Financial Literacy
and Personal Finance
Research Network

G53 Network Working Paper Series

Working Paper 2022-3

Evaluating Deliberative Competence: A Simple Method with an Application to Financial Choice

Sandro Ambuehl, B. Douglas Bernheim, and Annamaria Lusardi

MARCH 2022

www.g53network.org

Evaluating Deliberative Competence: A Simple Method with an Application to Financial Choice

Sandro Ambuehl, B. Douglas Bernheim, and Annamaria Lusardi*

March 2022

Abstract

We examine methods for evaluating opportunity-neutral interventions designed to improve the quality of decision making in settings where people imperfectly comprehend consequences. In an experiment involving financial education, conventional outcome metrics (financial literacy and directional changes in behavior) imply that two interventions, one with practice and feedback, one without, are equally beneficial even though only the first reduces average bias. We trace these evaluative failures to violations of implicit assumptions. We propose a simple intuitive outcome metric that properly differentiates between the interventions, and that is robustly interpretable as a measure of welfare loss even when consumers suffer from other biases.

*Ambuehl: University of Zurich, Department of Economics, Blüemlisalpstrasse 10, 8006 Zürich, Switzerland, sandro.ambuehl@econ.uzh.ch. Bernheim: Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, bernheim@stanford.edu. Lusardi: The George Washington University School of Business, 2201 G Street, NW, Suite 450E, Washington, DC 20052, alusardi@email.gwu.edu. This paper incorporates some material contained in previous working papers entitled "Financial education, financial competence, and consumer welfare" and "A method for evaluating the quality of financial decision making, with an application to financial education". We thank Stefano DellaVigna (the editor), Charles Sprenger, Steven Sheffrin, Glen Weyl, our referees, and participants at various conferences and seminars for helpful comments and suggestions. Fulya Ersoy and David Zuckerman provided excellent research assistance. We appreciate funding from the TIAA-CREF, the Alfred P. Sloan Foundation, and the Department of Economics at Stanford University. Experiment A was approved in Stanford IRB protocol 29615, Experiment B in University of Toronto REB protocol 34511.

1 Introduction

Mounting evidence documents the prevalence of low-quality decision making in an assortment of policy-relevant domains, such as household finance (Beshears et al., 2018), health insurance (Loewenstein et al., 2013), and the consumption of durable goods (Grubb and Osborne, 2015; Allcott and Taubinsky, 2015). An important class of corrective policies seek to address the resulting inefficiencies by influencing consumers' decisions without changing their opportunities.[1] Examples include (1) programs that seek, through education and training, to equip decision makers with the knowledge and skills necessary to process and interpret naturally occurring information, (2) efforts to "nudge" people toward better decisions through informational labels, disclosures, and other alterations in the framing of their decision tasks (also known as *choice architecture*, Thaler and Sunstein, 2009) and (3) measures that influence decision making by altering the nature or frequency of personal interactions with professional advisers, family members, or peers. Evaluating the welfare effects of such policies presents challenges. For example, if we proceed from the premise that certain decisions are misguided, we cannot simply defer indiscriminately to choices as in standard welfare economics. The purpose of this paper is to devise a simple method of evaluation that overcome these challenges.

While the approach we develop targets a variety of potential applications, for the sake of concreteness we focus on the welfare effects of opportunity-neutral interventions in the domain of personal financial decisions. Three features of this setting make it ideal for our purposes. First, because financial illiteracy is pervasive (see, e.g., Hastings et al., 2013; Lusardi and Mitchell, 2014), poor decision making is likely widespread. Second, it is easy to imagine opportunity-neutral interventions could potentially improve the quality of financial decision making. Third, financial education is already the subject of a vast academic literature (see Fernandes et al., 2014; Kaiser et al., 2021, for reviews), and has become a ubiquitous public policy tool.[2]

Two methods for evaluating financial education interventions dominate the literature (see Hastings et al., 2013; Lusardi and Mitchell, 2014; Beshears et al., 2018; Kaiser et al., 2021, for reviews). The first investigates whether an intervention changes behavior in a direction that counteracts a suspected bias. For example, on the assumption that people save too little, some studies ask whether financial education leads them to save more (e.g. Bernheim et al., 2001; Bernheim and Garrett, 2003; Cole and Shastry, 2010; Bruhn et al., 2014, 2016).[3] A second approach uses exam-style questions to assess comprehension of financial principles (Mandell, 2009; Mandell and Klein, 2009; Walstad et al., 2010; Carpena et al., 2011; Collins, 2013; Heinberg et al., 2014; Lusardi et al., 2015; Bruhn et al., 2014, 2016, e.g.). Both methods are easy to implement, but neither measures the quality of decision making directly, and the literature recognizes some of their limitations.[4] Still, if education both improves financial literacy and shifts behavior

[1] Other corrective policies alter consumers' opportunities, usually through taxes and subsidies (see for example Bernheim and Rangel, 2009; Gruber and Kőszegi, 2004; Farhi and Gabaix, 2020).

[2] Most OECD member countries, as well as India and China, have implemented policies to enhance financial education. Altogether, these initiatives seek to reach more than five billion people in sixty countries (OECD, 2015). The European Commission recently identified financial literacy as a key objective in its *capital markets union 2020 action plan* (European Commission, 2020) and, in cooperation with the OECD, has launched the development of a financial competence framework for its 27 member states (OECD, 2021). Additionally, in 2013, private non-profit organizations spent nearly half a billion dollars on the provision of financial education in the US alone (CFPB, 2013).

[3] Other directional biases include the tendency to choose low-deductible health insurance plans (Sydnor, 2010; Loewenstein et al., 2013), naïve diversification (Benartzi and Thaler, 2001), and investors' tendency to exit the market in downturns (Bennyhoff and Kinniry Jr, 2011).

[4] For example, in some instance, improvements in financial knowledge appear to leave behavior unchanged; see Bhargava and Conell-Price (2020). In other domains, people do not always employ the knowledge the possess; see Chetty et al. (2009), Taubinsky and Rees-Jones (2018).

in a presumptively beneficial direction, there is an understandable tendency to infer that decisions change for the right reasons, and consequently that consumers are better off.

The first part of this paper (Section 2) explores these evaluative metrics in the context of a laboratory experiment, and then develops our alternative. Experimental methods are widely used in research on the quality of decision making both in the lab and in the field because they provide the ability both to control the environment tightly and to collect the type of detailed measurements required for rigorous inference about welfare (see, e.g. Bernheim and Taubinsky, 2018; Beshears et al., 2018, for reviews).[5] We focus on a fundamental concept in personal finance: compound interest. Some of our subjects receive financial education pertaining to compound interest, while others do not. We use two educational interventions that closely follow a popular text on investing and differ from each other only in that one provides practice with individualized feedback, while the other does not. We elicit subjects' reservation prices for interest-bearing investments (*complex framing*), and for each corresponding future payment stated transparently (*simple framing*).[6] Given the well-documented tendency for people to underestimate exponential growth (*exponential growth bias*, Wagenaar and Sagaria, 1975; Eisenstein and Hoch, 2007; Stango and Zinman, 2009; Almenberg and Gerdes, 2012; Levy and Tasoff, 2016), there is a strong presumption that effective education would increase the demand for interest-bearing assets. Subjects also take an incentivized test on compound interest that requires them to apply, rather than merely to recite, the principles they learned.

According to both conventional metrics, the two interventions are effective: they improve performance on the exam-style questions, and they shift valuations in the presumptively beneficial direction. Moreover, in each case, their effects are nearly identical. Such analyses would therefore appear to imply that the basic intervention significantly improves the quality of decision making, and that the addition of practice and feedback neither amplifies nor diminishes these benefits.

Next, we demonstrate that the inferences one would draw from the conventional metrics are, in fact, misleading. Because subjects performed equivalent valuation tasks with both complex and simple framing, we can assess the magnitudes of the biases resulting from misapprehension of compound interest. A subject who understands interest compounding and who applies that knowledge when making decisions will exhibit the same valuation for an interest-bearing investment as for the future payment it yields. For subjects with poor comprehension, valuations will differ across the frames. Using these data, we show that the two interventions have profoundly different effects (see in particular Figure 3). Without practice and feedback, the basic intervention increases subjective valuations of interest-bearing assets across the board, even though subjects' initial biases are highly heterogeneous. It therefore leads some subjects to overcorrect for their initial biases, and it yields greater bias among those whose initial estimates of exponential growth

[5] A reasonable question is whether, with larger real-life stakes, people might deliberate more carefully and at greater length, and might be more likely to employ analytic tools or seek advice. In practice, most people make financial decisions without qualified assistance (Bernheim, 1998; Benartzi and Thaler, 1999; Lusardi and Mitchell, 2011), spend less than an hour making retirement saving decisions (Benartzi and Thaler, 1999), fail to develop explicit financial or retirement plans (Lusardi and Mitchell, 2011), rely on simple rules of thumb (Bernheim, 1994), and respond to non-substantive nudges (Madrian and Shea, 2001; Bernheim et al., 2015). Notably, the fraction of subjects reporting that they did not seek advice in our experiment (three quarters) matches experience in the field (Lusardi and Mitchell, 2011). See also Enke et al. (2021) for general evidence that large stakes do not remove cognitive biases.

[6] Specifically, we present interest-bearing investments as follows: "*We will invest $a in an account with X% interest per day. Interest is compounded daily. We will pay you the proceeds in t days.*" We present the corresponding future payment as follows: "*We will pay you $r in t days.*"

are roughly correct or excessive.[7] In contrast, with practice and feedback, the intervention reduces the degree to which subjects underestimate compound interest without increasing the prevalence or degree of overestimation. The benefits of education are therefore unambiguous with practice and feedback, but not without it.

The temptation to conclude, incorrectly, that the basic intervention (without practice and feedback) is unambiguously beneficial arises from the implicit assumption that, when an improvement in knowledge accompanies a directionally appropriate shift in behavior, the first causes the second. Further analysis involving two additional treatments falsifies that assumption. For one treatment, we remove most of the substantive material from the intervention and maintain only its rhetorical elements (e.g. "Albert Einstein is said to have called compound interest the most powerful force in the universe"). The resulting intervention largely reproduces the behavioral effect of the full intervention, but has almost no effect on measured comprehension. For the second treatment, we retain the substantive material and remove the rhetoric. That treatment roughly reproduces the effect of the original intervention on measured comprehension, but largely fails to alter behavior.[8] Thus, without practice and feedback, the effects of the educational intervention on test performance arise mainly from the substantive elements of instruction, but its behavioral effects arise mainly from the rhetorical elements. These findings explain why the intervention can enhance test performance and shift behavior in the "desired" direction without improving the quality of decision making overall.

Our critique of the conventional evaluative metrics points to a simple alternative: compute the average distance between simply and complexly framed valuations in equivalent tasks (our proposed measure of *deliberative competence*). This metric is the absolute value (or square) of the *price-metric bias* studied in Behavioral Public Economics (see Bernheim and Taubinsky, 2018). In the absence of second-best considerations, the price-metric bias is interpretable as the ideal price correction (tax or subsidy). As we explain, it has another interpretation that renders it useful for welfare analyses of opportunity-neutral policies: intuitively, the absolute size of the bias gauges the dollar denominated harm the consumer might suffer due to errors resulting from misunderstanding her options. Additional analysis of the data from our experiment reveals that our measure of deliberative competence captures the essential patterns that conventional evaluative metrics miss: it implies that the intervention is effective in experiment A, where it includes practice and feedback, but is not effective in experiment B, where these elements are absent.[9]

The second part of this paper (Section 3) supplements our intuitive justifications with formal analysis. Our objective is to determine whether deliberative competence, as we define it, is robustly interpretable as a measure of the welfare loss resulting from poor decision making (rather than merely as a measure of bias), and whether this interpretation survives consideration of biases outside the scope of analysis that may raise second-best issues (in the sense of Lipsey and Lancaster, 1956). Our approach proceeds in the spirit of recent research in Behavioral Public Economics, in that it evaluates naturalistic choices based on the decisions consumers would make if they properly understood the

[7] Goda et al. (2019) and Levy and Tasoff (2017) also document considerable heterogeneity with respect to the perceived benefits of compounding. Relatedly, Harrison et al. (2020) finds that increased take-up of seemingly beneficial products (index insurance) can lead to welfare losses due to subject heterogeneity.

[8] Related dissociations have been observed in other contexts. Enke and Zimmermann (2019) shows that many people tend to neglect correlations in decision making, despite knowing how to account for them when prompted explicitly. Taubinsky and Rees-Jones (2018) find that many consumers underreact to excise taxes, even though they can properly compute tax-inclusive prices when prompted to do so.

[9] This application underscores an attractive feature of our approach: it yields welfare measures at the *individual level*, and consequently can easily accommodate population heterogeneity with respect to preferences and decision-making defects. As Taubinsky and Rees-Jones (2018) have shown, behavioral welfare analyses that neglect heterogeneity are potentially subject to severe measurement biases.

relationships between options and consequences.[10] Here, the complexly framed tasks are naturalistic, and the simply framed tasks avoid misunderstandings associated with compound interest.

We begin by formalizing the interpretation of deliberative competence as a measure of welfare loss under the restrictive assumption that the consumer's decision-making apparatus involves no defects other than the misunderstandings that the pertinent policy interventions target. This is a strong assumption, and one might entertain legitimate concerns about the possibility that our normative benchmarks, simply framed choices, are susceptible to other biases. In our experiment, considerations such as present bias and excessive suspicion concerning the likelihood the experimenter will renege on promised future payments may suppress simply framed valuations. If the "overshooting" referenced above offsets these biases, one might be tempted to conclude that it is beneficial rather than harmful.[11] We reject this criticism on the grounds that it implicitly addresses the wrong question. Specifically, it introduces additional biases without considering the interventions policy makers might use to address them (e.g., for present bias, commitment options and corrective taxation).

Absent any practical constraints, one would account for second-best concerns by comprehensively analyzing all pertinent biases and all associated policy options comprehensively. But as a general matter (and certainly within the context of analyses such as ours), we lack the requisite comprehensive understanding of human behavior to proceed in this manner. Practical considerations necessitate a compartmentalized approach, which is precisely what one finds in the existing literature on Behavioral Public Economics.

Unfortunately, the literature does not offer a fully satisfactory theoretical framework for conducting compartmentalized welfare analysis. The typical paper focuses on one or two biases at a time and assumes (usually implicitly) that the process generating the observed data is otherwise free of pertinent defects – in effect, that any other biases are either orthogonal or have been resolved. In an analogous framework involving multiple distortions, Lipsey and Lancaster (1956) demonstrated that this approach to welfare analysis is conceptually fraught. Pasting together a collection of such solutions does not generally yield an overall optimum. We fill this gap in the literature by proposing the notion of *idealized welfare analysis*, which acknowledges that other biases infect the available data, but assumes that remedies for them will be forthcoming. In contrast to the usual compartmentalized approach, this perspective allows the analyst to arrive at an overall solution to multiple interacting problems by solving them one at a time.

We then demonstrate that deliberative competence is interpretable as a valid measure of the idealized welfare loss despite the presence of biases outside the scope of the analysis, even when relatively little is known about their nature and magnitudes. This result requires a restrictive separability assumption, but it encompasses many of the types of biases that might infect a study such as ours. We also address other concerns that could in principle limit the applicability of our approach, showing for example that a simple adjustment removes a class of potential confounds.

[10] For foundations, see Bernheim and Rangel (2009), Bernheim (2009), Bernheim (2016), Bernheim and Taubinsky (2018), Bernheim (2021). Bernheim and Taubinsky (2018) also survey applications.

[11] Similar issues arise elsewhere in Behavioral Public Economics. For example, as noted by Blumenthal and Volpp (2010), accurate information about the caloric content of food can exacerbate excessive consumption if people tend to overestimate calories. Likewise, according to Downs et al. (2009), "dietary information is likely to improve self-protective behavior only if existing biases encourage unhealthy eating, but the reverse is equally likely. When it comes to smoking, for example, there is evidence that smokers tend to overestimate the health risks, in which case providing risk information could undermine their motivation to quit."

This paper contributes to several literatures. It adds to the growing literatures in Behavioral Public Economics and Behavioral Welfare Economics (cited above) by showing that, under specified conditions, simple transformations of the price-metric bias are robustly interpretable as measures of the welfare loss from poor decision making, suitable for evaluating opportunity-neutral interventions. With respect to the literature on financial education (also cited above), it documents important limitations of conventional evaluative metrics and offers an easily implemented alternative. It also contributes to a smaller methodological literature on measures of decision-making quality. Alternative measures include the frequencies of (a) either dominated or dominant choices (Ernst et al., 2004; Calvet et al., 2007, 2009; Agarwal et al., 2009; Baltussen and Post, 2011; Choi et al., 2011; Aufenanger et al., 2016; Bhargava et al., 2017), or (b) WARP or GARP violations (Afriat, 1972; Choi et al., 2014; Echenique et al., 2011). Those methods are not easily adapted to the types of decisions that provide the focus for applications such as ours. Consistency-based methods are unlikely to detect problems arising from a consistent misunderstanding (e.g., of compound interest), and dominance-based methods are inapplicable when the best choice depends on preferences. See Bernheim and Taubinsky (2018) for further discussion.

2 Experimental evaluations of financial education interventions

For our application, we investigate the efficacy of two educational interventions aimed at improving intertemporal decision making by enhancing comprehension of compound interest. Interest compounding is one of the most fundamental concepts in finance, and a core topic in the vast majority of courses and books on the principles of financial decision making. We begin by describing the experiment (Section 2.1) and its implementation (Section 2.2). After evaluating the educational interventions based on conventional metrics (Section 2.3), we demonstrate that their implications are untenable and their underlying assumptions invalid (2.4). We then propose, intuitively justify, and implement a simple alternative outcome metric and show that it yields sensible conclusions. Section 3 provides formal theoretical foundations for this alternative approach.

2.1 Design

Our investigation involves two experiments. Both consist of three stages. First, subjects participate in a randomly assigned financial education intervention. Second, subjects complete paired valuation tasks. Third, subjects answer a battery of exam-style questions on compound interest. Table 1 presents a schematic overview. Additional detail concerning each stage follows.

Stage 1: Education intervention. Using material from a popular investment guide, *The Elements of Investing: Easy Lessons for Every Investor* by Malkiel and Ellis (2013), we produced an instructional video on compound interest that covers the topic through a narrated slide presentation.[12] The narration is verbatim from the text (with a few

[12] We chose this approach because existing research indicates that financial education videos are generally more effective than written text (Lusardi et al., 2015).

Table 1: Experiments overview

	Experiment A				Experiment B	
	Treatment	Control	Rhetoric-only	Subst.-only	Treatment	Control
Stage 1: Education intervention						
Topic	Compound interest	Index funds	Compound interest	Compound interest	Compound interest	Portfolio diversification
Rhetoric	✓	✓	✓		✓	✓
Substance	✓	✓		✓	✓	✓
Practice + feedback					✓	✓
Stage 2: Valuation tasks	✓	✓	✓	✓	✓	✓
Stage 3: Exam-style questions						
Interest compounding	✓	✓	✓	✓	✓	✓
Index funds	✓	✓	✓	✓		
Portfolio diversification					✓	✓

minor adjustments), while the slides summarize key points. Stylistically, the video resembles those offered through the educational internet platform *www.khanacademy.org.*[13]

For Experiment A, the video constitutes the entire Treatment. For Experiment B, we split the video into three parts, each of which is followed by practice problems with automated, individualized feedback. If a subject provides an answer that is mistaken but consistent with the calculation of simple interest, the intervention mentions the likely source of the error, explains why the answer is mistaken, and suggests how to get started with a calculation of the correct response. There are six practice questions in total. For the first four, subjects who do not answer correctly in three attempts move on to the next part of the Treatment. For the last two questions, subjects still receive feedback, but they must select the unique correct answer from 13 options before they can continue.[14] We reproduce the complete practice stage in Appendix F.

Subjects in the Control conditions for each experiment participate in interventions that are stylistically similar and that require comparable amounts of cognitive effort as the respective Treatments. In Experiment A, subjects watch a video based on a section about index funds from the same investment guide. Experiment B employs a Control condition concerning portfolio diversification that also contains practice problems with personalized feedback, based on Malkiel and McCue (1985), supplemented with material from Malkiel and Ellis (2013). Neither control intervention mentions compound interest or the time value of money.

In order to isolate the features of the intervention that affect test scores and behavior, Experiment A employs two additional treatments that dissect the main Treatment into its constituent parts. The complete educational module begins with a simple explanation of compound interest illustrated through a standard iterative calculation. The remainder of the text consists of two components:

[13] While the intervention is brief, it is important to bear in mind that financial education in the workplace is also brief. A meta-analysis by Fernandes et al. (2014) finds that the average financial education program involves only 9.7 hours of instruction. That time is divided among a long list of complex topics. For example, Skimmyhorn (2016) reports that a financial education program used by the U.S. military covers compound interest, the focus of our current study, along with a collection of several more complex topics – retirement concepts, the Thrift Savings Plan, military retirement programs, and investments – all within a single two-hour session.

[14] A single subject dropped out of the study when attempting to answer these questions.

(i) An explanation of a simple, memorable, and potentially valuable heuristic, the rule of 72, along with five illustrative applications.[15] The rule of 72 is a method for approximating an investment's doubling period; one can also use it to approximate the growth in an investment's value over a fixed holding period. It states that the percentage interest rate on an investment multiplied by the number of periods required for its value to double equals 72 (approximately).

(ii) Rhetorical material. The section opens with the observation that "Albert Einstein is said to have described compound interest as the most powerful force in the universe." It provides various anecdotes concerning small investments that grew to impressive sums (in some cases millions of dollars) over long time periods. These anecdotes do not include any computations, and hence are not helpful for understanding the mechanics of compound interest. It also explicitly exhorts readers to behave frugally, and characterizes compounding as a "miracle."

Subjects in the *Substance-Only* treatment view a video covering all of the substantive material, but omitting exhortations and atmospheric quotes.[16] In contrast, subjects in the *Rhetoric-Only* treatment view a video containing all of the rhetorical material, as well as the introductory explanation of compound interest, but omitting all material on the rule of 72.

Figure 1: Elicitation of valuations.

	You will get the specified dollar amount within two days from today	We will invest $\$a$ in an account with X% interest per day. Interest is compounded daily. We will pay you the proceeds in t days.
\$20	○	○
\$18	○	○
\$16	○	○
...	...	...
\$2	○	○
\$0	○	○

Notes: The figure displays the first decision list of a round with complex framing. Payment amounts on the left include all even dollar amounts between and including \$0 and \$20. For simply framed choices, the text on the upper right is replaced by "We will pay you R in t days."

Stage 2: Valuation tasks Subjects perform 10 pairs of valuation tasks. Each task elicits an equivalent current dollar value for a reward r to be received in either 36 or 72 days. With *simple framing*, we describe the reward as follows: "We will pay you $\$r$ in t days." With *complex framing*, we describe the same reward in terms of a return on an initial investment, as follows: "We will invest $\$a$ at an interest rate of X% per day. Interest is compounded daily. We will pay you the proceeds in t days." Subjects make two sets of choices pertaining to each future reward, one with simple framing, the other with complex framing. For each frame f, we elicit each subject j's immediate dollar equivalent of

[15] We used this particular investment guide in part because it teaches a useful quantitative heuristic. Some investment guides and educational interventions cover this topic without offering useful quantitative tools.

[16] In cases where it was impossible to remove sentences containing rhetorical material, we substituted neutral language. For instance, the first example of compounding presented in the original text is preceded by the transitional question, "Why is compounding so powerful?" In the Substance-Only-treatment, we substituted the question, "How does compounding work?"

a payment R received in t days, using the iterated multiple price list method with a resolution of \$0.20 (Andersen et al., 2006). Figure 1 presents an example of the first decision list of a round for a complexly framed prospect.

We randomize the order of the valuation tasks at the subject level. Subjects are not told that some of the tasks are substantively equivalent, and they typically do *not* perform equivalent simply and complexly framed tasks consecutively.

Table 2 lists the parameters t, X, a, and R used for the paired valuation tasks. We chose time horizons of 36 and 72 days to simplify applications of the rule of 72.[17] This feature increases the likelihood our study will find beneficial behavioral effects of the Treatment interventions.

Table 2: Decision problems.

Future Reward R	Investment Amount a	Daily Interest Rate X	Number of Doublings
		Duration: 72 days	
\$20	\$10	0.01	1
\$18	\$4.5	0.02	2
\$16	\$2	0.03	3
\$14	\$0.9	0.04	4
\$12	\$2	0.025	2.5
		Duration: 36 days	
\$20	\$10	0.02	1
\$18	\$4.5	0.04	2
\$16	\$2	0.06	3
\$14	\$0.9	0.08	4
\$12	\$2	0.05	2.5

Notes: *Number of doublings* is the number of times the initial investment doubles over the investment horizon according to the rule of 72. Final amounts are calculated using the rule of 72. Exact final amounts differ by no more than \$0.80, except for the 4% interest rate over 72 days, where the rule understates the future value by \$1.16. Our analysis controls for these differences.

Stage 3: Exam-style questions. The final stage of each experiment is an incentivized test consisting of the five questions about compound interest listed in Table 3, as well as five questions about the material covered in the video shown to the control group.[18] The exam-style questions have right and wrong answers which we specifically ask subjects to determine. These features distinguish the exam-style questions from the valuation tasks: for the latter, best choices depend on preferences, and we do not explicitly ask subjects to compute future values. Subjects are aware that test performance is incentivized prior to participating in the education intervention.

Payment All subjects know from the outset that they will be paid according to one randomly selected decision from the entire experiment with 75% probability, and according to their performance on the test with 25% probability, in which case they earn \$1 for each of the ten questions answered correctly. We guarantee subjects that they will receive

[17] We use two different time frames so subjects face a greater variety of decision problems, and hence are less likely to consider successive problems highly similar.

[18] The test questions for the material in the control interventions are in Appendix Table C.2. We randomize the order of all ten test questions at the subject level.

Table 3: Exam-style questions.

Q1. If the interest rate is 10% per year (interest is compounded yearly), how many years does it take until an investment doubles? *7 years, 7.2 years, 7.4 years, 7.8 years, 8 years*
Q2. If somebody tells you an investment should double in four years, what rate of return (per year) is he promising? *15%, 16%, 17%, 18%, 19%, 20%*
Q3. If the interest rate is 7% per year (interest is compounded yearly), about how long does it take until an investment has grown by a factor of four (i.e. is four times as large as it was originally)? *About 5 years* to *about 40 years*, in steps of 5 years.
Q4. Paul had invested his money into an account which paid 9% interest per year (interest is compounded yearly). After 8 years, he had $500. How big was the investment that Paul had made 8 years ago? *$200* to *$400* in steps of $10
Q5. If an investment grows at 8 percent per year (interest is compounded yearly), by how much has it grown after 4 years? *By 30%*, to *by 40%* in steps of one percentage point.

Notes: Questions were presented in random order and intermingled with the five questions concerning material covered in the respective control interventions.

payment within at most two days of the promised date. The instructions emphasize that subjects "should make every decision as if it is the one that counts, because it might be!" In addition to the incentive payment, each subject receives a completion payment of $10.

2.2 Implementation and preliminary analysis

We conducted both experiments through Amazon Mechanical Turk (AMT), an online labor market. Our primary reason for employing this subject pool is that the typical member has a poor understanding of compound interest.[19] Also, this group resembles the target populations for many financial education programs in terms of age and income. We ran Experiment A in eight sessions with a total of 504 subjects during April and May 2014 and Experiment B in two sessions with a total of 401 subjects in October 2018, all on weekday mornings. Each subject participated in exactly one treatment of exactly one experiment. We restricted participation to subjects who resided in the US and were at least 18 years of age. We took several precautionary measures to ensure that subjects were able to view the videos and that they would pay attention. We detail these measures and additional implementation details in Appendix C.

Each experiment begins with a two-and-a-half minute video recording of one of the authors (Bernheim) vouching that we will pay subjects exactly the amount we promise within at most two days of the promised date,[20] and ends with a number of non-incentivized questions about subjects' decision making.[21] We intentionally place no restriction

[19] We opted for this subject pool after pilot experiments at Stanford and at The Ohio State University indicated that the vast majority of student subjects at these universities correctly apply compound interest calculations.

[20] The video invites subjects to click a link to the authors' homepages so they can verify the authenticity of the video. Before participating in the main stages of the experiment, subjects also complete an unincentivized questionnaire concerning demographics, as well as a standard battery of five questions designed to assess financial literacy (Lusardi and Mitchell, 2009). We reproduce the five questions in Appendix Table C.1.

