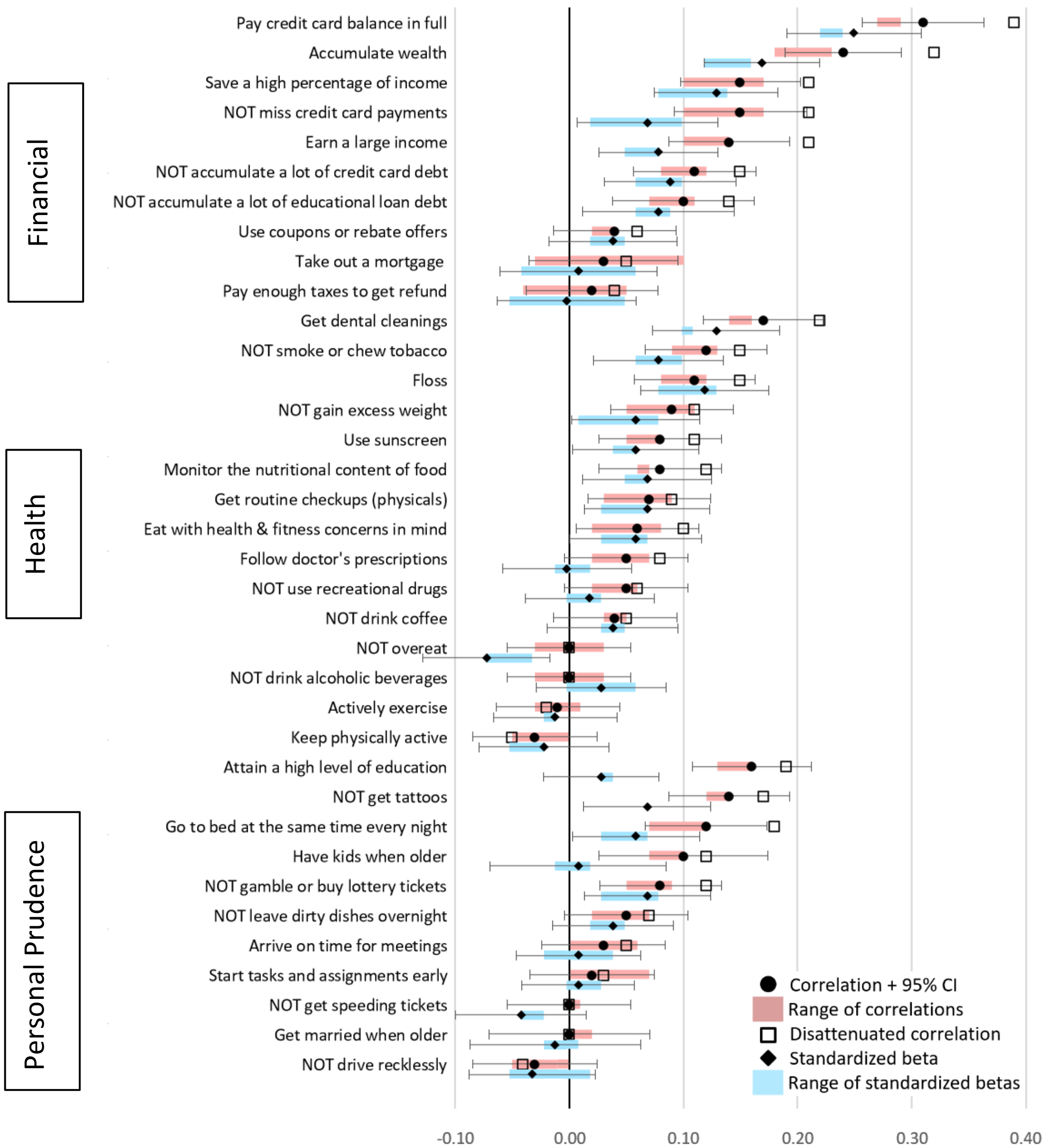


Figure 2. **Graphical depiction of associations.** Correlations are plotted as circles with 95% confidence intervals, while disattenuated correlations are plotted as squares. Standardized betas are plotted as diamonds with 95% confidence intervals. The pink and blue shaded bars represent the range of four possible correlations/standardized betas calculable using our two observations of each variable (wave 1-wave 1, 1-2, 2-1, and 2-2). See also Table 3.

Table 3. Study 1 measures of association between time preference and 36 self-reported behaviors. *** $p < .001$, ** $p < .01$, * $p < .05$. 1) Pearson correlations between our measure of time preference (higher scores = more patience/less discounting) and the 36 self-reported behaviors, with each variable averaged across two waves of data collection. 2) Range of correlations obtained from the four combinations of two waves of data (i.e., wave 1/2 time preferences × wave 1/2 behaviors) and, in parentheses, the number of times the relationship between time preference and the behaviors was statistically significant (at the $p < .05$ level). 3) Disattenuated correlations (Spearman, 1904), which account for the imperfect test-retest reliability of our measures (from Table 4), equal to the correlation coefficient divided by the product of the square root of the reliabilities of the relevant variables. 4) The standardized coefficients (i.e., betas) for time preference from 36 separate regressions predicting each behavior as a function of time preference while controlling for the 15 covariates. 5) The range of standardized coefficients, analogous to column two.

Domain	Behaviors	Test-retest reliability	Correlation	Across 4 correlation combinations	Disattenuated correlation	Standardized betas (controlling for 15 covariates)	Across 4 regression combinations
FINANCIAL	Pay credit card balance in full	0.86	0.31***	0.27 to 0.29 (4)	0.39	0.25***	0.22 to 0.24 (4)
	Accumulate wealth	0.77	0.24***	0.18 to 0.23 (4)	0.32	0.17***	0.12 to 0.16 (4)
	Save a high percentage of income	0.76	0.15***	0.10 to 0.17 (4)	0.21	0.13***	0.08 to 0.14 (4)
	NOT miss credit card payments	0.71	0.15***	0.10 to 0.17 (4)	0.21	0.07*	0.02 to 0.10 (2)
	Earn a large income	0.65	0.14***	0.10 to 0.14 (4)	0.21	0.08**	0.05 to 0.08 (3)
	NOT accumulate a lot of credit card debt	0.78	0.11***	0.08 to 0.12 (4)	0.15	0.09**	0.06 to 0.10 (3)
	NOT accumulate a lot of educational loan debt	0.78	0.10**	0.07 to 0.11 (3)	0.14	0.08*	0.04 to 0.09 (1)
	Use coupons or rebate offers	0.75	0.04	0.02 to 0.04 (0)	0.06	0.04	0.02 to 0.05 (0)
	Take out a mortgage	0.70	0.03	-0.03 to 0.10 (1)	0.05	0.01	-0.04 to 0.06 (0)
	Pay enough taxes to get refund	0.51	0.02	-0.04 to 0.05 (0)	0.04	0	-0.05 to 0.05 (0)
HEALTH	Get dental cleanings	0.84	0.17***	0.14 to 0.16 (4)	0.22	0.13***	0.10 to 0.11 (4)
	NOT smoke or chew tobacco	0.90	0.12***	0.09 to 0.13 (4)	0.15	0.08**	0.06 to 0.10 (4)
	Floss	0.82	0.11***	0.08 to 0.12 (4)	0.15	0.12***	0.08 to 0.13 (4)
	NOT gain excess weight	0.96	0.09**	0.05 to 0.11 (2)	0.11	0.06*	0.01 to 0.08 (2)
	Use sunscreen	0.75	0.08**	0.06 to 0.08 (4)	0.11	0.06*	0.04 to 0.06 (2)
	Monitor the nutritional content of food	0.62	0.08**	0.06 to 0.07 (4)	0.12	0.07*	0.05 to 0.07 (1)
	Get routine checkups (physicals)	0.78	0.07*	0.03 to 0.09 (3)	0.09	0.07*	0.03 to 0.07 (2)
	Eat with health & fitness concerns in mind	0.52	0.06*	0.02 to 0.08 (2)	0.1	0.06*	0.03 to 0.07 (2)
	Follow doctor's prescriptions	0.58	0.05	0.02 to 0.07 (1)	0.08	0	-0.01 to 0.02 (0)
	NOT use recreational drugs	0.87	0.05	0.02 to 0.06 (1)	0.06	0.02	0.00 to 0.03 (0)
	NOT drink coffee	0.89	0.04	0.03 to 0.05 (0)	0.05	0.04	0.03 to 0.05 (0)
	NOT overeat	0.63	0	-0.03 to 0.03 (0)	0.00	-0.07*	-0.07 to -0.03 (2)
	NOT drink alcoholic beverages	0.84	0	-0.03 to 0.03 (0)	0.00	0.03	0.00 to 0.06 (1)
	Actively exercise	0.71	-0.01	-0.03 to 0.01 (0)	-0.02	-0.01	-0.02 to -0.01 (0)
	Keep physically active	0.59	-0.03	-0.05 to 0.00 (0)	-0.05	-0.02	-0.05 to 0.02 (0)
PERSONAL	Attain a high level of education	0.96	0.16***	0.13 to 0.16 (4)	0.19	0.03	0.03 to 0.04 (0)
	NOT get tattoos	0.97	0.14***	0.12 to 0.14 (4)	0.17	0.07*	0.06 to 0.06 (4)
	Go to bed at the same time every night	0.65	0.12***	0.07 to 0.12 (4)	0.18	0.06*	0.03 to 0.07 (2)
	Have kids when older	0.97	0.10*	0.07 to 0.10 (3)	0.12	0.01	-0.01 to 0.02 (0)
	NOT gamble or buy lottery tickets	0.71	0.08**	0.05 to 0.09 (3)	0.12	0.07*	0.03 to 0.08 (2)
	NOT leave dirty dishes overnight	0.77	0.05	0.02 to 0.07 (2)	0.07	0.04	0.02 to 0.05 (0)
	Arrive on time for meetings	0.66	0.03	0.00 to 0.06 (1)	0.05	0.01	-0.02 to 0.04 (0)
	Start tasks and assignments early	0.61	0.02	-0.01 to 0.06 (1)	0.03	0.01	0.00 to 0.03 (0)
	NOT get speeding tickets	0.71	0	0.00 to 0.01 (0)	0	-0.04	-0.04 to -0.02 (0)
	Get married when older	0.93	0	0.00 to 0.02 (0)	0	-0.01	-0.02 to 0.01 (0)
	NOT drive recklessly	0.64	-0.03	-0.05 to 0.01 (0)	-0.04	-0.03	-0.05 to 0.02 (0)

Table 4. Study 1 test-retest reliabilities and correlations between predictors (each predictor averaged across two waves). *** $p < .001$, ** $p < .01$, * $p < .05$.

	Time preference	**Age**	**Parent Education**	**Gender (1 = Male)**	**Source (1 = MTurk)**	**Extraversion**	**Conscientiousness**	**Openness**	**Agreeableness**	**Neuroticism**	**Barratt Impulsiveness Scale**	**Financial Literacy**	**Numeracy-CRT**	**Tightwad-Spendthrift**	**Propensity to Plan**	**Risk Preference**
Test-Retest Reliability	0.70***	0.99***	0.94***	1.00***	1.00***	0.90***	0.86***	0.88***	0.87***	0.89***	0.83***	0.82***	0.81***	0.80***	0.75***	0.37***
Age	0.09**															
Parent Education	0.10***	-0.29***														
Gender (1=Male)	0.07**	0.02	0.02													
Source (1=MTurk)	-0.03	-0.44***	0.13***	0.13***												
Extraversion	-0.05	0.14***	-0.02	-0.06*	-0.27***											
Conscientiousness	0.04	0.25***	-0.10***	-0.04	-0.17***	0.31***										
Openness	-0.02	-0.07*	0.10***	0.04	0.08**	0.31***	0.19***									
Agreeableness	-0.02	0.20***	-0.09**	-0.11***	-0.19***	0.32***	0.45***	0.12***								
Neuroticism	-0.06*	-0.16***	0	-0.20***	0.03	-0.39***	-0.50***	-0.16***	-0.48***							
Barratt Impulsiveness Scale	-0.16***	-0.12***	-0.03	-0.08**	-0.01	-0.16***	-0.68***	-0.21***	-0.35***	0.49***						
Financial Literacy	0.25***	0.17***	0.15***	0.30***	0.21***	-0.05	0.05	0.05*	-0.12***	-0.13***	-0.21***					
Numeracy-CRT	0.21***	-0.19***	0.24***	0.30***	0.39***	-0.17***	-0.13***	0.06*	-0.22***	-0.03	-0.09**	0.57***				
Tightwad-Spendthrift	-0.20***	0	-0.02	-0.09**	-0.16***	0.18***	-0.17***	0	0	0.10***	0.38***	-0.14***	-0.14***			
Propensity to Plan	0.02	-0.04	0.01	-0.06*	0.02	0.19***	0.32***	0.22***	0.21***	-0.18***	-0.37***	-0.01	-0.08**	-0.19***		
Risk Preference	-0.07*	0	0.04	0.07**	-0.05	0.08**	0.01	0.04	0.05	-0.06*	0.01	-0.12***	-0.14***	0.08**	0.02	

Benchmark Two: Comparisons across behaviors

Following previous research (Bradford et al., 2017; Chabris et al., 2008), we classified the 36 behaviors into three domains: Financial, Health, and Personal Prudence (see Table 3). We created composite indices for each domain by (i) z-scoring across participants for each behavior and (ii) computing each participant's average z-score within each domain. In line with previous research, we found that time preference was a better predictor of these aggregate indices of behavior than it was for most individual behaviors (financial: $r(1306) = 0.27$, health: $r(1306) = 0.16$, personal: $r(1306) = 0.15$, ps < .001; the corresponding standardized betas from regressions on the these indices, using time preference and the other 15 covariates as predictors were 0.17, 0.10, and 0.06, ps < .05). Time preference was a better predictor of the financial index than the health and personal prudence indices (ts > 2.80, ps < .005), while being similarity predictive of those latter two ($t = -0.12$, $p = .90$). This finding is perhaps unsurprising because, following the prior literature, we measured time preference by offering people sooner and later monetary rewards. See Supplemental Material Section B for a discussion on the extent to which collapsing these behaviors into domains for the purpose of assessing their relationship with time preference might or might not be informative.

To examine differences in the predictive validity of time preference across domains, we ran three ANOVAs to compare each domains' (i) correlation coefficients, (ii) disattenuated correlation coefficients, and (iii) standardized regression coefficients, finding significant differences in all three specifications (Fs(2,35) = 3.73, 4.30, and 4.48; ps < .05). In particular, the associations with financial behaviors were greater than those for both health-related and personal behaviors. At the same time, these results were driven almost entirely by the large correlation for "pay credit card balance in full", which is an outlier (it is 2.05 interquartile ranges above the 75th percentile and 3.12 standard deviations above the average correlation)(Tukey, 1977). However, there was a large variation in predictive validity of time preference in each domain and participants' responses for behaviors within a given domain were only modestly correlated (average correlations for financial, health, and prudential behaviors were 0.09, 0.07, and 0.08).

And finally, an early goal of this project was to explain heterogeneity in these 36 correlations through moderation. In a separate norming study (reported in Supplemental Material Section A), we had participants rate each of these 36 behaviors on 23 differentiating characteristics. We then used the mean facet ratings for each behavior to predict the measures of association presented in Figure 2. Although this initial analysis yielded a few significant results, indicating *some* ability to predict what characteristics of behaviors made them more associated with time preference, further analyses of these relationships, unfortunately, reveals moderations that are difficult to interpret theoretically.

Thus, the predictive validity of time preference varies widely across the 36 behaviors in ways that we did not predict and, we suspect, might be difficult to predict. The estimates between time preference and behaviors were somewhat larger on average for the financial domain, than for the other two domains. This domain difference may be attributable to the fact that our measure of time preference, like most measures in the field, involve monetary outcomes. Aggregating behaviors into indices based on domains improved time preference's predictive validity, replicating previous research (Bradford et al., 2017; Chabris et al., 2008). That said, the intercorrelations of behaviors within a domain are small, suggesting that categorizing these behaviors by domain might not be justifiable. So, on our second benchmark, there is a lot of heterogeneity in correlation: It seems like time preference is a reasonably strong predictor of a few of the behaviors, but the reasons why it predicts well in some cases and not others have proven elusive.

Benchmark Three: Comparison of time preference with other covariates

We next assessed how well time preference and 15 covariates each predicted behaviors in two ways: (i) We counted how many times each variable was significant (at $p < .05$ level) across the 36 regressions reported in column 7 of Table 3; and (ii) We calculated the median absolute standardized betas across these regressions (see Supplemental Material Table S5). By both metrics, time preference ranks near the middle of our 16 predictor variables in terms of predictive validity. Supplemental Material Table S5 reports results from different multiple regression specifications for robustness, with similar results.

Note that we sampled these behaviors because they entail clear delayed consequences. So, we expected time preference to outperform other predictors on theoretical grounds (in aggregate), but its performance places it in the middle of the pack of predictors. Based on the number of times each variable was a significant predictor of behavior, time preference was outperformed by age, parent education, extraversion, and Barratt Impulsiveness scale. Based on median absolute standardized coefficients, age, extraversion, conscientiousness, Barratt

Impulsiveness Scale, financial literacy, and numeracy-CRT were more strongly associated with the 36 behaviors than time preference. If our only goal was to predict these 36 behaviors, we would have been better served to measure Barratt Impulsiveness, Extraversion, or age rather than to measure time preference.

Discussion

Based on Study 1's findings, we set out to examine whether the magnitudes of these correlations between time preference and these 36 behaviors were, in fact, predictable. In particular, we sought to elicit expert intuitions about these relationships because: (i) some of the correlations were smaller than we expected and, (ii) a clear explanation for the large heterogeneity in time preference's predictive validity across behaviors proved elusive to us. To that end, we ask researchers with expertise on time preference to forecast the size of these correlations in Study 2. Obtaining expert's intuitions about these associations also allows us to compare the observed correlations to a fourth yardstick which might be more reasonable, for some purposes, than a null hypothesis of zero correlation.

STUDY 2

For our fourth benchmark, we asked a group of experts to forecast the 36 correlations between time preference and self-reported behaviors from Study 1. Our motivation behind eliciting expert predictions is as follows: Many papers on time preference (including those we have authored) operate on the assumption that time preference is likely implicated in behaviors with delayed consequences and should therefore predict such behaviors. If we find that experts forecast these correlations to be larger than we found or poorly predict the relative magnitudes of these correlations across behaviors, we will have uncovered gaps in our understanding. Obtaining expert forecasts also helps us calibrate on whether our own surprise at many of the small correlations observed in Study 1 was idiosyncratic. More broadly, Study 2 was intended to capture expert intuitions about the relationship between time preference and behaviors, as these researchers have been exposed to a wealth of relevant data. In addition, expert intuitions guide which research questions and domains receive attention in time preference research.

Methods

Transparency and openness. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. All data are publicly available (see author's note). Data were analyzed using *R* version 3.6.2 (R Core Team, 2019) and the following packages: *tidyverse* version 1.3.0 (Wickham et al., 2019); *dplyr* version 1.0.7 (Wickham et al., 2021); *data.table* version 1.14.2 (Dowle & Srinivasan, 2021). This study's design and analyses were not pre-registered.

Participants. We aimed to recruit at least 40 academic researchers with expertise on time preference. We initially sent email invitations to 46 experts and then used snowball sampling (common in anthropology, criminology, and other fields), where we asked those experts to nominate other experts. We invited these additional 68 nominated experts in the same fashion as the original 46. Each expert received up to two reminders, at approximately 10 days and 20 days after the initial invitation. We stopped data collection after 90 days, when completion rates had slowed to a trickle, resulting in a total of 55 complete responses. Each expert received $100 for their participation. As an incentive for accuracy, we also donated $500 to the chosen charity of the expert who provided the best estimates (using the scoring criteria outlined below).

Procedure. After consenting to participate, these experts forecasted a series of correlations—between our measure of time preference and responses on each of 36 self-reported behaviors. Assessing the association between variables by directly asking experts for correlations is simpler, more accurate, and more consistent than asking for other measures of association (Clemen et al., 2000). We informed them that we ran a survey assessing these correlations and presented demographic information. We introduced them to the measure of time preference and explained that we scored it by simply counting the number of larger, later choices. We then presented them with a list of all 36 behaviors. We also reminded them of the scoring criteria for determining the forecasting competition winner (which had been specified to them in the invitation email). We would rank participants on two criteria: (i) the correlation between their predictions and the observed correlations in Study 1 in descending order, and (ii) the mean absolute deviation between their predictions and the actual correlations in ascending order. The winner would be the participant with the smallest sum of these two ranks.