[21] Appendix Table D.4 details the questions and responses by experiment and treatment.

on the use of resources such as calculators, the internet, or personal advice when making decisions, as subjects always have those options when making real-world decisions. Roughly a quarter of our subjects in Experiment A and less than eight percent in Experiment B report using such resources when completing the incentivized test, a fraction that does not vary meaningfully across treatments. Subjects could complete all valuation tasks at their own pace. Subjects in Experiment A took 62 minutes on average (s.d. 22 minutes) to complete the study, while those in Experiment B took 75 minutes (s.d. 25 minutes). The difference in completion times is attributable to the longer interventions in Experiment B.

On average, subjects earned \$22.86 in Experiment A and \$23.14 in Experiment B. In both experiments, earnings include a fixed \$10 participation fee and range from a low of \$10 to a high of \$30.47. In comparison, AMT participants typically earn about \$5 per hour (Mason and Suri, 2012; Hara et al., 2018).

Multiple switching. Any subject with coherent preferences will switch her choice from the immediate payment to the future reward at most once within a single price list. Hence, we informed subjects that "most people begin a decision list by preferring the option on the left and then switch to the option on the right." In practice, 9.7% of subjects (49 of 504) in Experiment A and 13.2% of subjects (53 of 401) in Experiment B switched two or more times in at least one price list. This number does not differ significantly across treatments ($p = 0.85$ and $p = 0.18$ for Experiments A and B, respectively). In laboratory studies of risky choices by undergraduate subjects (such as Holt and Laury, 2002), the corresponding figure typically falls in the range of 10 to 15%. Following the usual convention (see, for example, Harrison, Lau, Rutström and Sullivan, 2005), we focus attention on the 803 subjects who respected monotonicity.

Demographics. While our conclusions about the validity of methods for assessing decision quality do not depend on the demographic composition of our sample, we note that our participants are slightly more financially literate than other pools of U.S. subjects (see Lusardi and Mitchell, 2009, and Lusardi, 2011). There are also differences between the samples used for Experiments A and B, possibly due to changes in the composition of the AMT subject pool.[22] Subjects in Experiment A are on average five years younger, the fraction of women is 6.6 percentage points lower, and their median income is 30% lower. Appendix D.1 lists all demographic details. Due to the differences between the samples, we analyze each experiment separately, without making direct statistical comparisons. Appendix D.2 complements our main analysis by combining data across experiments. It shows that the relative performance of the two compound-interest interventions are driven by their differences rather than by divergences between the subject pools.

Randomization and attrition Randomization into treatments was successful. Of the 68 F-tests we perform to assess the differences in demographic characteristics across treatments and experiments (one for each characteristic and each experiment), two are significant at the 5%-level, and five more are significant at the 10% level (see Appendix D.1). These figures are well within the expected range.

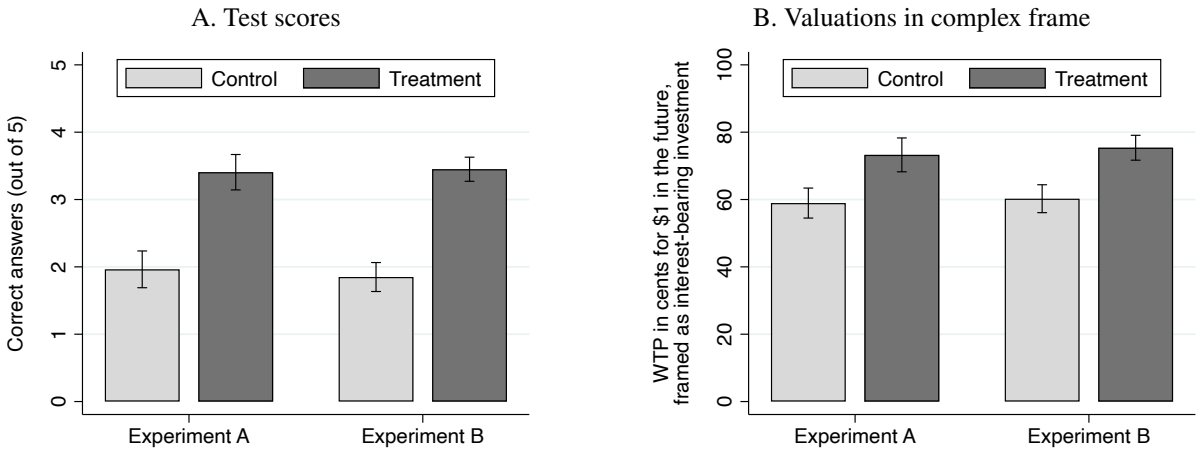
[22] In addition, for Experiment B, we required subjects to have at least 500 completed Human Intelligence Tasks (HITs) and at least a 98% approval rating.

Because we conduct the experiment over the internet, attrition is a possibility. In Experiment A, we find it to be negligible and unrelated to the treatments. Only four subjects who reached the stage at which they may have viewed a treatment video failed to complete the study (compared to 504 who completed it).[23] In Experiment B, two subjects in the Control condition and nine subjects in the Treatment condition began but did not finish the intervention. While these magnitudes are small compared to the sample size of 401 subjects, they differ statistically at the 5%-level.

2.3 Assessment based on conventional outcome measures

In this section, we analyze the effects of the Treatment interventions in each experiment using the two conventional metrics, exam-style questions that test comprehension, and directional effects on behavior. We defer the analysis of the Substance-Only and Rhetoric-Only treatments of Experiment A to Section 2.4.[24]

Figure 2: Assessment by conventional methods.



Notes: Panel A: Fraction of correct answers on the five exam-style questions about compound interest. Panel B: Mean valuation of complexly framed future payments, $r_c^{i,R,t}$, rescaled to range from 0 to 100 (percentage points) rather than 0 to 1. Whiskers indicate 95% confidence intervals. Panel B includes multiple observations per subject; standard errors are clustered on the subject level.

Effects on tested knowledge Panel A of Figure 2 shows how the treatments affect subjects' average scores on the five test questions pertaining to compound interest. In the Control condition of Experiment A, the average subject answers just under two of five, or 39%, of the questions correctly. The Treatment intervention increases the average score by roughly 29 percentage points (1.4 additional correct answers), to 68%.[25] The numbers in Experiment B are remarkably similar, with 37% correct responses in the Control condition and 69% in the Treatment condition.

Columns 1 and 2 of Table 4 summarize these results and present tests of equality between the Treatment and Control conditions using OLS regressions. The treatment effects in both experiments are highly statistically significant ($p < 0.01$). Columns 3 and 4 establish that the improvements in test performance are not due to effects of the Treatment

[23] A larger number of subjects quit before reaching that stage, but that type of attrition is necessarily independent of the treatment, and hence largely innocuous.

[24] All results reported in this section are robust with respect to various statistical controls and alternative specifications. For details, see Appendix D.

[25] See Appendix D.3 for the effects on individual test questions.

Table 4: Performance in exam-style test.

	(1)	(2)	(3)	(4)
	Test score compounding (out of 5)		Test score control module (out of 5)	
Experiment	A	B	A	B
Levels				
Control	1.963***	1.849***	3.284***	3.078***
	(0.140)	(0.110)	(0.114)	(0.101)
Treatment	3.406***	3.450***	2.226***	2.704***
	(0.135)	(0.092)	(0.092)	(0.098)
Difference	1.442***	1.601***	-1.058***	-0.374***
	(0.194)	(0.143)	(0.146)	(0.141)
Observations	215	348	215	348

Notes: Each column displays the coefficients of a separate OLS regression. Standard errors in parentheses. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

on general motivation or attention. If the effects were spurious, we would find comparable effects for subjects' scores on the five test questions concerning the Control interventions. On the contrary, in each experiment, subjects in the Treatment condition perform substantially worse on questions about the respective Control intervention than subjects in the Control condition. The differences of 1.06 and 0.37 questions in Experiments A and B, respectively, are again highly statistically significant ($p < 0.01$). The variation in effect size across experiments is attributable to the fact that the two experiments employ different Control conditions and different test questions about those conditions.

Table 5: Valuations in complexly framed tasks

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	58.949***	60.244***	56.401***	58.181***	61.496***	62.308***
	(2.275)	(2.115)	(2.336)	(2.134)	(2.374)	(2.169)
Treatment	73.261***	75.378***	70.629***	71.747***	75.893***	79.010***
	(2.569)	(1.886)	(2.718)	(1.904)	(2.621)	(2.001)
Difference	14.312***	15.134***	14.227***	13.566***	14.398***	16.702***
	(3.431)	(2.834)	(3.583)	(2.860)	(3.536)	(2.951)
Observations	2,150	3,480	1,075	1,740	1,075	1,740
Subjects	215	348	215	348	215	348

Notes: Valuations in the complex frame, $r_c^{i,R,t}$, in percentage points. Each column displays the coefficients of a separate OLS regression. Standard errors in parentheses, clustered by subject. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

Directional behavioral effects As we have noted, it is well-established that people, on average, tend to underestimate the power of compound interest. Consequently, following the approach sometimes adopted in the literature, one would

deem an intervention potentially welfare-improving if it causes subjects to state higher valuations for investments that involve compound interest. To make these comparisons, we express each valuation as a percentage of the associated future reward. Formally, letting $\tilde{r}_f^{i,R,t}$ denote subject i's valuation for the future reward R received with delay t when presented in frame $f \in \{s, c\}$, we define the normalized valuation as $r_f^{i,R,t} = \tilde{r}_f^{i,R,t}/R$.[26]

Panel B of Figure 2 displays normalized valuations for complexly framed prospects by experiment and condition. Averaged across reward amounts and timeframes, subjects in the Control condition in both experiments regard the interest-bearing investment as equivalent to an immediate payment of about 60% of the future reward amount. Subjects in the respective Treatment conditions state valuations that are about 15 percentage points higher on average. This pattern is precisely what one would expect if the Treatment in each experiment successfully counteracts exponential growth bias. Given the magnitude of the bias documented in the existing literature (Stango and Zinman, 2009), the size of the average treatment effect does not appear to raise concerns about systematic overcorrection. Following the standard approach in the literature, we would thus conclude that the Treatment interventions substantially increase the quality of financial decision making, and do so similarly across the two experiments.[27]

We formalize these findings by regressing normalized valuations in the complex frame, $r_c^{i,R,t}$, on treatment indicators, clustering standard errors by subject. Column 1 of Table 5 shows the results for Experiment A. In the Control condition, subjects value an investment that compounds to one future dollar at 59 cents on average. The Treatment condition raises average valuations by a substantial and statistically significant increment, to 73 cents ($p < 0.01$). The estimates for Experiment B, displayed in column 2, are remarkably similar. Subjects in the Control condition value a future dollar at 60 cents on average, and the Treatment condition increases this valuation by a substantial and statistically significant increment, to 75 cents ($p < 0.01$). Columns 3 - 6 of Table 5 replicate this analysis separately for decisions with 36-day delays and those with 72-day delays, respectively. The same conclusions follow in each instance.

Taken at face value, results based on these conventional metrics seem to imply that both versions of the educational intervention are highly beneficial: they shift behavior in the desired direction, apparently for the right reasons. Moreover, given the similarity of their effects, it is tempting to conclude that the provision of opportunities for practice and individualized feedback yield no incremental benefits.

2.4 Problems with assessments based on conventional outcome measures

Further analysis reveals that assessments based on conventional outcome metrics are misleading in this setting. Moreover, their failures here are traceable to problems that could plausibly arise in other applications.

Financial education interventions seek to ensure that people understand the consequences of their options when they make choices. In our experiment, simple framing simulates perfect education by making those consequences

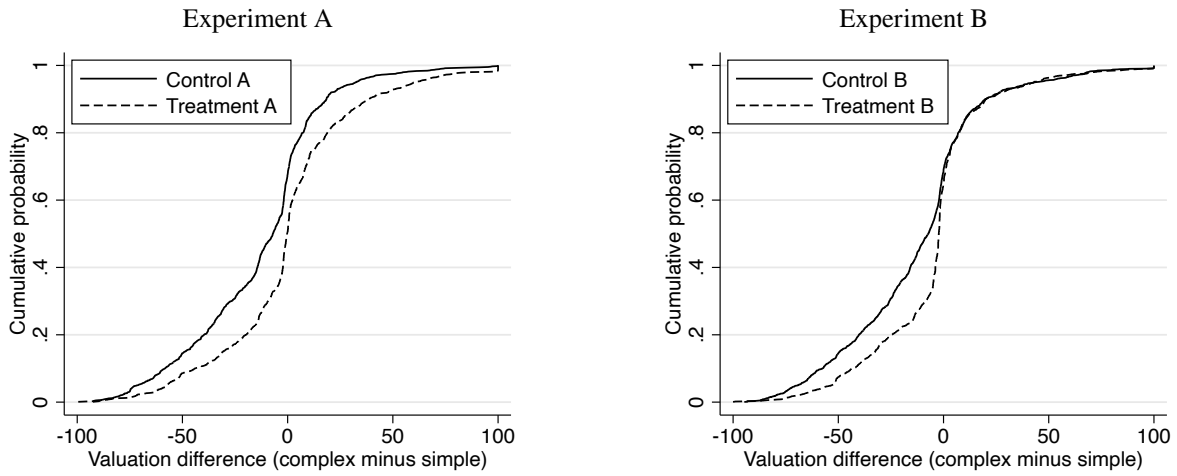
[26] Because we elicit valuations in multiple price lists, they are interval coded. Throughout, we use the midpoint of the pertinent interval for analysis. For further details on the iterated multiple price lists, see Appendix C.

[27] Some prominent studies do not directly observe behavior, but rely on self-reported data about behavior (e.g. Bruhn et al., 2016). Our experiments also elicit such data. At the end of the study, subjects report the number of decision problems for which they explicitly calculated future values, whether they have used the rule of 72 in complexly framed problems, whether they have used it in simply framed problems, and whether they have used external help to make their decisions. As we find in our analysis on directional behavioral effects, our analysis of self-reports suggests that both interventions improve decisions and that they are similarly effective. See Appendix D.4 for details.

transparent. The degree to which financial education closes the gap between complexly and simply framed valuations, $r_c^{i,R,t}$ and $r_s^{i,R,t}$, therefore provides a gauge of the intervention's success.

Figure 3 displays the empirical cumulative distribution function of the valuation difference $r_c^{i,R,t} - r_s^{i,R,t}$ for the Treatment and Control conditions in each experiment. In each Control condition, roughly 60% of subjects are afflicted by exponential growth bias ($r_c^{i,R,t} < r_s^{i,R,t}$), while a nontrivial fraction of subjects are well-calibrated, or overestimate compound interest ($r_c^{i,R,t} \geq r_s^{i,R,t}$).[28]

Figure 3: CDFs of valuation differences.



Notes: CDFs of valuation differences $r_c^{i,R,t} - r_s^{i,R,t}$ across framings for equivalent instruments in the Treatment and Control conditions of Experiments A and B, respectively. Mass to the left of zero indicates underestimation of compound interest (exponential growth bias). Mass to the right of zero indicates overestimation.

An effective intervention would increase valuations in complexly framed tasks for subjects who underestimate compound interest, and would *decrease* them for subjects who overestimate compound interest. As a result, the CDF of the valuation difference $r_c^{i,R,t} - r_s^{i,R,t}$ would become steeper around zero. We observe a much different effect in Experiment A: the Treatment shifts the entire CDF to the right. It increases valuations for complexly framed tasks across the board, irrespective of whether a subject initially underestimates compound interest, estimates it correctly, or overestimates it. This indiscriminate effect helps some subjects but hurts others. In sharp contrast, the effects of the Treatment in Experiment B are sensitive to subjects' initial bias: it increases valuations of complexly framed tasks only for subjects who would have underestimated compound interest, and not for those who would have estimated compound interest correctly, or who would have overestimated it.[29] Unlike the Experiment A Treatment, the Experiment B Treatment is therefore unambiguously successful. Yet neither of the conventional outcome metrics detects this difference.

Why do the conventional metrics fail to reveal the problems with the Experiment A Treatment? With respect to directional effects on behavior, the issues are obvious in light of Figure 3: asking whether the average change in

[28] Appendix D.7 tests and refutes the hypothesis that the overvaluation of complexly framed opportunities is solely attributable to noisy choice. A fraction of subjects in Goda et al. (2019) and in Levy and Tasoff (2016) also overestimate compound interest.

[29] In principle, a change in the valuation difference $r_c^{i,R,t} - r_s^{i,R,t}$ may reflect changes in either $r_c^{i,R,t}$ or $r_s^{i,R,t}$. We perform this decomposition in Section 3.5.

complexly framed valuations counters the typical initial bias ignores heterogeneity in both. Despite the fact that the average complexly framed valuation moves toward the average simply framed valuation, many subjects overshoot, and some apparently double down on their pre-existing biases.

The problems with knowledge assessments are more subtle. Improved performance on batteries of exam-style questions implies better financial decisions if (i) subjects employ that knowledge when making decisions, and (ii) changes in financial choices are only due to the additional knowledge, and do not involve mechanisms such as propaganda or brow-beating.[30] Further investigation falsifies these assumptions.

We evaluate assumptions (i) and (ii) directly by studying the effects of the Rhetoric-only and Substance-only interventions in Experiment A. First, we examine their effects on assessed comprehension by regressing test scores on each of the four treatments in Experiment A, clustering standard errors at the subject level. Column 1 of Table 6 displays the results. The Substance-Only intervention increases performance on test questions concerning compound interest by a similar amount as the full intervention, while the Rhetoric-Only intervention has a much smaller effect.[31] To verify that these results are not due to an effect of the Substance-Only intervention on general motivation, Column 2 uses performance on test questions about the Control intervention as the dependent variable. Indeed, subjects in the Control intervention perform significantly better on these questions than subjects in any other treatment. Accordingly, increases in performance on our assessment of comprehension arise through the expected mechanism (substantive instruction).

Table 6: Separate effects of rhetoric and substance in Experiment A.

VARIABLES	(1)	(2)	(3)	(4)
	Test scores on questions about		Valuations in frame	
	Treatment	Control	Complex	Simple
Levels				
Substance-Only	3.234***	1.945***	62.969***	72.273***
	(0.123)	(0.099)	(2.373)	(2.030)
Rhetoric-Only	2.455***	2.205***	77.538***	77.623***
	(0.146)	(0.095)	(2.785)	(2.119)
Treatment A	3.406***	2.226***	73.261***	72.657***
	(0.135)	(0.092)	(2.566)	(2.139)
Control A	1.963***	3.284***	58.949***	72.255***
	(0.140)	(0.114)	(2.272)	(2.089)
p-value of difference to Control				
Treatment A	0.000	0.000	0.000	0.893
Substance-Only	0.000	0.000	0.222	0.995
Rhetoric-Only	0.015	0.000	0.000	0.072
Observations	455	455	4,550	4,550
Subjects	455	455	455	455

Notes: Each column displays the coefficients of a separate OLS regression. Standard errors in parentheses, clustered by subject. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

[30]This approach also assumes that the effects of knowledge on the quality of decision making are monotonic, which may not be correct if "a little knowledge is a dangerous thing."

[31]The fact that the effect of that intervention is positive may be attributable to the inclusion of an example that illustrates the calculation of compound interest.

Next, we probe the connection between effects on comprehension and effects on behavior. Assumption (i) maintains that behavior responds to knowledge. Because we have just demonstrated that the Substance-Only treatment improves knowledge, this assumption implies that it ought to change valuations in complexly framed decision problems. Assumption (ii) maintains that behavior only responds to features of the intervention that enhance comprehension. Because we have just demonstrated that the Rhetoric-Only treatment has only a modest effect on knowledge, this assumption implies that it ought to change valuations in complexly framed decisions to a much smaller degree than the full intervention.

In fact, we find precisely the opposite. Column 3 displays the coefficients of a regression of complexly framed valuations on treatment indicators. The Substance-Only intervention does not have a statistically discernible effect on these valuations. In contrast, the Rhetoric-Only intervention yields effects comparable to those of the full intervention. Column 4 confirms that the effects of the interventions are largely confined to the complexly framed valuations; simply framed valuations are largely unaffected.[32]

We conclude that the financial education intervention largely impacts knowledge through its substantive component, but largely impacts behavior through its rhetorical component. In other words, behavioral effects are wholly disconnected from effects on comprehension. Thus, neither assumption (i) nor assumption (ii) holds in Experiment A. These findings invalidate inferences about the quality of decision making based on assessments of comprehension in Experiment A, and raise more general questions about the reliability of programmatic evaluations based on measures of knowledge and directional behavioral effects.

2.5 Assessments based on Deliberative Competence

In this subsection, we propose a simple alternative outcome metric, justify it intuitively, and use it to evaluate the various interventions studied in Experiments A and B. We demonstrate formally that this metric robustly captures welfare losses in Section 3.

Divergences between valuations in equivalent simply and complexly framed tasks are of concern because they imply that people may commit errors when deciding whether to purchase instruments. The greater the divergence, the greater the potential harm these errors cause. For a given valuation pair involving future reward R and delay t, we therefore gauge subject i's *deliberative competence* by computing the absolute value of the difference between the normalized complexly and simply framed valuations:

$$d_M^{i,R,t} = -\left| r_c^{i,R,t} - r_s^{i,R,t} \right|.$$

In the language of Behavioral Public Economics, $d_M^{i,R,t}$ is the absolute value of subject i's *price-metric bias* (Bernheim and Taubinsky, 2018). The sign convention used here aids interpretation because it associates larger numerical values

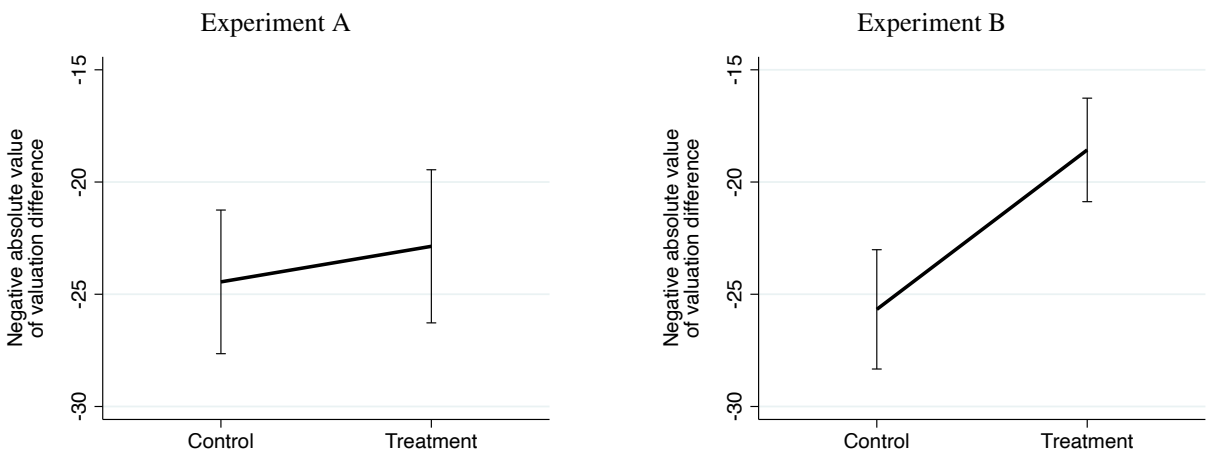
[32] Simply framed valuations are a bit higher for the Rhetoric-Only treatment than for the Control ($p < 0.1$). For the other three treatments, these valuations are essentially identical.

with higher quality decisions.[33] We average over tasks to arrive at the typical magnitude of decision error for subject i, and then average over subjects to obtain an overall measure of deliberative competence.

By definition, the absolute difference between $r_c^{i,R,t}$ and $r_s^{i,R,t}$ measures the consumer's valuation error. However, we are interested in measuring something else: the harm (or welfare loss) that such errors inflict on consumers in settings where they must decide whether to purchase instruments. Our theoretical analysis shows that the same absolute differential also measures the potential harm.

To understand this point, note that harm can befall the consumer either because she purchases an instrument when she should not, or because she fails to purchase it when she should. Suppose first that a complexly framed security paying R at time t is available for a price $p < \tilde{r}_s^{i,R,t}$, and that subject i chooses not to buy it because $\tilde{r}_c^{i,R,t} < p$. In that case, she foregoes a net gain of $\tilde{r}_s^{i,R,t} - p$. Subject to the constraint that the price remains between the two valuations (so the decisions differ), this expression reaches a maximum of $\tilde{r}_s^{i,R,t} - \tilde{r}_c^{i,R,t}$ at $p = \tilde{r}_c^{i,R,t}$. Next suppose that the same security is available for a price $p > \tilde{r}_s^{i,R,t}$, and that subject i chooses to buy it because $\tilde{r}_c^{i,R,t} > p$. In that case, she overpays by $p - \tilde{r}_s^{i,R,t}$. Subject to the constraint that the price remains between the two valuations (so the decisions differ), this expression reaches a maximum of $\tilde{r}_c^{i,R,t} - \tilde{r}_s^{i,R,t}$ at $p = \tilde{r}_c^{i,R,t}$. Thus, in both cases, $d_M^{i,R,t}$ represents the maximum loss to which the subject is exposed, normalized by R. Section 3 formalizes this logic and explores its robustness.

Figure 4: Assessment based on Deliberative Competence.



Notes: Mean Deliberative Competence, $d_M^{i,R,t} = -|r_c^{i,R,t} - r_s^{i,R,t}|$, with $r_f^{i,R,t}$ rescaled to range from 0 to 100 (percentage points) rather than 0 to 1. Whiskers represent 95% confidence intervals with standard errors clustered on the subject level.

Figure 4 shows the levels of deliberative competence across Control and Treatment groups in Experiments A and B, respectively, averaged over all valuation pairs and subjects. In each experiment, deliberative competence in the Control condition is approximately -25% (-24.4% in Experiment A and -25.7% in Experiment B). A key finding is that treatment effects differ dramatically across Experiments A and B. For Experiment A, the Treatment has no discernible effect on deliberative competence. In stark contrast, for Experiment B, the Treatment increases deliberative competence by a

[33] As discussed in Section 3, other distance metrics, such as $-\left(r_c^{i,R,t} - r_s^{i,R,t}\right)^2$, also have valid welfare interpretations. Appendix D.5 reproduces our analysis using the quadratic metric. Results are generally similar.

substantial and statistically significant increment, from -25.7% to -18.6%. Anticipating the results of Section 3, we can say that the Treatment in Experiment B eliminates 27.6 percent of the welfare loss associated with poor comprehension of consequences in the complex frame ($1 - 0.186/0.257 = 0.276$).[34] As a practical matter, this finding suggests that incorporating practice and feedback can have a dramatic impact on the efficacy of financial education, even when the feedback is fully automated and relatively simple. Thus, using deliberative competence rather than conventional outcome metrics reverses our conclusions concerning the role of practice and feedback.

Table 7: Deliberative Competence.

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-24.448***	-25.673***	-24.076***	-25.945***	-24.820***	-25.401***
	(1.633)	(1.357)	(1.744)	(1.428)	(1.682)	(1.393)
Treatment	-22.864***	-18.569***	-22.959***	-18.441***	-22.769***	-18.697***
	(1.741)	(1.177)	(1.799)	(1.208)	(1.886)	(1.253)
Substance-Only	-22.012***		-21.257***		-22.767***	
	(1.433)		(1.422)		(1.620)	
Rhetoric-Only	-19.797***		-20.191***		-19.403***	
	(1.406)		(1.458)		(1.506)	
p-value of difference to Control						
Treatment	0.507	0.000	0.656	0.000	0.418	0.000
Substance-Only	0.263		0.211		0.380	
Rhetoric-Only	0.031		0.088		0.017	
Observations	4,550	3,480	2,275	1,740	2,275	1,740
Subjects	455	348	455	348	455	348

Notes: Each column displays the coefficients of a separate OLS regression of deliberative competence on treatment indicators. Standard errors in parentheses, clustered by subject. Deliberative competence measured as $d_M^{i,R,t} = -|r_c^{i,R,t} - r_s^{i,R,t}|$, with $r_c^{i,R,t}$ rescaled to range from 0 to 100 (percentage points) rather than 0 to 1. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$

To evaluate treatment effects econometrically, we regress $d_M^{i,r,t}$ on a treatment indicator and cluster standard errors by subject, separately for each experiment. Panel A of Table 7 displays the results. Column 1 shows that the Treatment intervention in Experiment A increases deliberative competence by a small and statistically insignificant 1.5 percentage points. Column 2 exhibits the corresponding regression for Experiment B. It shows that while Control subjects display a level of deliberative competence similar to those in Experiment A, the Treatment intervention substantially raises deliberative competence. The increase of 7.1 percentage points is highly statistically significant ($p < 0.01$). Columns 3 and 4 replicate this analysis separately using the set of investments with 36-day delays, whereas Columns 5 and 6 do so for investments with 72-day delays. The same conclusions follow in both cases.