For the main task, experts sequentially viewed the original wording and possible responses for each of the 36 behavior questions in Study 1 in a randomized order. We also informed them of how each response was scored, whether the scores were log-scaled, and whether and how multiple questions were aggregated into a single measure. We asked experts to provide their best estimate of the correlation between our time preference measure and each of these 36 questions using a slider ranging from -1 to 1 (with 0.01 precision). On each screen, we reminded them that a higher score on the measure means more larger, later choices. Next, they answered questions about themselves: department(s) they held appointments in, knowledge of topics related to time preference, number of years spent studying related topics, and number of projects on related topics. We then asked, "How much confidence do you have in the estimates you made in this survey?" on a 5-point scale (1 = *None*, 5 = *A Lot*). Finally, we requested suggestions for other experts in time preference who would be good candidates for participating in the study.

Consensus Analysis. To ascertain whether there was enough consensus among experts to analyze their data in terms of one aggregate group (versus separate subgroups of experts), we performed a cultural consensus analysis (Romney et al., 1987). To do so, we ran a principal component analysis (PCA) across experts, using each expert's 36 forecasts. The PCA produced first and second eigenvalues of 15.58 and 4.48, with a ratio of 3.48, exceeds the threshold of 3 recommended to establish consensus (Weller, 2007). Hence, we proceeded with our aggregate analyses.

Results

Benchmark Four: Comparison to Expert Predictions

The 55 experts held academic appointments in departments including marketing (21), psychology (15), economics (11), decision science (8), organizational behavior/management (4), behavioral science (2), and a few others (4). More than half (29) have studied intertemporal choice and related topics for over 10 years and all but one have studied it for at least 3-5 years. The median expert had 3-5 projects (published and unpublished) in the area, with 12 reporting 5-10 projects and 12 reporting more than 10 projects. Despite their experience in the topic, nearly all experts reported "some" (26) or "a little" (25) confidence in their estimates, while four reported "a fair amount" and none reported "a lot" of confidence.

Aggregated Experts

Table 5 presents summary statistics for the expert predictions, their differences from the observed correlations in Study 1, and tests comparing expert predictions with the observed correlations. (Supplemental Material Table S6 reports analyses using the disattenuated correlations, with similar results.) The average expert prediction for the correlation between time preference and the 36 behaviors in our study was 0.11 (range from -0.04 to 0.27), although the observed average correlation was even smaller (0.08, range from -0.03 to 0.31).

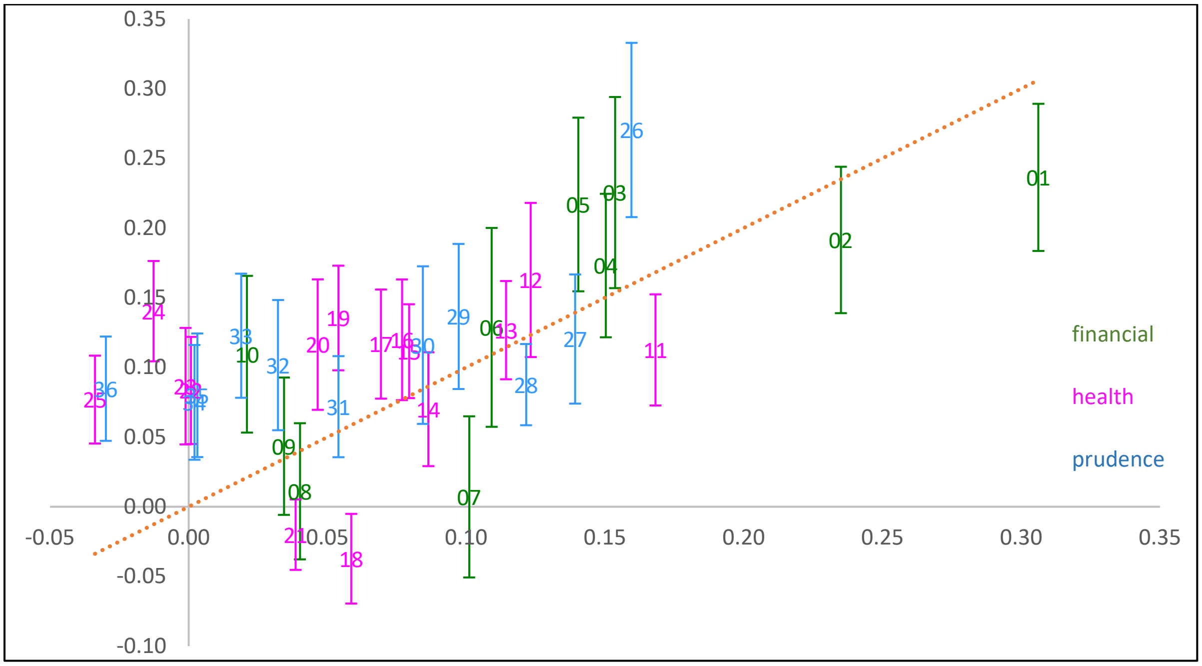


Figure 3. Scatterplot of expert forecasts (on the y-axis) versus the correlations observed in Study 1 (on the x-axis). Each point corresponds to 1 of 36 behaviors (numbers for each behavior refer to the order in which they are listed in Table 5) and error bars represent 95% confidence intervals. The 95% confidence intervals that fall above (below) the dashed line indicated that experts forecasted a higher (lower) degree of positive correlation than was observed in Study 1. Financial behaviors are presented in green, health behaviors in pink, and personal prudence behaviors in blue.

In general, experts overestimated the correlations between time preference and the behaviors. The average expert forecast significantly (at the $p < .05$ level) overestimated the correlations between time preference and behaviors for 16 of the behaviors and underestimated for 6. The average forecast was not significantly different from the observed correlations for the remaining 14 behaviors (although 10 of those were directionally overestimated). Figure 3 plots the relationship between the average forecasts and observed correlations. And at the individual level, 19 of 55 experts made forecasts that on average significantly overestimated the correlations, while 11 significantly underestimated.[7]

Along with some evidence for overprediction, we also found that the aggregate predictions were well-calibrated—the aggregated experts can predict which correlations are larger and smaller. Figure 3 depicts the high overall degree of correlation between the average expert forecasts and observed correlations across the 36 behaviors ($r = 0.60$, $t(34) = 4.36$, $p < .001$). Results using the median expert forecasts were similar ($r = 0.58$). However, it is important to note that the fact that average predicted were well-calibrated benefits from averaging over experts who did not always agree with each other (i.e., wisdom of the crowds).

Individual Experts

We next consider expert forecasts at the individual level. Figure 4 presents the distribution of correlations between each expert's forecasts and the observed correlations. The experts were mostly positively correlated between their forecasts and the observed correlations in Study 1 ($M = 0.28$, $SD = 0.22$, range = -.42 to .69).[8] Twenty-two experts (40%) made forecasts that were significantly correlated with the observed correlations (threshold for significance at $r = 0.33$). Consistent with wisdom of the crowds, only one individual expert made predictions that were more accurate than the aggregated expert predictions.

[7] Based on an anonymous reviewer's suggestion, we also examined whether there were any differences in the predictive accuracy of experts across the three different domains of behavior. The degree of correlation between the average expert forecasts and observed correlations was indeed highest for the 10 financial behaviors ($r = 0.75$), was reasonably high for the 11 prudential behaviors ($r = 0.64$), but was much lower for the 15 health behaviors ($r = 0.23$). However, due to the small sample sizes, no differences were significant ($p = .12$ for the difference between financial and health behaviors). In terms of mean absolute differences between predicted and observed correlations (MAD), the experts did not make better forecasts for the financial-domain behaviors than for other domains ($MAD_{financial} = 0.16$, $MAD_{health} = 0.12$, $MAD_{prudential} = 0.14$).

[8] Upon examination, we found that 2 of our 55 experts reported negative correlations for over half the behaviors. This raises the possibility that they may have miscomprehended the direction of our time preference measure. If we exclude these two experts, we find an average correlation of 0.30 between expert forecasts and absolute correlations, and a Mean Absolute Deviation (MAD) of 0.13.

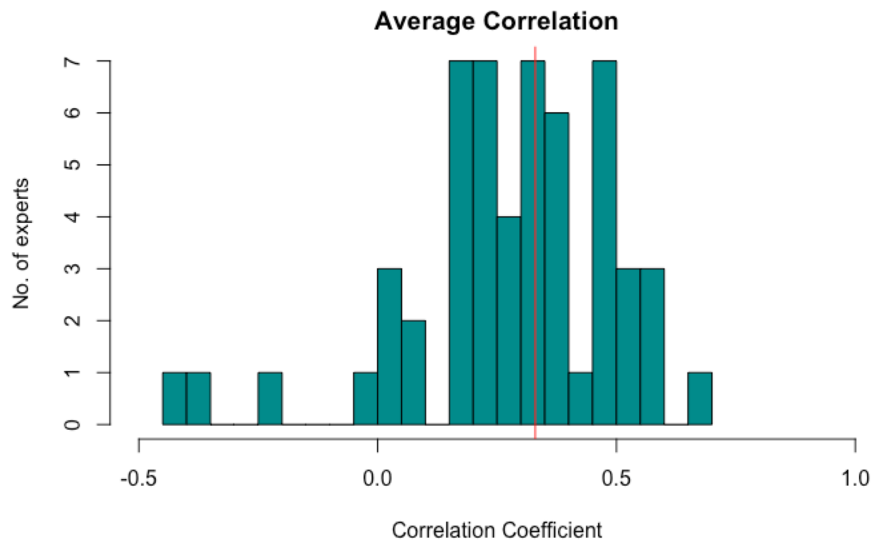


Figure 4. Expert study distribution of average correlations between forecasted correlations and observed correlations for the 36 experts. The red line at 0.33 denotes average correlations beyond the red line show estimates significantly greater than observed correlations.

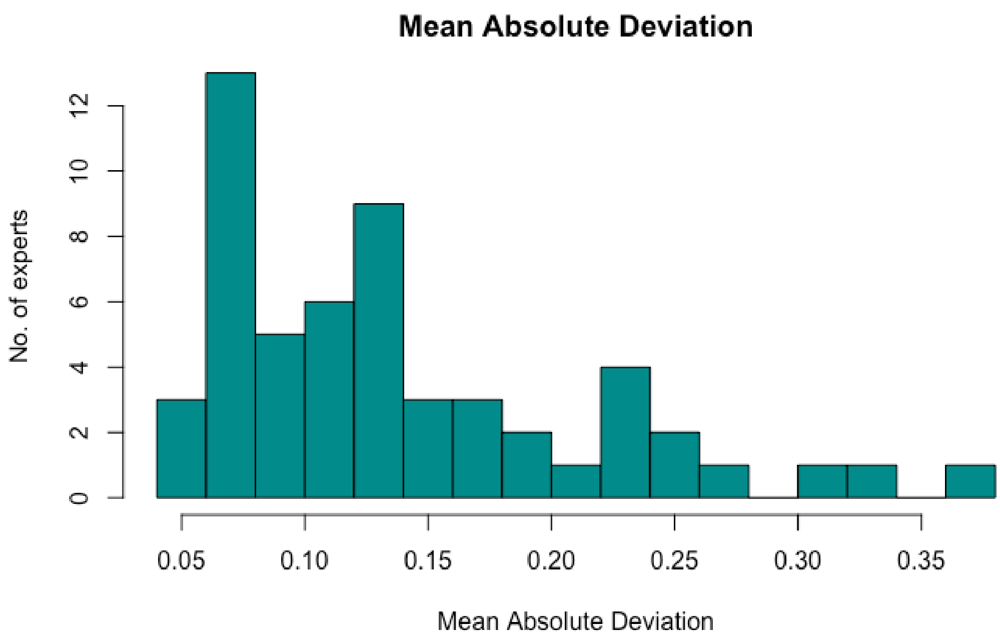


Figure 5. Study 2 distribution of the Mean Absolute Deviations between forecasted and observed correlations for the 55 experts.

We also examined the mean absolute deviation (MAD) between each expert's forecasts and the observed correlations (see Figure 5 and the last column of Table 1). The average MAD was 0.14 (*SD* = 0.08), which we interpret as small in an absolute sense but large in a relative sense. For comparison, 29 of the 36 behaviors had absolute observed correlations smaller than 0.14.

Finally, we ran a simple regression predicting expert performance as a function of their background and expertise, but no variable (including department, self-assessed knowledge, years studying time preference, number of projects, and forecast confidence) predicted either the correlation between their forecasts and observed forecasts or their MAD.

Discussion

In Study 2, we aimed to assess the predictive validity of time preference against our fourth benchmark: expert forecasts. We asked whether experts could predict the large heterogeneity in the association between time preference and behaviors that we observed in Study 1. The results paint a mixed picture of whether time preference researchers can determine which real-world intertemporal choice behaviors can be predicted using time preference. On the one hand, the average of experts' predictions tracked quite well which behaviors were more and less strongly correlated with time preference. Furthermore, 22 of 55 experts' predictions were significantly correlated with the observed correlations in Study 1 at $r > .33$.

On the other hand, the average expert significantly over-predicted nearly half of the correlations, suggesting that time preference researchers often believe that time preference is more predictive than we found in Study 1. Experts expected time preference to predict behaviors more consistently than it actually did, missing many of the behaviors that were not predicted well. Over-predictions tended to occur for behaviors that were not highly correlated with time preference—the 10 lowest observed correlations were all significantly over-predicted.

Despite their tendency to over-predict, our experts were not very optimistic about time preference's predictive validity, with an average predicted correlation of only 0.11 (which made the degree of over-prediction small, on average). So, experts expressed a lack of optimism regarding time preference's predictive validity along with an explicit lack of confidence in their forecasts. Some of this pessimism could be due to unfamiliarity with our novel time preference measure.

Our results suggest that we time preference researchers have more to learn about whether, when, and why time preference is relevant for a given behavior. That is, for a randomly-selected behavior, we are not especially likely to know how well time preference will relate to that behavior. We therefore suggest exercising caution in generalizing our findings to other domains on the assumption that they will generalize without actually studying those domains.

On the whole, our findings suggest that: (i) Experts are good, in the aggregate, and most are individually, at predicting the relative size of these relationships. (ii) When experts err, their forecasts are more often overpredictions—predicting a higher degree of positive association between time preference and behavior than the observed correlation. (iii) There was no detectable relationship between variables that might indicate level of expertise and good forecasting (as in DellaVigna & Pope, 2018). We also did not find a strong relationship between accuracy and confidence (as in Tsai et al., 2008).

Table 5. Expert predictions versus observed correlations in Study 1.

Domain	#	Behavior	Mean Prediction	Median Prediction	Observed Correlation	Mean Difference	*t*-statistic	Mean |Difference| (MAD)
Financial	1	Pay credit card balance in full	0.24	0.20	0.31	-0.07	**-2.59***	0.17
	2	Accumulate wealth	0.19	0.16	0.24	-0.04	-1.62	0.16
	3	Save a high percentage of income	0.23	0.21	0.15	0.07	**2.05***	0.18
	4	NOT miss credit card payments	0.17	0.18	0.15	0.02	0.87	0.15
	5	Earn a large income	0.22	0.20	0.14	0.08	**2.41***	0.17
	6	NOT accumulate a lot of credit card debt	0.13	0.15	0.11	0.02	0.54	0.20
	7	NOT accumulate a lot of educational loan debt	0.01	0	0.10	-0.09	**-3.19****	0.19
	8	Use coupons or rebate offers	0.01	0.02	0.04	-0.03	-1.16	0.13
	9	Take out a mortgage	0.04	0.02	0.03	0.01	0.36	0.14
	10	Pay enough taxes to get refund	0.11	0.07	0.02	0.09	**3.08****	0.14
Health	11	Get dental cleanings	0.11	0.1	0.17	-0.06	**-2.74****	0.13
	12	NOT consume nicotine	0.16	0.15	0.12	0.04	1.40	0.16
	13	Floss	0.13	0.1	0.11	0.01	0.69	0.09
	14	NOT gain excess weight	0.07	0.08	0.09	-0.02	-0.79	0.12
	15	Use sunscreen	0.11	0.09	0.08	0.03	1.88	0.1
	16	Monitor the nutritional content of food	0.12	0.1	0.08	0.04	1.94	0.12
	17	Get routine checkups (physicals)	0.12	0.09	0.07	0.05	**2.37***	0.11
	18	Follow a diet plan	-0.04	-0.02	0.06	-0.10	**-5.84*****	0.12
	19	Follow doctor's prescriptions	0.14	0.1	0.05	0.08	**4.26*****	0.11
	20	NOT use recreational drugs	0.12	0.14	0.05	0.07	**2.93****	0.15
	21	NOT drink coffee	-0.02	0	0.04	-0.06	**-4.53*****	0.08
	22	NOT overeat	0.08	0.08	0	0.08	**4.21*****	0.12
	23	NOT drink alcoholic beverages	0.09	0.09	0	0.09	**4.12*****	0.14
	24	Actively exercise	0.14	0.11	-0.01	0.15	**8.33*****	0.17
	25	Keep physically active	0.08	0.08	-0.03	0.11	**6.88*****	0.14
Personal Prudence	26	Attain a high level of education	0.27	0.3	0.16	0.11	**3.47****	0.20
	27	NOT get tattoos	0.12	0.1	0.14	-0.02	-0.80	0.13
	28	Go to bed at the same time every night	0.09	0.06	0.12	-0.03	**-2.28***	0.09
	29	Have kids when older	0.14	0.10	0.10	0.04	1.48	0.15
	30	NOT gamble or buy lottery tickets	0.12	0.10	0.08	0.03	1.10	0.16
	31	NOT leave dirty dishes overnight	0.07	0.04	0.05	0.02	0.96	0.10
	32	Arrive on time for meetings	0.10	0.10	0.03	0.07	**2.91****	0.14
	33	Start tasks and assignments early	0.12	0.11	0.02	0.10	**4.57*****	0.15
	34	NOT get speeding tickets	0.07	0.05	0	0.07	**3.47****	0.12
	35	Get married when older	0.08	0.05	0	0.08	**3.40****	0.13
	36	NOT drive recklessly	0.08	0.07	-0.03	0.11	**6.00*****	0.14
		Overall Mean	0.11	0.10	0.08	0.03	1.39	0.14

GENERAL DISCUSSION

Prior literature has implicated time preference in a wide range of behaviors of interest to social scientists, from the mundane to some of the most important decisions people make. In this paper, we present the most comprehensive examination to date of how well laboratory-derived estimates of time preference predict self-reported real-world behaviors. We approached this investigation by comparing time preference to four benchmarks: (i) zero, (ii) across behaviors, (iii) against other predictors, and (iv) expert forecasts. We find that how promising time preference is as a predictor depends on which of our four benchmarks we use.

Study 1 used the first three benchmarks. On the first benchmark, time preference is a significant predictor for about half the behaviors, even when controlling for 15 other relevant variables. However, we sampled these 36 behaviors precisely because we expected them to be related to time preference on theoretical grounds. So, although it predicts many of the behaviors, some may find it surprising that it did not predict more of these behaviors. Such surprise may be especially warranted considering that a correlation as small as 0.06 is significant in our study because of our large sample.

On the second benchmark, there was considerable heterogeneity in the association between time preference and the behaviors that was neither accounted for by domain (e.g., financial vs. health) nor easily explained by ratings of the behaviors on their psychological facets (see Supplemental Material Section A). It is possible that people do not possess one singular time preference. For example, Chapman (1996) found that time preferences for health and time were not correlated. Also, some have suggested that other important preferences, like risk sensitivity and loss aversion (Lejarraga & Hertwig, 2021; Stephens, 1981) may be dynamic and state-dependent, varying as a function of an organism's environment and metabolic needs. Time preferences might similarly differ within an individual across contexts (cf. Krefeld-Schwalb, Bartels, & Johnson, 2021) and mental states, and this intra-individual variability could undermine the predictive validity of time preference. Whatever the cause of the unexplainable heterogeneity of the predictive validity of time preference that we observe across behaviors, we think this reveals gaps in our knowledge of the moderators of these relationships. Time preference appears to be relevant for some behaviors and not very relevant for a host of others, and we do not know why.