Column 1 also shows that the Substance-Only intervention leaves deliberative competence unchanged, despite its positive effect on knowledge. While the Rhetoric-Only intervention produces a small improvement in deliberative

[34]These estimates are based on equal weighting of observed deliberative competence across subjects and decision problems. We discuss the implications of equal weighting of subjects in Section 3.6.

competence, we know from our earlier analysis that this effect does not flow from improved comprehension. Rather, the motivational rhetoric provides a nudge that proves fortuitously beneficial to a small degree. Columns 3 and 5 show that we obtain similar results for tasks with 36-day and 72-day delays.

3 Theoretical framework and supplemental empirical analyses

In Section 2.5, we provided an intuitive justification for using differentials between simply and complexly framed valuations of the same opportunities to evaluate policy interventions that seek to improve comprehension of the consequences that follow from choices. According to our empirical analysis, such metrics capture important aspects of policy effects that conventional outcome measures overlook. The objectives of this section are to demonstrate that deliberative competence, as we define it, is robustly interpretable as a measure of the welfare loss resulting from poor decision making, even when consumers suffer from additional biases that may raise second-best issues, and to clarify the types of settings to which these conclusions apply.

After describing the setting (Subsection 3.1), we articulate our welfare criteria (3.2). The most novel aspect of this discussion concerns our treatment of decision-making defects outside the scope of analysis. The infeasibility of comprehensive welfare analysis makes compartmentalized analyses, by which we mean the practice of focusing on one or two policies at a time, inevitable. Yet the literature offers no framework for conducting such analyses in a manner that avoids the conceptual problems identified by Lipsey and Lancaster (1956). To fill this gap, we propose and justify the concept of *idealized welfare analysis*. We formalize the welfare interpretation of our deliberative competence measures under the restrictive assumption that the consumer's decision-making apparatus involves no defects other than the misunderstandings that the pertinent policy interventions target (Section 3.3), as well as under less restrictive assumptions that permit us to acknowledge the existence of other biases (Section 3.4). We then discuss a method for detecting and adjusting for a class of potential confounds (Sections 3.5), as well as a strategy for addressing issues involving interpersonal aggregation (Section 3.6), both of which we implement using our experimental data.

3.1 Setting

Following the literature on Behavioral Public Economics (reviewed in Bernheim and Taubinsky, 2018), we distinguish between choice options and the preference-relevant outcomes they induce. For example, people make choices over financial portfolios but care about the patterns of returns those portfolios generate, rather than the portfolios themselves. Likewise, people make choices among over-the-counter medications but care about the health consequences those products yield, rather than the medications themselves. We refer to such choice options as *instruments* because they matter instrumentally rather than intrinsically. The distinction between an instrument and its intrinsically valued consequences is important because it allows for the possibility that people may not understand the implications of their choices, a phenomenon known in Behavioral Public Economics as *characterization failure* (Bernheim and Taubinsky, 2018).

Our method of measuring deliberative competence involves simple binary decisions: whether to purchase an instrument $I \in \mathcal{I}$ at a price p. We assume the consumer cares about current cash, m, and a vector y of future consequences (e.g., specifying available cash at each point in time and state of nature). She starts with current cash m_0; without loss of generality, we simplify notation by setting $m_0 = 0$. If she foregoes the instrument, she anticipates the consequences y_0. In frame f, she interprets the instrument I as offering consequences y_f. In other words, she thinks the incremental effect of the instrument will be $y_f - y_0$. We assume there is a simple way to frame the instrument, $f = s$, that presents its consequences transparently, thereby ensuring the consumer interprets them correctly. Thus, y_s represents the instrument's actual consequences. With naturalistic (complex, non-transparent) framing, $f = c$, she misinterprets the instrument as offering the outcome $y_c \neq y_s$.[35] In other words, she experiences a form of characterization failure. Our object is to evaluate policies designed to mitigate those characterization failures. Where we distinguish between policies explicitly, we use $y_{c\theta}$ to denote the consumer's beliefs in the complex frame about the instrument's consequences when policy θ is in place. In our experiment, the instrument is a claim on a future payment, simply framed tasks transparently specify the future payment associated with the available instrument, complexly framed tasks describe the future payment in terms of compound interest, and the policies of interest target comprehension of compounding.

We posit the existence of a utility function, $V(m, y_f)$, that rationalizes observed choices. By assumption, the characteristics of the instrument and the frame only impact the consumer's choices insofar as they determine her perceptions of consequences. Some of our analysis concerns cases in which the decision maker suffers from additional and possibly unknown valuation biases that the aforementioned simple framing does not remove. We will assume that, with these ancillary biases removed through other means, the utility function $U(m, y_f)$ would rationalize choices in the simple and complex frames. Under these assumptions, $U(m, y_s)$ plays the role of "true preferences."[36] We interpret V and U as indirect utility functions, in the sense that m and y could represent income rather than consumption. We assume throughout that these functions are continuous, that they are unbounded above and below in m, that they have well-defined and bounded first and second derivatives, and that their first derivatives with respect to m are strictly positive.

We define the consumers' reservation price (alternatively, *valuation*) given her interpretation of the instrument's consequences, y_f, assessed according to utility function $u \in \{U, V\}$, as the highest price at which she purchases the instrument (denoted $r^u(y_f)$). This valuation is given by the following condition:[37]

$$u(-r^u(y_f), y_f) = u(0, y_0) \tag{1}$$

[35] As noted in the text, this framework accommodates uncertainty, in the sense that we can interpret y as a vector of state-contingent consequences. Under that interpretation, it would appear that false beliefs must pertain to those consequences and not to their likelihoods. However, if baseline consequences (y_0) are either constant over states or independent of the instrument's returns, we can reformulate a setting with false beliefs about probabilities as one with correct beliefs about probabilities and false beliefs about state-contingent payments, which our framework subsumes. To take a simple example, imagine that a lottery pays x with probability $\frac{1}{3}$ and z with probability $\frac{2}{3}$, but the consumer falsely believes the probabilities are reversed. We can represent this setting as one in which the lottery yields a payment y_s^ω in state of nature $\omega \in \{1, 2, 3\}$, where $y_s^1 = x$, $y_s^2 = y_s^3 = z$, and the states are equally probable; furthermore, the consumer has correct beliefs about the probabilities of the underlying states but incorrectly believes the lottery pays $y_c^1 = y_c^2 = x$ and $y_c^3 = z$.

[36] To be clear, our approach does not require the existence of a single "true preference" relation. As we discuss in Section 3.4.1, it extends directly to cases in which welfare-relevant choices are context-dependent, and consequently admit multiple frame-dependent utility representations, as permitted in the welfare framework of Bernheim and Rangel (2009).

[37] Here, we assume that in case of indifference, the consumer purchases the instrument. Existence of r^u follows from continuity, monotonicity, and unboundedness.

3.2 Framework for welfare analysis

We consider a social planner who seeks to quantify the utility loss a consumer sustains from characterization failures in the domain targeted by the policies we aim to evaluate. The planner uses a utility function u (the *evaluation standard*) for this purpose. For reasons that will become apparent in Section 3.4, we allow for the possibility that u is either U or V.

3.2.1 Welfare criteria

We evaluate welfare treating the price p of an instrument I as a random realization. This perspective is appropriate in the context of our experiment because p is literally random. Of course, measuring welfare is important in an experiment such as ours not because we care about the experimental tasks per se, but because those tasks arguably distill the types of tradeoffs people may confront in practice. When interpreting our welfare measures, one can think of the randomness in p as representing uncertainty about the particular tradeoffs that will prove relevant to real-world decisions.

We adopt two approaches to aggregating welfare over possible realizations. The first envisions a welfarist planner concerned with the consumer's expected utility. We assume the future price p for the instrument I is distributed according to some atomless CDF H with density h. The loss, in terms of expected utility, associated with employing valuation r rather than the utility-maximizing valuation $r^u(y_s)$ is then:

$$l_W^u\left(r, y_s, H\right) = \begin{cases} \int_r^{r^u(y_s)} \left[u\left(-p, y_s\right) - u\left(0, y_0\right)\right] h(p)dp & \text{if } r \leq r^u(y_s) \\ \int_{r^u(y_s)}^{r} \left[u\left(0, y_0\right) - u\left(-p, y_s\right)\right] h(p)dp & \text{if } r \geq r^u(y_s) \end{cases} \tag{2}$$

The subscript W reminds us that this measure reflects a welfarist perspective.

The welfarist approach assumes the probability distribution of the price p is known. While it is known for any given experiment, it is hard to say which prices give rise to tradeoffs that are likely to emerge in practice. Treating the probability distribution as unknown, we can employ an alternative approach from Computer Science that has recently found its way into Economics: evaluate an option based on the maximum possible loss it could induce.[38] The maximal loss associated with using the valuation r rather than the utility-maximizing valuation $r^u(y_s)$ is:

$$l_M^u\left(r, y_s\right) = \begin{cases} \min_{p\in[r, r^u(y_s)]} \left[u\left(-p, y_s\right) - u\left(0, y_0\right)\right] & \text{if } r \leq r^u(y_s) \\ \min_{p\in[r^u(y_s), r]} \left[u\left(0, y_0\right) - u\left(-p, y_s\right)\right] & \text{if } r \geq r^u(y_s) \end{cases} \tag{3}$$

The subscript M reminds us that this function measures the maximal loss.

Because the scale of u is arbitrary, in both cases we convert to money-metric utility. Specifically, we divide the welfare loss by the perceived marginal utility of income according to u when the individual does not purchase the

[38] See, for example, Carroll (2019), who adopts this approach to study optimal mechanism design in settings where the distribution of agent characteristics is unknown. Other applications include decision theory (Machina and Siniscalchi, 2014) and public policy (Sunstein, 2020).

instrument:[39]

$$
\begin{aligned}
L_W^u\left(r, y_s, H\right) &= \frac{1}{u_m(0, y_0)} l_W^u\left(r, y_s, H\right), \\
L_M^u\left(r, y_s\right) &= \frac{1}{u_m(0, y_0)} l_M^u\left(r, y_s\right).
\end{aligned}
$$

3.2.2 Accounting for additional biases

Even if simple framing removes the type of characterization failure it is designed to address, other valuation biases may remain. In such cases, whether simply framed choices provide valid normative benchmarks (loosely speaking, whether they reveal "true preferences") is unclear, because the remaining biases may distort those choices. Formally, the issue is that the utility function V may diverge from U (other than in the sense of a monotonic transformation), which could affect choices in both frames.

As an illustration of the problem, in our experiment, simple framing eliminates the need to understand compound interest, but it does not address considerations such as present-focus or excessive suspicion that the experimenter will renege on promised payments.[40] If these considerations bias subjects' choices toward lower valuations of future prospects in *both* frames, then one might be tempted to conclude that the "overshooting" documented in Section 2.4 is actually beneficial.

Lipsey and Lancaster (1956) reasoned that the right way to handle these types of issues is to analyze all distortions (here, biases affecting the valuation $r^V(r_c)$) and all remedies for them simultaneously. We call this approach *comprehensive second-best analysis*. While conceptually attractive, it is infeasible in practice.[41] Indeed, the field of Behavioral Economics has not yet arrived at the comprehensive structural understanding of imperfections in human decision making that this approach requires.

Because of these practical challenges, the overwhelming preference of economists, as revealed by research in this area, is to analyze policies one (or two) at a time. The idea is to focus on solving individual pieces of the overall welfare puzzle, thereby progressively building up a complete solution. But the compartmentalization of policy analysis raises challenging conceptual issues. Indeed, any approach other than comprehensive second-best analysis necessarily involves conceptual compromises. Still, to the extent the requirement of tractability necessitates a compromise, we ought to make the best one possible. Table 8 lists three frameworks for compartmentalized welfare analysis. Two are familiar; the third is our proposal.

The first framework, which we label *narrow second-best welfare analysis*, involves a perspective that Lipsey and Lancaster (1956) attribute to Meade (1955). We imagine that there are many biases affecting the valuation $r^V(r_c)$ and corresponding policy options that target them. We evaluate one policy targeting one bias (or a small collection of policies targeting a small number of biases) accounting for the existence of the other biases, but holding the rest of the

[39] This adjustment approximates the reduction in income without the instrument that would produce the same utility decrement as the consumer's decision error. Alternatively, one could define L_e^u (for $e \in \{W, M\}$) as the *exact* reduction in income sufficient to produce that utility decrement: $u(0, y_0) - u(-L_e^u, y_0) = l_e^u$. Because we ultimately take limiting approximations, these two scaling methods yield the same results.

[40] Present focus may not be an issue in our experiment because the earliest payments involve a two-day delay.

[41] For example, while the general theoretical framework of Farhi and Gabaix (2020) encompasses such analyses, they acknowledge that empirical implementation would be a "momentous task" (p. 311).

Table 8: Frameworks for compartmentalized welfare analysis.

Framework	Evaluated choice	Evaluation standard	Welfare criterion
Narrow second-best welfare analysis	$r^V(y_{c\theta})$	U	$L_M^U\left(r^V(y_{c\theta}), y_s\right)$ or $L_W^U\left(r^V(y_{c\theta}), y_s, H\right)$
Myopic welfare analysis	$r^V(y_{c\theta})$	V	$L_M^V\left(r^V(y_{c\theta}), y_s\right)$ or $L_W^V\left(r^V(y_{c\theta}), y_s, H\right)$
Idealized welfare analysis	$r^U(y_{c\theta})$	U	$L_M^U\left(r^U(y_{c\theta}), y_s\right)$ or $L_W^U\left(r^U(y_{c\theta}), y_s, H\right)$

policies fixed, even if they are suboptimal. Within our setting, this approach quantifies the welfare cost of choosing $r^V(y_{c\theta})$ (which reflects all biases, as they exist under prevailing policies) rather than $r^U(y_s)$ (which reflects the absence of all biases), using U as the standard of evaluation. In other words, as shown in Table 8, the pertinent measures of the welfare loss for policy θ are $L_M^U\left(r^V(y_{c\theta}), y_s\right)$ and $L_W^U\left(r^V(y_{c\theta}), y_s, H\right)$.

Setting aside the enormous practical difficulty of identifying and properly accounting for all pertinent biases, this approach suffers from a critical conceptual flaw: if the policy of interest is merely one component of a broader effort to design interventions that improve the quality of decision making, then treating other policies as fixed constraints can yield incorrect conclusions. To illustrate, consider the subjects in our experiment who overstate the returns to investments paying compound interest. If they suffer from present bias, then, as suggested at the outset of this section, one might be tempted to conclude that any improvement in their understanding of compounding would cause them harm. And yet, they would clearly derive the greatest benefit from a combination of policies that address both biases, such as education along with commitment devices. Thus, insisting that a policy ought to compensate for all biases that may infect choice, and not merely for the one it is designed to address, can yield misleading conclusions concerning the policy of interest.

The second framework in Table 8, *myopic welfare analysis*, focuses on small number (usually just one) of biases and associated policy instruments while assuming, at least implicitly, that the consumer's decision-making apparatus is otherwise faultless, perhaps because effective remedies for other valuation biases are already in place. Within our setting, this approach calls for evaluating the loss from choosing $r^V(y_{c\theta})$, the naturalistic selection, rather than $r^V(y_s)$, according to the utility function V. In other words, as shown in Table 8, the pertinent measures of the welfare loss for policy θ are $L_M^V\left(r^V(y_{c\theta}), y_s\right)$ and $L_W^V\left(r^V(y_{c\theta}), y_s, H\right)$. This form of compartmentalized policy analysis dominates the literature. Examples of studies that focus on a single bias include O'Donoghue and Rabin (2006); Gruber and Kőszegi (2001, 2004); DellaVigna and Malmendier (2004). A few studies consider two biases (and associated policies) at a time; for instance Allcott et al. (2014). While such analyses highlight some important second-best considerations, they inevitably ignore others.

The problem with myopic welfare analysis is that the data are not in fact generated by an otherwise faultless decision process. As Lipsey and Lancaster (1956) pointed out in an analogous setting involving multiple distortions, there is no reason to think that solutions derived in this manner (which they called "piecemeal") are in fact desirable. Nor can one

coherently paste together solutions for different biases derived in this way, because each solution implicitly assumes that the other solutions are unnecessary.[42]

The third approach listed in Table 8, *idealized welfare analysis*, reformulates piecemeal policy analysis in a manner that avoids Lipsey and Lancaster's criticism. Viewing the analysis as part of a broader effort to diagnose and correct flaws in decision making, one evaluates policies designed to address a single bias (or a small collection of biases) under the assumption that effective remedies for other biases *will be* forthcoming. In contrast to myopic welfare analysis, which assumes that other biases have *already been* addressed, one must therefore account for the possibility that the observed data are infected by biases other than those the policy of interest seeks to target. Approaching each component problem in this manner, one can potentially arrive at a medley of compartmentalized solutions that complement each other and together achieve an overall solution.[43] For our setting, idealized welfare analysis gauges the welfare cost of choosing $r^U(y_{c\theta})$ rather than $r^U(y_s)$ (unobserved valuations purged of other biases), which potentially differ from $r^V(y_{c\theta})$ and $r^V(y_s)$ (observed valuations infected with other biases), using U as the evaluation standard. Thus, unlike myopic welfare analysis, it does not take observed simply framed valuations at face value as valid normative benchmarks. In the context of our experiment, the approach calls for assessing the loss associated with the complexly framed valuations we would elicit in an idealized setting that removes all valuation biases other than a misunderstanding of compound interest, based on simply framed valuations elicited in that same scenario. As shown in Table 8, the pertinent measures of the welfare loss for policy θ are $L_M^U\left(r^U(y_{c\theta}), y_s\right)$ and $L_W^U\left(r^U(y_{c\theta}), y_s, H\right)$.

A natural first reaction to this proposed framework is that, if practical limitations render comprehensive second-best welfare analysis infeasible, then they must also preclude idealized welfare analysis. After all, idealized welfare analysis, too, requires knowledge of U, which the data do not reveal. And yet, as we prove below, there are reasonably general conditions under which idealized and myopic welfare measures coincide (up to a multiplicative factor), at least within the specific context of analyses involving comparisons of valuations between simple and complex frames.

Anything short of comprehensive welfare analysis inevitably involves conceptual compromises, and idealized welfare analysis is no exception. One must assume that other biases will be fully resolved, rather than either partially resolved or left untreated, and that the policies targeting them will not directly impact the bias of interest – for instance, that a policy targeting present bias or false beliefs concerning experimenter reliability will not meaningfully improve comprehension of compound interest.

The fact that one can perform idealized welfare analysis without recovering the idealized utility function U – potentially without understanding or even knowing about other biases – is counterintuitive. After providing formal results, we therefore offer a simple illustrative example, which also shows that the local approximations used in our main results can perform well for large distortions.

[42] In the context of our experiment, for instance, pasting together a solution for poor comprehension, devised under the assumption of no remedy for present bias, with a solution for present bias, devised under the assumption of no remedy for poor comprehension, does not generally yield policies that complement each other well and produce a desirable overall outcome.

[43] By way of analogy, consider the problem of maximizing a function $f(x_1, ..., x_N)$. Let x^* denote the solution. Idealized welfare analysis is akin to maximizing $f(x_n, x^*_{-n})$ for each n.

3.2.3 Approximations

In the following subsections, we provide formal justifications for our metrics of deliberative competence. Our approach is, in effect, to employ first- and second-order Taylor series approximations for welfare losses. Within Public Economics, this is a standard strategy for quantifying inefficiencies resulting from market distortions. For example, it provides the foundation for the famous Harberger triangle (Harberger, 1964).

Formally, we define

$$y_f^\alpha \equiv \alpha y_f + (1-\alpha)\, y_0$$

This formulation allows us to rescale the consequences of the instrument as perceived in the simple and complex frames up and down by the factor α. As the scale of the instrument shrinks, the pertinent distribution of prices (which we write as H^α to reflect the dependence on α) presumably becomes more concentrated near zero. Absent information about the distribution of prices encountered in real-world analogs to our experimental tasks, we assume for the purpose of the welfarist metric that H^α is uniform in the pertinent region, with density $\frac{h}{\alpha}$.

Defining distance metrics $d_M(r, r') \equiv \left|r - r'\right|$ and $d_W(r, r') \equiv \frac{1}{2}\left(r - r'\right)^2$, we show that $d_M\left(r^u(y_c), r^u(y_s)\right)$ provides a first-order approximation for $L_M^u\left(r^u(y_c), y_s\right)$, and that $d_W\left(r^u(y_c), r^u(y_s)\right)$ provides a second-order approximation for $L_W^u\left(r^u(y_c), y_s, H\right)$ (hence our subscript convention).[44]

In taking limits, we hold the consumer's misunderstanding of the instrument fixed and let the scale of its consequences shrink. Another possibility is to hold the scale of the instrument's consequences fixed and let the magnitude of the misunderstanding shrink. Potentially, the latter approach may provide better approximations for large instruments, but worse approximations for large misunderstandings.[45] As shown in Appendix B, this alternative approach yields similar approximation results.

3.3 Welfare analysis when simple framing involves no biases

We begin by assuming that the decision maker suffers from no pertinent biases other than the types of characterization failures targeted by the policies we seek to evaluate, or that simple framing, by itself, eliminates all pertinent biases. In that case, $V = U$. Our first result shows that one can approximate welfare losses by measuring distances between paired valuations.

Proposition 1. *For all y_c and y_s satisfying $\left.\frac{dr^V(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} \neq \left.\frac{dr^V(y_s^\alpha)}{d\alpha}\right|_{\alpha=0}$, we have*

$$\lim_{\alpha\to 0}\left(\frac{L_M^V\left(r^V(y_c^\alpha), y_s^\alpha\right)}{d_M\left(r^V\left(y_c^\alpha\right), r^V\left(y_s^\alpha\right)\right)}\right) = 1 \tag{4}$$

and

[44] We use a second-order approximation for L_W^u rather than a first-order approximation because the first-order term is identically equal to zero.

[45] Letting the scale of the instrument's consequences shrink may not be the right strategy in settings where risk plays a central role. For example, if the consumer is an expected utility maximizer and background risk is either absent or uncorrelated with the instrument's returns, the approximation involves an extrapolation from a limit in which risk aversion vanishes. In such cases, the alternative approximation may be more useful.

$$\lim_{\alpha \to 0} \left(\frac{L_W^V \left(r^V(y_c^\alpha), y_s^\alpha, H^\alpha \right)}{d_W \left(r^V \left(y_c^\alpha \right), r^V \left(y_s^\alpha \right) \right)} \right) = h. \tag{5}$$

Equation (4) provides a formal foundation for treating $d_M \left(r^V(y_c), r^V(y_s) \right)$ as an approximate measure of the maximal money-metric welfare loss, $L_M^V \left(r^V(y_c), y_s \right)$. Suppose our objective is to evaluate two corrective policies, θ and θ', that lead the consumer to interpret the instrument as offering $y_{c\theta}$ and $y_{c\theta'}$, respectively. According to Proposition 1, $d_M \left(r^V(y_{c\theta}), r^V(y_s) \right) - d_M \left(r^V(y_{c\theta'}), r^V(y_s) \right)$ approximates the difference in the maximal money-metric welfare losses, $L_M^V \left(r^V(y_{c\theta}), y_s \right) - L_M^V \left(r^V(y_{c\theta'}), y_s \right)$. Similarly, for the welfarist criterion, equation (5) implies that $L_W^V \left(r^V(y_c), y_s, H \right)$ is approximately proportional to $d_W \left(r^V \left(y_c \right), r^V \left(y_s \right) \right)$, where the constant of proportionality is independent of the instrument's perceived consequences, and hence of the corrective policy. While this measure does not tell us the absolute level of the welfare loss associated with any policy, it does allow us to rank any two policies correctly, and to approximate the ratio of their associated money-metric welfare losses, $\frac{L_W^V(r^V(y_{c\theta}), y_s, H)}{L_W^V(r^V(y_{c\theta'}), y_s, H)}$, with $\frac{d_W(r^V(y_{c\theta}), r^V(y_s))}{d_W(r^V(y_{c\theta'}), r^V(y_s))}$.

Despite the fact that the expected welfare loss L_W^V and the maximal welfare loss L_M^V reflect different approaches to conceptualizing the loss function, they admit similar welfare approximations, inasmuch as $\left(r^V(y_c) - r^V(y_s) \right)^2$ is simply the square of $\left| r^V(y_c) - r^V(y_s) \right|$. While the expected loss reflects a more conventional approach, there are two reasons to prefer the maximal loss. A practical advantage is that $\left| r^V(y_c) - r^V(y_s) \right|$ is less sensitive to outliers than $\left(r^V(y_c) - r^V(y_s) \right)^2$. A conceptual advantage is that the maximal loss is directly interpretable as a dollar-value measure, while the expected loss is a rescaling of a dollar-value measure by the unknown density h. The latter remains useful, however, in that it correctly ranks policies and accurately measures the ratio of their benefits.

In Section 2.5, we averaged d_M over a collection of instruments. Proposition 1 also allows us to interpret such averages. Suppose our objective is to determine $L_M^V(I) \equiv \sum_{I \in \mathfrak{I}} \lambda_I L_M^V \left(r^V(y_c^I), y_s^I \right)$, where I includes the instruments of interest, and y_f^I denotes the perceived outcome for instrument I in frame f. The weight λ_I could reflect judgments of each instrument's relevance for real-world choices, or about the relative importance of an incremental dollar in the types of circumstances that are associated with the instrument's availability. Proposition 1 tells us that we can approximate $L_M^V(I)$ by computing $\sum_{I \in \mathfrak{I}} \lambda_I d_M \left(r^V(y_c^I), r^V(y_s^I) \right)$. Likewise, the proposition tells us that we can approximate $L_W^V(I)$ (similarly defined) up to a constant of proportionality (which allows us to rank policies and measure their relative benefits) by computing $\sum_{I \in \mathfrak{I}} h_I \lambda_I d_W \left(r^V(y_c^I), r^V(y_s^I) \right)$, where h_I is the density of the price distribution for instrument I. Because the pertinent price distributions are unknown, it is reasonable to treat h_I as either a constant, or at least uncorrelated with the welfare loss. In either case, computing $\sum_{I \in \mathfrak{I}} \lambda_I d_W \left(r^V(y_c^I), r^V(y_s^I) \right)$ provides the desired approximation. Aggregation over consumers raises related issues; see Section 3.6.

A limitation of Proposition 1 and other results to follow is that they pertain to binary choices rather than to selections from arbitrary menus. Because binary comparisons lie at the core of more complex choices, this feature is less of a limitation than one might think. Suppose the consumer must purchase an instrument from the set $M \cup \{0\}$, where 0 represents the status quo. As an example, different elements of this set might represent different quantities of the same instrument. Suppose further that we have assessed simply and complexly framed valuations for all members of M. Let-

ting M_L (M_H) represent the subset of instruments for which the consumer's valuation is too low (too high), the greatest possible welfare loss is then proportional to $\max_{I \in M_H} d_M(r^V(I,c), r^V(I,s)) + \max_{I \in M_L} d_M(r^V(I,c), r^V(I,s))$ (where $r^u(I,f)$ is shorthand for the valuation of instrument I in frame f and, by convention, $\max_{I \in \emptyset} L_M(I) \equiv 0$) – i.e., it is the loss incurred when the consumer purchases the most overvalued instrument at the worst possible price and foregoes the most undervalued instrument at the best price consistent with purchasing the first instrument.[46] Welfarist extensions appear more challenging due to the need for assumptions concerning the joint distribution of prices, but may be possible in cases where such assumptions are natural (for example, when an instrument is available in various quantities at a fixed unit price).

3.4 Welfare analysis when simple framing involves additional biases

3.4.1 The general result

We now extend our analysis to allow for arbitrary differences between U and V. The main result of this section shows that our measure of deliberative competence, $d_e\left(r^V(y_c^\alpha), r^V(y_s)\right)$, approximates the corresponding idealized welfare measure up to a multiplicative constant for each welfare perspective $e \in \{M, W\}$ if a certain separability assumption holds. The assumption requires that we can write the utility functions as $V(m, \varphi(y))$ and $U(m, \varphi(y))$ for some subutility function φ.

When y is a scalar (as in our experiment), the separability assumption involves no loss of generality. When y is multidimensional, it is restrictive, but it nevertheless subsumes important possibilities, including the following.

(i) Instruments are investments that yield financial returns in future periods and states of nature, which the decision maker misunderstands. Our objective is to evaluate policies that target those misunderstandings. A complicating factor is that the consumer may also be present-biased to an unknown degree. In that case, the utility functions may take the form $V(m,y) = v(m) + \beta\varphi(y)$ and $U(m,y) = v(m) + \varphi(y)$, where $\varphi(y)$ is δ-discounted utility and β is the unknown present-bias parameter.