On the third benchmark, time preference outperformed a number of relevant variables but also consistently underperformed compared to some others. On the one hand, we included relevant, widely-used variables from personality psychology and financial decision making, and so it is promising that time preference performed better than many of them. On the other hand, given these behaviors were specifically chosen to be related to time preference, one might expect time preference to perform better than most if not all other variables.

Study 2 introduced the final benchmark. At the individual level, and even more so at the group level, expert forecasts broadly aligned with observed correlations. However, experts' forecasted correlations were overall fairly small and offered with muted confidence; despite that, their forecasts systematically *over*estimated the degree of positive association between time preference and the behaviors. There was no detectable relationship between variables that might indicate level of expertise and more accurate forecasting (as in DellaVigna & Pope, 2018).

Implications

So, what should we conclude about how well measures of time preference predict behaviors? One's answer likely depends on (i) prior beliefs about the importance of time preference, (ii) implementational concerns, and (iii) whether one thinks studies like ours can meaningfully address how well time preference predicts behavior.

First, do we start with the assumption that time preference is an important and/or useful predictor of behavior? Most of us who do research on intertemporal choice believe that it is an important research domain, for many reasons. Hence, we are inclined to look for evidence of prediction "successes" as support for the importance of the research endeavor. Others likely start with different expectations and reach different conclusions of how good or bad our mixed bag of evidence is for the general predictive validity of time preference. Regardless of one's expectations, the associations we found between time preference and our wide range of behaviors were mostly small in magnitude and notably smaller than our experts' already moderate predictions.

Second, why are time preference's relationships with behavior important and what other predictors are available? Predicting behavior is an important endeavor, and time preference has been associated with decisions of major import, like mortgage choice (Atlas et al., 2017), retirement savings decisions (Angeletos et al., 2001), and smoking (e.g., Reynolds et al., 2004).

However, we found several covariates that were better predictors of the behaviors we measured, even though we specifically chose behaviors we expected to be related to time preference. At the same time, time preference presumably outperforms many more covariates that we did not measure (and some we did, like height—measured for BMI—which was a worse predictor for 29 of the 36 behaviors). And measures of time preference are more easily administered than measures of many other predictors—they can be collected remotely, using a few simple questions, and are fairly reliable. We are certainly not advocating that people stop collecting data on time preference as one of a set of predictors of important behaviors.

Finally, can studies like ours meaningfully address how well time preference predicts behavior? A handful of caveats apply to our investigation and to the majority of the literature we cited as precedent for our investigation. First, our data are correlational, with the usual caveats about inferring causality (e.g., reverse causation and omitted variables). Following prior literature and for ease of communication, we use variants of the word "predict" to describe the relationship between time preference and behaviors. However, the coefficients we reported are degrees of association; we do not (and should not) make strong causal claims about these relationships. Second, our measures were self-reports, which raises the usual concerns about how memory errors and social desirability influence participant responses. Third, our measure of time preference was not incentive-compatible, which introduces the possibility that participants might not have responded according to their true time preference (Harrison & Rutström, 2008). Then again, past research suggests that incentivized elicitation measures do not generally differ significantly from non-incentivized measures (Camerer & Hogarth, 1999; Starmer & Sugden, 1991). Fourth, following the prior literature, we sampled our list of behaviors and covariates to have the most relevance for the current investigation. In particular, the behaviors we chose all have delayed consequences, and this method of stimulus sampling influenced our results. However, our stimulus sampling was theoretically-motivated: We curated our list of behaviors either by adapting items from previous studies (Chabris et al., 2008; Reimers et al., 2009), or created new items such that they varied along 23 potentially informative characteristics (see Supplemental Material Section A).

With those caveats in mind, our goal was to provide the most comprehensive investigation to date of how well laboratory-derived measures of time preference predict real-world behaviors. We did so by sampling more behaviors, including more relevant variables as covariates, employing a test-retest design with a large sample. We also elicited forecasts of the size of these correlations by the group of people most well-positioned to make these predictions—researchers who think about and publish on the topics of intertemporal choice and time preference. We think the predictions of these experts are important as a benchmark against which the observed correlations can be compared. Perhaps more importantly, it is the predictions of experts like these that determine where we look for evidence on the role of time preference in behavior. We think it is telling that nearly all of our experts expressed low levels of confidence in their predictions. We hope that the mixed bag of results produced by our comprehensive investigation of the topic help guide whether and where these and other researchers use time preference as a metaphor for and/or predictor of behaviors in future research.

Author Contributions: DB conceived idea. YL and DB designed survey instruments and collected the data; SB and YL performed the analysis; DB, YL, and SB wrote the paper.

REFERENCES

Angeletos, G. M., Laibson, D., Repetto, A., Tobacman, J., & Weinberg, S. (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *Journal of Economic Perspectives*, *15*(3), 47–68.

Atlas, S. A., Johnson, E. J., & Payne, J. W. (2017). Time preferences and mortgage choice. *Journal of Marketing Research*, *54*(3), 415–429.

Barlow, P., Reeves, A., McKee, M., Galea, G., & Stuckler, D. (2016). Unhealthy diets, obesity and time discounting: a systematic literature review and network analysis. *Obesity Reviews*, *17*(9), 810–819.

Bradford, W. D. (2010). The association between individual time preferences and health maintenance habits. *Medical Decision Making*, *30*(1), 99–112.

Bradford, W. D., Courtemanche, C., Heutel, G., McAlvanah, P., & Ruhm, C. (2017). Time preferences and consumer behavior. *Journal of Risk and Uncertainty*, *55*(2–3), 119–145.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1–3), 7–42.

Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P., & Taubinsky, D. (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty*, *37*(2–3), 237–269.

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 771–791.

Chmielewski, M., & Kucker, S. C. (2020). An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results. *Social Psychological and Personality Science*, *11*(4), 464–473.

Clemen, R. T., Fischer, G. W., & Winkler, R. L. (2000). Assessing dependence: Some experimental results. *Management Science*, *46*(8), 1100–1115.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 (ed.)). Erlbaum.

DellaVigna, S., & Pope, D. (2018). What motivates effort? Evidence and expert forecasts. *The Review of Economic Studies*, *85*(2), 1029–1069.

Dixon, M. R., Marley, J., & Jacobs, E. A. (2003). Delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, *36*(4), 449–458.

Dowle, M., & Srinivasan, A. (2021). *data.table: Extension of "data.frame"* (1.14.2). https://cran.r-project.org/package=data.table

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(2), 939–944.

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, *234*(4), 281–295.

Ericson, K. M., & Laibson, D. (2019). Intertemporal choice. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 1–67). North Holland.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645-1692.

Farrell, P., & Fuchs, V. R. (1982). Schooling and health: the cigarette connection. *Journal of*

Health Economics, *1*(3), 217-230.
Fernandes, D., Lynch, John G., J., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, *60*(8), 1861–1883.
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.
Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy*, *127*(4), 1826–1863.
Harrell, F. E. J. (2019). *Hmisc: Harrell Miscellaneous* (4.3-0). https://cran.r-project.org/package=Hmisc
Harrison, G. W., & Rutström, E. E. (2008). Risk aversion in the laboratory. *Research in Experimental Economics*, *12*(8), 41–196.
Hughes, J. (2020). *reghelper: Helper Functions for Regression Analysis* (0.3.6). https://cran.r-project.org/package=reghelper
John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, *2*(1999), 102–138.
Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478.
Lejarraga, T., & Hertwig, R. (2021). Three Theories of Choice and Their Psychology of Losses. *Perspectives on Psychological Science*.
Li, Y., Baldassi, M., Johnson, E. J., & Weber, E. U. (2013). Complementary cognitive capabilities, economic decision making, and aging. *Psychology and Aging*, *28*(3), 595–613.
Li, Y., Gao, J., Enkavi, A. Z., Zaval, L., Weber, E. U., & Johnson, E. J. (2015). Sound credit scores and financial decisions despite cognitive aging. *Proceedings of the National Academy of Sciences*, *112*(1), 65–69.
Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*(1), 37–44.
Lynch, J. G. J., Netemeyer, R. G., Spiller, S. A., & Zammit, A. (2010). A generalizable scale of propensity to plan: The long and the short of planning for time and for money. *Journal of Consumer Research*, *37*(1), 108–128.
MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., & Munafò, M. R. (2011). Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology*, *216*(3), 305–321.
Meier, S., & Sprenger, C. D. (2012). Time discounting predicts creditworthiness. *Psychological Science*, *23*(1), 56–58.
Patton, J. H., Stanford, M. S., & Barratt, E. S. (1996). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774.
Petry, N. M. (2001). Pathological gamblers, with and without substance abuse disorders, discount delayed rewards at high rates. *Journal of Abnormal Psychology*, *110*(3), 482–487.
R Core Team. (2019). *R: A language and environment for statistical computing* (3.6.2). R Foundation for Statistical Computing. https://www.r-project.org/
Read, D., & Read, N. L. (2004). Time discounting over the lifespan. *Organizational Behavior and Human Decision Processes*, *94*(1), 22–32.

Reed, D. D., & Martens, B. K. (2011). Temporal discounting predicts student responsiveness to exchange delays in a classroom token system. *Journal of Applied Behavior Analysis*, *44*(1), 1–18.

Reimers, S., Maylor, E. A., Stewart, N., & Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Personality and Individual Differences*, *47*(8), 973–978.

Reynolds, B. (2006). A review of delay-discounting research with humans: relations to drug use and gambling. *Behavioural Pharmacology*, *17*(8), 651–667.

Reynolds, B., Richards, J. B., Horn, K., & Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes*, *65*(1), 35–42.

Rick, S. I., Cryder, C. E., & Loewenstein, G. (2008). Tightwads and spendthrifts. *Journal of Consumer Research*, *34*(6), 767–782.

Romney, A. K., Batchelder, W. H., & Weller, S. C. (1987). Recent applications of cultural consensus theory. *American Behavioral Scientist*, *31*(2), 163–177.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). *afex: Analysis of Factorial Experiments* (1.0.1). https://cran.r-project.org/package=afex

Spearman, C. (1904). The proof and measurement of association between two things. *Am J Psychol*, *15*, 72–101.

Starmer, C., & Sugden, R. (1991). Does the random-lottery incentive system elicit true preferences? An experimental investigation. *The American Economic Review*, *81*(4), 971–978.

Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences. *Animal Behaviour*, *29*(2), 628–629.

Sutter, M., Kocher, M. G., & Glätzle-Rützler, D. Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, *103*(1), 510–531.

Toubia, O., Johnson, E. J., Evgeniou, T., & Delquié, P. (2013). Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters. *Management Science*, *59*(3), 613–640.

Tsai, C. I., Klayman, J., & Hastie, R. (2008). Effects of amount of information on judgment accuracy and confidence. *Organizational Behavior and Human Decision Processes*, *107*(2), 97–105.

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131–160).

Urminsky, O., & Zauberman, G. (2015). The psychology of intertemporal preferences. *The Wiley Blackwell Handbook of Judgment and Decision Making*, *2*, 141–181.

Vuchinich, R. E., & Simpson, C. A. (1998). Hyperbolic temporal discounting in social drinkers and problem drinkers. *Experimental and Clinical Psychopharmacology*, *6*(3), 292–305.

Weller, S. C. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field Methods*, *19*(4), 339–368.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., …, & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Wickham, H., Romain, F., Lionel, H., & Kirill, M. (2021). *dplyr: A Grammar of Data Manipulation* (1.0.7). https://cran.r-project.org/package=dplyr

HOW WELL DO LABORATORY-DERIVED ESTIMATES OF TIME PREFERENCE PREDICT REAL-WORLD BEHAVIOR? COMPARISONS TO FOUR BENCHMARKS

SUPPLEMENTAL MATERIAL

Table S1. Logistic regression predicting retention (completing second wave of study) n = 1,576. Bolded means significant at a Bonferroni corrected α of 0.0031.

Predictor	*B*	*SE*	*z*	*p*	*Exp(B)*	*95% C.I. for EXP(B)*	
						LL	*UL*
Time Preference	-0.03	0.02	-1.29	0.20	0.97	0.93	1.02
Age	0.02	0.01	2.67	0.008	1.02	1.00	1.03
Parent Education	-0.11	0.04	-2.93	**0.003**	0.89	0.83	0.96
Gender = male	-0.39	0.16	-2.49	0.013	0.68	0.50	0.92
Source = MTurk	-0.66	0.18	-3.64	**<0.001**	0.52	0.36	0.74
Extraversion	0.00	0.01	-0.23	0.82	1.00	0.97	1.02
Conscientiousness	-0.02	0.01	-1.05	0.30	0.98	0.96	1.01
Openness	0.03	0.02	1.46	0.14	1.03	0.99	1.06
Agreeableness	0.01	0.01	0.95	0.34	1.01	0.99	1.04
Neuroticism	0.00	0.01	0.21	0.83	1.00	0.98	1.03
Barratt Impulsiveness Scale	0.00	0.01	-0.24	0.81	1.00	0.97	1.03
Financial Literacy	-0.01	0.03	-0.34	0.74	0.99	0.93	1.05
Numeracy-CRT	0.14	0.04	3.25	**0.001**	1.14	1.06	1.24
Tightwad-Spendthrift	-0.03	0.02	-1.76	0.079	0.97	0.94	1.00
Propensity to Plan	0.00	0.01	0.24	0.81	1.00	0.99	1.01
Risk Preference	-0.01	0.05	-0.13	0.89	0.99	0.89	1.10
Constant	1.79	0.88	2.03	0.043	6.00	1.06	33.86

Table S2. Time preference correlations with behavior and standardized coefficients using only (1) wave 1 data and (2) rank-ordered analyses.

Domain	Behaviors	Wave 1 Only Correlation	Wave 1 Only Standardized Coefficients	Rank-ordered Analyses Correlation	Rank-ordered Analyses Standardized Coefficient
Financial	Pay credit card balance in full	0.25***	0.20***	0.13***	0.13***
	Accumulate wealth	0.19***	0.14***	0.22***	0.15***
	Save a high percentage of income	0.10***	0.08**	0.20***	0.17***
	NOT miss credit card payments	0.17***	0.10***	0.03	0.02
	Earn a large income	0.08***	0.04	0.14***	0.08**
	NOT accumulate a lot of credit card debt	0.11***	0.09***	0.13***	0.10***
	NOT accumulate a lot of educational loan debt	0.09**	0.05	0.03	0.03
	Use coupons or rebate offers	0.04	0.03	0.04	0.02
	Take out a mortgage	-0.03	-0.04	-0.07*	-0.03
	Pay enough taxes on their paychecks to get refund	0.02	0	-0.01	0
Health	Get dental cleanings	0.16***	0.12***	0.18***	0.14***
	NOT consume nicotine	0.11***	0.09**	0.11***	0.07*
	Floss	0.08**	0.07**	0.11***	0.11***
	NOT gain excess weight	0.04	0.03	0.07*	0.06*
	Use sunscreen	0.06*	0.03	0.08**	0.05
	Monitor the nutritional content of food	0.06*	0.05*	0.09**	0.07*
	Get routine checkups (physicals)	0.07**	0.07*	0.07**	0.07**
	Follow a diet plan	0.02	0.03	0.05	0.06*
	Follow doctor's prescriptions	0.06*	0.02	0.06*	0.02
	NOT use recreational drugs	0.03	-0.01	0.05	0.02
	NOT drink coffee	0.03	0.02	0.03	0.04
	NOT overeat	-0.03	-0.06*	-0.01	-0.06*
	NOT drink alcoholic beverages	0.02	0.05*	0	0.04
	Actively exercise	0	0.01	-0.01	-0.02
	Keep physically active	-0.01	0	-0.04	-0.04
Personal Prudence	Attain a high level of education	0.12***	0.02	0.15***	0.04
	NOT get tattoos	0.14***	0.07**	0.15***	0.09**
	Go to bed at the same time every night	0.11***	0.07*	0.11***	0.05
	Have kids when older	0.07	0	0.03	0.03
	NOT gamble or buy lottery tickets	0.09***	0.08**	0.10***	0.09***
	NOT leave dirty dishes overnight	0.06*	0.05*	0.05*	0.05
	Arrive on time for meetings	0.06*	0.04	0.03	0
	Start tasks and assignments early	0	-0.01	0.01	0.01
	NOT get speeding tickets	0	-0.03	0.02	-0.01
	Get married when older	-0.02	-0.03	-0.02	0.02
	NOT drive recklessly	0.01	0	-0.05	-0.05

Table S3. Descriptive statistics for all behaviors in main study.

Domain	Behaviors	Reliability	Mean	Standard Deviation	Median	Mode
Financial	Pay credit card balance in full	0.86	3.34	1.47	3.5	5
	Accumulate wealth	0.77	2.55	0.94	2.5	2
	Save a high percentage of income	0.76	10.98	14.22	6.25	0
	Miss credit card payments	0.71	1.39	0.68	1	1
	Earn a large income	0.65	1.82	1.35	1.83	1.35
	Accumulate credit card debt	0.78	0.30	0.44	0.03	0
	Accumulate educational loan debt	0.78	2.53	1.12	2.5	1
	Use coupons or rebate offers	0.75	3.38	0.91	3.5	3
	Take out a mortgage	0.70	2.47	0.98	2.5	3
	Pay enough taxes to get refund	0.51	2.59	1.56	3	4
Health	Get dental cleanings	0.84	3.34	1.41	3.5	5
	Consume nicotine	0.90	2.15	1.88	1	1
	Floss	0.82	2.26	1.08	2	1
	Gain excess weight (BMI)	0.96	27.82	7.13	26.23	25.74
	Use sunscreen	0.75	3.07	1.15	3	3
	Monitor the nutritional content of food	0.62	3.25	1.64	3	2
	Get routine checkups (physicals)	0.78	0.65	0.37	0.75	1
	Follow a diet plan	0.52	0.15	0.24	0	0
	Follow doctor's prescriptions	0.58	3.44	0.66	3.5	4
	Use recreational drugs	0.87	1.42	1.00	1	1
	Drink coffee	0.89	3.76	1.85	4.5	5
	Overeat	0.63	2.30	0.78	2	2
	Drink alcoholic beverages	0.84	2.69	1.23	2.5	2
	Actively exercise	0.71	0.70	0.42	0.77	0
	Keep physically active	0.59	1.58	0.42	1.59	1.48
Personal Prudence	Attain a high level of education	0.96	4.56	1.78	5	6
	Get tattoos	0.97	0.13	0.27	0	0
	Go to bed at the same time every night	0.65	3.13	0.79	3	4
	Have kids when older	0.97	4.03	1.10	4.05	4.05
	Gamble or buy lottery tickets	0.71	1.72	0.85	1.5	1
	Leave dirty dishes overnight	0.77	2.51	0.94	2.5	3
	Arrive on time for meetings	0.66	3.66	0.54	4	4
	Start tasks and assignments early	0.61	3.13	0.77	3	4
	Get speeding tickets	0.71	0.24	0.56	0	0
	Get married when older	0.93	24.30	5.52	23.50	21
	Drive recklessly	0.64	1.98	0.75	2	2

Table S4. ORIV (Obviously Related Instrumental Variables; Gillen, Snowberg, & Yariv, 2019) correlations and multiple regression standardized betas.

Behaviors	**ORIV correlation**	**bootstrap standard error**	***z***	***p*-level**	**bootstrapped 95% CI**		**ORIV beta**	**bootstrap standard error**	***z***	***p*-level**	**bootstrapped 95% CI**	
Pay credit card balance in full	0.291	0.026	11.31	0	0.241	0.342	0.235	0.028	8.42	0	0.18	0.29
Accumulate wealth	0.228	0.026	8.66	0	0.177	0.280	0.159	0.025	6.31	0	0.11	0.21
Save a high percentage of income	0.149	0.030	5.06	0	0.092	0.207	0.126	0.029	4.29	0	0.07	0.18
Miss credit card payments	0.152	0.028	5.47	0	0.097	0.206	0.065	0.030	2.15	0.032	0.01	0.12
Earn a large income	0.145	0.029	4.98	0	0.088	0.202	0.075	0.028	2.7	0.007	0.02	0.13
Accumulate credit card debt	0.107	0.026	4.18	0	0.057	0.157	0.085	0.028	3.03	0.002	0.03	0.14
Accumulate educational loan debt	0.102	0.031	3.32	0.001	0.042	0.163	0.072	0.034	2.14	0.032	0.01	0.14
Use coupons or rebate offers	0.041	0.027	1.5	0.134	-0.012	0.094	0.042	0.028	1.52	0.13	-0.01	0.10
Take out a mortgage	0.040	0.037	1.07	0.285	-0.033	0.113	0.017	0.039	0.42	0.673	-0.06	0.09
Pay enough taxes to get refund	0.019	0.034	0.56	0.577	-0.048	0.086	0.002	0.036	0.06	0.955	-0.07	0.07
Get dental cleanings	0.160	0.025	6.3	0	0.110	0.210	0.112	0.026	4.3	0	0.06	0.16
Consume nicotine	0.116	0.024	4.87	0	0.070	0.163	0.077	0.025	3.03	0.002	0.03	0.13
Floss	0.109	0.027	4.1	0	0.057	0.162	0.114	0.028	4.13	0	0.06	0.17
Gain excess weight	0.079	0.024	3.23	0.001	0.031	0.126	0.051	0.027	1.87	0.062	0.00	0.10
Use sunscreen	0.078	0.027	2.89	0.004	0.025	0.131	0.054	0.027	2.02	0.044	0.00	0.11
Monitor the nutritional content of food	0.079	0.029	2.71	0.007	0.022	0.137	0.067	0.030	2.22	0.027	0.01	0.13
Get routine checkups (physicals)	0.069	0.027	2.57	0.01	0.016	0.121	0.064	0.026	2.45	0.014	0.01	0.12
Follow a diet plan	0.065	0.032	2.03	0.043	0.002	0.127	0.072	0.034	2.09	0.037	0.00	0.14
Follow doctor's prescriptions	0.057	0.029	1.97	0.049	0.000	0.114	0.006	0.030	0.2	0.838	-0.05	0.07
Use recreational drugs	0.040	0.026	1.56	0.118	-0.010	0.091	0.008	0.026	0.3	0.763	-0.04	0.06
Drink coffee	0.036	0.026	1.36	0.174	-0.016	0.087	0.041	0.027	1.49	0.136	-0.01	0.09
Overeat	0.001	0.028	0.04	0.967	-0.054	0.056	-0.068	0.028	-2.43	0.015	-0.12	-0.01
Drink alcoholic beverages	-0.001	0.027	-0.05	0.958	-0.054	0.051	0.039	0.027	1.45	0.147	-0.01	0.09
Actively exercise	-0.013	0.027	-0.47	0.638	-0.066	0.040	-0.011	0.027	-0.41	0.68	-0.06	0.04
Keep physically active	-0.037	0.029	-1.29	0.197	-0.093	0.019	-0.022	0.029	-0.76	0.449	-0.08	0.04
Attain a high level of education	0.146	0.024	6.09	0	0.099	0.194	0.027	0.023	1.16	0.246	-0.02	0.07
Get tattoos	0.128	0.024	5.3	0	0.080	0.175	0.059	0.025	2.4	0.016	0.01	0.11
Go to bed at the same time every night	0.126	0.027	4.69	0	0.073	0.179	0.051	0.028	1.84	0.066	0.00	0.11
Have kids when older	0.088	0.033	2.63	0.008	0.023	0.154	0.006	0.035	0.18	0.856	-0.06	0.07
Gamble or buy lottery tickets	0.084	0.029	2.94	0.003	0.028	0.140	0.067	0.029	2.35	0.019	0.01	0.12
Leave dirty dishes overnight	0.054	0.027	1.96	0.05	0.000	0.107	0.045	0.026	1.71	0.087	-0.01	0.10
Arrive on time for meetings	0.036	0.028	1.29	0.198	-0.019	0.090	0.011	0.028	0.38	0.705	-0.04	0.07
Start tasks and assignments early	0.018	0.028	0.65	0.516	-0.037	0.073	0.008	0.026	0.32	0.75	-0.04	0.06
Get speeding tickets	0.002	0.029	0.07	0.946	-0.054	0.058	-0.041	0.032	-1.29	0.198	-0.10	0.02
Get married when older	0.009	0.033	0.27	0.784	-0.055	0.073	-0.001	0.033	-0.02	0.985	-0.06	0.06
Drive recklessly	-0.028	0.027	-1.01	0.31	-0.081	0.026	-0.030	0.028	-1.07	0.286	-0.09	0.03

Table S5. Predictive validity of 16 predictor variables across 36 behaviors: 1) Number of times the predictors meet different *p*-value benchmarks and 2) median absolute standard beta for each predictor across 36 regressions with all 16 predictors.

Variable	**Benchmark**					**Percentage**	**Median Absolute**
	$p < .001$	$.001 \leq p < .01$	$.01 \leq p < .05$	$.05 \leq p < .1$	$p \geq .1$	$p < .05$	**Standard Beta**
Time preference	5	3	11	0	17	52.8%	0.059
Age	15	2	3	1	15	55.6%	0.090
Parent Education	11	4	4	2	15	52.8%	0.058
Gender	7	5	5	5	14	47.2%	0.053
Source	7	3	5	1	20	41.7%	0.046
Extraversion	12	4	7	2	11	63.9%	0.078
Conscientiousness	4	4	4	3	21	33.3%	0.059
Openness	7	2	6	3	18	41.7%	0.049
Agreeableness	1	5	5	3	22	30.6%	0.040
Neuroticism	3	0	3	2	28	16.7%	0.032
Barratt Impulsiveness Scale	11	6	3	1	15	55.6%	0.104
Financial Literacy	9	3	4	3	17	44.4%	0.069
Numeracy-CRT	3	2	6	5	20	30.6%	0.059
Tightwad-Spendthrift	6	3	4	2	21	36.1%	0.044
Propensity to Plan	2	4	5	2	23	30.6%	0.037
Risk Preference	0	6	3	3	24	25.0%	0.027

Table S6. Expert predictions compared to disattenuated correlations.

Domain	Behavior	Mean Prediction	Disattenuated Correlation	Mean Difference	*t*-statistic	MAD
Financial	Pay credit card balance in full	0.24	0.39	-0.15	**-5.70*** **	0.21
	Accumulate wealth	0.19	0.32	-0.13	**-4.80*** **	0.20
	Save a high percentage of income	0.23	0.21	0.02	0.44	0.18
	NOT miss credit card payments	0.17	0.21	-0.04	-1.41	0.15
	Earn a large income	0.22	0.21	0.01	0.22	0.16
	NOT accumulate a lot of credit card debt	0.13	0.15	-0.02	-0.58	0.19
	NOT accumulate a lot of educational loan debt	0.01	-0.14	0.15	**6.06*** **	0.20
	Use coupons or rebate offers	0.01	0.06	-0.02	-0.66	0.14
	Take out a mortgage	0.04	0.05	0.06	**2.07***	0.14
	Pay enough taxes to get refund	0.11	0.04	-0.03	-1.12	0.16
Health	Get dental cleanings	0.11	0.22	-0.11	**-5.28*** **	0.16
	NOT consume nicotine	0.16	0.15	0.01	0.45	0.16
	Floss	0.13	0.15	-0.02	-1.29	0.10
	NOT gain excess weight	0.07	0.11	-0.04	-1.92	0.12
	Use sunscreen	0.11	0.11	0.00	0.10	0.10
	Monitor the nutritional content of food	0.12	0.12	0.00	-0.01	0.12
	Get routine checkups (physicals)	0.12	0.09	0.03	1.34	0.11
	Follow a diet plan	-0.04	0.10	-0.14	**-8.36*** **	0.15
	Follow doctor's prescriptions	0.14	0.08	0.06	**2.90** **	0.10
	NOT use recreational drugs	0.12	0.06	0.06	**2.36***	0.14
	NOT drink coffee	-0.02	0.05	-0.07	**-5.43*** **	0.09
	NOT overeat	0.08	0.00	0.08	**4.25*** **	0.12
	NOT drink alcoholic beverages	0.09	0.00	0.09	**4.06*** **	0.14
	Actively exercise	0.14	-0.02	0.16	**8.73*** **	0.17
	Keep physically active	0.08	-0.05	0.13	**7.89*** **	0.15
Personal Prudence	Attain a high level of education	0.27	0.19	0.08	**2.52***	0.19
	NOT get tattoos	0.12	0.17	-0.05	**-2.10***	0.15
	Go to bed at the same time every night	0.09	0.18	-0.09	**-6.21*** **	0.12
	Have kids when older	0.14	0.12	0.02	0.62	0.15
	NOT gamble or buy lottery tickets	0.12	0.12	0.00	-0.14	0.16
	NOT leave dirty dishes overnight	0.07	0.07	0.00	0.10	0.10
	Arrive on time for meetings	0.10	0.05	0.05	**2.17***	0.13
	Start tasks and assignments early	0.12	0.03	0.09	**4.08*** **	0.14
	NOT get speeding tickets	0.07	0.00	0.07	**3.57*** **	0.13
	Get married when older	0.08	0.00	0.08	**3.54*** **	0.13
	NOT drive recklessly	0.08	-0.04	0.12	**6.53*** **	0.15
	Overall Mean	0.11	0.10	0.01		0.14

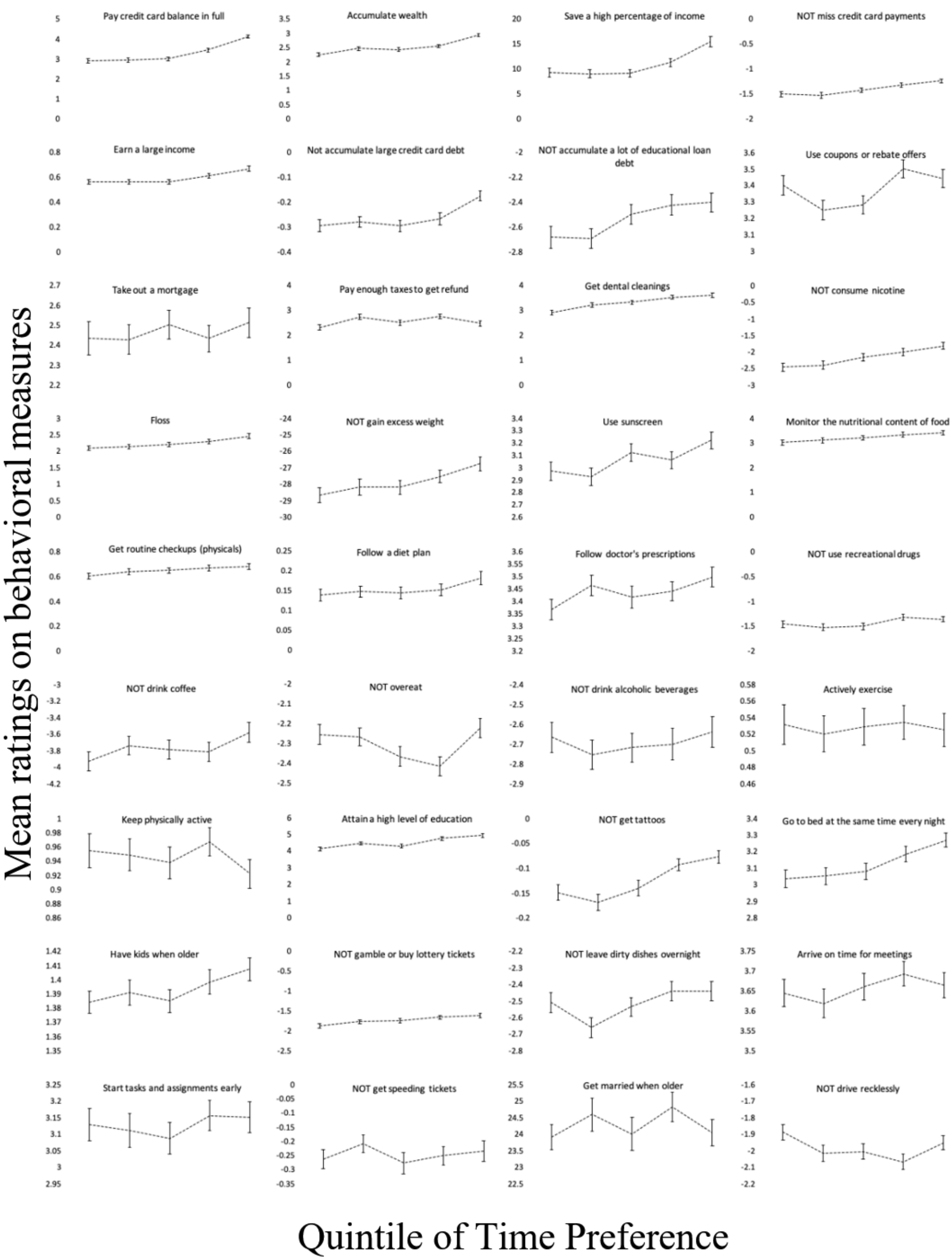


Figure S1. Mean ratings on each of the 36 behaviors (on the y-axis) versus quintiles of time preference (on the x-axis). Error bars represent standard errors,

SECTION A: PILOT STUDY OF 23 CHARACTERISTICS THAT DIFFERENTIATE THE 36 BEHAVIORS USED IN MAIN STUDY

An early goal of this project, and its primary motivation initially, was to better understand how time preference enters into decision making across a variety of behaviors. To explore this question, we aimed to see how differences across these behaviors on various characteristics—which we call "facets"—moderate how much the behaviors were correlated with estimates of time preference. Compare, for example, the decision about how much of your current income to save versus the decision about whether to have a slice of cake at an office birthday, and consider how they vary on three facets: (i) "tradeoffiness"—whether the rationale for the behavior involves a weighing of current versus future consumption (saving = high; cake = low), (ii) "urginess"—whether the behavior involves suppressing an immediate, short-term goal (saving = low; cake = high), and (iii) internal locus of control—whether the behavior is deliberately controlled by the actor making a decision versus a reaction to the environment that is out of the actor's control (saving = high; cake at a party = potentially low). We initially expected that behaviors that were high in tradeoffiness, internal locus of control, or low in urginess would be where time preferences were most relevant and would, thus, be better predicted by them.

Table A1. *List of 23 psychological characteristics (i.e., facets) we expected behaviors to vary on.*

Facet	***Wording in norming study***
future is uncertain	we do not know what will happen—the future is unpredictable.
goals will change	we believe our personal goals will change over time.
helps achieve a goal	we want to make progress toward a personal goal.
other goals higher priority	our other personal goals receive higher priority.
enough goal progress	we are doing well on the relevant personal goal.
considering tradeoffs	we are considering the potential tradeoffs.
has consequences	it feels like it will have potential costs and/or benefits.
consequences do not affect ME	we believe the potential costs and/or benefits will NOT affect us.
has long-term consequences	we appreciate that it has a long-term impact on our future well-being.
feels good now	it makes us feel good NOW.
will feel good in the future	we believe it will make us feel good in the FUTURE.
following a rule	we are following a general rule (and not making a deliberate choice).
requires vigilance, self-control	it requires constant vigilance and/or or effort.
actively choose to do it	we choose to—this behavior is controllable.
have a habit	we have fallen into a habit (rather than because we are making a conscious choice each time we do so).
happens by chance	it happens by accident or chance.
gut instinct	it is done instinctively (following gut reactions), rather than through careful deliberation.
self-signaling and self-presentation	we want to convey (to ourselves or others) something important about ourselves—we want to express what kind of a person we are.
product of the environment	the situations that we find ourselves in strongly influence our behaviors—our behaviors are a product of the environment.
descriptive norm	most people do that—it feels like the normal thing to do.
prescriptive norm	experts would recommend it—it's the "smart" thing to do.
not worth thinking about	it is not worth thinking about—the costs and/or benefits are too small.
following a plan	we are following through on a plan.

Early discussions with colleagues and other intertemporal choice researchers substantially expanded our set of facets, which in turn led us to expand the variety of behaviors we sampled (to better differentiate the ratings of facets across behaviors). This back-and-forth introduced even more factors and behaviors, until we had 23 facets and 36 behaviors. (See Table A1 for list of 23 facets and Table 1 for list of 36 behaviors).

Our aim was to predict the degree of association between each of 36 behaviors and time preference using each of 23 facet (which differentiate the behaviors) as moderators. To do this, we first ran a pilot study in which people would rate each of the 36 behaviors in terms of whether each facet was related to each behavior. That is, we measured people's lay theories for why they and others are patient or not on a variety of behaviors.

Methods

Participants

We recruited 142 participants from a market research firm, with an ultimate goal of 100 participants in the final sample. Due to the length of the study (data collection took from three to five hours for each participant), each participant answered questions for 12 of the 36 behaviors (counterbalanced) in each of three waves over the course of a week in April 2013. Participants who completed each wave receive an invitation to the next wave the following day. Of the 142 participants who started the study, 120 (85%) completed the first wave and we invited back 108 (90%) participants who passed at least 3 attention check questions embedded in the survey (see below for more details). Of these 108 participants invited to the second round, 103 (95%) completed the second wave and 102 (94%) completed the third wave.

Materials

Participants rated whether each of 23 facets was relevant for each of 36 behaviors—both doing and not doing the behavior. Specifically, on each screen, participants were asked to rate their agreement (1 = strongly disagree; 5 = strongly agree) with statements of the following form, with two examples shown below. The bold emphasis was not in the original but shows the repeated behavior stem, which is the only thing that varied between each behavior block.

"When people (myself and others) **save a high percentage of their income**, it is because…

1. …we choose to do so—this behavior is controllable."

… [21 more differentiating characteristics]

23. …we want to make progress toward a personal goal"

"When people (myself and others) do NOT **floss**, it is because…

1. … we want to make progress toward a personal goal."

… [21 more differentiating characteristics]

23. … it happens by accident or chance."

Within each wave, we randomized the order of the 12 behavior blocks, but the 23 facet ratings for *doing* the behavior always appeared on a separate page before the facet ratings for *NOT* doing the behavior. We also randomized the order of the 23 facet ratings in each block of questions. For four behavior blocks in each wave, we also randomly embedded among the facets an attention check of the form "When people floss, it is because if you are reading this, please check the button for strongly agree" for a total 12 attention check questions divided across three waves.

Results

Sample Characteristics. To be conservative, we excluded 29 participants who failed on any of the 12 attention check questions or who straight-lined responses (i.e., answered all 23 facet ratings the same in a question block) on any of these 72 sets of facet ratings, which left an analysis sample of 73 participants. Analyses with the full dataset are qualitatively similar.

These 73 participants (26 males) were aged 20 to 85 (M = 49.7, SD = 14.2) and had a median education level of "some college but did not finish" (37% had 4 year college degree or more).

Descriptive Statistics. Participants took about 2 minutes to complete each of the 72 blocks of 23 facet ratings, with no difference between doing a behavior and not doing a behavior ($M_{\text{log\_do\_time}}$ = 4.81 vs. $M_{\text{log\_do NOT\_time}}$ = 4.79, paired t-test p = 0.22). Tables A2a and A2b present average facet ratings for doing and not doing each behavior, respectively. Ratings for doing and not doing a behavior had an average alpha of 0.46 (range = 0.09 to 0.89), suggesting that participants generally treated them as different questions.

Table A2. Average ratings of agreement for whether (a) *doing* and (b) *not doing* each behavior is because of each of 23 facets.