(ii) Instruments are insurance contracts that yield future payments contingent on the occurrence of a particular event. The consumer misunderstands the formulas that determine those payments, and our objective is to evaluate policies that target those misunderstandings. A complicating factor is that the consumer may misjudge the probability of the event to an unknown degree. In that case, the utility functions may take the form $V(m,y) = v(m) + \tilde{\pi}\varphi(y) + (1-\tilde{\pi})u_0$ and $U(m,y) = v(m) + \pi\varphi(y) + (1-\pi)u_0$, where $\tilde{\pi}$ is the unknown perceived probability, π is the true probability, $\varphi(y)$ is expected utility conditional on the event, and u_0 is expected utility conditional on the event's absence.

[46] Suppose the consumer purchases I at price p when she should have purchased I' at price p'. We then have $r^u(I,c) - p \geq \max\{0, r^u(I',c) - p'\}$ and $r^u(I',s) - p' \geq \max\{0, r^u(I,s) - p\}$. For any given p, the loss is maximized by setting p' as low as possible, consistent with these inequalities. Ignoring the non-negativity constraints for the moment, we have $p' = r^u(I',c) - r^u(I,c) + p$. Setting $p = r^u(I,c)$ and $p' = r^u(I',c)$ satisfies non-negativity. The total loss, which we approximate as $\left[r^u(I',s) - p'\right] + \left[p - r^(I,s)\right]$, is then $\left[r^u(I',s) - r^u(I',c)\right] + [r^u(I,c) - r^u(I,s)]$. Choosing I and I' to be the instruments with, respectively, the greatest overvaluation and the greatest undervaluation maximizes this expression. Note that this argument subsumes the case in which either I or I' is 0.

(iii) Instruments are medications that impact various aspects of future health in ways consumers misunderstand. Our objective is to evaluate policies that target those misunderstandings. A complicating factor is that the consumer may also be imperfectly attentive to health consequences. In that case, the pertinent utility functions might take the assumed forms, with $\varphi(y)$ representing a composite health good, and with the discrepancy between V and U reflecting inattention.

(iv) An additional complicating factor in any of the preceding examples is that the method used for eliciting valuations may activate biases. For example, multiple price lists can induce "end-point" effects (Andersen et al., 2006) and, more generally, framing can "anchor" subjects' assessments (e.g. Ariely et al., 2003). In such cases, we might have $U(m, y) = \nu(m) + \varphi(y)$ and $V(m, y) = \mu(m) + \xi(\varphi(y))$ for functions μ and ξ.

While this list is not exhaustive, it is important to understand that the separability assumption limits potential applications by ruling out biases that impact tradeoffs among the elements of the vector y. For example, treating departures from exponential discounting as biases, our setting accommodates quasi-hyperbolic discounting but excludes hyperbolic discounting.[47]

The following result shows, in effect, that the idealized welfare loss, $L_M^U\left(r^U(y_c), y_s\right)$ or $L_W^U\left(r^U(y_c), y_s, H\right)$, and the corresponding measure of deliberative competence, $d_e\left(r^V(y_c), r^V(y_s)\right)$ for $e \in \{W, M\}$, are approximately proportional (in the sense that they become exactly proportional in the limit as $\alpha \to 0$), where the factor of proportionality does not depend on the decision maker's misperceptions of consequences.

Proposition 2. *For each $e \in \{W, M\}$, there exists a strictly positive constant, K_e, such that for all y_c and y_s satisfying $\left.\frac{dr^u(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} \neq \left.\frac{dr^u(y_s^\alpha)}{d\alpha}\right|_{\alpha=0}$, we have*

$$\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_c^\alpha), y_s^\alpha\right)}{d_e\left(r^V\left(y_c^\alpha\right), r^V\left(y_s^\alpha\right)\right)}\right) = K_M \tag{6}$$

$$\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{d_e\left(r^V\left(y_c^\alpha\right), r^V\left(y_s^\alpha\right)\right)}\right) = K_W \tag{7}$$

Proposition 2 (for idealized welfare analysis) is weaker than Proposition 1 (for myopic welfare analysis) in only one respect: it may be the case that $K_M \neq 1$. Even with that qualification, several important conclusions follow.

First, we can conduct useful welfare analysis in a compartmentalized fashion *even when we do not have estimates of the parameters governing other biases outside the scope of the analysis*.[48] Specifically, analogously to Proposition 1 (which concerned the myopic welfare losses $L_M^V\left(r^V(y_c), y_s\right)$ and $L_W^V\left(r^V(y_c), y_s, H\right)$), even though the bias-free utility function U is potentially unknown and unobservable, we can infer the ranking of corrective policies according to the

[47] In the case of dynamic choices, biases can affect a given decision either directly or indirectly through future choices that shape the consequences of the current choice. Our method is suitable for measuring the direct effects. Specifically, it allows us to gauge the magnitudes of mistakes people make in one-shot decisions, with future consequences fixed experimentally. Such decisions are the building blocks of dynamic decisions, in the sense that dynamic problems require the consumer to solve one-shot problems where the future consequences are fixed at the continuation solutions.

[48] Naturally, if we can measure the parameters governing the other biases, we can compute the factors of proportionality, $K_e(y_s)$, and then estimate the absolute welfare effects of various corrective policies by making simple multiplicative adjustments. Such exercises would still fall within the scope of idealized welfare analysis rather than comprehensive second-best analysis, in that they would use estimates of other biases to "correct" the data before evaluating the policy of interest, without explicitly modeling the ancillary remedies that would achieve those corrections.

idealized welfare loss, $L_M^U\left(r^U(y_c), y_s\right)$ or $L_W^U\left(r^U(y_c), y_s, H\right)$, by ranking them according to the corresponding index of deliberative competence, $d_e\left(r^V(y_c), r^V(y_s)\right)$, which is measurable. We can even gauge the percentage differences between the dollar-equivalents of the benefits flowing from these policies. We can also aggregate over instruments, subject to the same qualifications as in Section 3.3. Thus, for our experiment, columns 1 and 2 of Table 7 show that the intervention in Experiment A reduces the idealized welfare loss associated with poor comprehension of compound interest by only 6.5% (that is, $1 - 22.864/24.448 = 0.065$, s.e. $= 0.095$), while the intervention in Experiment B reduces it by 27.7% (that is, $1 - 18.569/25.673 = 0.277$, s.e. $= 0.060$).

Second, $d_e\left(r^V(y_c), r^V(y_s)\right)$ remains a valid approximation of idealized welfare (up to an unknown multiplicative scalar) even when the analyst is unaware that certain types of other biases exist. This robustness property is reassuring.

Third, to the extent our method is robust to certain known and unknown biases (up to a multiplicative scalar), it is also robust to normative ambiguity about such biases (i.e., the possibility that both U and U' may reflect legitimate normative perspectives, as permitted in the welfare framework of Bernheim and Rangel, 2009). Hence, we can assess the effect of a policy θ on the welfare loss from mischaracterization within the framework of idealized welfare economics without needing to take a stance on thorny issues such as the normative relevance of present bias.

3.4.2 An example

Proposition 2 may initially strike the reader as counterintuitive, because it allows us to rank corrective policies according to an unknown function. The following simple example makes the logic more transparent.

For the purpose of this example, we interpret m as current consumption, y as vector describing state-contingent consumption in future periods, and I as a financial instrument that enables the consumer to convert the former into the latter. The consumer's beliefs about the instrument are mistaken in two ways. First, due to complex framing, she thinks the instrument promises to deliver $y_c - y_0$ rather than $y_s - y_0$. Second, she entertains the objectively groundless suspicion that the instrument is a scam that will pay nothing, leaving her future outcomes unchanged with probability $1 - \pi$. We also assume that immediate and future consequences, m and y, enter decision utility additively through separate functions, $v(\cdot)$ and $\varphi(\cdot)$, respectively, where $v(m) = m$. Finally, we assume the consumer discounts future expected utility at the rate $\beta\delta$, where β measures the extent of present bias: $V(m, y) = m + \beta\delta\varphi(y)$, and $U(m, y) = m + \delta\varphi(y)$. Our objective is to evaluate the idealized welfare effects of policies that seek to improve the consumer's understanding of promised consequences (i.e., to reduce the discrepancy between y_s and y_c), treating excessive skepticism and present focus as ancillary biases, of which we may or may not be aware.

In frame f (with the ancillary biases present), the consumer decides whether to purchase the instrument by comparing the perceived utility derived from purchase, $V(m_0 - p, y_f) = m_0 - p + \pi\beta\delta\varphi(y_f) + (1 - \pi)\beta\delta\varphi(y_0)$, with the perceived utility derived from no purchase, $V(m_0, y_0) = m_0 + \beta\delta\varphi(y_0)$. The consumer's reservation price, $r^V(y_f)$, is the value of p that equates these expressions: $r^V(y_f) = \pi\beta\delta\left(\varphi(y_f) - \varphi(y_0)\right)$. Similarly, $r^U(y_f) = \delta\left(\varphi(y_f) - \varphi(y_0)\right)$. In the case where $\varphi(y_s) > \varphi(y_c)$, we then have

$$L_M^U\left(r^U(y_c), r^U(y_s)\right) = U\left(m_0 - r^U(y_c), \varphi(y_s)\right) - U\left(m_0, \varphi(y_0)\right)$$

$$= \left[m_0 - r^U(y_c) + \delta\varphi(y_s)\right] - \left[m_0 - \delta\varphi(y_0)\right] = \delta\left[\varphi(y_s) - \varphi(y_c)\right]$$

We also have

$$d_M\left(r^V(y_c), r^V(y_s)\right) = r^V(y_s) - r^V(y_c) = \pi\beta\delta\left[\varphi(y_s) - \varphi(y_c)\right]$$

Putting these observations together yields

$$L_M^U\left(r^U(y_c), y_s\right) = \frac{1}{\pi\beta} d_M\left(r^V(y_c), r^V(y_s)\right)$$

A similar argument applies when $\varphi(y_s) < \varphi(y_c)$. Thus, $K_M(y_s) = \frac{1}{\pi\beta}$. If H is uniform within the relevant interval, we can likewise show that $L_W^U\left(r^U(y_c), y_s, H\right)$ is exactly proportional to $d_W\left(r^V(y_c), r^V(y_s)\right)$.

The implications of this analysis merit emphasis. Suppose an analyst evaluates educational interventions based on d_M or d_W, but is unaware that consumers suffer from either present bias or false beliefs about experimenter reliability. The analysis would still yield valid measures of idealized welfare effects, up to an unknown constant of proportionality. Comparisons of relative benefits across policies would remain valid.

A notable feature of this example is that the relationship between the idealized welfare loss and our measure of deliberative competence is exact. Accordingly, the example shows that our approximations can yield accurate answers even when consequences and misunderstandings are both large.

3.5 Policy-induced confounds

So far, we have assumed that neither V nor U depends directly on the policies we seek to compare (e.g., θ and θ')—in other words, that there are no policy-induced confounds. This assumption is responsible for the invariance of the multiplicative constant, K_e, across policies, which these comparisons exploit. Yet, it is possible that an intervention changes V, causing the consumer's simply framed valuations to vary with the policy. For example, if an educational intervention targeting the mechanics of interest compounding exhorts people to be patient, simply framed valuations for delayed payments might increase.

This concern arises in our experiment. Table 9 presents regressions describing valuations in the simple frame, separately for each timeframe. Columns 1 and 3 show that, in Experiment A, there are no treatment effects on these valuations. In each condition, subjects value a dollar received with a 72-day delay at roughly 70 cents, and a dollar received with a 32-day delay at roughly 75 cents.[49] In contrast, a treatment effect is present in Experiment B. The valuations of subjects in the Control condition are similar to those elicited from subjects in Experiment A, as columns 2 and 4 show. Average simply framed valuations in the Treatment condition, however, are higher by 8.2 and 7.4 cents for the 72-day and 36-day timeframes, respectively.

How should an analyst proceed if the data reveal these types of confounds? As we show next, a simple adjustment to our measure of deliberative competence preserves all the implications of Proposition 2.

[49] These magnitudes are typical for studies that elicit time preferences over short horizons (Frederick et al., 2002). Not only do subjects discount the future heavily, but the discount function decreases much less steeply for longer delays than for short delay. Part of the explanation may involve perceptions of experimenter reliability.

Table 9: Valuations in simply framed tasks

	(1)	(2)	(3)	(4)
Delay in Days	72		36	
Experiment	A	B	A	B
Levels				
Control	68.637***	69.698***	75.874***	75.343***
	(2.246)	(1.766)	(2.015)	(1.628)
Treatment	69.513***	77.908***	75.801***	82.791***
	(2.265)	(1.707)	(2.076)	(1.451)
Difference	0.876	8.210***	-0.073	7.448***
	(3.190)	(2.456)	(2.893)	(2.181)
Observations	1,075	1,740	1,075	1,740
Subjects	215	348	215	348

Notes: Valuations in simply framed tasks, $r_s^{i,R,t}$, rescaled to range from 0 to 100 (percentage points) rather than 0 to 1. Each column displays the coefficients of a separate OLS regression. Standard errors in parentheses, clustered by subject. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

Formally, we allow the indirect utility function V to depend on policies θ directly rather than only through the subjects' misperception of the instrument's consequences, y. That is, we write $V(m, \varphi(y), \theta)$. Because θ influences behavior through valuation biases, we assume it does not directly affect the bias-free utility function, $U(m, \varphi(y))$. Notationally, we include θ as an argument of r^V to accommodate the confounding framing effects.

Our next proposition identifies the adjustment to measured deliberative competence required to restore comparability across policies.

Proposition 3. *There exist a strictly positive function, $\kappa(y_s)$, such that for all values of $y_{c\theta}$ and y_s satisfying $\left.\frac{dr^u(y_{c\theta}^\alpha)}{d\alpha}\right|_{\alpha=0} \neq \left.\frac{dr^u(y_s^\alpha)}{d\alpha}\right|_{\alpha=0}$, and all policies θ, we have*

$$\lim_{\alpha\to 0}\left[\left(\frac{L_M^U\left(r^U(y_{c\theta}^\alpha), y_s^\alpha\right)}{d_M\left(r^V\left(y_{c\theta}^\alpha,\theta\right), r^V\left(y_s^\alpha,\theta\right)\right)}\right) - \frac{\alpha\kappa(y_s)}{r^V(y_s^\alpha,\theta)}\right] = 0 \tag{8}$$

and

$$\lim_{\alpha\to 0}\left[\left(\frac{L_W^U\left(r^U(y_{c\theta}^\alpha), y_s^\alpha, H^\alpha\right)}{d_W\left(r^V\left(y_{c\theta}^\alpha,\theta\right), r^V\left(y_s^\alpha,\theta\right)\right)}\right) - h\left(\frac{\alpha\kappa(y_s)}{r^V(y_s^\alpha,\theta)}\right)^2\right] = 0 \tag{9}$$

Furthermore, when y is a scalar, $\kappa(y_s) = (y_s - y_0)\kappa^*$ for some strictly positive constant κ^*.

According to this proposition, we can address the confound simply by rescaling our measures of deliberative competence. Specifically, using the proposition to generate approximations for the case of $\alpha = 1$, we conclude that $L_M^U\left(r^U(y_{c\theta}), y_s\right)$ is proportional to $\frac{d_M(r^V(y_{c\theta},\theta),r^V(y_s,\theta))}{r^V(y_s,\theta)}$, and that $L_W^U\left(r^U(y_{c\theta}), y_s, H\right)$ is proportional to $\frac{d_W(r^V(y_{c\theta},\theta),r^V(y_s,\theta))}{(r^V(y_s,\theta))^2}$, where the constants of proportionality are independent of y_c, and consequently of the policy θ.

The fact that the constants of proportionality depend on y_s (through $\kappa(y_s)$) introduces a qualification when aggregating over instruments. However, for the case in which y_s is a scalar, the proposition implies that $L_M^U\left(r^U(y_{c\theta}), y_s\right)$ is approximately proportional to $\frac{d_M(r^V(y_{c\theta},\theta),r^V(y_s,\theta))}{r^V(y_s,\theta)/(y_s-y_0)}$, and that $L_W^U\left(r^U(y_{c\theta}), y_s, H\right)$ is approximately proportional to $\frac{d_W(r^V(y_{c\theta},\theta),r^V(y_s,\theta))}{(r^V(y_s,\theta)/(y_s-y_0))^2}$, where the constants of proportionality are also independent of y_s (apart from the possibility that

y_s is related to h). In other words, to restore all the implications of Proposition 2, we simply rescale the distance metrics by $r^V(y_s, \theta)/(y_s - y_0)$ and $\left(r^V(y_s, \theta)/(y_s - y_0)\right)^2$, respectively.

We use this proposition to reevaluate the data from our experiment.[50] The correction effectively changes the units of measurement: the differences in Table 7 reflect subjects' valuations of future prospects in terms of immediate payments, while division by simply framed valuations reflects subjects valuations of future prospects in terms of future payments.[51] Table 10 reproduces the results of Table 7 with the corrected measure of financial competence. Columns 1 and 2 average across both timeframes; the remaining columns display estimates separately within each timeframe. We continue to observe large and highly statistically significant improvements in deliberative competence in Experiment B; the estimated relative improvement is now $1 - 25.491/37.541 = 0.321$ (s.e. 0.058). Treatment effects in Experiment A remain insubstantial and statistically distinguishable from zero.

Table 10: Deliberative Competence corrected for policy-induced framing effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-34.639***	-37.541***	-36.246***	-39.326***	-33.032***	-35.756***
	(2.115)	(1.906)	(2.535)	(2.127)	(2.053)	(2.001)
Treatment	-37.204***	-25.491***	-38.363***	-26.293***	-36.044***	-24.689***
	(3.787)	(1.775)	(3.882)	(1.865)	(4.202)	(1.882)
Substance-Only	-37.606***		-40.610***		-34.602***	
	(4.710)		(6.550)		(3.210)	
Rhetoric-Only	-29.728***		-32.242***		-27.214***	
	(2.427)		(2.808)		(2.337)	
p-value of difference to Control						
Treatment	0.555	0.000	0.648	0.000	0.520	0.000
Substance-Only	0.566		0.535		0.680	
Rhetoric-Only	0.128		0.290		0.062	
Observations	4,550	3,470	2,275	1,735	2,275	1,735
Subjects	455	347	455	347	455	347

Notes: Each column displays the coefficients of a separate OLS regression of deliberative competence $d_M^{i,R,t} = -|r_c^{i,R,t} - r_s^{i,R,t}|$, corrected for the change in simply framed valuations, on treatment indicators. One subject in experiment B revealed valuations of 0 in each simply framed task. Because the correction entails dividing by the within-subject mean of simply framed valuations, we exclude this subject from the analysis. $^{***}p < 0.01, ^{***}p < 0.05, ^{*}p < 0.1$.

[50] Recall that our empirical analysis involves normalized valuations, which equal the raw valuations divided by the future payments. To mitigate noise in the simply framed elicitations, we average over each subject's simply framed valuations for each timeframe and use the resulting mean to correct complexly framed choices by that subject in that timeframe. Note that the normalized future payoff, $y_s(I) - y_0$, is 1. Consequently, when we rescale the distance metrics by $r^V(y_s, \theta)$ and $\left(r^V(y_s, \theta)\right)^2$, respectively, we are also rescaling by $r^V(y_s, \theta)/(y_s - y_0)$ and $\left(r^V(y_s, \theta)/(y_s - y_0)\right)^2$, which gives us comparability across instruments (as noted above).

[51] In principle, we could have elicited valuations in terms of future payments rather than immediate payments directly, but then each complexly framed decision task would have been a "math problem" with a single correct answer. Because preferences may distort cognitive reasoning (see, e.g., Kunda, 1990), it is important to pose tasks in which they play meaningful roles.

3.6 Interpersonal aggregation

Equipped with a set of welfare weights that reflect the relative social value of an extra dollar received by each individual, one can meaningfully aggregate money-metric welfare losses over members of a population. That standard procedure becomes more challenging in the current context because, unless we adopt the maximal loss criterion and assume away biases outside the scope of the analysis, we measure dollar-equivalent welfare losses up to a constant of proportionality that can potentially vary from one person to the next. To focus on that issue, we will assume for simplicity that the social value of an extra dollar is the same for everyone.

Proposition 5 tells us that, as an approximation, the welfare loss for individual j given perspective $e \in \{W, M\}$ (for some specified instrument and policy), abbreviated $L_e^U(j)$, is approximately equal to measured deliberative competence for individual j, abbreviated $d_e^V(j)$, times a constant $K_e(j)$. It follows that the aggregate (average) welfare loss is $\overline{L}_e^U = \overline{d}_e \overline{K}_e + \text{cov}\left(d_e, K_e\right)$, where we use bars to denote means across individuals for each variable.

Our objective is to compare policies according to their average idealized welfare losses, $\overline{L}_e^U$. Assume for the moment that decision-making defects within the scope of the analysis are uncorrelated with defects outside the scope of the analysis – in other words, $\text{cov}\left(d_e^V, K_e\right) = 0$. In that case, the invariance of each $K_e(j)$, and hence of $\overline{K}_e$, across corrective policies implies that the change in the aggregate welfare loss ($\overline{L}_e^U$) is a fixed multiple of the change in average deliberative competence ($\overline{d}_e$), which we can measure.

Unfortunately, biases tend to be correlated across domains (Dean and Ortoleva, 2019; Stango and Zinman, 2019; Chapman et al., 2018).[52] The correlation is also unlikely to be stable across policies or to vary in proportion to $\overline{d}_e$. To illustrate, suppose present bias is associated with greater underestimation (less overestimation) of compound interest. A policy that shifts the entire population from underestimation to overestimation will reverse the correlation between $d_e^V(j)$ and the ancillary bias.

One strategy for addressing potentially problematic correlations between the biases of interest and other decision-making defects is to disaggregate into more homogeneous subsamples, and to evaluate whether results are robust with respect to subsample weights. While we obviously cannot disaggregate based on the magnitudes of unobserved biases, we can exploit the fact that such biases tend to impact simply framed valuations. For example, in our experiment, considerations such as present bias and false beliefs about experimenter reliability would suppress valuations of future rewards.

Accordingly, we sort subjects into quartiles based on their mean valuations of simply framed prospects (separately for each treatment group to account for the fact that the Treatment intervention in Experiment B affected those valuations). For each experiment, we then regress measured deliberative competence on a Treatment dummy, an indicator for each quartile of simply framed valuations, and an interaction between those indicators and the Treatment dummy. We cluster standard errors on the subject level.

Table 11 displays the results. Within each quartile, the Treatment effect is larger in Experiment B than in Experiment A by several percentage points. For Experiment B, the benefit of the intervention is roughly zero for subjects in the

[52] In the case of $e = W$, the constant of proportionality $K_W(j)$ incorporates the density parameter $h\left(r^V(y_s)\right)$ for individual j. There is no particular reason to think this term is correlated with either $d_W^V(j)$ or with the other components of $K_W(j)$. To the extent these correlations are zero, heterogeneity in the density term does not impact $\bar{L}_W^U$.

Table 11: Deliberative Competence by quartiles of simply framed valuations

Experiment A

VARIABLE	Deliberative Competence			
	Quartile simply framed valuation			
	1	2	3	4
Levels				
Control	-17.007***	-21.693***	-26.589***	-32.216***
	(2.037)	(2.596)	(3.410)	(3.893)
Treatment	-24.967***	-22.345***	-22.855***	-21.367***
	(4.484)	(2.779)	(2.559)	(3.772)
Effect	-7.961	-0.652	3.734	10.849**
	(4.925)	(3.803)	(4.263)	(5.421)
Observations	2,150			
Subjects	215			

Experiment B

VARIABLE	Deliberative Competence			
	Quartile simply framed valuation			
	1	2	3	4
Levels				
Control	-17.734***	-25.920***	-25.706***	-33.156***
	(1.893)	(2.073)	(2.414)	(3.605)
Treatment	-17.818***	-21.239***	-17.215***	-18.018***
	(2.155)	(2.392)	(2.254)	(2.547)
Effect	-0.085	4.681	8.492**	15.138***
	(2.869)	(3.165)	(3.302)	(4.414)
Observations	3,480			
Subjects	348			

Notes: Effect of treatments on deliberative competence, $d_M^{i,R,t} = -|r_c^{i,R,t} - r_s^{i,R,t}|$, by quartiles of simply framed valuations, rescaled to range from 0 to 100 (percentage points) rather than 0 to 1. Each panel presents the output of a single OLS regression. Standard errors in parentheses, clustered by subject. $***p < 0.01$,$***p < 0.05$,$*p < 0.1$.

bottom quartile, positive for all other quartiles, and statistically significant for the top two quartiles. For Experiment A, the benefit is negative and substantial (though not statistically significant at conventional levels) for those in the lowest quartile, slightly negative for those in the second quartile, and positive for the top two quartiles, but statistically significant only for the top quartile. We conclude that the Treatment intervention in Experiment B has a (weakly) positive effect on deliberative competence regardless of the welfare weights assigned to different subjects based on a proxy for biases outside the scope of the analysis, whereas the Treatment intervention in Experiment A has mixed effects, clearly benefiting only those with the highest simply framed valuations. Because the Treatment intervention in Experiment B has a larger positive effect within each quartile than the intervention in Experiment A, it is unlikely that our conclusions concerning the benefits of practice and individualized feedback are driven by a spurious relationship between welfare weights and extraneous biases.

4 Conclusion

We have developed a simple method for evaluating opportunity-neutral policies that seek to improve the quality of decision making. It involves simple transformations of the price-metric bias, which are interpretable robustly as measures of the welfare loss resulting from poor comprehension of consequences. We illustrated its advantages over conventional evaluative metrics in an application involving financial education.

The method lends itself to many other applications. Maintaining a focus on financial education, one could use it to evaluate training pertaining to loans and leases, rent-to-own agreements, portfolio diversification, contracts with multi-part tariffs, and insurance products. One could also evaluate alternatives to education, such as peer advice (as we show in Ambuehl et al., 2018), professional guidance, disclosure rules, and informational labels.

Outside the financial domain, our method offers a potential avenue for evaluating a variety of opportunity-neutral policies that target specific biases. As in other recent work in Behavioral Public Economics (reviewed in Bernheim and Taubinsky, 2018), the main requirement is that one can contrive versions of valuation tasks that plausibly remove the bias of interest. This requirement is most easily satisfied when the bias pertains to a specific misconception about the consequences of particular actions, and where one can describe the options in terms of their consequences without triggering the misconception. For example, to evaluate the benefits of "traffic light" labels concerning energy efficiency for household appliances, "simple framing" might entail full and correct information concerning energy costs (as in Allcott and Taubinsky, 2015). One can imagine similar applications involving nutritional labels and other health advisories.

Naturally, the approach has limitations, particularly regarding robustness with respect to biases outside the scope of the analysis, which we highlighted in our theoretical discussion. Idealized welfare analysis overcomes the Lipsey-Lancaster critique of "piecemeal" welfare analysis by making a couple of conceptual compromises, which we have stated explicitly. It also involves a separability assumption that may or may not be plausible in the context of any given application. A practical limitation is that implementation requires the analyst to observe an individual's simply and complexly framed valuations for the same instrument. Because naturally occurring processes do not generally produce such data, potential applications are limited to experiments. While the current investigation involves a laboratory experiment with stylized tasks, applications involving field experiments with more naturalistic tasks are also feasible.

References

Afriat, Sidney N., "Efficiency Estimation of Production Functions," *International Economic Review*, 1972, *13* (3), 568–98.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson, "The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation," *Brookings Papers on Economic Activity*, 2009, *Fall*, 51–101.

Allcott, Hunt and Dmitry Taubinsky, "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 2015, *105* (8), 2501–38.

_ , Sendhil Mullainathan, and Dmitry Taubinsky, "Energy policy with externalities and internalities," *Journal of Public Economics*, 2014, *112*, 72–88.

Almenberg, Johan and Christer Gerdes, "Exponential Growth Bias and Financial Literacy," *Applied Economics Letters*, 2012, *19* (17), 1693–696.

Ambuehl, Sandro, B Douglas Bernheim, Fulya Ersoy, and Donna Harris, "Peer Advice on Financial Decisions: A case of the blind leading the blind?," *NBER Working Paper*, 2018, *w25034*.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutstrom, "Elicitation using Multiple Price List Formats," *Experimental Economics*, 2006, *9*, 383–405.

Ariely, Dan, George Loewenstein, and Drazen Prelec, "âCoherent arbitrarinessâ: Stable demand curves without stable preferences," *The Quarterly Journal of Economics*, 2003, *118* (1), 73–106.

Aufenanger, Tobias, Friedemann Richter, and Matthias Wrede, "Measuring Decision-Making Ability in the Evaluation of Financial Literacy Education Programs," *Unpublished Manuscript*, 2016.

Baltussen, Guido and Gerrit T. Post, "Irrational Diversification: An Examination of Individual Portfolio Choice," *Journal of Financial and Quantitative Analysis*, 2011, *5*, 1463–491.

Benartzi, Shlomo and Richard H Thaler, "Risk aversion or myopia? Choices in repeated gambles and retirement investments," *Management science*, 1999, *45* (3), 364–381.

_ and Richard H. Thaler, "Naive diversification strategies in defined contribution saving plans.," *American Economic Review*, 2001, pp. 79–98.

Bennyhoff, Donald G and Francis M Kinniry Jr, "Advisor's Alpha," *Vanguard research paper (December). https://advisors. vanguard. com/iwe/pdf/ICRAA. pdf*, 2011.

Bernheim, B. Douglas, "Personal Saving, Information, and Economic Literacy: New Directions for Public Policy," in "Tax Policy for Economic Growth in the 1990s" American Council for Capital Formation Washington, DC 1994, pp. 53–78.

_ , "Financial Illiteracy, Education, and Retirement Saving," in Olivia S. Mitchell and Sylvester J. Schieber, eds., *Living with Defined Contribution Pensions. Remaking Responsibility for Retirement.*, University of Pennsylvania Press, 1998, chapter 3.

_ , "Behavioral Welfare Economics," *Journal of the European Economic Association*, 2009, *7* (2-3), 267–319.

_ , "The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics," *Journal of Benefit-Cost Analysis*, 2016, *7* (1), 12–68.

Bernheim, B Douglas, "In defense of behavioral welfare economics," *Journal of Economic Methodology*, 2021, *28* (4), 385–400.

Bernheim, B. Douglas and Antonio Rangel, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 2009, *124* (1), 51–104.

_ and Daniel M. Garrett, "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households," *Journal of Public Economics*, 2003, *87*, 1487–519.

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral public economics," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 381–516.

Bernheim, B. Douglas, Daniel M. Garrett, and Dean M. Maki, "Education and Saving: The Long-Term Effects of High School Financial Curriculum Mandates," *Journal of Public Economics*, 2001, *80*, 435–65.

_ , Igor Popov, and Andrey Fradkin, "The Welfare Economics of Default Options in 401(k) Plans," *American Economic Review*, 2015, *105* (9), 2798–837.

Beshears, John, James J Choi, David Laibson, and Brigitte C Madrian, "Behavioral household finance," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 177–276.

Bhargava, Saurabh and Lynn Conell-Price, "Serenity Now, Save Later? Evidence on Retirement Savings Puzzles from a 401 (k) Field Experiment," *Unpublished*, 2020.

_ , **George Loewenstein, and Justin Sydnor**, "Choose to lose: Health plan choices from a menu with dominated option," *The Quarterly Journal of Economics*, 2017, *132* (3), 1319–1372.

Blumenthal, Karen and Kevin G Volpp, "Enhancing the effectiveness of food labeling in restaurants," *JAMA*, 2010, *303* (6), 553–554.

Bruhn, Miriam, Gabriel Lara Ibarra, and David McKenzie, "The minimal impact of a large-scale financial education program in Mexico City," *Journal of Development Economics*, 2014, *108*, 184–189.

_ , **Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia**, "The impact of high school financial education: Evidence from a large-scale evaluation in Brazil," *American Economic Journal: Applied Economics*, 2016, *8* (4), 256–95.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 2007, *115* (5), 707–47.

_ , _ , **and** _ , "Measuring the Financial Sophistication of Households," *American Economic Review*, 2009, *99* (2), 393–98.

Carpena, Fenella, Shawn Cole, Jeremy Shapiro, and Bilal Zia, "Unpacking the Causal Chain of Financial Literacy," *The World Bank Policy Research Working Paper*, 2011, *5798.*

Carroll, Gabriel, "Robustness in mechanism design and contracting," *Annual Review of Economics*, 2019, *11*, 139–166.

CFPB, "Navigating the Market: A comparison of spending on financial education and financial marketing," *Washington, DC: Consumer Financial Protection Bureau*, 2013.

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer, "Econographics," *NBER Working Paper*, 2018.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–177.

Choi, James J., David Laibson, and Brigitte C. Madrian, "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans," *Review of Economics and Statistics*, 2011, *93* (3), 748–63.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman, "Who is (More) Rational?," *American Economic Review*, 2014, *104* (6), 1518–550.

Cole, Shawn and Gauri Kartini Shastry, "Is High School the Right Time to Teach Self-control? The Effect of Financial Education and Mathematics Courses on Savings Behavior," *Unpublished Manuscript*, 2010.

Collins, J.M., "The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study," *Journal of Economic Behavior & Organization*, 2013, *95*, 146–58.

Dean, Mark and Pietro Ortoleva, "The empirical relationship between nonstandard economic behaviors," *Proceedings of the National Academy of Sciences*, 2019, *116* (33), 16262–16267.

DellaVigna, Stefano and Ulrike Malmendier, "Contract design and self-control: Theory and evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Downs, Julie S, George Loewenstein, and Jessica Wisdom, "Strategies for promoting healthier food choices," *American Economic Review*, 2009, *99* (2), 159–64.

Echenique, Federico, Sangmok Lee, and Matthew Shum, "The Money Pump as a Measure of Revealed Preference Violations," *Journal of Political Economy*, 2011, *119* (6), 1201–223.

Eisenstein, Eric M. and Stephen J. Hoch, "Intuitive Compounding: Framing, Temporal Perspective, and Expertise," *Unpublished Manuscript*, Dec 2007.

Enke, Benjamin and Florian Zimmermann, "Correlation neglect in belief formation," *The Review of Economic Studies*, 2019, *86* (1), 313–332.

_ , **Uri Gneezy, Brian Hall, David Martin, Vadim Nelidov, Theo Offerman, and Jeroen van de Ven**, "Cognitive Biases: Mistakes or Missing Stakes?," *The Review of Economics and Statistics*, 2021, pp. 1–45.

Ernst, Keith, John Farris, and Uriah King, "Quantifying the Economic Cost of Predatory Payday Lending," Technical Report, Center for Responsible Lending 2004.

European Commission, "A Capital Markets Union for people and businesses-new action plan," `https://ec.europa.eu/info/business-economy-euro/growth-and-investment/capital-markets-union/capital-markets-union-2020-action-plan_en` 2020.

Farhi, Emmanuel and Xavier Gabaix, "Optimal taxation with behavioral agents," *American Economic Review*, 2020, *110* (1), 298–336.

Fernandes, Daniel, John G Lynch Jr., and Richard G Netemeyer, "Financial Literacy, Financial Education, and Downstream Financial Behaviors," *Management Science*, 2014, *60* (8), 1861–883.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Goda, Gopi Shah, Matthew Levy, Colleen Flaherty Manchester, Aaron Sojourner, and Joshua Tasoff, "Predicting retirement savings using survey measures of exponential-growth bias and present bias," *Economic Inquiry*, 2019, *57* (3), 1636–1658.

Grubb, Michael D and Matthew Osborne, "Cellular service demand: Biased beliefs, learning, and bill shock," *American Economic Review*, 2015, *105* (1), 234–71.

Gruber, Jonathan and Botond Kőszegi, "Is addiction "rational"? Theory and evidence," *The Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

_ and _ , "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes," *Journal of Public Economics*, 2004, *88* (9-10), 1959–1987.

Hara, Kotaro, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P Bigham, "A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk," in "Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems" ACM 2018, p. 449.

Harberger, Arnold C, "The measurement of waste," *The American Economic Review*, 1964, *54* (3), 58–76.

Harrison, Glenn, Karlijn Morsink, and Mark Schneider, "Do No Harm? The Welfare Consequences of Behavioral Interventions," *Unpublished*, 2020.

Harrison, Glenn W, Morten Igel Lau, E Elisabet Rutström, and Melonie B Sullivan, "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues," *Field Experiments in Economics*, 2005, *10*, 125–218.

Hastings, Justine S., Brigitte C. Madrian, and William L. Skimmyhorn, "Financial Literacy, Financial Education, and Economic Outcomes," *Annual Review of Economics*, 2013, *5*, 347–73.

Heinberg, Aileen, Angela Hung, Arie Kapteyn, Annamaria Lusardi, Anya Savikhin Samek, and Joanne Yoong, "Five Steps to Planning Success: Experimental Evidence from U.S. Households," *Oxford Review of Economic Policy*, 2014, *30* (4), 697–724.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92* (5), 1644–655.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban, "Financial education affects financial knowledge and downstream behaviors," *Journal of Financial Economics*, 2021.

Kunda, Ziva, "The Case for Motivated Reasoning," *Psychological bulletin*, 1990, *108* (3), 480–98.

Levy, Matthew and Joshua Tasoff, "Exponential Growth Bias and Lifecycle Consumption," *Journal of the European Economic Association*, 2016, *14* (3), 545–83.

Levy, Matthew R. and Joshua Tasoff, "Exponential-Growth Bias and Overconfidence," *Journal of Economic Psychology*, 2017, *58*, 1–14.

Lipsey, Richard G and Kelvin Lancaster, "The general theory of second best," *The Review of Economic Studies*, 1956, *24* (1), 11–32.

Loewenstein, George, Joelle Y Friedman, Barbara McGill, Sarah Ahmad, Suzanne Linck, Stacey Sinkula, John Beshears, James J Choi, Jonathan Kolstad, David Laibson et al., "Consumers' misunderstanding of health insurance," *Journal of Health Economics*, 2013, *32* (5), 850–862.

Lusardi, Annamaria, "Americans' Financial Capability," *NBER Working Paper*, 2011, *17103.*

_ and Olivia Mitchell, "Financial Literacy and Planning: Implications for Retirement Well-being," in Annamaria Lusardi and Olivia S Mitchell, eds., *Financial Literacy. Implications for Retirement Security and the Financial Marketplace*, Oxford University Press, 2011, pp. 17–39.

_ and _ , "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 2014, *52* (1), 1–44.

_ and Olivia S Mitchell, "How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement," *NBER Working Paper*, 2009, *15350.*

_ , Anya Samek, Arie Kapteyn, Lewis Glinert, Angela Hung, and Aileen Heinberg, "Visual Tools and Narratives: New Ways to Improve Financial Literacy," *Journal of Pension Economics and Finance*, 2015, pp. 1–27.

Machina, Mark J and Marciano Siniscalchi, "Ambiguity and ambiguity aversion," in "Handbook of the Economics of Risk and Uncertainty," Vol. 1, Elsevier, 2014, pp. 729–807.

Madrian, Brigitte C and Dennis F Shea, "The power of suggestion: Inertia in 401 (k) participation and savings behavior," *The Quarterly Journal of Economics*, 2001, *116* (4), 1149–1187.

Malkiel, Burt G. and Charles D. Ellis, *The Elements of Investing: Easy Lessons for Every Investor*, New Jersey: Wiley, 2013.

Malkiel, Burton Gordon and Kerin McCue, *A random walk down Wall Street*, Norton New York, 1985.

Mandell, Lewis, "The Financial Literacy of Young American Adults: Results of the 2008 National Jump$tart Coalition Survey of High School Seniors and College Students," *Jump$tart Coalition*, 2009, *Washington, D.C.*

_ **and Linda Schmid Klein**, "The Impact of Financial Literacy Education on Subsequent Financial Behavior," *Journal of Financial Counseling and Planning*, 2009, *20* (1), 15–24.

Mason, Winter and Siddarth Suri, "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavior Research Methods*, 2012, *44* (1), 1–23.

Meade, James Edward, *The theory of customs unions*, North-Holland Publishing, 1955.

O'Donoghue, Ted and Matthew Rabin, "Optimal sin taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

OECD, "National Strategies for Financial Education," OECD/INFE Policy Handbook 2015.

_ , "Launch of the EC/OECD-INFE project to develop a financial competence framework for the EU," `https://www.oecd.org/financial/education/launch-eu-financial-competence-framework.htm` April 2021.

Skimmyhorn, William, "Assessing financial education: Evidence from boot camp," *American Economic Journal: Economic Policy*, 2016, *8* (2), 322–43.

Stango, Victor and Jonathan Zinman, "Exponential Growth Bias and Household Finance," *The Journal of Finance*, 2009, *64* (6), 2807–849.

_ **and** _ , "We are all behavioral, more or less: Measuring and using consumer-level behavioral sufficient statistics," *NBER Working Paper*, 2019.

Sunstein, Cass R, "Maximin," *Yale Journal on Regulation*, 2020.

Sydnor, Justin, "(Over) insuring modest risks," *American Economic Journal: Applied Economics*, 2010, *2* (4), 177–99.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *The Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Thaler, Richard H and Cass R Sunstein, *Nudge: Improving decisions about health, wealth, and happiness*, Penguin, 2009.

van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie, "Financial Literacy and Stock Market Participation," *Journal of Financial Economics*, 2011, *101* (2), 449–72.

Wagenaar, William M. and Sabato D. Sagaria, "Misperception of Exponential Growth," *Perception and Psychophysics*, 1975, *18* (6), 416–22.

Walstad, William B., Ken Rebeck, and Richard A. MacDonald, "The Effects of Financial Education on the Financial Knowledge of High School Students," *Journal of Consumer Affairs*, 2010, *44* (2), 336–57.

ONLINE-APPENDIX
NOT FOR PUBLICATION

Evaluating Deliberative Competence:
A Simple Method with an Application to Financial Choice

Sandro Ambuehl, B. Douglas Bernheim, Annamaria Lusardi

Table of Contents

A Proofs of Propositions 1-3 **1**

A.1 Proof of Proposition 1 1

A.2 Proof of Proposition 2 3

A.3 Proof of Proposition 3 5

B Additional theoretical results **6**

B.1 Proof of Proposition 4 7

B.2 Proof of Proposition 5 9

B.3 Proof of Proposition 6 11

C Experiment details **12**

D Additional Data Analysis **16**

D.1 Demographics 16

D.2 Main results controlling for demographics 16

D.3 Effects on individual test questions 18

D.4 Self-reported behavior 19

D.5 Analysis using the welfarist measure of Deliberative Competence 20

D.6 Deliberative Competence with correction for policy-induced confounds based on the approach of Appendix B 23

D.7 Valuation difference compared to noise in the simple frame 26

E Instructions **27**

F Practice problems with personalized feedback **43**

A Proofs of Propositions 1-3

A.1 Proof of Proposition 1

In proving this proposition, we write the evaluation standard as $u \in V, U$, rather than as V, so that we can invoke this result for both V and U in the proof of Proposition 2.

Step 1: Proof of equation (4).

For $u \in \{U, V\}$, we can rewrite equation (4) as follows:

$$\lim_{\alpha \to 0} \left(\frac{l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right)}{u_m(0, y_0) \left| r^u \left(y_c^\alpha \right) - r^u \left(y_s^\alpha \right) \right|} \right) = 1. \tag{10}$$

The numerator and denominator both converge to zero as $\alpha \to 0$, so we apply L'Hôpital's rule to the terms $\frac{l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right)}{u_m(0, y_0)}$ and $\left| r^u \left(y_c^\alpha \right) - r^u \left(y_s^\alpha \right) \right|$.

Suppose first that $\left. \frac{\partial r^u(y_s^\alpha)}{\partial \alpha} \right|_{\alpha=0} - \left. \frac{\partial r^u(y_c^\alpha)}{\partial \alpha} \right|_{\alpha=0} > 0$, so that $r^u \left(y_s^\alpha \right) > r^u \left(y_c^\alpha \right)$ for small α. In that case we can replace $\left| r^u \left(y_c^\alpha \right) - r^u \left(y_s^\alpha \right) \right|$ in the denominator with $r^u \left(y_s^\alpha \right) - r^u \left(y_c^\alpha \right)$.

By definition,

$$l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right) = u \left(-r^u \left(y_c^\alpha \right), y_s^\alpha \right) - u \left(0, y_0 \right).$$

It follows that

$$\frac{d}{d\alpha} \left(\frac{l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right)}{u_m(0, y_0)} \right) = -\frac{u_m \left(-r^u \left(y_c^\alpha \right), y_s^\alpha \right)}{u_m(0, y_0)} \frac{\partial r^u \left(y_c^\alpha \right)}{\partial \alpha} + \frac{u_y \left(-r^u \left(y_c^\alpha \right), y_s^\alpha \right) \left(y_s - y_0 \right)}{u_m(0, y_0)}$$

(Note that the gradient of u with respect to y, denoted u_y, is a row vector and $(y_s - y_0)$ is a column vector). Using the fact that

$$\frac{d r^u \left(y_s^\alpha \right)}{d\alpha} = \frac{u_y \left(-r^u \left(y_s^\alpha \right), y_s^\alpha \right) \left(y_s - y_0 \right)}{u_m \left(-r^u \left(y_s^\alpha \right), y_s^\alpha \right)},$$

we have

$$\frac{d}{d\alpha} \left(\frac{l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right)}{u_m(0, y_0)} \right) = -\frac{u_m \left(-r^u \left(y_c^\alpha \right), y_s^\alpha \right)}{u_m \left(0, y_0 \right)} \frac{\partial r^u \left(y_c^\alpha \right)}{\partial \alpha}$$

$$+ \frac{u_y \left(-r^u \left(y_c^\alpha \right), y_s^\alpha \right) \left(y_s - y_0 \right)}{u_y \left(-r^u \left(y_s^\alpha \right), y_s^\alpha \right) \left(y_s - y_0 \right)} \frac{u_m \left(-r^u \left(y_s^\alpha \right), y_s^\alpha \right)}{u_m(0, y_0)} \frac{\partial r^u \left(y_s^\alpha \right)}{\partial \alpha}$$

Taking limits, we have

$$\lim_{\alpha \to 0} \frac{d}{d\alpha} \left(\frac{l_M^u \left(r^u(y_c^\alpha), y_s^\alpha \right))}{u_m(0, y_0)} \right) = \left. \frac{\partial r^u \left(y_s^\alpha \right)}{\partial \alpha} \right|_{\alpha=0} - \left. \frac{\partial r^u \left(y_c^\alpha \right)}{\partial \alpha} \right|_{\alpha=0}$$

The desired conclusion then follows directly from l'Hôpital's rule.

An analogous argument for the case of $\left.\frac{\partial r^u(y_s^\alpha)}{\partial \alpha}\right|_{\alpha=0} - \left.\frac{\partial r^u(y_c^\alpha)}{\partial \alpha}\right|_{\alpha=0} < 0$, which implies $r^u\left(y_s^\alpha\right) > r^u\left(y_c^\alpha\right)$ for small α, completes the proof of equation (10).

Step 2: Proof of equation (5).

For $u \in \{U, V\}$, we can rewrite equation (5) as follows:

$$\lim_{\alpha\to 0}\left(\frac{\alpha l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{\frac{1}{2}\left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)^2 hu_m(0, y_0)}\right) = 1. \tag{11}$$

The numerator and denominator both converge to zero as $\alpha \to 0$, so we apply L'Hôpital's rule to the terms $\frac{\alpha l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{hu_m(0,y_0)}$ and $\frac{1}{2}\left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)^2$. As we show below, the derivatives of both terms also converge to zero as $\alpha \to 0$, so two applications of the rule are required.

We begin with $\frac{1}{2}\left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)^2$. We have:

$$\frac{d}{d\alpha}\frac{1}{2}\left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)^2 = \left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)\left(\frac{dr^u\left(y_c^\alpha\right)}{d\alpha} - \frac{dr^u\left(y_s^\alpha\right)}{d\alpha}\right),$$

Given our assumptions concerning bounds on the derivatives of the utility functions, $\frac{dr^u\left(y_f^\alpha\right)}{d\alpha}$ is bounded. Consequently, this expression converges to zero as $\alpha \to 0$ (because both valuations converge to zero). So we consider the second derivative:

$$\frac{d^2}{d\alpha^2}\frac{1}{2}\left(r^u\left(y_c^\alpha\right) - r^u\left(y_s^\alpha\right)\right)^2 = \left(\frac{dr^u\left(y_c^\alpha\right)}{d\alpha} - \frac{dr^u\left(y_s^\alpha\right)}{d\alpha}\right)^2 + \left(r^u\left(y_s^\alpha\right) - r^u\left(y_c^\alpha\right)\right)\left(\frac{d^2r^u\left(y_c^\alpha\right)}{d\alpha^2} - \frac{d^2r^u\left(y_s^\alpha\right)}{d\alpha^2}\right)$$

It is straightforward to verify that $\frac{d^2r^u\left(y_f^\alpha\right)}{d\alpha^2}$ is bounded under our assumptions. It follows that

$$\lim_{\alpha\to 0}\frac{1}{2}\frac{d^2}{d\alpha^2}\left[\left(r^u\left(y_s^\alpha\right)\right)^2 - \left(r^u\left(y_c^\alpha\right)\right)^2\right] = \left(\lim_{\alpha\to 0}\frac{dr^u\left(y_s^\alpha\right)}{d\alpha} - \lim_{\alpha\to 0}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha}\right)^2, \tag{12}$$

which is non-zero by assumption.

Next consider $\frac{\alpha l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{hu_m(0,y_0)}$. Suppose first that $\left.\frac{dr^u(y_s^\alpha)}{d\alpha}\right|_{\alpha=0} - \left.\frac{\partial r^u(y_c^\alpha)}{\partial \alpha}\right|_{\alpha=0} > 0$, so that $r^u\left(y_s^\alpha\right) > r^u\left(y_c^\alpha\right)$ for small α. By definition,

$$\frac{\alpha}{h} l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right) = \int_{r^u(y_c^\alpha)}^{r^u(y_s^\alpha)}\left[u\left(-p, y_s^\alpha\right) - u\left(0, y_0\right)\right] dp$$

Noting that the integrand is zero at the upper limit of integration, it follows that

$$\frac{d}{d\alpha}\left[\frac{\alpha}{h} l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)\right] = -\left[u\left(-r^u\left(y_c^\alpha\right), y_s^\alpha\right) - u\left(0, y_0\right)\right]\frac{dr^u\left(y_c^\alpha\right)}{d\alpha} + \int_{r^u(y_c^\alpha)}^{r^u(y_s^\alpha)}\left[u_y\left(-p, y_s^\alpha\right)\left(y_s - y_0\right)\right] dp$$

Given our assumptions concerning bounds on the derivatives of the utility functions, $\frac{dr^u(y_c^\alpha)}{d\alpha}$ and $u_y\left(-p, y_s^\alpha\right)(y_s - y_0)$ are bounded. Consequently, both terms of this expression converge to zero as $\alpha \to 0$. We therefore take the second derivative:

$$\begin{aligned}
\frac{d^2}{d\alpha^2}\left[\frac{\alpha}{h} l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)\right] &= -\left[u\left(-r^u\left(y_c^\alpha\right), y_s^\alpha\right) - u\left(0, y_0\right)\right]\frac{d^2 r^u(y_c^\alpha)}{d\alpha^2} + u_m\left(-r^u\left(y_c^\alpha\right), y_s^\alpha\right)\left(\frac{dr^u(y_c^\alpha)}{d\alpha}\right)^2 \\
&\quad -2u_y\left(-r^u\left(y_c^\alpha\right), y_s^\alpha\right)\left(y_s - y_0\right)\frac{dr^u(y_c^\alpha)}{d\alpha} + u_y\left(-r^u\left(y_s^\alpha\right), y_s^\alpha\right)(y_s - y_0)\frac{dr^u(y_s^\alpha)}{d\alpha} \\
&\quad + \int_{r^u(y_c^\alpha)}^{r^u(y_s^\alpha)} \frac{d}{d\alpha}\left[u_y\left(-p, y_s^\alpha\right)(y_s - y_0)\right] dp
\end{aligned}$$

Recall that

$$\frac{dr^u\left(y_s^\alpha\right)}{d\alpha} = \frac{u_y\left(-r^u\left(y_s^\alpha\right), y_s^\alpha\right)\left(y_s - y_0\right)}{u_m\left(-r^u\left(y_s^\alpha\right), y_s^\alpha\right)},$$

We use this expression to substitute for $u_y\left(-r^u\left(y_s^\alpha\right), y_s^\alpha\right)\left(y_s - y_0\right)$ in the fourth term. We also multiply the third term by $\frac{dr^u(y_s^\alpha)}{d\alpha}\frac{u_m(-r^u(y_s^\alpha), y_s^\alpha)}{u_y(-r^u(y_s^\alpha), y_s^\alpha)(y_s - y_0)} = 1$. Then we take limits as $\alpha \to 0$. The boundedness of $\frac{d^2 r^u\left(y_f^\alpha\right)}{d\alpha^2}$ (noted above) guarantees that the first term vanishes. The boundedness of the first and second derivatives of the utility functions guarantees that the last term vanishes. We thus obtain:

$$\begin{aligned}
\lim_{\alpha\to 0}\frac{d^2}{d\alpha^2}\left[\frac{\alpha}{h} l_W^u\left(r^u(y_c^\alpha), y_s^\alpha, H^\alpha\right)\right] &= u_m\left(0, y_0\right)\left[\left(\lim_{\alpha\to 0}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha}\right)^2 - 2\left(\lim_{\alpha\to 0}\frac{dr^u\left(y_s^\alpha\right)}{d\alpha}\right)\left(\lim_{\alpha\to 0}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha}\right) + \left(\lim_{\alpha\to 0}\frac{dr^u\left(y_s^\alpha\right)}{d\alpha}\right)^2\right] \\
&= u_m\left(0, y_0\right)\left[\lim_{\alpha\to 0}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha} - \lim_{\alpha\to 0}\frac{dr^u\left(y_s^\alpha\right)}{d\alpha}\right]^2 \qquad (13)
\end{aligned}$$

Equation (11) follows immediately from equations (12) and (13).

■

A.2 Proof of Proposition 2

Step 1. We claim that there exists a constant K such that, for all y_c and y_s,

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right) - r^U\left(y_s^\alpha\right)}{r^V\left(y_c^\alpha\right) - r^V\left(y_s^\alpha\right)}\right) = K. \qquad (14)$$

Because the numerator and denominator both converge to zero, we apply L'Hopital's rule:

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right) - r^U\left(y_s^\alpha\right)}{r^V\left(y_c^\alpha\right) - r^V\left(y_s^\alpha\right)}\right) = \frac{\lim_{\alpha\to 0}\frac{dr^U(y_c^\alpha)}{d\alpha} - \lim_{\alpha\to 0}\frac{dr^U(y_s^\alpha)}{d\alpha}}{\lim_{\alpha\to 0}\frac{dr^V(y_c^\alpha)}{d\alpha} - \lim_{\alpha\to 0}\frac{dr^V(y_s^\alpha)}{d\alpha}}$$

Recall that $r^u(y)$ is defined by the following equation:

$$u\left(-r^u(y), \varphi(y)\right) = u\left(0, \varphi\left(y_0\right)\right)$$

Differentiating implicitly with respect to α and evaluating at y_f^α , we obtain:

$$\frac{dr^u\left(y_f^\alpha\right)}{d\alpha} = \frac{u_\varphi\left(-r^u\left(y_f^\alpha\right), \varphi\left(y_f^\alpha\right)\right)}{u_m\left(-r^u\left(y_f^\alpha\right), \varphi\left(y_f^\alpha\right)\right)}\left(\nabla\varphi\left(y_f^\alpha\right)\cdot\left(y_f - y_0\right)\right)$$

Using the fact that $\lim_{\alpha\to 0} r^u\left(y_f^\alpha\right) = 0$, it follows that

$$\lim_{\alpha\to 0}\frac{\partial r^u\left(y_f^\alpha\right)}{\partial\alpha} = \frac{u_\varphi\left(0, \varphi\left(y_0\right)\right)}{u_m\left(0, \varphi\left(y_0\right)\right)}\left(\nabla\varphi\left(y_0\right)\cdot\left(y_f - y_0\right)\right)$$

Therefore,

$$\lim_{\alpha\to 0}\frac{\partial r^u\left(y_c^\alpha\right)}{\partial\alpha} - \lim_{\alpha\to 0}\frac{\partial r^u\left(y_s^\alpha\right)}{\partial\alpha} = \frac{u_\varphi\left(0, \varphi\left(y_0\right)\right)}{u_m\left(0, \varphi\left(y_0\right)\right)}\left(\nabla\varphi\left(y_0\right)\cdot\left(y_c - y_s\right)\right)$$

Accordingly, we have

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right) - r^U\left(y_s^\alpha\right)}{r^V\left(y_c^\alpha\right) - r^V\left(y_s^\alpha\right)}\right) = \frac{U_\varphi\left(0, \varphi\left(y_0\right)\right)V_m\left(0, \varphi\left(y_0\right)\right)}{U_m\left(0, \varphi\left(y_0\right)\right)V_\varphi\left(0, \varphi\left(y_0\right)\right)} \equiv K > 0$$

As claimed, K does not depend on y_s or y_c (even though the limits of the numerator and denominator do).