Table A2a. Facet ratings for *doing* behavior	future is uncertain	goals will change	helps achieve a goal	other goals higher priority	enough goal progress	considering tradeoffs	has consequences	consequences do not affect ME	has long-term consequences	feels good now	will feel good in the future	following a rule	requires vigilance, self-control	actively choose to do it	have a habit	happens by chance	gut instinct	self-signaling and self-	product of the environment	descriptive norm	prescriptive norm	not worth thinking about	following a plan
gain excess weight	2.67	2.41	1.70	3.32	1.74	2.12	2.48	3.12	1.84	2.92	1.59	2.38	2.18	3.45	4.00	2.53	3.27	1.77	3.52	1.93	1.30	2.36	1.73
overeat	2.73	2.27	1.73	2.74	1.93	2.15	2.40	3.56	1.84	4.30	1.92	2.30	2.25	4.16	3.96	2.27	3.59	1.99	3.44	2.38	1.33	2.67	1.68
get dental cleanings	2.95	2.40	3.95	2.18	3.68	4.27	4.42	1.99	4.63	4.00	4.55	3.18	3.74	4.63	2.85	1.40	2.29	3.78	3.30	3.88	4.71	1.75	4.07
follow doctor's instructions for prescriptions	3.62	2.71	4.14	2.42	3.56	4.14	4.36	1.88	4.63	4.01	4.51	3.27	3.79	4.55	2.66	1.47	2.85	3.04	3.11	4.15	4.66	2.01	4.26
use recreational drugs	2.78	2.14	1.64	1.93	1.66	2.21	2.96	3.58	1.60	4.53	2.44	1.99	2.36	4.16	3.77	1.84	2.85	2.84	3.47	2.18	1.26	2.21	1.86
drive recklessly	2.60	2.11	2.21	3.03	1.90	2.25	2.53	3.82	1.67	3.62	1.79	2.23	2.18	4.47	3.58	2.01	3.19	2.81	3.30	2.22	1.40	2.40	1.85
accumulate a lot of credit card debt	3.00	2.78	2.42	3.70	1.99	2.34	2.85	3.64	2.03	3.75	2.16	2.47	2.26	3.92	3.70	2.30	2.97	2.48	3.30	2.71	1.29	2.15	2.11
gamble or buy lottery tickets	3.82	2.78	3.22	2.11	2.44	3.67	4.01	2.68	2.81	4.07	3.95	2.26	2.66	4.42	3.19	1.70	2.73	2.34	3.19	2.86	1.68	2.26	2.84
leave dirty dishes overnight	2.25	2.21	2.15	4.22	2.11	2.81	2.26	3.47	1.74	3.30	1.73	2.11	1.93	4.51	3.78	2.52	2.92	1.88	3.15	2.44	1.44	3.42	2.01
attain a high level of education	3.73	3.29	4.66	2.30	4.15	4.19	4.44	1.64	4.67	3.67	4.68	2.45	4.48	4.55	2.08	1.48	2.04	4.21	3.23	3.19	4.44	1.53	4.63
get married too early	3.45	2.64	3.26	2.38	2.77	2.88	3.62	3.44	2.88	4.33	3.71	1.99	2.89	4.42	2.18	2.22	3.32	3.25	3.53	2.29	1.74	2.34	3.23
get tattoos	2.44	2.08	2.84	2.00	2.62	2.36	3.04	3.49	2.37	4.29	3.36	2.22	2.23	4.59	2.62	1.78	2.70	4.40	3.66	2.75	1.56	2.56	3.26
keep physically active	3.26	3.05	4.38	2.32	3.96	4.19	4.47	1.86	4.64	4.29	4.68	2.49	4.33	4.48	2.88	1.64	2.55	3.99	3.19	3.22	4.45	1.82	4.27
eat with health and fitness concerns in mind	3.29	2.84	4.44	2.29	4.01	4.34	4.37	1.75	4.62	4.11	4.70	2.55	4.40	4.51	2.56	1.51	2.21	4.01	3.12	3.19	4.62	1.70	4.36
get routine check-ups (physicals)	3.88	2.85	4.00	2.33	3.71	4.03	4.27	1.84	4.60	3.74	4.41	3.08	3.95	4.56	2.58	1.42	2.64	3.07	3.10	3.89	4.51	1.68	4.21
drink coffee	2.29	2.22	2.49	2.26	2.51	2.75	3.32	2.92	2.42	4.38	2.85	2.41	2.10	4.45	3.75	1.60	3.10	2.55	3.16	3.74	2.38	3.00	2.48
smoke or use chewing tobacco	2.90	2.27	1.62	2.30	1.74	2.22	2.30	3.74	1.74	4.26	1.92	2.05	2.21	4.21	4.15	1.63	3.03	2.66	3.42	2.45	1.21	2.42	1.79
save a high percentage of their income	4.29	3.59	4.56	2.51	4.25	4.30	4.42	1.75	4.66	3.52	4.55	2.47	4.41	4.47	2.53	1.53	2.01	3.63	3.07	2.78	4.42	1.53	4.49
miss credit card payments	2.37	2.32	1.85	3.96	1.68	2.33	2.74	2.74	1.81	2.30	1.49	1.92	1.84	3.59	2.88	3.05	2.10	1.59	3.11	1.77	1.22	2.10	1.78
use coupons or rebate offers	3.04	2.77	3.92	2.42	3.48	4.11	4.38	1.71	3.75	4.15	3.89	2.37	4.07	4.60	2.77	1.59	2.08	3.11	3.15	3.08	4.12	1.71	4.10
arrive on-time for meetings	3.04	2.67	4.14	2.42	3.95	3.97	4.30	1.93	3.97	3.95	3.68	3.26	3.97	4.58	2.89	1.59	2.77	4.40	3.23	4.03	4.32	1.88	4.19
earn a large income	3.26	3.10	4.25	2.36	4.25	3.55	4.23	1.92	4.33	4.19	4.36	2.51	4.05	3.47	2.26	2.14	2.22	3.90	3.18	2.56	3.70	1.77	4.14
have kids too early	2.84	2.62	2.52	2.62	2.34	2.42	3.07	3.25	2.42	3.38	2.75	2.18	2.89	3.81	2.70	3.82	3.33	2.51	3.42	2.36	1.60	2.30	2.44
take out a mortgage to buy a home	3.03	2.99	4.33	2.47	4.07	4.11	4.32	2.11	4.36	3.99	4.34	2.27	3.93	4.48	1.85	1.41	1.92	3.55	3.11	3.89	3.75	1.68	4.38
actively exercise	3.21	3.04	4.52	2.36	4.03	4.26	4.44	1.71	4.64	4.25	4.64	2.51	4.44	4.62	2.99	1.40	2.21	4.04	3.25	3.16	4.52	1.64	4.42
monitor the nutrition content of the food	3.25	2.84	4.40	2.59	3.88	4.15	4.37	1.62	4.47	3.75	4.41	2.44	4.34	4.59	2.70	1.49	2.22	3.56	3.14	2.77	4.44	1.81	4.25
floss	3.03	2.56	4.12	2.23	3.79	4.01	4.37	1.71	4.52	4.00	4.51	3.08	4.12	4.66	3.34	1.42	2.55	3.51	3.19	3.45	4.67	1.93	4.05
drink alcoholic beverages	2.77	2.42	1.86	2.07	2.23	2.40	2.84	3.36	1.89	4.44	2.45	2.16	2.12	4.47	3.52	1.70	2.92	3.15	3.82	3.59	1.78	2.59	2.07
use sunscreen	3.67	2.59	3.59	2.26	3.48	4.25	4.45	1.68	4.49	3.58	4.38	2.85	4.03	4.63	2.84	1.47	2.67	3.14	3.67	3.81	4.70	1.95	3.78
accumulate wealth	3.88	3.34	4.37	2.26	4.19	4.04	4.44	1.89	4.51	4.12	4.55	2.79	4.22	3.90	2.53	2.05	2.08	3.97	3.25	2.84	4.08	1.75	4.33
pay credit card balances in full	3.70	3.08	4.41	2.29	4.23	4.34	4.44	1.79	4.52	4.32	4.56	2.77	4.07	4.59	2.56	1.45	2.34	3.84	3.07	2.95	4.62	1.78	4.52
go to bed the same time every night	2.88	2.42	3.75	2.52	3.64	3.74	4.16	1.97	4.14	3.95	4.12	2.90	3.25	4.40	3.79	1.81	3.25	2.95	3.27	3.44	4.01	2.23	3.99
start long-term tasks or assignments early	3.78	2.93	4.38	2.51	3.89	4.23	4.37	2.05	4.12	3.85	4.37	2.47	4.16	4.48	2.66	1.59	2.29	3.99	3.18	2.90	4.16	1.95	4.37
pay enough taxes that they get a refund	3.25	2.77	3.51	2.33	3.36	3.85	3.89	2.03	3.67	3.00	4.01	3.08	2.84	4.00	2.85	2.08	2.49	2.85	2.88	3.48	3.38	2.34	3.84
get speeding tickets	2.36	1.93	1.77	3.15	1.58	2.19	2.52	3.16	1.64	2.08	1.44	2.16	1.96	3.90	3.16	2.95	2.79	2.04	3.25	1.99	1.33	2.23	1.73
take out educational loans	3.11	3.16	4.49	2.59	3.32	4.01	4.21	2.15	4.22	3.21	4.11	2.37	3.36	4.37	2.11	1.49	2.10	3.21	3.49	3.29	3.34	1.74	4.29

Table A2b. Facet ratings for *NOT doing* behavior	future is uncertain	goals will change	helps achieve a goal	other goals higher priority	enough goal progress	considering tradeoffs	has consequences	consequences do not affect ME	has long-term consequences	feels good now	will feel good in the future	following a rule	requires vigilance, self-control	actively choose to do it	have a habit	happens by chance	gut instinct	self-signaling and self-	product of the environment	descriptive norm	prescriptive norm	not worth thinking about	following a plan
gain excess weight	3.14	2.97	4.25	2.70	3.97	4.14	4.32	1.99	4.59	4.05	4.56	2.85	4.12	4.22	2.55	1.73	2.53	4.04	3.25	3.25	4.52	1.93	4.19
overeat	2.99	2.90	4.23	3.29	3.75	4.01	4.23	1.96	4.51	3.82	4.41	3.08	3.81	4.49	2.68	1.60	2.59	3.81	3.22	3.37	4.49	1.96	4.03
get dental cleanings	2.48	2.37	1.81	3.68	1.88	2.42	2.60	3.47	1.70	2.49	1.67	1.93	2.00	4.21	3.32	2.00	2.53	1.82	3.05	1.97	1.26	2.66	1.85
follow doctor’s instructions for prescriptions	2.62	2.56	1.93	3.33	2.26	2.82	2.84	3.70	1.97	2.48	2.05	2.08	2.32	4.32	3.00	2.49	2.55	2.16	2.90	1.77	1.34	2.62	1.93
use recreational drugs	3.53	3.04	4.26	3.99	3.88	4.16	4.16	2.18	4.75	3.73	4.37	2.82	3.03	4.55	2.23	1.36	2.55	4.15	3.19	3.73	4.47	2.27	3.97
drive recklessly	3.56	2.60	3.73	2.86	3.63	4.12	4.12	1.86	4.48	3.74	3.90	3.29	3.78	4.60	3.16	1.53	3.12	3.82	3.18	3.97	4.41	1.84	3.53
accumulate a lot of credit card debt	4.00	3.29	4.27	3.38	4.12	4.25	4.12	2.04	4.47	3.75	4.37	2.86	4.08	4.42	2.63	1.63	2.26	3.68	3.05	3.25	4.51	1.89	4.25
gamble or buy lottery tickets	3.40	2.89	3.66	4.19	3.30	4.00	3.47	3.12	3.93	3.15	3.40	2.53	2.36	4.52	2.36	1.52	2.33	3.23	2.96	2.85	3.68	3.26	3.34
leave dirty dishes overnight	2.89	2.59	3.62	2.38	3.45	3.79	3.89	2.16	3.38	4.05	4.07	3.16	3.78	4.58	3.44	1.64	2.89	3.75	3.18	3.88	3.90	2.42	3.82
attain a high level of education	3.07	2.95	2.41	3.88	2.48	2.70	2.73	3.38	2.01	3.00	1.90	2.15	1.89	3.99	3.00	1.89	2.77	2.15	3.36	2.37	1.44	2.44	2.29
get married too early	3.74	3.84	4.00	4.25	3.62	4.03	3.93	2.56	4.15	3.22	3.67	2.81	2.81	4.42	2.12	1.88	2.34	3.41	3.21	3.42	3.82	2.36	4.08
get tattoos	3.33	3.63	3.37	3.99	3.27	4.10	3.73	2.30	4.21	3.38	3.74	2.56	1.99	4.64	2.08	1.47	2.44	3.86	3.19	3.60	3.82	2.74	3.33
keep physically active	2.55	2.74	1.82	3.97	2.05	2.48	2.34	3.36	1.88	3.21	1.75	2.21	2.25	4.21	4.05	2.04	2.85	2.01	3.32	2.70	1.37	2.60	1.88
eat with health and fitness concerns in mind	2.88	2.59	1.90	3.58	2.03	2.27	2.30	3.60	1.68	3.60	1.77	2.36	2.10	4.25	3.84	1.95	3.18	2.08	3.44	2.75	1.40	2.84	1.78
get routine check-ups (physicals)	2.68	2.47	1.93	3.92	2.37	2.53	2.51	3.49	1.89	2.47	1.77	2.12	2.26	4.16	3.49	1.92	2.42	1.93	3.15	2.10	1.48	2.88	1.90
drink coffee	2.48	2.42	3.04	3.16	2.95	3.49	3.34	2.51	3.47	3.03	3.32	2.23	2.30	4.56	2.84	1.82	2.60	2.86	2.86	2.51	3.08	2.92	3.30
smoke or use chewing tobacco	3.27	2.64	3.97	3.45	3.81	4.21	4.25	2.00	4.70	4.10	4.59	2.71	3.00	4.66	2.27	1.49	2.59	3.92	3.01	3.41	4.52	2.01	3.89
save a high percentage of their income	3.00	2.82	2.16	3.95	2.19	2.66	2.44	3.11	1.85	3.41	1.63	2.18	2.15	3.78	3.45	2.00	2.77	2.05	3.36	2.53	1.44	2.45	1.88
miss credit card payments	3.30	2.81	4.42	2.27	4.27	4.11	4.49	1.78	4.62	4.07	4.47	3.21	4.22	4.55	2.78	1.42	2.79	3.85	3.08	4.01	4.67	1.90	4.32
use coupons or rebate offers	2.47	2.45	1.93	3.73	2.51	2.41	2.07	3.59	1.82	2.18	1.78	2.23	2.12	4.19	3.40	2.18	2.78	2.34	2.84	2.58	1.62	3.73	1.92
arrive on-time for meetings	2.51	2.22	1.79	3.75	1.90	2.05	2.18	3.37	1.79	2.29	1.64	2.14	2.04	3.92	3.53	2.88	2.59	1.93	3.04	2.00	1.34	2.53	1.84
earn a large income	2.41	2.59	2.04	3.14	1.96	2.23	2.44	2.21	2.10	1.74	1.70	2.12	1.95	2.75	2.68	2.68	2.26	1.97	3.38	2.36	1.41	1.89	1.81
have kids too early	3.66	3.70	4.26	4.04	3.78	4.15	3.93	2.27	4.21	3.59	3.85	2.63	3.90	4.34	2.33	2.01	2.32	3.48	3.33	3.26	3.82	2.14	4.27
take out a mortgage to buy a home	3.71	3.66	3.00	3.85	2.67	3.59	3.19	2.81	3.21	2.79	2.60	2.10	2.38	4.07	2.36	1.82	2.10	2.29	3.21	2.25	2.14	2.26	2.92
actively exercise	2.59	2.58	1.84	3.88	2.08	2.34	2.22	3.32	1.85	3.07	1.64	2.10	2.19	4.27	3.90	1.93	2.81	2.00	3.30	2.62	1.40	2.60	1.84
monitor the nutrition content of the food	2.62	2.60	1.85	3.60	2.38	2.25	2.30	3.74	1.81	2.90	1.75	2.49	2.34	4.18	3.82	2.10	3.08	2.00	3.18	3.07	1.64	3.36	1.85
floss	2.48	2.27	1.74	3.52	1.86	2.22	2.11	3.53	1.62	2.36	1.62	2.10	2.15	4.29	3.95	2.23	2.85	1.79	2.92	2.29	1.40	2.78	1.71
drink alcoholic beverages	3.14	2.79	3.89	4.10	3.77	4.16	4.03	2.04	4.37	3.71	4.27	2.40	3.27	4.62	2.52	1.56	2.25	3.93	3.30	2.56	4.10	2.26	3.99
use sunscreen	2.52	2.36	2.14	3.41	1.95	2.34	2.37	3.49	1.75	2.66	1.79	2.01	2.30	4.41	3.29	2.30	2.89	2.12	2.88	2.14	1.38	2.85	2.01
accumulate wealth	2.77	2.67	2.16	3.55	2.07	2.36	2.48	2.78	1.88	2.27	1.63	2.33	2.16	3.07	3.16	2.40	2.64	2.05	3.56	2.37	1.53	2.22	1.90
pay credit card balances in full	2.82	2.71	2.11	3.88	1.85	2.45	2.60	3.36	1.86	2.88	1.74	2.22	2.12	3.84	3.33	2.00	2.51	1.99	3.08	3.01	1.45	2.59	2.04
go to bed the same time every night	2.88	2.68	2.23	3.68	2.53	2.66	2.56	3.45	1.96	3.37	2.11	2.51	2.14	4.07	3.42	2.78	3.11	2.23	3.56	2.96	1.67	2.99	2.23
start long-term tasks or assignments early	2.63	2.77	1.89	4.01	2.23	2.62	2.44	3.60	1.85	3.34	1.82	2.27	2.26	4.27	3.62	2.04	2.85	1.97	3.41	2.74	1.60	3.11	2.08
pay enough taxes that they get a refund	2.89	2.75	2.63	3.18	2.37	2.96	2.84	3.32	2.27	3.48	2.08	2.40	2.23	3.95	2.89	2.36	2.62	2.00	2.99	2.36	2.16	2.73	2.49
get speeding tickets	2.89	2.62	3.68	3.03	3.64	4.11	4.11	1.88	4.36	3.90	4.04	3.26	4.08	4.42	2.67	1.92	2.75	3.81	3.19	3.60	4.42	1.77	3.81
take out educational loans	3.58	3.23	2.96	3.67	3.23	3.64	3.52	2.55	3.45	3.15	2.96	2.10	2.66	4.23	2.07	1.59	1.93	2.85	3.08	2.40	2.58	2.19	3.29

Table A3. Correlations between facet ratings.