Step 2. Proof of equations (6) and (7).

First consider $e = M$. We have

$$\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_c^\alpha), y_s^\alpha\right)}{d_M\left(r^V\left(y_c^\alpha\right), r^V\left(y_s^\alpha\right)\right)}\right) = \lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_c^\alpha), y_s^\alpha\right)}{d_M\left(r^U\left(y_c^\alpha\right), r^U\left(y_s^\alpha\right)\right)}\right)\lim_{\alpha\to 0}\frac{d_M\left(r_c^U\left(y_c^\alpha\right), r_s^U\left(y_s^\alpha\right)\right)}{d_M\left(r_c^V\left(y_c^\alpha\right), r_s^V\left(y_s^\alpha\right)\right)}$$

From Proposition 1, we know that the first term after the equals sign is unity. For the second term, we have

$$\lim_{\alpha\to 0}\frac{d_M\left(r_c^U\left(y_c^\alpha\right), r_s^U\left(y_s^\alpha\right)\right)}{d_M\left(r_c^V\left(y_c^\alpha\right), r_s^V\left(y_s^\alpha\right)\right)} = \lim_{\alpha\to 0}\left|\frac{r_c^U\left(y_c^\alpha\right) - r_s^U\left(y_s^\alpha\right)}{r_c^V\left(y_c^\alpha\right) - r_s^V\left(y_s^\alpha\right)}\right| = \left|\lim_{\alpha\to 0}\frac{r_c^U\left(y_c^\alpha\right) - r_s^U\left(y_s^\alpha\right)}{r_c^V\left(y_c^\alpha\right) - r_s^V\left(y_s^\alpha\right)}\right| = K$$

Consequently, equation (6) holds for $K_M = K$.

Now consider $e = W$. We have

$$\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{d_W\left(r^V\left(y_c^\alpha\right), r^V\left(y_s^\alpha\right)\right)}\right) = \lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_c^\alpha), y_s^\alpha, H^\alpha\right)}{d_W\left(r^U\left(y_c^\alpha\right), r^U\left(y_s^\alpha\right)\right)}\right)\lim_{\alpha\to 0}\frac{d_W\left(r_c^U\left(y_c^\alpha\right), r_s^U\left(y_s^\alpha\right)\right)}{d_W\left(r_c^V\left(y_c^\alpha\right), r_s^V\left(y_s^\alpha\right)\right)}$$

From Proposition 1, we know that the first term after the equals sign is h. For the second term, we have

$$\lim_{\alpha\to 0}\frac{d_W\left(r_c^U\left(y_c^\alpha\right),r_s^U\left(y_s^\alpha\right)\right)}{d_W\left(r_c^V\left(y_c^\alpha\right),r_s^V\left(y_s^\alpha\right)\right)}=\lim_{\alpha\to 0}\left(\frac{r_c^U\left(y_c^\alpha\right)-r_s^U\left(y_s^\alpha\right)}{r_c^V\left(y_c^\alpha\right)-r_s^V\left(y_s^\alpha\right)}\right)^2=\left(\lim_{\alpha\to 0}\frac{r_c^U\left(y_c^\alpha\right)-r_s^U\left(y_s^\alpha\right)}{r_c^V\left(y_c^\alpha\right)-r_s^V\left(y_s^\alpha\right)}\right)^2=K^2$$

Consequently, equation (7) holds for $K_W = hK^2$.

■

A.3 Proof of Proposition 3

Define

$$\kappa(y_s)\equiv\left(\frac{U_\varphi\left(0,\varphi\left(y_0\right)\right)\nabla\varphi\left(y_0\right)\cdot\left(y_s-y_0\right)}{U_m\left(0,\varphi\left(y_0\right)\right)}\right)$$

First consider $e = M$. Proceeding as in the proof of Proposition 2, we have

$$\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_{c\theta}^\alpha),y_s^\alpha\right)}{d_M\left(r^V\left(y_{c\theta}^\alpha,\theta\right),r^V\left(y_s^\alpha,\theta\right)\right)}\right)=\frac{U_\varphi\left(0,\varphi\left(y_0\right)\right)V_m\left(0,\varphi\left(y_0\right),\theta\right)}{U_m\left(0,\varphi\left(y_0\right)\right)V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)}$$

Furthermore,

$$\lim_{\alpha\to 0}\left(\frac{\alpha\kappa(y_s)}{r^V(y_s^\alpha,\theta)}\right)=\lim_{\alpha\to 0}\left(\frac{r^V(y_s^\alpha,\theta)}{\alpha}\right)^{-1}\kappa(y_s)=\left(\left.\frac{dr^V(y_s^\alpha,\theta)}{d\alpha}\right|_{\alpha=0}\right)^{-1}\kappa(y_s)$$

$$=\left(\frac{V_m\left(0,\varphi\left(y_0\right),\theta\right)}{V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)\nabla\varphi\left(y_0\right)\cdot\left(y_s-y_0\right)}\right)\kappa(y_s)=\frac{U_\varphi\left(0,\varphi\left(y_0\right)\right)V_m\left(0,\varphi\left(y_0\right),\theta\right)}{U_m\left(0,\varphi\left(y_0\right)\right)V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)}$$

Equation (8) is thereby verified.

Now consider $e = W$. Proceeding as in the proof of Proposition 2, we have

$$\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_{c\theta}^\alpha),y_s^\alpha,H^\alpha\right)}{d_W\left(r^V\left(y_{c\theta}^\alpha,\theta\right),r^V\left(y_s^\alpha,\theta\right)\right)}\right)=h\left(\frac{U_\varphi\left(0,\varphi\left(y_0\right)\right)V_m\left(0,\varphi\left(y_0\right),\theta\right)}{U_m\left(0,\varphi\left(y_0\right)\right)V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)}\right)^2$$

Furthermore,

$$\lim_{\alpha\to 0}h\left(\frac{\alpha\kappa(y_s)}{r^V(y_s^\alpha,\theta)}\right)^2=\lim_{\alpha\to 0}h\left(\frac{r^V(y_s^\alpha,\theta)}{\alpha}\right)^{-2}\left(\kappa(y_s)\right)^2=\left(\left.\frac{dr^V(y_s^\alpha,\theta)}{d\alpha}\right|_{\alpha=0}\right)^{-2}\left(\kappa(y_s)\right)^2$$

$$=h\left(\frac{V_m\left(0,\varphi\left(y_0\right),\theta\right)\kappa(y_s)}{V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)\nabla\varphi\left(y_0\right)\cdot\left(y_s-y_0\right)}\right)^2=h\left(\frac{U_\varphi\left(0,\varphi\left(y_0\right)\right)V_m\left(0,\varphi\left(y_0\right),\theta\right)}{U_m\left(0,\varphi\left(y_0\right)\right)V_\varphi\left(0,\varphi\left(y_0\right),\theta\right)}\right)^2$$

Equation (9) is thereby verified.

Finally, note that when y is a scalar, we have $\kappa\left(y_s\right) = \left(y_s - y_0\right)\kappa^*$, where

$$\kappa^* \equiv \left(\frac{U_\varphi\left(0, \varphi\left(y_0\right)\right)\varphi'\left(y_0\right)}{U_m\left(0, \varphi\left(y_0\right)\right)}\right).$$

■

B Additional theoretical results

For the approximations in the main test, we held the the magnitude of the misunderstanding fixed and let the consequences of the instrument shrink. Another possibility is to hold the scale of the instrument constant and let the magnitude of the misunderstanding shrink. As we noted in the text, the latter approach may provide better approximations for large instruments (which may be particularly useful in settings where risk preferences play a central role), but worse approximations for large misunderstandings. This section develops this alternative approach.

Formally, we define $y_c^\alpha = \alpha y_c + (1-\alpha)y_s$, and study the relationship between the welfare loss and the valuations $r^V\left(y_c^\alpha\right)$ and $r^V(y_s)$ as α shrinks to zero. Because the scale of the instrument remains fixed as we vary α, we take the distribution of prices, H, to be fixed as well. We also assume that the density, $h(p)$, is bounded and differentiable with a bounded derivative. This alternative formulation requires a slight modification of the welfarist criterion. As before, we convert l_e^u to money-metric utility (for $e \in \{M, W\}$), but in this case we divide by $u_m\left(-r^u(y_s), y_s\right)$.

We now state and discuss counterparts for the three propositions in the main text. Proofs follow.

Proposition 4. *For all y_c satisfying $\left.\frac{dr^V(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} \neq 0$, we have*

$$\lim_{\alpha\to 0}\left(\frac{L_M^V\left(r^V(y_c^\alpha), y_s\right)}{d_M\left(r^V(y_c^\alpha), r^V(y_s)\right)}\right) = 1 \tag{15}$$

and

$$\lim_{\alpha\to 0}\left(\frac{L_W^V\left(r^V(y_c^\alpha), y_s, H\right)}{d_W\left(r^V(y_c^\alpha), r^V(y_s)\right)}\right) = h\left(r^V(y_s)\right) \tag{16}$$

Two distinctions between Propositions 1 and 4 merit discussion. First, the constant of proportionality in equation (16) involves $h\left(r^V(y_s)\right)$ rather than the scalar h. In either case, to aggregate over instruments, one must assume that the pertinent density is uncorrelated with measured deliberative competence. Second, for Proposition 4, the money-metric scaling factor $u_m\left(-r^u(y_s), y_s\right)$ also depends on the instrument through y_s. Aggregation over instruments therefore requires either that the differences in the marginal utility of income are small (e.g., because curvature is modest over the relevant range), or that measured deliberative competence is uncorrelated with simply framed valuations across instruments.

Proposition 5. *For each $e \in \{W, M\}$, there exists a strictly positive function $K_e(y_s)$, such that for all y_c satisfying $\left.\frac{dr^V(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} \neq 0$, we have*

$$\lim_{\alpha \to 0} \left(\frac{L_M^U\left(r^U\left(y_c^\alpha\right), y_s\right)}{d_M\left(r^V\left(y_c^\alpha\right), r^V\left(y_s\right)\right)} \right) = K_M\left(y_s\right) \tag{17}$$

$$\lim_{\alpha \to 0} \left(\frac{L_W^U\left(r^U\left(y_c^\alpha\right), y_s, H\right)}{d_W\left(r^V\left(y_c^\alpha\right), r^V\left(y_s\right)\right)} \right) = K_W\left(y_s\right) \tag{18}$$

An important difference between Propositions 2 and 5 is that, for the latter, the constants of proportionality depend on the instrument (even with no differences in the density term across instruments for the welfarist criterion). Once again, this feature introduces some qualifications with respect to aggregating over instruments for the alternative approach.

Our next proposition concerns potential policy-induced confounds. It identifies an adjustment to measured deliberative competence required to restore comparability across policies when they potentially impact simply framed valuations. It involves a function $\rho^V\left(y_s, \theta\right)$, defined as follows. Let $y_s^\beta = \beta y_s + (1-\beta) y_0$, where β is a scalar. Then $\rho^V\left(y_s, \theta\right) = \left.\frac{dr^V(y_s^\beta, \theta)}{d\beta}\right|_{\beta=1}$. In words, $\rho^V\left(y_s, \theta\right)$ is the amount by which the simply framed valuation changes as we scale up the instrument's consequences. Critically, the value of this term is easily inferred from the type of data we collected in our experiment.

Proposition 6. *For each $e \in \{W, M\}$, there exists a strictly positive function $\kappa_e(y_s)$ such that, for all $y_{c\theta}$ satisfying $\left.\frac{dr^V(y_c^\alpha, \theta)}{d\alpha}\right|_{\alpha=0} \neq 0$ and policies θ, we have*

$$\lim_{\alpha \to 0} \left(\frac{L_M^U\left(r^U(y_c^\alpha), y_s\right)}{d_M^V\left(r^V\left(y_{c\theta}^\alpha, \theta\right), r^V\left(y_s, \theta\right)\right)} \right) = \frac{\kappa_M(y_s)}{\rho^V\left(y_s, \theta\right)} \tag{19}$$

and

$$\lim_{\alpha \to 0} \left(\frac{L_W^U\left(r^U(y_c^\alpha), y_s, H\right)}{d_W^V\left(r^V\left(y_{c\theta}^\alpha, \theta\right), r^V\left(y_s, \theta\right)\right)} \right) = \frac{\kappa_W(y_s)}{\left[\rho^V\left(y_s, \theta\right)\right]^2} \tag{20}$$

According to this proposition, we can address the confound simply by rescaling our measures of deliberative competence. In particular, once we divide $d_M^V\left(r^V\left(y_{c\theta}, \theta\right), r^V\left(y_s, \theta\right)\right)$ by $\rho^V\left(y_s, \theta\right)$, and $d_W^V\left(r^V\left(y_{c\theta}, \theta\right),\ r^V\left(y_s, \theta\right)\right)$ by $\left[\rho^V\left(y_s, \theta\right)\right]^2$, the main implications of Proposition 5 follow.

Section D.6 uses Proposition 6 to correct for policy-induced framing effects and shows that our empirical results are qualitatively unchanged.

B.1 Proof of Proposition 4

In proving this proposition, we write the evaluation standard as $u \in \{V, U\}$, rather than as V, so that we can invoke this result for both V and U in the proof of Proposition 5.

The proof involves two steps.

Step 1: Proof of equation (15).

Suppose first that $\left.\frac{dr^u(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} < 0$. Hence, for all sufficiently small α we have that $r^u\left(y_s\right) > r^u\left(y_c^\alpha\right)$ and that $r^u\left(y_c^\alpha\right)$ is strictly increasing. In this case the denominator of the expression in equation (15) is $r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)$. Because $L_M^u\left(r^u(y_c^\alpha), y_s\right)$ (the numerator) and $r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)$ both converge to zero as $\alpha \to 0$, we apply l'Hopital's rule.

For the numerator, the facts that u is strictly increasing in its first component, and that $r^u\left(y_c^\alpha\right)$ is strictly increasing, imply

$$L_M^u\left(r^u(y_c^\alpha), y_s\right) = \max_{p\in[r^u(y_c^\alpha), r^u(y_s)]}\left[\frac{u\left(-p, y_s\right) - u\left(0, y_0\right)}{u_m(-r^u(y_s), y_s)}\right] = \frac{u\left(-r^u\left(y_c^\alpha\right), y_s\right) - u\left(0, y_0\right)}{u_m(-r^u(y_s), y_s)}$$

It follows that

$$\frac{dL_M^u\left(r^u(y_c^\alpha), y_s\right)}{d\alpha} = -\frac{u_m\left(-r^u\left(y_c^\alpha\right), y_s\right)}{u_m(-r^u(y_s), y_s)}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha},$$

and accordingly that

$$\lim_{\alpha\to 0}\frac{dL_M^u\left(r^u(y_c^\alpha), y_s\right)}{d\alpha} = -\lim_{\alpha\to 0}\frac{dr^u\left(y_c^\alpha\right)}{d\alpha}$$

For the denominator, the derivative is $-\frac{dr^u(y_c^\alpha)}{d\alpha}$, so the limit is the same as for the numerator. Equation (15) follows (in this case) immediately from L'Hopital's rule.

Now suppose $\left.\frac{dr^u(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} > 0$, so that $r^u\left(y_s\right) > r^u\left(y_c^\alpha\right)$ for small α, in which case the denominator of the expression in equation (15) is $r^u\left(y_c^\alpha\right) - r^u\left(y_s\right)$. In this case, for the numerator, we have

$$L_M^u\left(r^u(y_c^\alpha), y_s\right) = \max_{p\in[r^u(y_s), r^u(y_c^\alpha)]}\left[\frac{u\left(0, y_0\right) - u\left(-p, y_s\right)}{u_m(-r^u(y_s), y_s)}\right] = \frac{u\left(0, y_0\right) - u\left(-r^u\left(y_c^\alpha\right), y_s\right)}{u_m(-r^u(y_s), y_s)}$$

The remainder of the argument is the same, except that the signs of the numerator and denominator are both switched, so the limiting ratio of derivatives is again unity.

Step 2: Proof of equation (16).

Suppose first $\left.\frac{dr^u(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} < 0$, so that $r^u\left(y_s\right) > r^u\left(y_c^\alpha\right)$ for small α. Because $L_W^u\left(r^u(y_c^\alpha), y_s, H\right)$ (the numerator of the expression in equation (16)) and $d_W\left(r^u\left(y_c^\alpha\right), r^u\left(y_s\right)\right)$ (the denominator) both converge to zero as $\alpha \to 0$, we apply l'Hopital's rule.

Consider the numerator, $L_W^u\left(r^u(y_c^\alpha), y_s, H\right)$. By definition,

$$L_W^u\left(r^u(y_c^\alpha), y_s, H\right) = \frac{1}{u_m(-r^u(y_s), y_s)}\int_{r^u(y_c^\alpha)}^{r^u(y_s)}\left[u\left(-p, y_s\right) - u\left(0, y_0\right)\right]h(p)dp$$

It follows that

$$\frac{d}{d\alpha}L_W^u\left(r^u(y_c^\alpha), y_s, H\right) = -\left[\frac{u\left(-r^u\left(y_c^\alpha\right), y_s\right) - u\left(0, y_0\right)}{u_m(-r^u(y_s), y_s)}\right]\frac{dr^u\left(y_c^\alpha\right)}{d\alpha}h\left(r^u\left(y_c^\alpha\right)\right)$$

Given our assumptions concerning bounds on the density and derivatives of the utility functions, $\frac{dr^u(y_c^\alpha)}{d\alpha}$, $\frac{1}{u_m(-r^u(y_s), y_s)}$, and $h\left(r^u\left(y_c^\alpha\right)\right)$ are all bounded, so this expression converges to zero as $\alpha \to 0$. Anticipating the same property for the

denominator, we take the second derivative:

$$\frac{d^2}{d\alpha^2} L_W^u \left(r^u(y_c^\alpha), y_s, H\right) = -\left[\frac{u\left(-r^u\left(y_c^\alpha\right), y_s\right) - u\left(0, y_0\right)}{u_m(-r^u(y_s), y_s)}\right]\left[\frac{d^2 r^u\left(y_c^\alpha\right)}{d\alpha^2} h\left(r^u\left(y_c^\alpha\right)\right) + \frac{d r^u\left(y_c^\alpha\right)}{d\alpha}\frac{d h\left(r^u\left(y_c^\alpha\right)\right)}{d\alpha}\right]$$

$$+\frac{u_m\left(-r^u\left(y_c^\alpha\right), y_s\right)}{u_m(-r^u(y_s), y_s)}\left(\frac{d r^u\left(y_c^\alpha\right)}{d\alpha}\right)^2 h\left(r^u\left(y_c^\alpha\right)\right)$$

As before, our assumptions imply that $\frac{dr^u(y_c^\alpha)}{d\alpha}$, $\frac{1}{u_m(-r^u(y_s),y_s)}$, and $h\left(r^u\left(y_c^\alpha\right)\right)$ are bounded. Boundedness of the derivative of the density then implies that $\frac{dh(r^u(y_c^\alpha))}{d\alpha}$ is bounded. It is also straightforward to verify that $\frac{d^2 r^u(y_c^\alpha)}{d\alpha^2}$ is bounded under our assumptions. It follows that, as $\alpha \to 0$, the first term converges to zero. Because the ratio in the second term converges to unity, we have:

$$\lim_{\alpha\to 0}\frac{d^2}{d\alpha^2} L_W^u \left(r^u(y_c^\alpha), y_s, H\right) = h\left(r^u\left(y_s\right)\right)\lim_{\alpha\to 0}\left(\frac{d r^u\left(y_c^\alpha\right)}{d\alpha}\right)^2$$

Now consider the denominator, $\frac{1}{2}\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)^2$. Taking the derivative yields

$$\frac{d}{d\alpha}\left[\frac{1}{2}\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)^2\right] = -\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)\frac{d r^u\left(y_c^\alpha\right)}{d\alpha}$$

Because $\frac{dr^u(y_c^\alpha)}{d\alpha}$ is bounded, this expression converges to 0 as $\alpha \to 0$, as claimed above. Taking the second derivative yields

$$\frac{d^2}{d\alpha^2}\left[\frac{1}{2}\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)^2\right] = -\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)\frac{d^2 r^u\left(y_c^\alpha\right)}{d\alpha^2} + \left(\frac{d r^u\left(y_c^\alpha\right)}{d\alpha}\right)^2$$

As $\alpha \to 0$, the first term converges to zero (given the boundedness of the second derivative noted above), so we have

$$\lim_{\alpha\to 0}\frac{d^2}{d\alpha^2}\left[\frac{1}{2}\left(r^u\left(y_s\right) - r^u\left(y_c^\alpha\right)\right)^2\right] = \lim_{\alpha\to 0}\left(\frac{d r^u\left(y_c^\alpha\right)}{d\alpha}\right)^2$$

Thus, the ratio of the limit of second derivatives for the numerator and the denominator is $h\left(r^u\left(y_s\right)\right)$. Equation (16) follows (in this case) immediately from L'Hopital's rule.

Now suppose $\left.\frac{dr^u(y_c^\alpha)}{d\alpha}\right|_{\alpha=0} > 0$, so that $r^u\left(y_s\right) > r^u\left(y_c^\alpha\right)$ for small α. The same analysis of the denominator applies. For the numerator, the limits of the integral change places, as do the two terms in the integrand. As a result, the first and second derivatives are unchanged. Consequently, the same arguments deliver equation (16) in this case as well.

■

B.2 Proof of Proposition 5

The proof involves two steps.

Step 1: We claim that there exists a constant $K(y_s)$ such that, for all y_c and y_s,

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right)=K(y_s). \tag{21}$$

Because the numerator and denominator of the preceding expression both converge to zero, we apply L'Hopital's rule:

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right)=\frac{\lim_{\alpha\to 0}\frac{dr^U(y_c^\alpha)}{d\alpha}}{\lim_{\alpha\to 0}\frac{dr^V(y_c^\alpha)}{d\alpha}}$$

Recall that $r^V(y)$ is defined by the following equation:

$$V\left(-r^V(y),\varphi(y)\right)=V\left(0,\varphi\left(y_0\right)\right)$$

Evaluating $r^V(y)$ at y_c^α and differentiating implicitly with respect to α, we obtain:

$$\frac{dr^V\left(y_c^\alpha\right)}{d\alpha}=\frac{V_\varphi\left(-r^V\left(y_c^\alpha\right),\varphi\left(y_c^\alpha\right)\right)}{V_m\left(-r^V\left(y_c^\alpha\right),\varphi\left(y_c^\alpha\right)\right)}\left(\nabla\varphi\left(y_c^\alpha\right)\cdot\left(y_c-y_s\right)\right)$$

Using the fact that $\lim_{\alpha\to 0} r^V\left(y_c^\alpha\right)=r^V(y_s)$, it follows that

$$\lim_{\alpha\to 0}\frac{dr^V\left(y_c^\alpha\right)}{d\alpha}=\frac{V_\varphi\left(-r^V(y_s),\varphi\left(y_0\right)\right)}{V_m\left(-r^V(y_s),\varphi\left(y_0\right)\right)}\left(\nabla\varphi\left(y_0\right)\cdot\left(y_c-y_s\right)\right)$$

Repeating these calculations for U, we obtain

$$\lim_{\alpha\to 0}\frac{dr^U\left(y_c^\alpha\right)}{d\alpha}=\frac{U_\varphi\left(-r^U(y_s),\varphi\left(y_0\right)\right)}{U_m\left(-r^U(y_s),\varphi\left(y_0\right)\right)}\left(\nabla\varphi\left(y_0\right)\cdot\left(y_c-y_s\right)\right)$$

Accordingly, we have

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right)=\frac{U_\varphi\left(-r^U(y_s),\varphi\left(y_s\right)\right)V_m\left(-r^V(y_s),\varphi\left(y_s\right)\right)}{U_m\left(-r^U(y_s),\varphi\left(y_s\right)\right)V_\varphi\left(-r^V(y_s),\varphi\left(y_s\right)\right)}\equiv K(y_s)>0$$

As claimed, K does not depend on y_c (even though the limits of the derivatives of the numerator and denominator with respect to α in equation (21) do).

Step 2: Proof of equations (17) and (18).

Consider $e=M$. We have

$$\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_c^\alpha),y_s\right)}{d_M\left(r^V\left(y_c^\alpha\right),r^V\left(y_s\right)\right)}\right)=\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_c^\alpha),y_s\right)}{d_M\left(r^U\left(y_c^\alpha\right),r^U\left(y_s\right)\right)}\right)\lim_{\alpha\to 0}\left(\frac{d_M\left(r^U\left(y_c^\alpha\right),r^U\left(y_s\right)\right)}{d_M\left(r^V\left(y_c^\alpha\right),r^V\left(y_s\right)\right)}\right)$$

From Proposition 1 (which holds for U as well as V), we know that the first limit after the equals sign converges to unity. For the second term, we have

$$\lim_{\alpha\to 0}\frac{\left|r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)\right|}{\left|r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)\right|}=\lim_{\alpha\to 0}\left|\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right|=\left|\lim_{\alpha\to 0}\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right|=K(y_s)$$

Therefore, equation (17) holds for $K_M(y_s)=K(y_s)$.

Now consider $e=W$. We have

$$\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_c^\alpha),y_s,H\right)}{d_W\left(r^V\left(y_c^\alpha\right),r^V\left(y_s\right)\right)}\right)=\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_c^\alpha),y_s,H\right)}{d_W\left(r^U\left(y_c^\alpha\right),r^U\left(y_s\right)\right)}\right)\lim_{\alpha\to 0}\left(\frac{d_W\left(r^U\left(y_c^\alpha\right),r^U\left(y_s\right)\right)}{d_W\left(r^V\left(y_c^\alpha\right),r^V\left(y_s\right)\right)}\right)$$

From Proposition 1, we know that the limit after the equals sign converges to $h\left(r^U(y_s)\right)$. For the second term, we have

$$\lim_{\alpha\to 0}\frac{\left(r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)\right)^2}{\left(r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)\right)^2}=\lim_{\alpha\to 0}\left(\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right)^2=\left(\lim_{\alpha\to 0}\frac{r^U\left(y_c^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_c^\alpha\right)-r^V\left(y_s\right)}\right)^2=\left[K\left(y_s\right)\right]^2$$

Therefore, equation (18) holds for $K_W\left(y_s\right)=h\left(r^U(y_s)\right)\left[K\left(y_s\right)\right]^2$.

■

B.3 Proof of Proposition 6

Proceeding as in the proof of Proposition 5, Step 1, we see that, for any given θ,

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_{c\theta}^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_{c\theta}^\alpha,\theta\right)-r^V\left(y_s,\theta\right)}\right)=\frac{U_\varphi\left(-r^U(y_s),\varphi\left(y_s\right)\right)V_m\left(-r^V(y_s,\theta),\varphi\left(y_s\right),\theta\right)}{U_m\left(-r^U(y_s),\varphi\left(y_s\right)\right)V_\varphi\left(-r^V(y_s,\theta),\varphi\left(y_s\right),\theta\right)} \tag{22}$$

Moreover, it is straightforward to check that

$$\rho^V\left(y_s,\theta\right)=\left.\frac{\partial r^V\left(y_s^\beta,\theta\right)}{\partial\beta}\right|_{\beta=1}=\frac{V_\varphi\left(-r^V(y_s,\theta),\varphi\left(y_s\right),\theta\right)}{V_m\left(-r^V(y_s,\theta),\varphi\left(y_s\right),\theta\right)}\left(\nabla\varphi\left(y_s\right)\cdot\left(y_s-y_0\right)\right) \tag{23}$$

Defining $\kappa(y_s)\equiv\frac{U_\varphi(-r^U(y_s),\varphi(y_s))}{U_m(-r^U(y_s),\varphi(y_s))}\nabla\varphi\left(y_s\right)\cdot\left(y_s-y_0\right)$, equations (22) and (23) then imply

$$\lim_{\alpha\to 0}\left(\frac{r^U\left(y_{c\theta}^\alpha\right)-r^U\left(y_s\right)}{r^V\left(y_{c\theta}^\alpha,\theta\right)-r^V\left(y_s,\theta\right)}\right)=\frac{\kappa(y_s)}{\rho^V\left(y_s,\theta\right)} \tag{24}$$

Consider $e=M$. Analogously to Step 2 of Proposition 5, we have

$$\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_{c\theta}^\alpha),y_s\right)}{d_M\left(r^V\left(y_{c\theta}^\alpha,\theta\right),r^V\left(y_s,\theta\right)\right)}\right)=\lim_{\alpha\to 0}\left(\frac{L_M^U\left(r^U(y_{c\theta}^\alpha),y_s\right)}{d_M\left(r^U\left(y_{c\theta}^\alpha\right),r^U\left(y_s\right)\right)}\right)\lim_{\alpha\to 0}\left(\frac{d_M\left(r^U\left(y_{c\theta}^\alpha\right),r^U\left(y_s\right)\right)}{d_M\left(r^V\left(y_{c\theta}^\alpha,\theta\right),r^V\left(y_s,\theta\right)\right)}\right) \tag{25}$$

Proposition 4, as applied to U, is unchanged. Therefore, the first term on the right equals 1. Using the same arguments as in Step 2 of Proposition 5, equation (19) follows directly from equations (24) and (25), with $\kappa_M(y_s) = \kappa(y_s)$.

Now consider $e = W$. Analogously to Step 2 of Proposition 5, we have

$$\lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_{c\theta}^\alpha), y_s, H\right)}{d_W\left(r^V\left(y_{c\theta}^\alpha, \theta\right), r^V\left(y_s, \theta\right)\right)}\right) = \lim_{\alpha\to 0}\left(\frac{L_W^U\left(r^U(y_{c\theta}^\alpha), y_s, H\right)}{d_W\left(r^U\left(y_{c\theta}^\alpha\right), r^U\left(y_s\right)\right)}\right)\lim_{\alpha\to 0}\left(\frac{d_W\left(r^U\left(y_{c\theta}^\alpha\right), r^U\left(y_s\right)\right)}{d_W\left(r^V\left(y_{c\theta}^\alpha, \theta\right), r^V\left(y_s, \theta\right)\right)}\right) \tag{26}$$

Again, Proposition 4, as applied to U, is unchanged. Therefore, the first term on the right equals $h\left(r^U\left(y_s\right)\right)$. Using the same arguments as in Step 2 of Proposition 5, equation (20) follows directly from equations (24) and (26), with $\kappa_W(y_s) = \left[\kappa(y_s)\right]^2 h\left(r^U(y_s)\right)$.

■

C Experiment details

In this section we detail the implementation of the experiment. Screenshots of the instructions and the experimental interface are in Appendix E.

Amazon Mechanical Turk Workers log on to AMT through an interface that displays a list of *Human Intelligence Tasks* (HITs), each with a title, an estimated duration, and an estimated remuneration rate. Other HITs include taking surveys, categorizing images, writing product descriptions, and identifying performers on music recordings.

To ensure that subjects were *technically* able to view the videos, we told them at the outset of the study that access to youtube.com was required. We also asked them to reproduce the last word spoken in the welcome video, and the last word of the title slide of whichever treatment video they viewed. Subjects who were not able to complete these tasks correctly were not allowed to continue with the study. The videos were embedded in the survey so that subjects could not find the other treatment videos used in this study.

We ensured that each subject participated in our study only once using the unique identifying numbers assigned by AMT.[53] A subject can only receive payment for participation in the study if she correctly provides this information, and hence has no incentive for misrepresentation.

Initial Financial Literacy Before participating in the main stages of the experiment, subjects completed the unincentivized financial literacy test in Table C.1. This test of financial literacy originated with Lusardi and Mitchell (2009) and van Rooij et al. (2011), and has been used in many other studies (Lusardi and Mitchell, 2014).

Attention to the Video Before subjects watched the treatment video, we informed them that, with 25% probability, their earnings would be entirely determined by their performance on a test,[54] and that 'to be able to answer the questions

[53] Nonetheless, one subject managed to participate in our study twice. Both times, this subject exhibited multiple switching points, and hence is excluded from all analyses.

[54] Hastings et al. (2013) criticize most existing studies that use such test scores as outcome measures on the grounds that the tests are unincentivized. One of the few exceptions is Levy and Tasoff (2016).

in the test, you need to both understand and know the contents of the video.' We also explained that the video could help them make better decisions both during the experiment and in real life, inasmuch as it was made by 'internationally recognized academic experts on financial decision making.' Finally, we disabled the *continue* button for the duration of the video.

Iterated Multiple Price List Each line of each price list was a binary choice between the future reward and a specified dollar amount to be received no more than two days after completion of the experiment. For the first price list, the immediate payment varied from \$0 to \$20 in increments of \$2. For the second price list, it varied from $\$x$ to $\$(x+1.8)$ in increments of \$0.20, where $x+2$ is the smallest amount chosen over the future reward in the first list. (See appendix E for screenshots of the computer interface.) If a subjects' payment was determined according to a price list, the randomization over lines proceeded as follows. A line was randomly selected from the first price list. If that line did not correspond to x (defined above), it was implemented. Otherwise, a random line from the second price list was selected, and the decision for that line was implemented. With this procedure, truthful revelation of preferences is optimal.

Questionnaire Questions concerning decision strategies employed the following wording. Use of the rule of 72 in complexly framed problems: "Sometimes in this experiment, you were given a choice such as 'We will invest \$10 in an account with 1% interest per week. Interest is compounded weekly. We will pay you the proceeds in 72 days.' When deciding about this choice, did you use the rule of 72?"[55] Use of the rule of 72 in simply framed problems: "Sometimes in this experiment, you were given a choice such as 'We will pay you \$20 in 36 days.' When deciding about such a choice, did you use the rule of 72?" In both cases, subjects answered either "Yes", "No", or "I don't know the rule of 72." Number of problems for which the future reward was calculated explicitly: "In total, you were given 10 rounds in which one of the options was something like 'we will invest \$... in an account with ...% interest per day. Interest is compounded daily. We will pay you the proceeds in... days.' Out of these 10 rounds, how many times did you explicitly calculate the money amount that this investment would yield within the specified time?" Subjects responded by selecting an integer between 0 and 10. Use of external help on the test: "When you completed the test about the video on financial investing, did you use external resources (such as other websites, books, etc.) to find the right answers?" Subjects answered either "Yes" or "No."

We also asked subjects how much attention they had paid to their choices, how much attention they had paid to the video, whether they had any suggestions about the study, and whether they had experienced any technical difficulties. The overwhelming majority of subjects reported the highest level of attention in answer to both questions—a finding we interpret with caution.

[55] The survey question incorrectly described the interest rate as pertaining to a week rather than a day. We believe the meaning of the question was nevertheless clear despite this typo.

Table C.1: Financial Literacy questionnaire.

FL1. Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow? *More than $102 (92.03%), Exactly $102 (4.61%), Less than $102 (1.99%), Do not know (1.37%)*
FL2. Suppose you had $100 in a savings account and the interest rate is 20 percent per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total? *More than $200 (75.84%), Exactly $200 (19.68%), Less than $200 (2.74%), Do not know (1.74%)*
FL3. Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, how much would you be able to buy with the money in this account? *More than today (7.35%), exactly the same (4.98%), less than today (84.43%), do not know (3.24%)*
FL4. Assume a friend inherits $10,000 today and his sibling inherits $10,000 3 years from now. Who is richer because of the inheritance? *My friend (56.79%), his sibling (6.72%), they are equally rich (29.89%), do not know (6.60%)*
FL5. Suppose that in the year 2015, your income has doubled and prices of all goods have doubled too. In 2015, how much will you be able to buy with your income? *More than today (4.23%), the same (90.16%), less than today (4.73%), do not know (0.87%)*

Notes: Numbers in brackets indicate the percentage of subjects who chose a given answer. The fraction of correct answers to each question only differs slightly across experiments and treatments, see Table D.1.

Test questions in Control interventions Panels A and B of Table C.2 display the test questions about the Control intervention in Experiments A and B, respectively. The two experiments involve different sets of questions about their respective Control interventions because the Control interventions differ. We decided to use a different Control intervention for Experiment B because the Control intervention in Experiment A is largely descriptive, and hence is not well-suited to incorporating practice questions with individualized feedback.

Table C.2: Test questions concerning the Control interventions.

Experiment A

Which of the following quotes is attributed to Benjamin Franklin?
Compound interest is the most powerful force in the universe; Youth is wasted on the young; Money makes money. And the money that money makes, makes money;

Which quote is attributed to the author Upton Sinclair?
Only liars manage always to be out of the market during bad times and in during good times; It is difficult to get a man to understand something when his salary depends upon his not understanding it; There are three classes of people who do not believe that markets work: the Cubans, the North Koreans, and active managers; Nobody knows more than the market.

What percentage of mutual funds tends to be outperformed by the market (S&P 500 Index) each year?
between 10 and 30% between 30 and 50% between 50 and 70% between 70 and 90%

What is an "indexing" investment strategy?
Buying index funds, which hold assets that have been indexed as particularly profitable by financial experts; Buying index funds, which hold stocks of companies that provide information about the stock market as a whole (stock market indices); Buying index funds, which hold the market portfolio; Buying index funds, which hold optimally diversified, custom tailored portfolios.

Professional investors as a whole are responsible for what percentage of stock market trading?
30%; 50%; 70%; 90%.

Experiment B

In order to limit your risk, you might invest in which of the following pairs of stocks?
Microsoft and Google; General Motors and Chrysler; Coca-Cola and Pepsi; General Motors and Microsoft; Facebook and Twitter.

We would expect the degree of relation between the returns of Coca-Cola stock and the returns of Pepsi stock to be closest to _____? [-1 means perfect negative relation and +1 means perfect positive relation]:
-0.7; -0.3; 0; 0.3; 0.7.

Considering a long time period (for example 10 or 20 years), which asset normally gives the highest return?
Savings accounts; Corporate bonds; Government bonds; T-Bills; Stocks.

Normally, which asset displays the highest fluctuations over time?
Savings accounts; Corporate bonds; Government bonds; T-Bills; Stocks.

A degree of relation of _____between two assets will NOT help reduce your risk.
1; 0.5; 0; -0.5; -1.

D Additional Data Analysis

D.1 Demographics

Table D.1 presents detailed demographics of our subject pool by treatment, as well as their initial financial literacy.[56] Column 5 lists data for the representative US citizen. Demographic variables are taken from the 2010 US Census. Employment variables are for April 2014, and come from the Bureau of Labor Statistics. Financial literacy scores are from Lusardi (2011), and from the 2012 FED bulletin for stock holdings.[57] (Representative data on financial literacy only exist for questions FL1 and FL3.) For empty cells, no representative data are available. Column 6 reports, for each variable, the p-value of an F-test for differences across treatments. The number of significant differences is well within the range we would expect given the number of tests performed.

As reported in section 2.2, our sample is poorer, better educated, and more likely to live in larger households than the average US citizen. While the incidence of full-time employment in our sample mirrors that of the general population, the fraction of respondents who classify themselves as employed part-time is double that of the general population. Our subjects are also disproportionately male and white, younger, slightly more urban, and more likely to have never been married than the representative US citizen.

D.2 Main results controlling for demographics

Table D.2 presents our main results in a regression that includes data from both experiments and controls for demographics. Demographic controls consist of all variables listed in Table D.1, except for the summary statistics "FL1-FL3 all correct" and "FL1-FL5 all correct". For brevity, we pool across the timeframes.

In each case we see that coefficient estimates are barely changed in comparison to the estimates in the main text, which do not control for demographic characteristics. We conclude that the differences between experiments A and B reflect differences in the interventions rather than differences in subject characteristics.

[56] These statistics only include subjects who did not exhibit multiple switching points in any of the price lists.

[57] J. Bricker, A. B. Kennickell K. B. Moore, and J. Sabelhaus, 2012, *Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances*, Federal Reserve Bulletin, 98(2).

[58] In our survey, household income is interval coded. The values stated are the midpoints of the median intervals.

[59] Percentage of civilian noninstitutional population that is full-time employed.

[60] Percentage of civilian noninstitutional population that is part-time employed.

[61] Our questionnaire included the option "Prefer not to say". The three subjects who chose this response are not accounted for in this table.

Table D.1: Demographics and financial literacy.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Experiment A					Experiment B			
Treatment	Treat.	Cont.	Subst. only	Rhet. only	p-value	Treat.	Cont.	p-value	US
FL1	0.93	0.92	0.92	0.95	0.83	0.90	0.92	0.59	0.65
FL2	0.81	0.73	0.73	0.71	0.32	0.75	0.80	0.29	-
FL3	0.82	0.82	0.83	0.85	0.93	0.88	0.85	0.38	0.64
FL4	0.58	0.64	0.50	0.59	0.17	0.53	0.59	0.31	-
FL5	0.96	0.90	0.87	0.91	0.09*	0.91	0.88	0.30	-
FL1 - FL3 all correct	0.71	0.63	0.62	0.62	0.44	0.68	0.70	0.72	-
FL1 - FL5 all correct	0.47	0.45	0.34	0.40	0.20	0.40	0.42	0.60	-
Male	0.57	0.57	0.61	0.50	0.40	0.47	0.53	0.28	0.49
Age (median)	28	32	29	29	0.09*	36	34	0.11	37.2
Household Income (median)[58]	45,000	35,000	45,000	35,000	0.06**	45,000	57,500	0.15	53,046
Race									
African-american	0.08	0.06	0.08	0.04	0.68	0.05	0.10	0.10*	0.13
Asian	0.08	0.11	0.12	0.05	0.22	0.06	0.06	0.90	0.05
Caucasian	0.81	0.72	0.72	0.77	0.34	0.82	0.79	0.41	0.63
Hispanic	0.03	0.07	0.03	0.10	0.06**	0.04	0.04	0.88	0.17
Other	0.01	0.04	0.05	0.04	0.44	0.02	0.01	0.37	0.02
Education									
Less than high school	0.01	0.00	0.00	0.00	0.35	0.00	0.00	1.00	0.14
High school	0.13	0.12	0.15	0.14	0.92	0.14	0.13	0.71	0.31
Vocational / technical	0.08	0.08	0.08	0.03	0.29	0.07	0.06	0.56	0.09
Some college	0.35	0.37	0.33	0.44	0.34	0.31	0.34	0.67	0.19
College	0.39	0.37	0.38	0.34	0.90	0.37	0.37	0.97	0.18
Graduate degree	0.05	0.06	0.07	0.05	0.88	0.11	0.11	0.88	0.09
Employment									
Full time employed	0.50	0.50	0.48	0.43	0.70	0.61	0.68	0.16	0.48[59]
Part time employed	0.21	0.23	0.26	0.27	0.72	0.18	0.16	0.60	0.11[60]
Marital Status[61]									
Married	0.28	0.30	0.32	0.29	0.94	0.46	0.47	0.80	0.27
Widowed	0.00	0.00	0.00	0.00	1.00	0.01	0.01	0.53	0.56
Divorced	0.07	0.05	0.04	0.04	0.80	0.04	0.06	0.37	0.06
Never married	0.64	0.65	0.64	0.64	1.00	0.49	0.46	0.69	0.10
Urban / Rural									
Urban and suburban	0.17	0.17	0.11	0.17	0.48	0.25	0.17	0.06*	0.81
Rural	0.83	0.83	0.89	0.83	0.48	0.75	0.83	0.06*	0.19
Household size									
1	0.18	0.13	0.11	0.19	0.26	0.10	0.13	0.33	0.22
2	0.22	0.24	0.25	0.24	0.95	0.27	0.23	0.42	0.36
3	0.14	0.19	0.17	0.22	0.46	0.18	0.22	0.35	0.17
4 or more	0.46	0.44	0.47	0.35	0.23	0.45	0.41	0.50	0.26
Owns stocks	0.16	0.23	0.20	0.23	0.54	0.43	0.39	0.44	0.15
N	109	106	128	112	-	169	179	-	-

Notes: The sample includes all subjects who completed the study and did not exhibit any multiple switching points. Column 5 presents the *p*-values of an *F*-test for joint equality of the coefficients listed in columns 1 - 4. Column 8 presents the *p*-value of a *t*-test for joint equality of the coefficients in columns 6 - 7. Column 9 lists comparison values for the representative US citizen whenever available. See text for data sources.

Table D.2: Main results controlling for demographics in joint analysis of both experiments.

VARIABLES	(1) Test scores on module	(2)	(3) Valuation in frame	(4)	(5) Deliberative Competence	(6)
	Treatment	Control	Complex	Simple		
Correction for changes in valuations in simple frame					No	Yes
Difference to Control B						
Treatment B	1.640***	-0.375***	15.067***	7.242***	7.901***	12.883***
	(0.139)	(0.124)	(2.790)	(2.164)	(1.773)	(2.563)
Treatment A	1.528***	-0.660***	11.928***	1.522	1.133	-1.283
	(0.197)	(0.158)	(3.969)	(3.189)	(2.521)	(4.652)
Control A	0.117	0.405**	-1.504	1.121	0.832	3.151
	(0.201)	(0.174)	(3.740)	(3.059)	(2.430)	(3.309)
Substance-Only	1.400***	-0.886***	3.406	2.158	2.525	-0.418
	(0.190)	(0.164)	(3.789)	(3.086)	(2.258)	(5.042)
Rhetoric-Only	0.608***	-0.651***	16.741***	6.611**	5.247**	7.755**
	(0.210)	(0.165)	(4.082)	(3.148)	(2.259)	(3.386)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
p-value						
Control A = *Treatment A*	0.000	0.000	0.000	0.890	0.894	0.277
Observations	803	803	8,030	8,030	8,030	8,020
Subjects	803	803	803	803	803	802

Notes: Each column corresponds to a separate regression. Deliberative Competence measured as $d_M^{i,R,t} = -\left|r_c^{i,R,t} - r_s^{i,R,t}\right|$. Column 6 omits the subject for whom $r_S^{i,R,t} = 0$ for all instruments. Standard errors in parentheses, clustered on the subject level. $^{***}p < 0.01, ^{***}p < 0.05, ^{*}p < 0.1$.

D.3 Effects on individual test questions

We analyze the effect of the treatments on answers to individual test questions in table D.3. The test questions differ by how closely they follow the material in the education intervention, and by how easily they are answered without knowledge of the rule of 72.

Q1 is the only question for which the answer was explicitly given in the education video (including in the Substance-Only treatment but not in the Rhetoric-Only treatment). The video also discussed an example that is similar, but not identical, to Q2.[62]

The remaining questions require more flexible thinking. Q3 and Q4 can easily be answered with the rule of 72. Knowledge of this rule, however, is not necessary to answer these questions correctly. Q3 can be answered by iteratively multiplying a starting value with 1.07, and counting the number of iterations required for the amount to increase to the desired value. Likewise, Q4 can be answered by calculating the factor by which an investment grows within 8 years at 9 percent interest (either iteratively, or using the compound interest formula), and then dividing 500 by this number. Q5 is a standard compound interest calculation, and parallels the calculations that need to be made in the complexly framed decision problems.

[62]The example is: "To double your money in 10 years, what rate of return do you need? The answer: 10 times $X = 72$, so $X = 7.2$ percent."

Table D.3 displays the treatment effects on the success rates for each of these questions. Baseline rates of correct answers are highly similar across the two experiments. Moreover, in both experiments, the significant effect of the Full and Substance-Only treatments on the total score derive from questions Q1, Q2, and Q5. The fact that performance in Q5 increased in these treatments is reassuring, as it demonstrates that the increase in test scores is at least partly due to subjects' increased ability to analyze previously unseen problems properly. Moreover, while treatment effects are similar across the experiments for questions Q1 to Q4, the treatment effect on Q5 in Experiments B is more than double that in Experiment A, tentatively hinting at our finding that our intervention in Experiment B is more effective than that in Experiment A.

Table D.3: Fraction of correct responses on individual questions in the test about the Treatment intervention.

Question	Q1	Q2	Q3	Q4	Q5
			Experiment A		
Treatment effects					
Treatment	0.566***	0.619***	0.062	0.021	0.174***
	(0.054)	(0.053)	(0.068)	(0.068)	(0.067)
Substance-Only	0.584***	0.592***	-0.037	0.023	0.109*
	(0.051)	(0.053)	(0.065)	(0.065)	(0.065)
Rhetoric-Only	0.072	0.191***	0.067	0.114*	0.050
	(0.065)	(0.061)	(0.067)	(0.067)	(0.067)
Level in Control	0.330***	0.220***	0.514***	0.422***	0.477***
	(0.045)	(0.040)	(0.048)	(0.047)	(0.048)
Observations	455	455	455	455	455
			Experiment B		
Treatment effect	0.559***	0.696***	0.045	-0.075	0.375***
	(0.042)	(0.038)	(0.054)	(0.052)	(0.048)
Level in Control	0.346***	0.168***	0.464***	0.436***	0.436***
	(0.036)	(0.028)	(0.037)	(0.037)	(0.037)
Observations	348	348	348	348	348

Notes: Standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

D.4 Self-reported behavior

The study ends with a brief non-incentivized questionnaire. We ask subjects whether they had used the rule of 72 in the complexly framed problems, and whether they had used it in the simply framed problems. We also elicit the number of complexly framed valuation tasks for which subjects explicitly calculated the future value of the investment, and ask whether they obtained help when taking the test on compound interest. The questionnaire also addresses a small number of additional issues.

Subjects in the Control condition report similar numbers of decisions for which they engaged in explicit calculations in each of the experiments (6.4 and 6.7 in Experiments A and B, respectively). The Treatment condition significantly

increases that number, by 1.7 problems in Experiment A (column 1, $p < 0.01$) and by 2.6 problems in Experiment B (column 2, $p < 0.01$). The treatment effect on the fraction of subjects reporting to have used the rule of 72 in their decision making in complexly framed decisions does not differ substantially across experiments (57.9% and 60.5% in Experiments A and B, respectively). There is, however, a difference in levels. Only 12.8% of subjects in the Control condition of Experiment A report using the rule, whereas 31.8% of subjects in the Control condition of Experiment B do so.

As expected, the fraction of subjects reporting to have used the rule of 72 for simply framed problems is substantially smaller; averaging 9.2% and 22.3% in the Control conditions of Experiments A and B, respectively. In both experiments the Treatment condition increases the frequency of such reports, but does so almost twice as much in Experiment A (by 17.2 percentage points) than in Experiment B (by 9.6 percentage points). Finally, when asked about the use of external help with the test questions at the end of the experiment, we do not find treatment effects in either experiment, although the fraction of subjects reporting the use of such help exceeds 20% in Experiment A, whereas it is lower than 8% in Experiment B.

Unlike performance on test scores and directional behavioral changes, these self-reported behaviors suggest that the effects of the Treatment interventions differ across the experiments, though that interpretation is complicated by the fact that baseline levels differ across the experiments. Like the conventional measures, however, data on self-reported behavior suggest that the Treatment interventions are effective in either experiment.

Table D.4: Self-reported behavior.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-report	Engages in explicit calculation		Uses of rule of 72 in complex frame		Uses of rule of 72 in simple frame		External help with test	
Experiment	A	B	A	B	A	B	A	B
Levels								
Control	6.404***	6.693***	0.128***	0.318***	0.092***	0.223***	0.220***	0.078***
	(0.377)	(0.277)	(0.032)	(0.035)	(0.028)	(0.031)	(0.040)	(0.020)
Treatment	8.142***	9.331***	0.708***	0.923***	0.264***	0.320***	0.208***	0.059***
	(0.342)	(0.209)	(0.044)	(0.021)	(0.043)	(0.036)	(0.040)	(0.018)
Difference	1.738***	2.639***	0.579***	0.605***	0.172***	0.096**	-0.013	-0.019
	(0.509)	(0.347)	(0.055)	(0.041)	(0.051)	(0.048)	(0.056)	(0.027)
Observations	215	348	215	348	215	348	215	348

Notes: Each column displays the coefficients of a separate OLS regression. Standard errors in parentheses. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

D.5 Analysis using the welfarist measure of Deliberative Competence

Here, we present our main empirical results on Deliberative Competence using the measure $d_W^{i,R,t} = -\left(r_c^{i,R,t} - r_s^{i,R,t}\right)^2$ which approximates the average rather than the maximal loss from characterization failure.

Table D.5 replicates Table 7 with this alternative measure. Panel A shows that the intervention in Experiment A leaves $d_W^{i,R,t}$ nearly unchanged on average whereas the intervention in Experiment B leads to a substantial and statistically highly significant increase both in the pooled sample and separately within each timeframe.

Panel B applies Proposition 6 to correct for changes in valuations in the simple frame.[63] Again, we find that the intervention in Experiment A, if anything, harms subjects, whereas the intervention in Experiment B significantly increases their welfare. The relative magnitude of these effects differs from those in Table 7. Here, we find that the harm caused by the intervention in Experiment A is of a similar magnitude as the benefits caused by the intervention in Experiment B, whereas in Panel B of Table 7 the benefits of the intervention in Experiment B exceed the magnitude of the harm in Experiment A severalfold. One reason for this divergence is the stronger sensitivity of $d_W^{i,R,t}$ to large valuation differences, $r_C^{i,R,t} - r_S^{i,R,t}$.

[63] Proceeding as in Table 7, we account for noise in the elicitation of valuations in the simple frame. Specifically, we calculate the mean valuation for simply framed choices for each timeframe and use the square of the resulting average as the correction factor. Moreover, by our normalization, $y_{sI} - y_0 = 1$ for all instruments I.

Table D.5: Deliberative Competence based on expected welfare loss.

A. Deliberative Competence: Welfarist Measure

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-31.635***	-34.235***	-30.287***	-33.938***	-32.983***	-34.531***
	(2.329)	(2.064)	(2.462)	(2.199)	(2.504)	(2.116)
Treatment	-34.017***	-28.382***	-33.157***	-26.907***	-34.877***	-29.858***
	(3.110)	(2.067)	(3.192)	(2.050)	(3.471)	(2.297)
Substance-Only	-29.025***		-27.255***		-30.795***	
	(2.174)		(2.190)		(2.398)	
Rhetoric-Only	-31.431***		-31.640***		-31.223***	
	(2.763)		(2.907)		(2.877)	
p-value of difference to Control						
Treatment	0.540	0.046	0.477	0.020	0.658	0.136
Substance-Only	0.413		0.358		0.528	
Rhetoric-Only	0.955		0.723		0.645	
Observations	4,550	3,480	2,275	1,740	2,275	1,740
Subjects	455	348	455	348	455	348

B. Deliberative Competence: Welfarist Measure corrected for changes in simply framed valuations

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-63.503***	-74.623***	-68.950***	-79.013***	-58.055***	-70.234***
	(5.603)	(6.236)	(8.005)	(7.043)	(4.839)	(7.107)
Treatment	-97.289***	-50.961***	-97.543***	-52.627***	-97.035***	-49.295***
	(16.557)	(4.683)	(16.502)	(5.118)	(19.657)	(4.651)
Substance-Only	-114.092***		-149.662**		-78.521***	
	(40.816)		(67.113)		(15.593)	
Rhetoric-Only	-66.779***		-75.521***		-58.036***	
	(7.662)		(9.977)		(6.560)	
p-value of difference to Control						
Treatment	0.054	0.003	0.120	0.003	0.055	0.014
Substance-Only	0.220		0.233		0.211	
Rhetoric-Only	0.730		0.608		0.998	
Observations	4,550	3,470	2,275	1,735	2,275	1,735
Subjects	455	347	455	347	455	347

Notes: Each column displays the coefficients of a separate OLS regression of the welfarist measure of Deliberative Competence, $d_W^{i,R,t} = -\left(r_c^{i,R,t} - r_s^{i,R,t}\right)^2$, on treatment indicators. Standard errors in parentheses, clustered by subject. The reason for the smaller number of observations in Panel B in Experiment B is one subject who consistently made choices consistent with a valuation of zero in the simple frame. As the correction consists in dividing by simply framed valuations, this subject is excluded from that analysis. $^{***}p < 0.01, ^{***}p < 0.05, ^{*}p < 0.1$.

Table D.6 replicates Table 11 using using the welfarist measure of Deliberative Competence, $d_W^{i,R,t} = -\left(r_c^{i,R,t} - r_s^{i,R,t}\right)^2$. As in Table 11, we find that the intervention in Experiment A harms subjects in the lowest quartile of valuations in the simple frame ($p < 0.1$), and has beneficial effects for subjects in the highest quartile ($p < 0.01$). Also paralleling the result in Table 11, the intervention in Experiment B does not harm subjects in any quartile, but has substantially positive effects for subjects in the second-highest ($p < 0.1$) and highest quartiles ($p < 0.01$). We conclude that our inferences regarding practice and feedback are not driven by an unintended relationship between welfare weights and biases that impact simply framed valuations, also based on the welfarist measure.

Table D.6: Welfarist measure of Deliberative Competence by quartiles of simply framed valuations.