	Do1. future is uncertain	Do2. goals will change	Do3. helps achieve a goal	Do4. other goals higher priority	Do5. enough goal progress	Do6. considering tradeoffs	Do7. has consequences	Do8. consequences do not affect ME	Do9. has long-term consequences	Do10. feels good now	Do11. will feel good in the future	Do12. following a rule	Do13. requires vigilance, self-control	Do14. actively choose to do it	Do15. have a habit	Do16. happens by chance	Do17. gut instinct	Do18. self-signaling and self-presentation	Do19. product of the environment	Do20. descriptive norm	Do21. prescriptive norm	Do22. not worth thinking about	Do23. following a plan	Not1. future is uncertain	Not2. goals will change	Not3. helps achieve a goal	Not4. other goals higher priority	Not5. enough goal progress	Not6. considering tradeoffs	Not7. has consequences	Not8. consequences do not affect ME	Not9. has long-term consequences	Not10. feels good now	Not11. will feel good in the future	Not12. following a rule	Not13. requires vigilance, self-control	Not14. actively choose to do it	Not15. have a habit	Not16. happens by chance	Not17. gut instinct	Not18. self-signaling and self-presentation	Not19. product of the environment	Not20. descriptive norm	Not21. prescriptive norm	Not22. not worth thinking about
do2	0.43																																												
do3	0.38	0.38																																											
do4	0.09	0.28	-0.04																																										
do5	0.34	0.39	0.75	-0.04																																									
do6	0.34	0.32	0.65	-0.06	0.57																																								
do7	0.30	0.26	0.63	-0.12	0.56	0.60																																							
do8	-0.10	-0.04	-0.43	0.25	-0.40	-0.40	-0.42																																						
do9	0.37	0.34	0.77	-0.13	0.70	0.67	0.67	-0.52																																					
do10	0.17	0.17	0.20	-0.05	0.23	0.14	0.19	0.09	0.15																																				
do11	0.37	0.29	0.74	-0.19	0.66	0.62	0.65	-0.43	0.77	0.26																																			
do12	0.18	0.19	0.19	0.08	0.19	0.17	0.17	-0.02	0.23	0.05	0.18																																		
do13	0.30	0.32	0.60	-0.12	0.56	0.54	0.51	-0.38	0.62	0.11	0.57	0.17																																	
do14	0.07	0.01	0.20	-0.14	0.16	0.26	0.23	-0.06	0.21	0.22	0.26	-0.05	0.18																																
do15	-0.07	-0.02	-0.28	0.15	-0.24	-0.22	-0.23	0.33	-0.27	0.06	-0.28	0.22	-0.18	-0.06																															
do16	-0.05	0.03	-0.27	0.19	-0.21	-0.23	-0.27	0.24	-0.29	-0.18	-0.32	0.03	-0.22	-0.33	0.18																														
do17	0.04	0.06	-0.18	0.10	-0.17	-0.20	-0.17	0.30	-0.21	0.10	-0.18	0.30	-0.16	-0.07	0.45	0.25																													
do18	0.32	0.32	0.53	-0.07	0.53	0.38	0.40	-0.16	0.46	0.34	0.51	0.21	0.42	0.18	-0.13	-0.18	-0.02																												
do19	0.20	0.23	0.03	0.13	0.05	0.00	0.01	0.13	-0.01	0.15	-0.01	0.14	0.01	0.00	0.22	0.11	0.26	0.21																											
do20	0.18	0.23	0.33	0.01	0.34	0.32	0.30	-0.11	0.36	0.22	0.36	0.37	0.24	0.16	0.09	-0.14	0.16	0.31	0.16																										
do21	0.31	0.27	0.72	-0.14	0.68	0.63	0.62	-0.54	0.81	0.15	0.72	0.28	0.61	0.21	-0.22	-0.34	-0.19	0.46	-0.04	0.43																									
do22	-0.06	0.02	-0.23	0.25	-0.18	-0.22	-0.24	0.46	-0.30	0.03	-0.24	0.07	-0.22	-0.03	0.28	0.21	0.28	-0.10	0.06	0.04	-0.27																								
do23	0.36	0.36	0.82	-0.11	0.73	0.64	0.62	-0.45	0.77	0.19	0.75	0.20	0.61	0.22	-0.29	-0.32	-0.21	0.54	0.00	0.34	0.74	-0.25																							
not1	0.18	0.22	-0.02	0.17	-0.05	-0.06	-0.09	0.25	-0.11	0.03	-0.10	0.00	-0.06	-0.03	0.10	0.16	0.15	0.04	0.23	0.02	-0.20	0.11	-0.07																						
not2	0.17	0.27	0.04	0.14	0.03	-0.02	-0.06	0.19	-0.02	0.03	-0.03	0.02	0.02	-0.05	0.03	0.15	0.12	0.14	0.23	0.04	-0.12	0.10	0.01	0.50																					
not3	-0.06	0.00	-0.47	0.23	-0.45	-0.43	-0.41	0.48	-0.55	-0.05	-0.50	-0.14	-0.44	-0.15	0.22	0.30	0.20	-0.26	0.18	-0.22	-0.63	0.25	-0.50	0.36	0.35																				
not4	0.18	0.08	0.22	-0.18	0.17	0.14	0.22	-0.10	0.19	0.15	0.25	0.01	0.16	0.19	-0.04	-0.12	-0.02	0.21	0.09	0.12	0.15	-0.08	0.21	0.17	0.25	0.03																			
not5	-0.03	0.04	-0.40	0.25	-0.40	-0.36	-0.36	0.43	-0.46	-0.07	-0.44	-0.12	-0.38	-0.14	0.19	0.28	0.18	-0.23	0.17	-0.16	-0.54	0.22	-0.44	0.35	0.37	0.72	0.04																		
not6	-0.10	-0.07	-0.40	0.16	-0.40	-0.36	-0.33	0.46	-0.47	-0.04	-0.41	-0.12	-0.37	-0.05	0.20	0.20	0.16	-0.24	0.12	-0.16	-0.55	0.27	-0.42	0.34	0.32	0.61	0.07	0.54																	
not7	-0.09	-0.07	-0.40	0.19	-0.38	-0.35	-0.36	0.41	-0.48	-0.05	-0.44	-0.16	-0.35	-0.11	0.17	0.22	0.16	-0.23	0.14	-0.18	-0.54	0.21	-0.43	0.31	0.26	0.61	-0.01	0.52	0.58																
not8	0.26	0.17	0.42	-0.10	0.37	0.42	0.37	-0.34	0.46	0.01	0.42	0.20	0.34	0.18	-0.13	-0.20	-0.03	0.25	0.01	0.23	0.48	-0.13	0.44	-0.02	0.05	-0.32	0.28	-0.23	-0.27	-0.31															
not9	-0.16	-0.09	-0.57	0.20	-0.54	-0.52	-0.49	0.56	-0.66	-0.03	-0.59	-0.18	-0.49	-0.17	0.23	0.29	0.22	-0.31	0.15	-0.25	-0.74	0.25	-0.58	0.37	0.28	0.72	-0.03	0.64	0.65	0.64	-0.42														
not10	0.01	0.04	-0.21	0.16	-0.20	-0.16	-0.21	0.22	-0.26	-0.17	-0.25	-0.07	-0.13	-0.01	0.18	0.16	0.12	-0.13	0.13	-0.14	-0.31	0.13	-0.22	0.25	0.22	0.41	0.05	0.40	0.33	0.35	0.00	0.36													
not11	-0.17	-0.11	-0.58	0.23	-0.55	-0.52	-0.52	0.53	-0.67	-0.06	-0.62	-0.19	-0.51	-0.15	0.26	0.30	0.23	-0.32	0.13	-0.25	-0.74	0.30	-0.60	0.33	0.27	0.75	-0.04	0.66	0.63	0.64	-0.37	0.79	0.47												
not12	-0.01	0.04	-0.23	0.17	-0.19	-0.17	-0.25	0.28	-0.26	-0.06	-0.27	0.05	-0.16	-0.14	0.24	0.24	0.20	-0.11	0.10	-0.10	-0.27	0.17	-0.24	0.22	0.21	0.28	-0.03	0.27	0.21	0.25	-0.07	0.29	0.20	0.32											
not13	-0.03	0.01	-0.33	0.26	-0.29	-0.32	-0.33	0.32	-0.37	-0.06	-0.39	-0.03	-0.32	-0.18	0.20	0.30	0.21	-0.22	0.11	-0.12	-0.41	0.15	-0.37	0.20	0.19	0.49	-0.09	0.45	0.34	0.35	-0.23	0.44	0.18	0.46	0.22										
not14	-0.08	-0.14	-0.19	-0.01	-0.22	-0.09	-0.13	0.17	-0.20	0.04	-0.17	-0.17	-0.18	0.36	0.09	-0.10	0.02	-0.13	0.03	-0.05	-0.21	0.07	-0.19	0.08	0.09	0.19	0.14	0.19	0.27	0.18	0.05	0.24	0.23	0.25	0.03	0.11									
not15	0.14	0.12	0.24	0.00	0.26	0.22	0.21	-0.21	0.28	0.04	0.20	0.20	0.26	0.10	0.11	-0.04	0.05	0.19	0.08	0.14	0.33	-0.08	0.24	-0.02	0.00	-0.25	0.11	-0.19	-0.23	-0.24	0.32	-0.33	0.02	-0.26	0.16	-0.04	0.00								
not16	0.13	0.14	0.20	-0.01	0.24	0.16	0.15	-0.14	0.22	-0.01	0.17	0.21	0.21	-0.11	0.03	0.14	0.10	0.17	0.02	0.18	0.25	0.02	0.22	0.01	0.03	-0.20	0.03	-0.17	-0.13	-0.18	0.20	-0.24	-0.09	-0.23	0.07	-0.10	-0.25	0.22							
not17	0.10	0.11	0.04	0.06	0.06	0.03	-0.01	0.02	0.05	-0.02	0.01	0.16	0.08	-0.03	0.17	0.09	0.20	0.08	0.16	0.03	0.08	0.04	0.02	0.11	0.08	-0.06	0.02	-0.02	-0.07	-0.02	0.17	-0.07	0.13	-0.03	0.33	0.02	-0.01	0.39	0.20						
not18	-0.03	0.02	-0.43	0.25	-0.39	-0.40	-0.40	0.47	-0.51	-0.02	-0.47	-0.12	-0.39	-0.13	0.27	0.31	0.25	-0.16	0.24	-0.18	-0.58	0.27	-0.46	0.38	0.32	0.66	0.01	0.62	0.54	0.55	-0.26	0.67	0.45	0.70	0.33	0.40	0.20	-0.19	-0.13	0.04					
not19	0.22	0.18	0.09	0.06	0.10	0.04	0.04	0.00	0.06	0.03	0.04	0.06	0.11	-0.06	0.10	0.09	0.14	0.18	0.41	0.04	0.02	-0.08	0.07	0.22	0.26	0.13	0.15	0.14	0.07	0.06	0.11	0.04	0.17	0.06	0.12	0.12	-0.02	0.21	0.11	0.20	0.17				
not20	-0.06	0.02	-0.26	0.19	-0.23	-0.24	-0.29	0.27	-0.33	-0.11	-0.32	-0.09	-0.21	-0.10	0.16	0.26	0.13	-0.13	0.09	-0.25	-0.37	0.16	-0.30	0.25	0.28	0.37	-0.02	0.36	0.29	0.32	-0.07	0.38	0.39	0.42	0.41	0.22	0.11	0.05	-0.05	0.23	0.42	0.09			
not21	-0.21	-0.16	-0.65	0.22	-0.60	-0.55	-0.56	0.55	-0.72	-0.10	-0.66	-0.20	-0.56	-0.18	0.26	0.34	0.23	-0.37	0.10	-0.31	-0.81	0.29	-0.66	0.30	0.21	0.70	-0.10	0.63	0.59	0.59	-0.41	0.79	0.38	0.79	0.34	0.46	0.23	-0.28	-0.24	-0.03	0.66	0.02	0.44		
not22	0.10	0.09	0.20	-0.04	0.19	0.23	0.20	-0.14	0.20	0.00	0.21	0.11	0.16	0.12	-0.02	-0.07	0.01	0.12	-0.03	0.13	0.22	0.06	0.21	0.04	0.08	-0.16	0.27	-0.07	-0.08	-0.19	0.51	-0.21	0.03	-0.16	0.02	-0.16	0.04	0.23	0.17	0.18	-0.11	0.01	0.05	-0.19	
not23	-0.10	-0.02	-0.46	0.25	-0.45	-0.42	-0.43	0.50	-0.55	-0.05	-0.51	-0.14	-0.43	-0.13	0.21	0.33	0.20	-0.27	0.14	-0.21	-0.64	0.26	-0.48	0.36	0.35	0.79	-0.01	0.70	0.61	0.62	-0.32	0.74	0.44	0.76	0.28	0.46	0.21	-0.26	-0.20	-0.07	0.67	0.08	0.38	0.71	-0.16

Factor Analysis. Although each of the 23 facets was chosen to minimize redundancy, Table A3 shows that many of them nonetheless shared considerable variance. It is therefore possible that a dimension reduction technique such as factor analysis could organize the facets to better understand the relationships between related facets. To do this, we used Mplus to perform factor analysis using the higher ratings of the two ratings (max[do, not]) that participants provided for each behavior-facet pair, since we reasoned that the higher rating for doing versus not doing a behavior is more indicative of how relevant that facet is for the behavior. Each facet was therefore rated across 36 behaviors by 73 participants, and we controlled for participant-level variation in two ways, (1) clustering by participant, and (2) including participant factors in a multilevel analysis. Because we are interested in categorizing the behaviors according to the facets, the raters are incidental, so we also performed a third (single-level) factor analysis on the data averaged across participants.

Finally, we tried multidimensional scaling techniques and found similar difficulty in collapsing the 23 facets into meaningful dimensions.

Moderation Analysis. The most important question this study aimed to answer is how these facet ratings relate to observed correlations in Study 1. Are some facets able to differentiate which behaviors were more correlated with time preferences—i.e., do they moderate the relationship between time preference and behavior? Table A4 shows, for each facet, the correlation between the average facet ratings for each behavior and the association between that behavior and time preference in Study 1. We repeated this for the facet ratings for doing each behavior, for not doing each behavior, and for the greater of these two ratings (max column).

As shown in Table A4, facet ratings mostly did not moderate the observed correlations between time preferences and self-reports of behavior. Only two facet ratings for doing a behavior significantly moderated the observed correlation, as did three facet ratings for the greater of the do and NOT do ratings.

On the other hand, nine of the average facet ratings significantly moderated the average expert predicted correlations in Study 2. This larger number of significant moderations with the expert predictions is consistent with the fact that both sets of ratings were derived from individuals providing their intuitions about what drives these behaviors.

More importantly, these cases of statistical moderation do not lend themselves to a clear theoretical explanation. For example, the strongest moderation, for the facet "enough goal progress" is depicted in Figure A1a below. A priori, we expected that behaviors that were rated highly on this factor would be predicted poorly by time preference, yet we found the opposite. In almost every case of moderation, we were left revising our a priori predictions about the direction and/or strength these moderations.

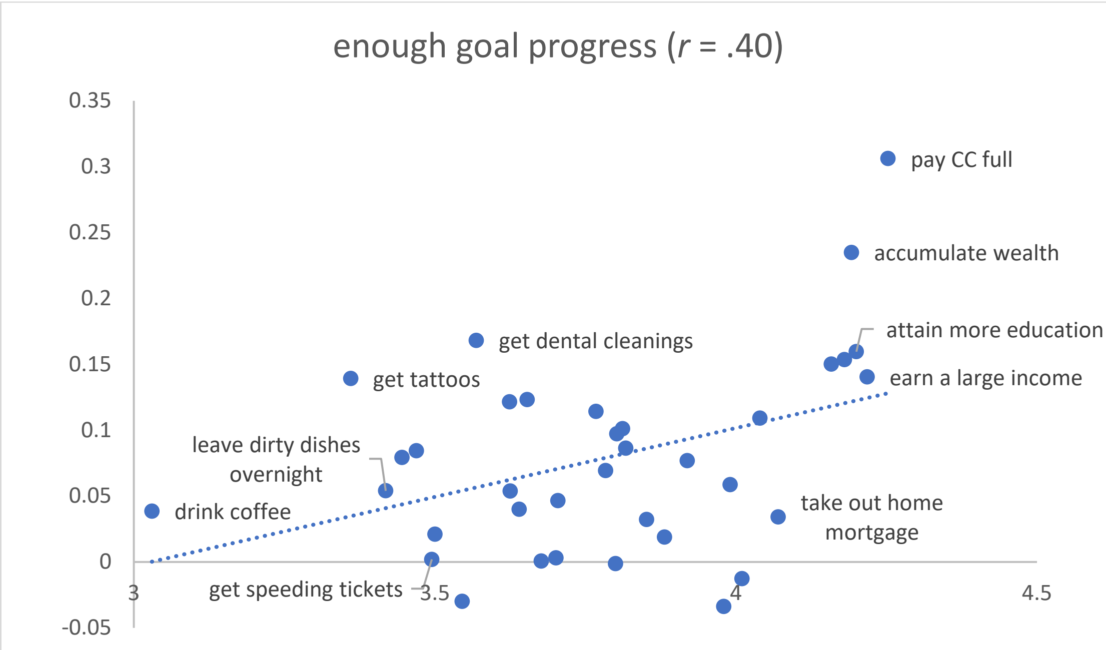
Discussion

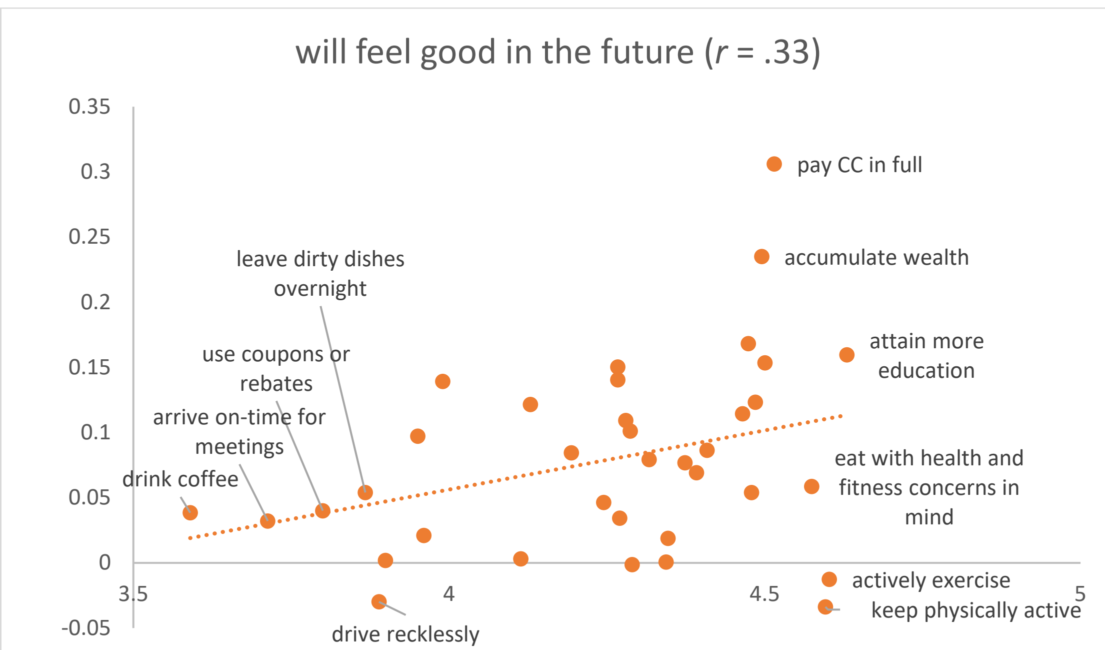
Rather than strain to interpret these findings and report them in the main text, we present the moderation analysis table below and will post the raw data with the hope that someone else can make something of this rich set of norming data, either in a reanalysis of our data or in new, independent research.

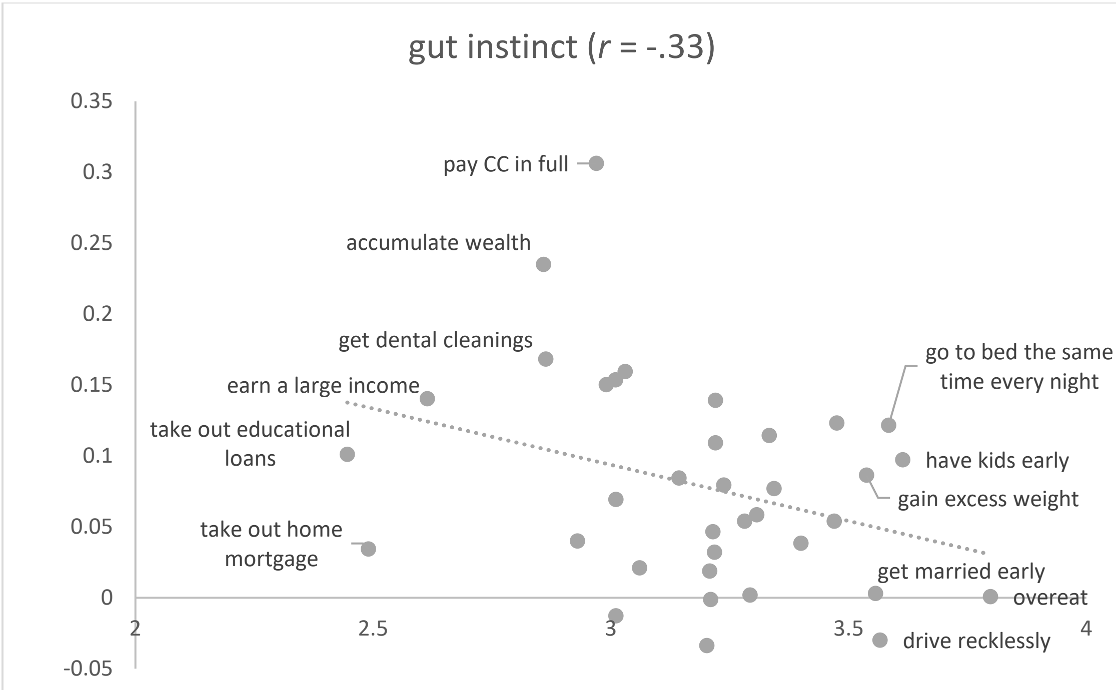
Table A4. Moderation analysis for how much facet ratings for doing or NOT doing each behavior—and each participant's higher of these two ratings for each behavior-facet pair—correlates with the observed correlations between time preference and the behavior in Study 1 and with the average predicted correlations in Study 2. (For 36 data points, $r = 0.33$ is the threshold for significance at 0.05 level, and $r = 0.43$ for the 0.01 level.)

Facet	Correlation with observed correlations			Correlation with expert predictions		
	Do	**Not**	**Max(Do,Not)**	**Do**	**Not**	**Max(Do,Not)**
future is uncertain	0.29	-0.03	0.26	0.45**	-0.04	0.42*
goals will change	0.33*	0.00	0.26	0.38*	-0.01	0.34*
helps achieve a goal	0.24	-0.16	0.25	0.18	-0.11	0.38*
other goals higher priority	-0.06	0.01	0.07	-0.06	0.10	0.18
enough goal progress	0.28	-0.22	0.40*	0.24	-0.17	0.54**
considering tradeoffs	0.21	-0.22	0.04	0.12	-0.18	0.11
has consequences	0.23	-0.16	0.19	0.15	-0.12	0.34*
consequences do not affect ME	-0.26	0.09	-0.29	-0.18	0.04	-0.19
has long-term consequences	0.27	-0.21	0.23	0.20	-0.17	0.38*
feels good now	0.00	-0.24	-0.13	-0.02	-0.16	-0.05
will feel good in the future	0.24	-0.21	0.33*	0.18	-0.19	0.44**
following a rule	0.20	-0.23	-0.03	0.12	-0.10	0.12
requires vigilance, self-control	0.21	-0.24	0.10	0.24	-0.13	0.29
actively choose to do it	-0.14	-0.45	-0.28	-0.18	-0.44	-0.25
have a habit	-0.22	-0.05	-0.15	-0.20	0.03	-0.08
happens by chance	-0.04	0.12	0.07	0.04	0.14	0.18
gut instinct	-0.35*	-0.13	-0.33*	-0.24	-0.03	-0.17
self-signaling and self-presentation	0.19	-0.21	0.07	0.24	-0.17	0.28
product of the environment	-0.12	0.09	-0.04	-0.13	0.22	0.05
descriptive norm	-0.04	-0.10	-0.28	-0.17	-0.04	-0.32
prescriptive norm	0.22	-0.22	0.12	0.15	-0.16	0.28
not worth thinking about	-0.29	-0.03	-0.22	-0.32	-0.22	-0.40*
following a plan	0.26	-0.20	0.26	0.18	-0.18	0.34*

Figure A1. Correlations between average facet ratings and the observed correlations in Study 1 facets with significant moderations.






gut instinct (r = -.33)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
2
2.5
3
3.5
4
pay CC in full
accumulate wealth
get dental cleanings
earn a large income
take out educational loans
take out home mortgage
go to bed the same time every night
have kids early
gain excess weight
get married early
overeat
drive recklessly

SECTION B: ADDITIONAL ANALYSES

Collapsing behaviors by domain

In the manuscript, we collapse behaviors across domains to assess their relationship with time preference. However, previous research offers mixed evidence about whether measures of time preference differ across domains (Chapman & Elstein, 1995; Ioannou & Sadeh, 2016; Bradford et al., 2017)(27)(28)(7). In the present research, we noted that our four largest positive correlations (for "pay credit card balances in full", "accumulate wealth", "get dental cleanings", and "education," *r*s = .31, .24, .17, and .16) span the three domains.

As an additional analysis, we generated 10,000 bootstrapped samples of 10 behaviors without replacement (the same number as number of financial behaviors) and computed the correlations between the mean z-score of the 10 behaviors and time preference. We then repeated the process with 15 and 11 behaviors respectively (corresponding to the number of behaviors in health and personal prudence domains). In all 3 analyses, the bootstrapped 95% confidence intervals included the actual correlations between the mean z-scores of behaviors in a given domain and time preference, suggesting no significant difference between behaviors in these domains and randomly sampled behaviors with respect to their associations with time preference.

Multicollinearity

As shown in Table 4, all of our predictor variables correlated with at least a few other variables. However, none of the correlations were large enough to pose a problem with multicollinearity (as assessed by variance inflation factors) in the multiple regressions we report. That said, as a result of these correlations, the way that variance is partialed across predictors depends on which variables are included in a multiple regression. For this reason, although our main analyses included all 15 covariates, we ran six other multiple regression specifications for robustness. These regressions included time preference plus: (i) demographics, (ii) Big Five personality variables, (iii) financial decision making variables, (iv) demographics and Big Five, (v) demographics and financial decision making variables, and (vi) Big Five and financial decision making variables. We used the same two metrics to compare time preference and the 15 covariates across these seven regression specifications (see tables B1 & B2). Based on median absolute standardized coefficient, age, conscientiousness, Barratt Impulsiveness Scale, financial literacy, and numeracy-CRT always performed better than time preference, across all regression specifications. Gender, openness, agreeableness, neuroticism, propensity to plan, and risk preference always performed worse. Based on numbers of times each variable was a significant predictor of behavior, age, extraversion, and Barratt Impulsiveness scale always outperformed time preference for each regression specification. Openness, agreeableness, neuroticism, propensity to plan, and risk preference always performed worse. So, looking across all the regression specifications we ran, time preference is again near the middle of the pack with respect to predictive validity.

Table B1. Rank ordering of the predictors based on the average number of behaviors they predicted significantly (at the $p < .05$ threshold) across the 7 regression specifications.

Rank	Predictor	Score on the index
1	Barratt Impulsiveness	23.75
2	Extraversion	23
3	Age	22
4	Eduparent	19.75
5	Discchoiceraw	18.57
6	Conscientiousness	18.25
7	Fin. Lit.	17.5
8	Male	17.25
9	Mturk	15.75
10	Numeracy-CRT	15.75
11	Openness	14.25
12	Tightwad-Spenthrift	14
13	Agreeableness	12.75
14	Prop. to plan	12.25
15	Risk Pref.	11.25
16	Neuroticism	9.75

Table B2. Rank ordering of the predictors based on the median absolute beta across the 7 regression specifications.

Rank	Predictor	Score on the index
1	Barratt Impulsiveness	0.12
2	Age	0.09
3	Extraversion	0.08
4	Conscientiousness	0.08
5	Fin. Lit.	0.07
6	Numeracy-CRT	0.07
7	Eduparent	0.06
8	Discchoiceraw	0.06
9	Male	0.05
10	Openness	0.05
11	Mturk	0.05
12	Tightwad-Spenthrift	0.05
13	Agreeableness	0.05
14	Prop. to plan	0.04
15	Neuroticism	0.04
16	Risk Pref.	0.03

Analyses Controlling for Age, Income, and Gender

Table B3. Standardized beta for time preference from regressions with age, income, and gender as controls, along with the zero order correlations.

Domain	Behaviors	Standardized Beta for Time preference	Zero Order Correlatic for Time preference
FINANCIAL	Pay credit card balance in full	0.29***	0.31***
	Accumulate wealth	0.18***	0.24***
	Save a high percentage of income	0.14***	0.15***
	NOT miss credit card payments	0.14***	0.15***
	NOT accumulate a lot of credit card debt	0.13***	0.11***
	NOT accumulate a lot of educational loan debt	0.09**	0.10**
	Use coupons or rebate offers	0.03	0.04
	Take out a mortgage	-0.01	0.03
	Pay enough taxes to get refund	0.01	0.02
HEALTH	Get dental cleanings	0.12***	0.17***
	NOT smoke or chew tobacco	0.14***	0.12***
	Floss	0.09***	0.11***
	NOT gain excess weight	0.09**	0.09**
	Use sunscreen	0.08**	0.08**
	Monitor the nutritional content of food	0.07*	0.08**
	Get routine checkups (physicals)	0.03	0.07*
	Eat with health & fitness concerns in mind	0.06*	0.06*
	Follow doctor's prescriptions	0.03	0.05
	NOT use recreational drugs	0.04	0.05
	NOT drink coffee	0.07*	0.04
	NOT overeat	-0.01	0.00
	NOT drink alcoholic beverages	0.03	0.00
	Actively exercise	-0.04	-0.01
	Keep physically active	-0.04	-0.03
PERSONAL PRUDENCE	Attain a high level of education	0.12***	0.16***
	NOT get tattoos	0.11***	0.14***
	Go to bed at the same time every night	0.08	0.12***
	Have kids when older	0.04	0.10*
	NOT gamble or buy lottery tickets	0.12***	0.08**
	NOT leave dirty dishes overnight	0.01	0.05
	Arrive on time for meetings	0.01	0.03
	Start tasks and assignments early	-0.02	0.02
	NOT get speeding tickets	0.00	0.00
	Get married when older	-0.03	0.00
	NOT drive recklessly	-0.03	-0.03

Analyses Using DEEP Time Parameter Estimates

Table B4. Behavior correlations and standardized betas (from regressions with 15 other predictors) for three measures of time preference: the 12-item measure, and the beta and delta parameters estimated from DEEP Time.

		Correlations			**Standardized betas**		
	Behaviors	**12-item measure**	**DEEP Beta**	**DEEP Delta**	**12-item measure**	**DEEP Beta**	**DEEP Delta**
FINANCIAL	Pay credit card balance in full	0.31***	0.01	0.28***	0.25***	0.01	0.23***
	Accumulate wealth	0.24***	0.02	0.22***	0.17***	0.02	0.17***
	Save a high percentage of income	0.15***	-0.12***	0.17***	0.13***	-0.09**	0.14***
	NOT miss credit card payments	0.15***	0.04	0.13***	0.07*	0.02	0.07*
	Earn a large income	0.14***	-0.09***	0.12***	0.08**	-0.09***	0.07**
	NOT accumulate a lot of credit card debt	0.11***	-0.05	0.09***	0.09**	-0.05	0.07*
	NOT accumulate a lot of educational loan debt	0.10**	0.10**	0.11***	0.08*	0.13***	0.09**
	Use coupons or rebate offers	0.04	-0.04	0.01	0.04	0	0.03
	Take out a mortgage	0.03	-0.07*	0.02	0.01	-0.04	0.01
	Pay enough taxes to get refund	0.02	0.13***	0.01	0	0.10**	0.01
HEALTH	Get dental cleanings	0.17***	0.03	0.17***	0.13***	0.05	0.14***
	NOT smoke or chew tobacco	0.12***	0.03	0.13***	0.08**	0	0.10***
	Floss	0.11***	0.05*	0.10***	0.12***	0.12***	0.11***
	NOT gain excess weight	0.09**	0	0.11***	0.06*	0.01	0.09**
	Use sunscreen	0.08**	-0.01	0.07*	0.06*	-0.02	0.07*
	Monitor the nutritional content of food	0.08**	0	0.08**	0.07*	0.02	0.09**
	Get routine checkups (physicals)	0.07*	0.01	0.05*	0.07*	0.07*	0.06*
	Eat with health & fitness concerns in mind	0.06*	0.03	0.07*	0.06*	0.03	0.08**
	Follow doctor's prescriptions	0.05	0.03	0.03	0	0.05	0
	NOT use recreational drugs	0.05	-0.01	0.05	0.02	-0.01	0.03
	NOT drink coffee	0.04	0.01	0.04	0.04	0	0.04
	NOT overeat	0	-0.04	0	-0.07*	-0.04	-0.06
	NOT drink alcoholic beverages	0	-0.06*	-0.02	0.03	-0.04	0.01
	Actively exercise	-0.01	-0.09**	-0.01	-0.01	-0.02	0
	Keep physically active	-0.03	-0.04	-0.05	-0.02	0	-0.03
PERSONAL PRUDENCE	Attain a high level of education	0.16***	0.06*	0.18***	0.03	-0.03	0.06*
	NOT get tattoos	0.14***	-0.01	0.13***	0.07*	-0.03	0.05
	Go to bed at the same time every night	0.12***	0	0.11***	0.06*	-0.02	0.06*
	Have kids when older	0.10*	0	0.15***	0.01	-0.08*	0.06
	NOT gamble or buy lottery tickets	0.08**	0.13***	0.09**	0.07*	0.11***	0.08**
	NOT leave dirty dishes overnight	0.05	-0.13***	0.04	0.04	-0.07**	0.03
	Arrive on time for meetings	0.03	-0.05	0	0.01	-0.04	0
	Start tasks and assignments early	0.02	-0.09**	0	0.01	-0.02	0
	NOT get speeding tickets	0	-0.01	0.01	-0.04	-0.05	-0.01
	Get married when older	0	-0.02	0.06	-0.01	-0.02	0.04
	NOT drive recklessly	-0.03	-0.11***	-0.05	-0.03	-0.06*	-0.04

SECTION C: SURVEY MATERIALS

Below, we present the:
1) Questions used for behaviors along with scores associated with each response.
2) Any special scoring, including transformations if applicable.
3) Question source.

FINANCIAL BEHAVIORS

1. Pay credit card balance in full:

Question: If you have any credit cards, over the past two years, how often have you paid your credit card bill in full, as opposed to paying less than the full amount? (Paying in full means carrying no debt to the next month's bill.)

- Never pay in full (1)
- Rarely pay in full (2)
- Pay in full about half of the time (3)
- Usually pay in full (4)
- Always pay in full (5)
- I do not have any credit cards (coded as missing)

Source: Chabris et al. (2008)

2. Accumulate wealth:

Questions: Compared to your friends who are close to you in age, how much wealth have you accumulated? (Wealth includes retirement savings, stocks, bonds, and mutual funds you own, money in bank accounts, the value of your home minus the mortgage, etc.)

- Less than all (1)
- Less than most (2)
- About average (3)
- More than most (4)
- More than all of my friends (5)

Compared to the other members of your family in your generation—brothers, sisters, and cousins close to your age-how much wealth have you accumulated?

- Less than all (1)
- Less than most (2)
- About average (3)

- More than most (4)
- More than all the other members of my family in my generation (5)

Scoring: average of responses on both questions
Source: Chabris et al. (2008)

3. Save a high percentage of income

Question: Over the past three years, what percentage of your income have you saved? (Please include savings into retirement plans and any other form of savings that you do.)
[Free Quantitative Response from 0 to 100]

Source: Chabris et al. (2008)

4. NOT miss credit card payments

Question: If you have any credit cards, over the past two years how many times were you charged a late fee for making a credit card payment after the deadline?

- Never (1)
- 1-2 times (2)
- 3-4 times (3)
- 5 or more times (4)
- I do not have any credit cards (coded as missing)

Scoring: reverse-scored
Source: Chabris et al. (2008)

5. Earn a large income

Question: What is your **personal** (not household) pre-tax income?

- Less than $10,000 (1)
- $10,000 to $19,999 (2)
- $20,000 to $29,999 (3)
- $30,000 to $39,999 (4)
- $40,000 to $49,999 (5)
- $50,000 to $59,999 (6)
- $60,000 to $69,999 (7)
- $70,000 to $79,999 (8)
- $80,000 to $89,999 (9)
- $90,000 to $99,999 (10)
- $100,000 to $109,999 (11)
- $110,000 to $119,999 (12)
- $120,000 to $129,999 (13)
- $140,000 to $149,999 (15)
- $150,000 to $159,999 (16)
- $160,000 to $169,999 (17)
- $170,000 to $179,999 (18)
- $180,000 to $189,999 (19)
- $190,000 to $199,999 (20)
- $200,000 to $209,999 (21)
- $210,000 to $219,999 (22)
- $220,000 to $229,999 (23)
- $230,000 to $239,999 (24)
- $240,000 to $249,999 (25)
- More than $250,000 (26)
- I prefer not to answer (coded as missing)

Scoring: log-scaled – log(x)
Source: Other

6. NOT accumulate a lot of credit card debt

Question: How much credit card debt do you currently have (total, across all of your credit cards)?

- None, I pay off my credit card in full every month (0)
- Less than $100 (.05)
- $100 to $500 (.1)
- $500 to $1,000 (.5)
- $1,000 to $2,000 (1)
- $2,000 to $5,000 (2)
- $5,000 to $10,000 (5)
- $10,000 to $25,000 (10)
- $25,000 to $50,000 (25)
- $50,000 to $100,000 (50)
- $100,000 or more (100)
- I prefer not to answer (coded as missing)

Scoring: log-scaled, and then reverse scored – log(x+1)
Source: Other

7. NOT accumulate a lot of educational loan debt

Questions: Compared to your friends who are close to you in age, how much have you taken out in loans for education (e.g., student loans or loans to cover job training or certification)?

- Less than all of my friends (1)
- Less than most of my friends (2)
- About average with my friends (3)
- More than most of my friends (4)
- More than all of my friends (5)
- Not applicable or none (0)
- I prefer not to answer (coded as missing)

Compared to the other members of your family in your generation—brothers, sisters, and cousins close to your age—how much have you taken out in loans for education (e.g., student loans or loans to cover job training or certification)?

- Less than all members of my family in my generation (1)
- Less than most members of my family in my generation (2)
- About average with members of my family in my generation (3)
- More than most members of my family in my generation (4)
- More than all members of my family in my generation (5)
- Not applicable or none (0)
- I prefer not to answer (coded as missing)

Scoring: average of responses on both questions, then reverse-scored; participants selecting "Not applicable or none" for either of these two questions were excluded to calculate the correlation.
Source: Other

8. Use coupons or rebate offers

Question: To what extent do you use coupons or rebate offers when you shop?

- Never (1)
- Rarely (2)
- Sometimes (3)
- Usually (4)
- Always (5)
- I prefer not to answer (coded as missing)

Source: Other

9. Take out a mortgage

Questions: Compared to your friends who are close to you in age, how much money have you taken out for mortgage(s) to buy a home or homes?

- Less than all of my friends (1)
- Less than most of my friends (2)
- About average with my friends (3)
- More than most of my friends (4)
- More than all of my friends (5)
- Not applicable or none (0)
- I prefer not to answer (coded as missing)

Compared to the other members of your family in your generation—brothers, sisters, and cousins close to your age—how much money have you taken out for mortgage(s) to buy a home or homes?

- Less than all members of my family in my generation (1)
- Less than most members of my family in my generation (2)
- About average with members of my family in my generation (3)
- More than most members of my family in my generation (4)
- More than all members of my family in my generation (5)
- Not applicable or none (0)
- I prefer not to answer (coded as missing)

Scoring: average of responses on both questions; participants selecting "Not applicable or none" for either of the above two questions were excluded to calculate the correlation.
Source: Other

10. Pay enough taxes to get refund

Question: If you owe taxes on your income or salary, how much do you withhold (pay) with each paycheck?

- None. I pay my taxes quarterly or at the end of the year. (0)
- A lot less than enough. I owe a lot of taxes at the end of the year. (1)
- Less than enough. I owe some taxes at the end of the year. (2)
- Just enough. I pay enough taxes and get a small or no refund at the end of the year. (3)
- More than enough. I get some tax refund. (4)
- A lot more than enough. I get a big tax refund. (5)
- Other – text entry box (coded as missing)
- Not applicable. I don’t pay taxes (coded as missing)
- I prefer not to answer (coded as missing)

Source: Other

HEALTH BEHAVIORS

11. Get dental cleanings

Question: How often do you visit your dentist for a check-up?

- Never (1)
- Less than once every two years (2)

- Once every two years (3)
- Once per year (4)
- Two or more times a year (5)

Source: Chabris et al. (2008)

12. NOT consume nicotine

Question: How would you describe your intake of nicotine—how often do you consume it?

- Never (1)
- Rarely (2)
- Monthly (3)
- Weekly (4)
- Daily (5)
- More than once a day (6)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Reimers et al. (2009)

13. Floss

Question: How often do you floss your teeth?

- Rarely or never (1)
- Once or twice each week (2)
- Most days each week (3)
- At least once per day (4)

Source: Chabris et al. (2008)

14. Not gain excess weight

Questions: What is your height (in feet and inches)? [Free Numerical Response]
What is your weight (in pounds)? [Free Numerical Response]

Scoring: Based on responses on the above two questions, participants' Body Mass Index (BMI) was calculated, and then the BMI was reverse-scaled.

Body Mass Index = (weight in kilograms)/(height in meters)2

$= kg/m^2$
1 pound ~ 0.45 kg
1 inch ~ 0.03 m

Source: Reimers et al. (2009)

15. Use sunscreen

Question: How often do you use sunscreen when exposed to harsh sunlight?