Experiment A

VARIABLE	Deliberative Competence			
	Quartile simply framed valuation			
	1	2	3	4
Levels				
Control	-17.117***	-27.762***	-34.273***	-46.826***
	(2.880)	(4.659)	(3.696)	(5.064)
Treatment	-33.360	-32.323	-33.912	-36.446
	(8.733)	(5.078)	(3.533)	(6.432)
Effect	-16.243*	-4.561	0.361	10.380
	(9.196)	(6.891)	(5.114)	(8.187)
Observations	2,150			
Subjects	215			

Experiment B

VARIABLE	Deliberative Competence			
	Quartile simply framed valuation			
	1	2	3	4
Levels				
Control	-16.939***	-33.560***	-35.883***	-50.172***
	(2.719)	(4.078)	(3.257)	(4.574)
Treatment	-20.018	-30.052	-30.076	-33.268
	(3.634)	(3.306)	(3.922)	(5.111)
Effect	-3.079	3.508	5.807	16.905**
	(4.539)	(5.250)	(5.099)	(6.859)
Observations	3,480			
Subjects	348			

Notes: Effect of treatments on the welfarist measure of Deliberative Competence, $d_W^{i,R,t} = -\left(r_c^{i,R,t} - r_s^{i,R,t}\right)^2$, by quartiles of simply framed valuations. Each panel presents the output of a single OLS regression. Standard errors in parentheses, clustered by subject. $^{***}p < 0.01$,$^{***}p < 0.05$,$^{*}p < 0.1$.

D.6 Deliberative Competence with correction for policy-induced confounds based on the approach of Appendix B

In this section, we provide estimates of Deliberative Competence corrected for policy-induced framing effects using the alternative approximation approach developed in Section B.

We estimate the correction factor ρ by regressing valuations in the simple frame on the future amount the subject can receive from purchasing the instrument. We run a separate OLS regression for each timeframe and for each treatment within each Experiment. The slope coefficient is a consistent estimate for $\rho^V\left(y_s, \theta\right)$; in each case, we divide the absolute valuation difference by this estimate.

Panel A of Table D.7 applies this correction factor to replicate Table 7. Panel B performs parallel analysis using the welfarist measure d_W and the corresponding correction factor ρ^2. In each case we find large and statistically highly significant treatment effects for Experiment B but not for Experiment A.

Table D.7: Deliberative Competence corrected for policy-induced confounds based on the approach of Appendix B

A. Deliberative Competence, d_M

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-36.875***	-38.517***	-39.765***	-40.982***	-33.985***	-36.052***
	(2.472)	(2.036)	(2.880)	(2.255)	(2.304)	(1.978)
Treatment	-31.591***	-26.576***	-33.291***	-27.800***	-29.891***	-25.353***
	(2.402)	(1.684)	(2.609)	(1.821)	(2.476)	(1.699)
Substance-Only	-30.356***		-30.824***		-29.888***	
	(1.972)		(2.062)		(2.127)	
Rhetoric-Only	-27.375***		-29.278***		-25.471***	
	(1.941)		(2.115)		(1.977)	
p-value of difference to Control						
Treatment	0.126	0.000	0.096	0.000	0.227	0.000
Substance-Only	0.040		0.012		0.192	
Rhetoric-Only	0.003		0.004		0.005	
Observations	4,550	3,480	2,275	1,740	2,275	1,740
Subjects	455	348	455	348	455	348

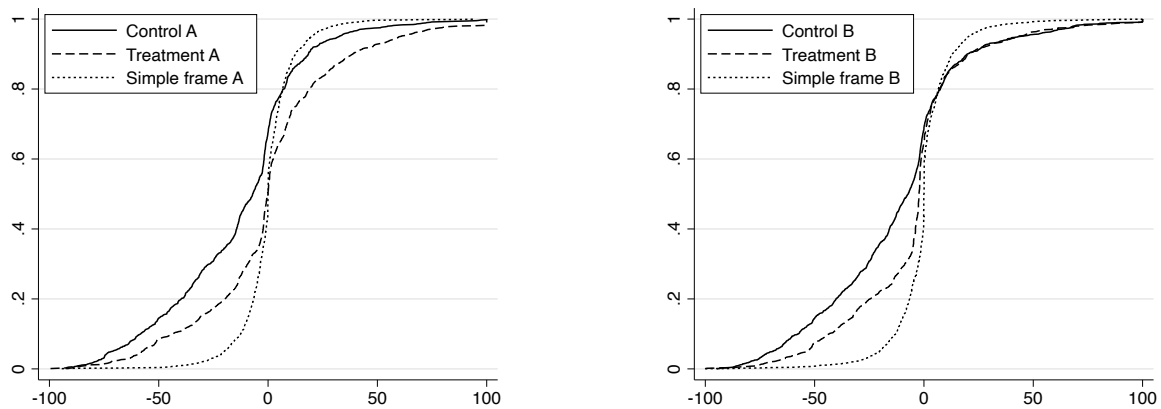
B. Deliberative Competence, d_W

	(1)	(2)	(3)	(4)	(5)	(6)
Delay in days	both	both	72	72	36	36
Experiment	A	B	A	B	A	B
Levels						
Control	-72.230***	-77.119***	-82.620***	-84.680***	-61.840***	-69.557***
	(5.363)	(4.666)	(6.715)	(5.487)	(4.695)	(4.263)
Treatment	-64.912***	-58.021***	-69.717***	-61.145***	-60.106***	-54.898***
	(5.925)	(4.223)	(6.712)	(4.658)	(5.982)	(4.224)
Substance-Only	-55.189***		-57.307***		-53.071***	
	(4.140)		(4.605)		(4.133)	
Rhetoric-Only	-60.167***		-66.526***		-53.808***	
	(5.291)		(6.112)		(4.959)	
p-value of difference to Control						
Treatment	0.360	0.003	0.175	0.001	0.820	0.015
Substance-Only	0.012		0.002		0.162	
Rhetoric-Only	0.110		0.077		0.240	
Observations	4,550	3,480	2,275	1,740	2,275	1,740
Subjects	455	348	455	348	455	348

Notes: Each column displays the coefficients of a separate OLS regression of Deliberative Competence, on treatment indicators. Deliberative Competence is corrected for the change in simply framed valuations using the approximation strategy developed in Section B. Panel A uses the maximal loss measure of Deliberative Competence, d_M; Panel B uses the welfarist measure d_W.

D.7 Valuation difference compared to noise in the simple frame

Figure D.1: Replication of CDFs from Figure 3 with a measure of the distribution of noise in the simple frame superimposed.



Experiment A Experiment B

Here we investigate the possibility that the measured overestimation of compound interest exhibited by a fraction of our subjects is solely attributable to elicitation noise. To test the hypothesis, we estimate the amount of noise in simply framed decisions. We then compare the frequency of excessive valuations we would observe based on that amount of noise alone to the frequency of excessive valuations we actually observe in the complex frame. Specifically, we calculate, for each subject and each timeframe $t \in \{36, 72\}$, the values $\Delta_1^t = r_S^{j,20,t} - r_S^{j,18,t}$, $\Delta_2^t = r_S^{j,18,t} - r_S^{j,20,t}$ $\Delta_3^t = r_S^{j,16,t} - r_S^{j,14,t}$, and $\Delta_4^t = r_S^{j,14,t} - r_S^{j,16,t}$ (recall that valuations $r_S^{i,R,t}$ involve the normalization of the future value to \$1). Figure D.1 superimposes the CDF of Δ_k^t (pooled across $k \in \{1, .., 4\}$ and $t \in \{36, 72\}$) on the CDF of the valuation difference from Figure 3. If excessive valuations in the complex frame were solely attributable to elicitation noise, the CDF of Δ_k^t should coincide with the CDF of $r_C^{i,R,t}$ to the left of zero. By contrast, we find substantial differences between these curves in each experiment, and especially for the treatment in Experiment A. Accordingly, elicitation noise alone cannot explain the overestimation of compound interest in either condition of either experiment.

E Instructions

This is a research study run by the department of economics at Stanford University.

IMPORTANT

This study may take up to ONE AND A HALF HOURS to complete. Please start this study only if you do have that much time in a single session.

If you do not complete the study, or if the HIT times out on you, we will not be able to pay you. (The HIT is set to time out in 3 hours.)

You will earn $10 just for completing this study. In addition, you will receive up to $20, depending on the decisions you make in this study.

Do not start this study if you do not have access to youtube.com. Some browsers will block embedded videos. Please make sure your browser will display them, as you may otherwise not be able to complete this study.

Click here to start the study: https://stanforduniversity.qualtrics.com/SE/?SID=SV_0GPXNo1f9TX5YIR

Provide the survey code here:

WELCOME

This is a research study run by the department of economics at Stanford University.

IMPORTANT

This study may take up to ONE AND A HALF HOURS to complete. Please start this study only if you do have that much time in a single session.

If you do not complete the study, or if the HIT times out on you, we will not be able to pay you. (The HIT is set to time out in 3 hours.)

You will earn $10 just for completing this study. In addition, you will receive up to $20, depending on the decisions you make in this study.

Do not start this study if you do not have access to youtube.com. Some browsers will block embedded videos. Please make sure your browser will display them.

By clicking the button below, you consent to participating in this research study.

Questions, Concerns, or Complaints: If you have any questions, concerns or complaints about this research study, its procedures, risks and benefits, you should ask the Protocol Director, Sandro Ambuehl, sambuehl@stanford.edu

Independent contact: If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Stanford Institutional Review Board (IRB) to speak to someone independent of the research team at (650)-723-2480 or toll free at 1-866-680-2906. You can also write to the Stanford IRB, Stanford University, Stanford, CA 94305-5401

>>

[Some browsers will ask you whether you want to display this content. Please click "display all content".]

[There should be a video here. If if does not load, please click here]

Links to researchers' personal homepages

Professor B. Douglas Bernheim

Sandro Ambuehl

To continue, please enter the LAST word that Doug Bernheim said in this video. A continue button will appear after the duration of the video.

>>

Before we start this study, we would like to ask you a few questions about yourself. Please answer these questions truthfully. Your answers will not affect your payment from this experiment.

What is your gender?

- male
- female

What is your age?

What is your ethinicity?

- African-American
- Asian
- Caucasian
- Hispanic
- Other

Please indicate the highest level of education you completed.

- Elementary School
- Middle School
- High School or equivalent
- Vocational/Technical School (2 year)
- Some College
- College Graduate (4 year)
- Master's Degree (MS)
- Doctoral Degree (PhD)
- Professional Degree (MD, JD, etc.)

What is your current marital status?

- Divorced
- Living with another
- Married
- Separated
- Single
- Widowed
- Prefer not to say

Please answer the following questions as well as you can. Your answers to these questions will not affect your payment from this study.

Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, how much would you be able to buy with the money in this account?

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

Suppose you had $100 in a savings account and the interest rate is 20 percent per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total?

- ◯ More than $200
- ◯ Exactly $200
- ◯ Less than $200
- ◯ Do not know

Assume a friend inherits $10,000 today and his sibling inherits $10,000 3 years from now. Who is richer because of the inheritance?

- ◯ My friend
- ◯ His sibling
- ◯ They are equally rich
- ◯ Do not know

Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

- ◯ More than $102
- ◯ Exactly $102
- ◯ Less than $102
- ◯ Do not know

Suppose that in the year 2015, your income has doubled and prices of all goods have doubled too. In 2015, how much will you be able to buy with your income?

- ◯ More than today
- ◯ The same
- ◯ Less than today
- ◯ Do not know

>>

You will now watch a

12-MINUTE VIDEO ABOUT FINANCIAL INVESTING.

Please follow this video carefully.

Please watch the ENTIRE video.

(a "continue" button will appear after 12 minutes.)

Doing so will be useful to you for three reasons:

1. *TEST with PAYMENT FOR CORRECT ANSWERS.*

Your earnings from this experiment may be entirely determined by a test on this video. **The final part of this experiment is a test about the contents of this video. There is a one in four chance that your earnings from this experiment are wholly determined by your performance in this test. The test has 10 questions. For each question you answer correctly, you will receive $1 within at most two days from today. For each question you answer incorrectly, you will receive $0. To be able to answer the questions in the test, you need to both *understand* and *know* the contents of the video. You may scroll back to watch parts of the video multiple times if you wish.**

2. *REMAINDER OF THIS STUDY.*

The video may help you with your decisions in the remainder of this experiment.

In each remaining part of this experiment, you will make financial investment decisions. There is a three in four chance that one of these decisions wholly determines your earnings from this experiment.

3. REAL LIFE

The video may help you with your decisions in real life.

This video was made by internationally recognized academic experts on financial decision making (Burton G. Malkiel, Charles D. Ellis, and B. Douglas Bernheim). This video may help you make financial decisions in your life in general.

>>

PLEASE FOLLOW THIS VIDEO CAREFULLY

PLEASE WATCH THE ENTIRE VIDEO

[Some browsers will ask you whether you want to display this content. Please click "display all content".]

[There should be a video here. If it does not load, please click here.]

To continue, enter the FOURTH word of the FIRST slide of this video. A continue button will appear after the duration of the video.

>>

PLEASE READ THESE INSTRUCTIONS CAREFULLY

The remainder of this experiment consists of 20 rounds of decision making.

Your payment may be determined entirely by ONE RANDOMLY CHOSEN decision you make in this part of the experiment.

This will happen with a three in four chance. Otherwise, your payment is determined by your performance in the test about the video you just watched.

Hence, you should make every decision as if it is the one that counts, because it might be!

>>

PLEASE READ THESE INSTRUCTIONS CAREFULLY

In each round, you will be presented with two lists. The first list will be like the following:

	you will get the specified dollar amount within two days from today	Option X
$20	○	○
$18	○	○
$16	○	○
$14	○	○
$12	○	○
$10	○	○
$8	○	○
$6	○	○
$4	○	○
$2	○	○
$0	○	○

Option X will vary from round to round. For instance, option X may be "get $15 in 8 weeks".

PLEASE READ THESE INSTRUCTIONS CAREFULLY

Our payment procedure is designed such that it is in your best interest to choose, on each line of each decision list, the option you genuinely prefer.

Here's why: You'll get exactly what you chose, for one randomly drawn decision.

Read this paragraph if you want to know more details.

Question: When will I be paid according to the first decision list, and when will I be paid according to the second decision list in a round?

Answer: Suppose you filled in the *first* decision list of a round as follows:

YOU WILL NOW MAKE YOUR DECISIONS

It is in your best interest to choose as you genuinely prefer. Please think about your choices carefully.

There are no right or wrong choices!

>>

Please choose, on each line, the option you genuinely prefer.

If you pick the option on the LEFT,

you will get the specified dollar amount within two days from today.

If you pick the option on the RIGHT,

we will invest $4.50 in an account with 2% interest per day. Interest is compounded daily. We will pay you the proceeds in 72 days.

You may switch from left to right at most once.

This is the
first
decision list for these options.

	you will get the specified dollar amount within two days from today	we will invest $4.50 in an account with 2% interest per day. Interest is compounded daily. We will pay you the proceeds in 72 days.
$20	○	○
$18	○	○
$16	○	○
$14	○	○
$12	○	○
$10	○	○
$8	○	○
$6	○	○
$4	○	○
$2	○	○
$0	○	○

>>

Please choose, on each line, the option you genuinely prefer.

If you pick the option on the LEFT,

you will get the specified dollar amount within two days from today.

If you pick the option on the RIGHT,

we will invest $4.50 in an account with 2% interest per day. Interest is compounded daily. We will pay you the proceeds in 72 days.

You may switch from left to right at most once.

This is the
second
decision list for these options.

	you will get the specified dollar amount within two days from today	we will invest $4.50 in an account with 2% interest per day. Interest is compounded daily. We will pay you the proceeds in 72 days.
$ 9.8	○	○
$ 9.6	○	○
$ 9.4	○	○
$ 9.2	○	○
$ 9	○	○
$ 8.8	○	○
$ 8.6	○	○
$ 8.4	○	○
$ 8.2	○	○
$ 8	○	○

>>

TEST

You will now participate in a test about the video you have watched at the beginning of the experiment. The test has 10 questions.

There is a one in four chance that your earnings from this study are entirely determined by your performance in this test.

IF you are randomly chosen to be paid according to this test, THEN: For each question you answer correctly, you will earn $1. For each question you answer incorrectly, you will earn $0. You will be paid within at most two days from today.

>>

What is an "indexing" investment strategy?

- Buying index funds, which hold assets that have been indexed as particularly profitable by financial experts
- Buying index funds, which hold stocks of companies that provide information about the stock market as a whole (stock market indices)
- Buying index funds, which hold the market portfolio
- Buying index funds, which hold optimally diversified, custom tailored portfolios

Paul had invested his money into an account which paid 9% interest per year (interest is compounded yearly). After 8 years, he had $500. How big was the investment that Paul had made 8 years ago?

- $200
- $210
- $220
- $230
- $240
- $250
- $260
- $270
- $280
- $290
- $300
- $310
- $320
- $330
- $340
- $350
- $360
- $370
- $380
- $390
- $400

if the interest rate is 10% per year (interest is compounded yearly), how many years does it take until an investment doubles?

- 7 years
- 7.2 years
- 7.4 years
- 7.8 years
- 8 years

If an investment grows at 8 percent per year (interest is compounded yearly), by how much has it grown after 4 years?

- by 30%
- by 31%
- by 32%
- by 33%
- by 34%
- by 35%
- by 36%
- by 37%
- by 38%
- by 39%
- by 40%

Please answer the following questions truthfully. Your answers to these questions DO NOT AFFECT YOUR PAYMENT for this study.

How much attention did you pay to your choices?

- I paid quite a bit of attention for all of my choices.
- For some choices I paid attention, for others I didn't pay much attention
- I clicked through most of the choices without paying much attention.

At the beginning of the experiment, we asked you to watch a video about financial investing. Please indicate which of the following describes your situation best

- I watched the entire video, and paid close attention
- I watched the entire video, but sometimes didn't pay attention
- I skipped parts of the video, because I already knew the material
- I skipped parts of the video, because it was boring (but I did not already know the material)
- I did not watch the video.

Sometimes in this experiment, you were given a choice such as "We will invest $10 in an account with 1% interest per day. Interest is compounded weekly. We will pay you the proceeds in 72 days." When deciding about this choice, did you use the rule of 72?

- Yes
- No
- I don't know the rule of 72

Sometimes in this experiment, you were given a choice such as "We will pay you $20 in 36 days." When deciding about such a choice, did you use the rule of 72?

- Yes
- No
- I don't know the rule of 72

In total, you were given 10 rounds in which one of the options was something like "we will invest $... in an account with ...% interest per week. Interest is compounded weekly. We will pay you the proceeds in … days". Out of these 10 rounds, how many times did you explicitly calculate the money amount that this investment would yield within the specified time?

When you completed the test about the video on financial investing, did you use external resources (such as other websites, books, etc.) to find the right answers?

- Yes
- No

Do you have any suggestions for us about this experiment?

Did you experience any technical difficulties with this study?

>>

F Practice problems with personalized feedback

Education Intervention Part 1: `https://www.youtube.com/watch?v=EnFVLiM1dTs`

Part 1 Practice Question

If you invest $100 at 2% (compounded yearly), how much will be in your account after 36 years?

- ○ $102
- ○ $172
- ○ $200
- ○ $202
- ○ $300
- ○ $302
- ○ $400
- ○ $402

BACK NEXT

If the answer is correct in the first trial:

Great job! Watch the next part of the video to hone your skills even more.

BACK NEXT

If the answer is incorrect in the first trial:

Hmm, that's not quite right.

Please try again. The rule of 72 will help!

If you invest $100 at 2% (compounded yearly), how much will be in your account after 36 years?

- ○ $102
- ○ $172
- ○ $200
- ○ $202
- ○ $300
- ○ $302
- ○ $400
- ○ $402

BACK NEXT

If the answer is correct in the second time:

Nice! You got it this time!

Please watch the next part of the video to hone your skills even more!

BACK NEXT

If the answer is incorrect in the second trial:

Hmm, that's still not quite right.

Watch the next part of the video, so you see how you can get a good idea about how compound interest works.

BACK NEXT

PLEASE FOLLOW THIS VIDEO CAREFULLY

PLEASE WATCH THE ENTIRE VIDEO

The next button will appear automatically when the video ends (after time equal to the duration of the video passes.).

You can stop the video by clicking on it once and make it full screen by clicking on it twice.

If you reload the page, you will again need to wait for the next button. So please do not close your web browser or reload the page unless it is necessary.

[Some browsers will ask you whether you want to display this content. Please click "display all content".]

Education Intervention Part 2: `https://youtu.be/3pjkVdOXMlk`

Part 2 Practice Questions

Now you try:

$100 is invested at 9% for 32 years, compounded yearly. How much will be in the account after these 32 years?

- ○ $100
- ○ $200
- ○ $388
- ○ $400
- ○ $600
- ○ $800
- ○ $1200
- ○ $1600

BACK NEXT

If the answer is correct in the first trial:

This is correct!

Please click next to move on the next part of the video.

BACK NEXT

If the answer is incorrect in the first trial:

Subjects see one of the following explanations depending on their previous answer and they re-attempt the question.

If the subject selected answer $100:

You selected $100. That's not quite right.

You start out with $100. Then you get 9% interest each year! Hence after 32 years, you will have MORE than $100!

Please watch the video again to understand how much you will have.

BACK NEXT

If the subject selected answer $200:

You selected $200. That's not quite right.

You probably remembered from the example above that at 9%, an investment doubles in 8 years.

Thus, the $100 double to $200 after 8 years.

These $200 then double to $400 in the next 8 years. (That is, until year 16).

In the next 8 years, from year 16 to year 24, these $400 double to $800!

Then, in the next 8 years, from year 24 to year 32, it will double again.

Please give it another try.

BACK NEXT

If the subject selected answer $388:

You selected $388. That's not quite right.

You probably got this because you thought you'd get 32 times the interest of 9% on your $100, which is $9.

But, starting from the second year, you also get interest on the interest you earned!

Here's how: You do start out with $100. In the first year, you get 9% interest. That's $9. You start the second year with $109 in your account. Your interest in the second year is 9% of $109, which is MORE than $9. In fact, you'll get 9% of $109, which is $9.80.

Please watch the video again, so you'll understand how compound interest works.

BACK NEXT

If the subject selected answer $400:

You selected $400. That's not quite right.

In this question, the $100 are invested for 32 years, not just for 16 years, as in the example above.

Please give it another try.

BACK NEXT

If the subject selected answer $600:

You selected $600. That's not quite right.

You probably remembered from the example above that at 9%, an investment doubles in 8 years.

Thus, the $100 double to $200 after 8 years.

These $200 then double to $400 in the next 8 years. (That is, until year 16).

In the next 8 years, from year 16 to year 24, these $400 double to $800!

Then, in the next 8 years, from year 24 to year 32, it will double again.

Please give it another try.

BACK NEXT

If the subject selected answer $800:

You selected $800. That's not quite right.

You probably remembered from the example above that at 9%, an investment doubles in 8 years.

Thus, the $100 double to $200 after 8 years.

These $200 then double to $400 in the next 8 years. (That is, until year 16).

In the next 8 years, from year 16 to year 24, these $400 double to $800!

Then, in the next 8 years, from year 24 to year 32, it will double again.

Please give it another try.

BACK NEXT

If the subject selected answer $1200:

You selected $1200. That's not quite right.

You probably remembered from the example above that at 9%, an investment doubles in 8 years.

Thus, the $100 double to $200 after 8 years.

These $200 then double to $400 in the next 8 years. (That is, until year 16).

In the next 8 years, from year 16 to year 24, these $400 double to $800!

Then, in the next 8 years, from year 24 to year 32, it will double again.

Please give it another try.

BACK NEXT

Re-attempt the question:

Please try again:

$100 is invested at 9% for 32 years, compounded yearly. How much will be in the account after these 32 years?

- ○ $100
- ○ $200
- ○ $388
- ○ $400
- ○ $600
- ○ $800
- ○ $1200
- ○ $1600

BACK NEXT

If the answer is correct in the second time:

This is correct!

Please click next to move on the next part of the video.

BACK NEXT

If the answer is incorrect in the second trial:

Hmm, that's still not quite right.

But let's move to the next part of the video.

BACK NEXT

The next button will appear automatically when the video ends (after time equal to the duration of the video passes.).

You can stop the video by clicking on it once and make it full screen by clicking on it twice.

If you reload the page, you will again need to wait for the next button. So please do not close your web browser or reload the page unless it is necessary.

Education Intervention, part 3: `https://youtu.be/kjPYqcZNzPI`

Practice Questions at the end of the Intervention

Thanks for watching this video!

We'll now ask you to solve a bunch of problems on your own. We'll first walk you through in steps, and then it's up to you to find the right steps.

These questions are still a part of the education. They don't count for money, but you need to get them right so you can continue with the survey.

BACK NEXT

Question 1(a)

You invest $50 at 8%. Eventually, we want to know how much will be in your account after 27 years. But we'll get there in three easy steps.

1. How long does it take for the money to double at 8%?
2. How many times does it double in 27 years?
3. Hence, how much will be in the account after it doubles that many times?

So let's start with the first one of these.

How many years does it take for this investment to double?

BACK NEXT

If the answer to part (a) is incorrect in the first trial:

Your answer isn't quite correct. Remember: The rule of 72 says

percentage interest rate X number of years it takes for the investment to double = 72

Please try again: You invest $50 at 8%. How many years does it take for this investment to double?

BACK NEXT

If the answer to part (a) is incorrect in the second trial:

That's still not quite correct.

Here's how you can do it correctly:

The rule of 72 says that

percentage interest rate *times* the number of years it takes for the investment to double = 72

or, in mathematical notation,

X × Y = 72

In this problem, the percentage interest rate is 8%. Hence you just need to know: 8 times *what* equals 72?

That's how long it takes for the investment to double!

Enter your answer below.

BACK NEXT

If the answer to part (a) is correct in the first trial or later trials:

Great, you've got it!

Question 1(b)

We're still looking at that $50 invested at 8%. As you've figured out, at 8%, the investment doubles in 9 years.

Remember the three steps?

1. How long does it take for the money to double at 8%?
2. How many times does it double in 27 years?
3. Hence, how much will be in the account after it doubles that many times?

We now tackle the second step:

How many times does this investment double over the course of 27 years?

- ○ once
- ○ twice
- ○ three times
- ○ four times
- ○ five times
- ○ six times
- ○ seven times
- ○ eight times
- ○ nine times
- ○ ten times

If the answer to part (b) is incorrect in the first trial:

Unfortunately, that's not quite right.

As you've figured out, the investment doubles in 9 years. It doubles in every 9 years over the course of 27 years!

Hence, 9 times *what* equals 27?

The answer to this question tells you how many times the investment doubles!

Please choose one of the answers below.

- ○ once
- ○ twice
- ○ three times
- ○ four times
- ○ five times
- ○ six times
- ○ seven times
- ○ eight times
- ○ nine times
- ○ ten times

BACK NEXT

If the answer to part (b) is correct in the first trial or later trials:

Nice job!

Now to the last one of the three steps.

1. How long does it take for the money to double at 8%?
2. How many times does it double in 27 years?
3. Hence, how much will be in the account after it doubles that many times?

You figured out that over the course of 27 years, your $50, invested at 8% double three times.

Hence, how much will be in your account after 27 years?

- ○ $50
- ○ $100
- ○ $150
- ○ $200
- ○ $250
- ○ $300
- ○ $350
- ○ $400
- ○ $450
- ○ $500
- ○ $600
- ○ $700
- ○ $800

If the answer to part (c) is incorrect in the first trial:

Oops, that is not quite right.

In the first 9 years, your investment doubles by $50 and is then worth $100. In the second 9 years, these entire $100 double again. So after the second 9 years (that is after 18 years), you have $200.

So, how much will you have after 27 years?

- ○ $50
- ○ $100
- ○ $150
- ○ $200
- ○ $250
- ○ $300
- ○ $350
- ○ $400
- ○ $450
- ○ $500
- ○ $600
- ○ $700
- ○ $800

BACK NEXT

If the answer to part (c) is correct in the first trial or later trials:

Awesome job!

Now it's up to you to go through the steps in the right order.

Let's try this example:

Question 2

You invest $100 at 6%. How much will be in your account after 24 years?

- ○ $100
- ○ $106
- ○ $148
- ○ $200
- ○ $288
- ○ $300
- ○ $306
- ○ $400
- ○ $406
- ○ $500
- ○ $506
- ○ $600
- ○ $606

If the answer to Question 2 is incorrect in the first trial:

Oops, that's not quite right. Remember the three steps for using the rule of 72:

1. How long does it take for the money to double?
2. How many times will it double over the years?
3. Hence, how much will be in the account after it doubles that many times?

Give it another shot:

Question 2

You invest $100 at 6%. How much will be in your account after 24 years?

- ○ $100
- ○ $106
- ○ $148
- ○ $200
- ○ $288
- ○ $300
- ○ $306
- ○ $400
- ○ $406
- ○ $500
- ○ $506
- ○ $600
- ○ $606

If the answer to Question 2 is correct in the first trial or later trials:

Great! Thanks for paying attention to this education module. Before moving forward, we would like to ask you a question about the education module. Your answer to this question will not affect your payments from this study.

BACK NEXT