- Never (1)
- Rarely (2)
- Sometimes (3)
- Usually (4)
- Always (5)
- I prefer not to answer (coded as missing)

Source: Other

16. Monitor the nutritional content of food

Question: To what extent do you monitor your diet in terms of caloric, fat, carbohydrate, cholesterol, and/or sodium intake?

- Never (1)
- Rarely (2)
- Monthly (3)
- Weekly (4)
- Daily (5)
- More than once a day (6)
- Every time I eat anything (7)
- I prefer not to answer (coded as missing)

Source: Other

17. Get routine checkups (physicals)

How often do you visit a doctor for routine check-ups (physicals)?

- Never (0)
- Less than once every four years (1)

- Once every two to four years (2)
- Once (or more) a year (3)
- I prefer not to answer (coded as missing)

Source: Other

18. Follow a diet plan

Are you currently following a specific diet plan?

- Yes, very strictly (3)
- Yes, but not very strictly (2)
- No (1)

Source: Chabris et al. (2008)

19. Follow doctor's prescriptions

Question: When your doctor gives you a prescription to fill at the drugstore (excluding birth control), do you follow it exactly (for example, by going to the drugstore, picking up the medication, taking all of the medication on schedule, and finishing the entire prescription)?

- Rarely or never (1)
- Sometimes (2)
- Usually (3)
- Always (4)

Source: Chabris et al. (2008)

20. NOT use recreational drugs

Question: How would you describe your intake of recreational drugs (e.g., marijuana)—how often do you consume them?

- Never (1)
- Rarely (2)
- Monthly (3)
- Weekly (4)
- Daily (5)
- More than once a day (6)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Reimers et al. (2009)

21. NOT drink coffee

Question: How would you describe your intake of coffee—how often do you consume it?

- Never (1)
- Rarely (2)
- Monthly (3)
- Weekly (4)
- Daily (5)
- More than once a day (6)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Reimers et al. (2009)

22. NOT overeat

Question: In a typical week, how often do you eat more than you think you should eat?

- No meals (1)
- Few meals (2)
- Some meals (3)
- Most meals (4)
- Every meal (5)

Scoring: reverse-scored
Source: Chabris et al. (2008)

23. NOT drink alcoholic beverages

Question: How would you describe your intake of alcohol—how often do you consume it?

- Never (1)
- Rarely (2)
- Monthly (3)
- Weekly (4)
- Daily (5)

- More than once a day (6)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Reimers et al. (2009)

24. Actively exercise

Question: How many of those hours represent exercise primarily intended to improve or maintain your health or fitness? (Please answer in hours) [Free Numerical Response]

Scoring: Free response inputs by participants (log-scaled) – log(x+1)
Source: Chabris et al. (2008)

25. Keep physically active

Question: How many hours per week are you physically active (for example, walking, working around the house, working out)? [Free Numerical Response]

Scoring: Free response inputs by participants (log-scaled) – log(x+1)
Source: Chabris et al. (2008)

PERSONAL PRUDENCE BEHAVIORS

26. Attain a high level of education

Question: What is the highest level of education you have completed?

- Some high school or less (1)
- High school diploma or equivalent (2)
- Some college but did not finish (3)
- Currently in college (4)
- Associate's degree or equivalent (5)
- Bachelor's degree or equivalent (6)
- Master's degree (e.g., M.S., M.B.A.) (7)
- Professional degree (e.g., M.D., J.D.) (8)
- Doctoral degree (e.g., Ph.D.) (9)
- Don’t know (coded as missing)

Source: Other

27. NOT get tattoos

Question: How many permanent tattoos do you have, if any?

- None, and no intentions of getting any (0)
- None, but plan to get one or more (0.5)
- Tattoos that are usually covered with clothing (write how many) [Free response *x*]
- Tattoos that are usually visible to others (write how many) [Free response *y*]
- I prefer not to answer (coded as missing)

Scoring: 0, 0.5, or the sum of *x* and *y* (log-scaled and reverse-scored – log(x+1))
Source: Other

28. Go to bed at the same time every night

Question: Do you go to bed the same time every night?

- Never (1)
- Rarely (2)
- Sometimes (3)
- Usually or Always (4)
- I prefer not to answer (coded as missing)

Source: Other

29. Have kids when older

Question: If you have children, at what AGE did you have your first child?

- Not applicable (no children) (*missing*)
- I had my first child at age: [Free Numerical Response]
- I prefer not to answer (coded as missing)

Scoring: For participants that input an age in the text entry, the response was recorded as the age number. For participants selecting Not applicable (no children), the response was initially recorded as 0, but the data were treated as missing. Free-response entries (log-scaled) – log(x)
Source: Other

30. NOT gamble or buy lottery tickets

Question: On average, how many days per month do you gamble money, including visiting casinos, buying lottery tickets, betting on sports, playing poker, etc?

- Never (1)
- Sometimes but rarely (2)

- 2–5 days per month (3)
- 6–10 days per month (4)
- More than 10 days per month (5)

Scoring: reverse-scored
Source: Chabris et al. (2008)

31. NOT leave dirty dishes overnight

Question: How often do you leave dirty dishes overnight?

- Never (1)
- Rarely (2)
- Sometimes (3)
- Usually or Always (4)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Other

32. Arrive on time for meetings

Question: To what extent are you on time to appointments, engagements, or meetings (both personal- and business-related)?

- I am sometimes or often very late (1)
- I am often a little late (2)
- I am sometimes a little late (3)
- I am always on time or early (4)
- I prefer not to answer (coded as missing)

Source: Other

33. Start tasks and assignments early

Question: When given a long-term assignment or task, when do you tend to start it?

- A day or less before the deadline (1)
- A few days before the deadline (3)
- A few days after it's assigned (4)
- Immediately, or on the first day it's assigned (5)

- I prefer not to answer (coded as missing)

Source: Other

34. Not get speeding tickets

Question: How many speeding tickets (or something similar) have you received in the last 5 years?

- 0 (0)
- 1 (1)
- 2 (2)
- 3 or more (3)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Other

35. Get married when older

Question: If you are currently or have been married, at what AGE did you first get married?

- Not applicable (never married) (*missing*)
- I was first married at age: [Free Numerical Response]
- I prefer not to answer (coded as missing)

Scoring: For participants that input an age in the text entry, the response was recorded as the age number. For participants selecting Not applicable (never married), the response was initially recorded as 0 but the data were treated as missing.
Source: Other

36. Not drive recklessly

Question: How often do you drive in a way that your driver's education teacher would consider "reckless"? (for example: driving more than 10mph over the speed limit, speeding up at a yellow light, weaving through traffic, using your phone while driving, not fully stopping at a stop sign, not stopping to turn right on red, etc.)

- Never (1)
- Rarely (2)
- Sometimes (3)
- Usually or Always (4)
- I prefer not to answer (coded as missing)

Scoring: reverse-scored
Source: Other

SECTION D: NOTES ON OUR 12(14)-ITEM MEASURE OF TIME PREFERENCE

The battery consists of one or two parts: a block of choice trials, and optionally, for more information, a block of yoked matching trials. Choice trials 3 and 13 are foils used as attention checks. The easiest index to compute is derived by first checking that respondents passed the attention check trials 3 and 13, and then simply counting the number of later, larger choices that people chose among the remaining 12 items.

Augmented version. For researchers who are interested in computing specific parameters from models of time preference, one way to do so using these items is to present the block of matching responses to participants and use the matching responses that are consistent with each participant's preceding corresponding choice (i.e., not considering her inconsistent responses) to compute, for example, a one-parameter hyperbolic discount rate (k, Mazur, 1987), the discount factor (δ) and present bias (β) parameters assumed by the quasi-hyperbolic model of discounting proposed by Laibson (1997), area under the curve (AUC, Myerson et al. 2010)—an index of total discounting that makes no assumptions about the mathematical form of discounting functions, or the parameters of many other models, subject to the modeling assumptions the researcher is willing to make (for an introduction to many popular models of time preference, see Doyle, 2013).

In our experience, respondents who are responding in conditions promoting high involvement (e.g., students participating in a quiet lab) are able to provide a high proportion (approximately 80%) of matching responses that are consistent with their choices, but respondents who are responding in more variable conditions (e.g., online participants) provide fewer matching responses that are consistent with their choices (with performance as low as 30%). For this reason, we tend to use the 14-choice items measure, exclude participants who fail a foil trial, and often omit the matching responses or, if included, we interpret the matching responses (and parameter estimates resulting from those matching responses) with extreme caution.

Materials

Choice Items

Imagine that you can choose which of two sums of money you'd like to receive, one available sooner and the other available later.
For each choice below, please indicate which of these two payments you would prefer to receive. Imagine that each payment is guaranteed to occur when promised.

1.	\$816 in six months	——OR——	\$860 in nine months
2.	\$213 today	——OR——	\$281 in two years
3.	\$791 today	——OR——	\$777 in one month
4.	\$457 today	——OR——	\$551 in six months
5.	\$1064 today	——OR——	\$1153 in one month
6.	\$600 today	——OR——	\$611 in one month
7.	\$816 in six months	——OR——	\$1028 in one year
8.	\$816 today	——OR——	\$5440 in one year

9.	$840 in six months	——OR——	$10,125 in two and a half years
10.	$777 today	——OR——	$791 in one month
11.	$816 today	——OR——	$860 in three months
12.	$400 in six months	——OR——	440 in one and a half years
13.	$621 in six months	——OR——	$670 in six months
14.	$504 today	——OR——	$524 in one month

Matching Items

Please fill in the blanks so that the sooner payment and later payment are equally attractive to you—in other words, so that you would not care whether you received the sooner or the later payment.

1. $213 today is as attractive as $__________ in two years.
2. $__________ today is as attractive as $791 in one month.
3. $__________ in six months is as attractive as $1028 in one year.
4. $__________ today is as attractive as $524 in one month.
5. $1064 today is as attractive as $__________ in one month.
6. $816 today is as attractive as $__________ in one year.
7. $840 in six months is as attractive as $__________ in two and a half years.
8. $__________ today is as attractive as $611 in one month.
9. $816 today is as attractive as $__________ in three months.
10. $457 today is as attractive as $__________ in six months.
11. $__________ in six months is as attractive as $440 in one and a half years.
12. $__________ in six months is as attractive as $860 in nine months.

Scoring Key for Matching Items

Choice Item	Matching Item	SS	LL	Chose SS, Matching Value	Chose LL, Matching Value
1	12	$816 in 6 months	$860 in 9 months	SS ≤ 816	860 ≥ SS ≥ 816
2	1	$213 today	$281 in 2 years	LL ≥ 281	213 ≤ LL ≤ 281
3	n/a	$791 today	$777 in 1 month	Pass Check	Fail Check
4	10	$457 today	$551 in 6 months	LL ≥ 551	457 ≤ LL ≤ 551
5	5	$1064 today	$1153 in 1 month	LL ≥ 1153	1064 ≤ LL ≤ 1153
6	8	$600 today	$611 in 1 month	SS ≤ 600	611 ≥ SS ≥ 600
7	3	$816 in 6 months	$1028 in 1 year	SS ≤ 816	1028 ≥ SS ≥ 816
8	6	$816 today	$5440 in 1 year	LL ≥ 5440	816 ≤ LL ≤ 5440
9	7	$840 in 6 months	$10,125 in 2.5 years	LL ≥ 10125	840 ≤ LL ≤ 10125
10	2	$777 today	$791 in 1 month	SS ≤ 777	791 ≥ SS ≥ 777
11	9	$816 today	$860 in 3 months	LL ≥ 860	816 ≤ LL ≤ 860
12	11	$400 in 6 months	440 in 1.5 years	SS ≤ 400	440 ≥ SS ≥ 400
13	n/a	$621 in 6 months	$670 in 6 months	Fail Check	Pass Check
14	4	$504 today	$524 in 1 month	SS ≤ 504	524 ≥ SS ≥ 504

Development of the battery

To develop this measure, we presented participants with batteries of smaller, sooner versus larger, later choice items, and fit two parameter item response theory (IRT) models to the batteries of items, keeping track of the two parameters—"difficulty" and "discrimination" that the models fit for each item in each battery. Two parameter IRT assumes that respondents' answers to each item (e.g., marking a correct answer on an exam, or in our case, choosing the larger, later option) are a function of how much of some underlying trait a person possesses (e.g., competency on an exam, or in our case, "patience") and how well an item differentiates between people who have more or less of the underlying trait. Some items are more difficult to answer correctly on an exam, and in our case, some trials require more patience to wait for the larger, later alternative—so, IRT models assign a "difficulty" score for each item. Also, in the framework of IRT models, some items do very well at discriminating between people who are close in competence (in our case, "patience"), whereas for others, a respondent's choice on a given trial has less to do with how competent or patient they are. So, IRT models assign a "discrimination" score for each item. (For much more detail, see DeMars, 2010 and Embretson & Reise, 2013.)

We tested, across 22 batteries of items, a total of 232 different smaller, sooner versus larger, later choice items. The batteries consisted of as few as nine and as many as 18 items, and each was viewed by between 160 and 210 MTurk participants. The smaller, sooner amounts and periods ranged from $40 to $4000 at times from now to 6 months from now, larger, later amounts ranged from $41 to $42,405 periods at times from 1 month to 5.5 years from now. We started with 54 items that varied the smaller amount, the interval between the sooner and later time, and the implied discount factor for the item and, over iterative construction of new batteries of items, we resampled from previously tested items that yielded high discrimination (by keeping track of the median discrimination parameter over multiple batteries of testing; we also tracked the median difficulty parameter) and developed new items that were aimed at maximizing how well these new items would differentiate between more patient and less patient people. So, each new battery was a combination of old, highly-discriminating items, and new test items.

The twelve items retained for the final measure had been included in three to seven prior batteries of items and were selected with three criteria in mind. First was coverage over time periods: We wanted a third of our trials to include only the first time period (now vs. 1 month from now), to exclude the first time period and vary the later time period (6 months from now vs. 9 months, 1 year, 1.5 years, and 2.5 years from now), to include the first time period and vary the later time period (now vs. 3 months, 6 months, 1 year, and 2 years from now). Second was coverage across "difficulty": We wanted a third of our trials to present larger, later choice options that would be chosen by only the most patient people (i.e., high difficulty items), by people near the middle of the range of patience (medium), and by all but the most impatient people (low difficulty items). And third, with those two guiding motivations in mind, we selected items that had the highest median discrimination in the previous testing periods. We ended up with four possible finalist batteries of twelve items that balanced these considerations.

We then tested the final four batteries again with new respondents and selected the battery that had the best information plot—as assumed and fit by IRT, a plot of the overall information in the test. There is a tension between including highly discriminating items (which provide a lot of information about whether a person is more or less patient than a specific point along the continuum of the underlying trait) the location of those highly-discriminating items (ideally, they would span much of the continuum), and including other items and including other items that provide less specific information over a wider range. So, we selected the final battery of twelve based on the IRT information plots of the four finalists, as these plots convey information about the regions of the underlying trait that the test will be most informative about.

Subsequent testing of our final twelve-item measure suggested that researchers are well-served to include "foil" items within the measure (e.g., there are fewer inconsistent responses when people who fail these foil trials are removed). So, we now advocate for the use of a 14-item measure (the original 12 items plus two attention-check foils, which are $791 today vs. $777 in one month and $621 in six months vs. $670 in six months), and for presenting these 14 items in a random order.

References Cited

DeMars, Christine (2010) Item Response Theory: Understanding Statistics Measurement. Oxford University Press, Oxford.

Doyle, John R. (2012) "Survey of time preference, delay discounting models." *Delay Discounting Models (April 20, 2012)*.

Embretson, Susan E., and Steven P. Reise. *Item response theory for psychologists*. Psychology Press, 2000.

Laibson, David. (1997) "Hyperbolic discounting and golden eggs." *Quarterly Journal of Economics* 112(2): 443-477.

Mazur, James E. (1987) "An adjusting procedure for studying delayed reinforcement." M.L. Commons, J.E. Mazur, J.A. Nevin, H. Rachlin (Eds.), *Quantitative analyses of behavior, Vol. 5. The effect of delay and of intervening events on reinforcement value*. Lawrence Erlbaum Associates, Inc. 55-73.

Myerson, Joel, Leonard Green, and Missaka Warusawitharana. "Area under the curve as a measure of discounting." *Journal of the Experimental Analysis of Behavior* 76.2 (2001): 235-243.

SECTION E: SUPPLEMENTAL STUDY

Three groups of participants recruited from Amazon Mechanical Turk rated the 36 behaviors in terms of whether the behavior or its opposite was "better for you in the long run" (*N* = 100→93), "more prudent" (*N* = 100→88), or "more responsible" (*N* = 102→86), as below:

Which of these two is (better for you in the long run) [more prudent] {more responsible}?

- saving a substantial amount of your income so you can use it later
- NOT saving a substantial amount of your income so you can use it later

All thirty seven questions appeared on the same page—we presented the "drink coffee" item twice as an attention check (the number of participants who passed the check appear after the arrows above). The order of the questions was randomized for each participant. In addition, because of the different possible interpretations of "prudent" and "responsible," we gave participants the Cambridge dictionary definitions for these terms, as follows: By "prudent," we mean "showing good judgment in avoiding risks and uncertainties; careful," as in the sentence "It's always prudent to read a contract carefully before signing it." By "responsible," we mean "having good judgment and the ability to act correctly and make decisions on your own," as in the sentence "We want to be responsible citizens." (Note: our "get married when older" and "have children when older" items were presented to these participants as "NOT get married too early" and "NOT have children too early.")

Table E1 shows the results, presented as the proportion indicating that the behavior (rather than its opposite) is what is better for you in the long run, more prudent, or more responsible, sorted by the rate of endorsement. Behaviors that were positively endorsed by fewer than 50% of participants were reverse-scaled in Study 1 analyses.

Table E1. Percentage participants in supplemental study who agreed each behavior was better for you in the long run, more prudent, and more responsible to do than not do.

Behaviors	Long Run	Prudent	Responsible	Average
Keep physically active	98%	94%	99%	97.0%
Actively exercise	99%	94%	98%	97.0%
Get dental cleanings	99%	94%	97%	96.6%
Eat with health & fitness concerns in mind	100%	97%	93%	96.5%
Start tasks and assignments early	98%	93%	95%	95.5%
Follow doctor's prescriptions	96%	93%	97%	95.1%
Monitor the nutritional content of food	96%	94%	95%	95.1%
Attain a high level of education	98%	94%	93%	95.1%
Get routine checkups (physicals)	99%	93%	93%	95.0%
Save a high percentage of income	100%	94%	91%	95.0%
Use sunscreen	97%	93%	94%	94.7%
Arrive on time for meetings	97%	94%	93%	94.7%
Pay credit card balance in full	98%	91%	94%	94.3%
Use coupons or rebate offers	96%	94%	92%	94.0%
Accumulate wealth	97%	88%	97%	93.6%
Go to bed at the same time every night	91%	97%	90%	92.5%
Floss	96%	91%	90%	92.0%
Earn a large income	100%	90%	86%	91.9%
Have kids when older	94%	89%	83%	88.2%
Get married when older	92%	90%	78%	86.7%
Pay enough taxes to get refund	83%	82%	85%	83.2%
Take out a mortgage	49%	51%	53%	51.4%

Reverse-Scaled Behaviors	Long Run	Prudent	Responsible	Average
Drink coffee	45%	32%	38%	38.5%
Accumulate a lot of educational loan debt	28%	38%	41%	35.4%
Get tattoos	23%	16%	20%	19.4%
Drive recklessly	14%	11%	26%	17.0%
Drink alcoholic beverages	16%	13%	16%	15.0%
Use recreational drugs	13%	10%	19%	13.9%
Get speeding tickets	8%	10%	23%	13.7%
Gamble or buy lottery tickets	9%	14%	19%	13.6%
Accumulate a lot of credit card debt	9%	10%	21%	13.3%
Overeat	9%	11%	19%	12.9%
Leave dirty dishes overnight	5%	10%	21%	12.2%
Smoke or chew tobacco	6%	10%	20%	12.1%
Gain excess weight	8%	8%	19%	11.4%
Miss credit card payments	5%	11%	16%	11.0%