CAGE

Evaluating the Sunk Cost Effect

CAGE working paper no.475

May 2020 (Revised April 2021)

David Ronayne
Daniel Sgroi
Anthony Tuckwell

WARWICK
THE UNIVERSITY OF WARWICK

UKRI
Economic and Social Research Council

Evaluating the Sunk Cost Effect*

David Ronayne†, Daniel Sgroi‡ and Anthony Tuckwell§

First version: May 15, 2020
This version: March 29, 2021
Forthcoming in the ***Journal of Economic Behavior & Organization***

Abstract

We provide experimental evidence of behavior consistent with the sunk cost effect. Subjects who earned a lottery via a real-effort task were given an opportunity to switch to a dominant lottery; 23% chose to stick with their dominated lottery. The endowment effect accounts for roughly only one third of the effect. Subjects' capacity for cognitive reflection is a significant determinant of sunk cost behavior. We also find stocks of knowledge or experience (crystallized intelligence) predict sunk cost behavior, rather than algorithmic thinking (fluid intelligence) or the personality trait of openness. We construct and validate a scale, the "SCE-8", which encompasses many resources individuals can spend, and offers researchers an efficient way to measure susceptibility to the sunk cost effect. *JEL: D91, C83, C90*

Keywords: sunk cost effect, sunk cost fallacy, endowment effect, cognitive ability, fluid intelligence, crystallized intelligence, reflective thinking, randomized controlled trial, online experiment, online survey, psychological scales, scale validation, Raven's progressive matrices, international cognitive ability resource, cognitive reflection test, openness.

*We are grateful for helpful discussions with Dan Bernhardt, Gordon Brown, Armin Falk, David Gill, Jack Knetsch, Andrew Oswald, Nattavudh Powdthavee, Chris Roth, Richard Thaler, Bertil Tungodden, Peyton Young and Alexey Zakharov. This work was supported by the ESRC CAGE Centre [grant number RES-626-28-0001]. Ethical approval was granted by the University of Oxford (ECONCIA19-20-08) and the University of Warwick (HSS 83/18-19). The experiment is registered at the AEA RCT registry `https://doi.org/10.1257/rct.4083-2.0`.

†ESMT Berlin. `david.ronayne@esmt.org`

‡University of Warwick and IZA Bonn. `daniel.sgroi@warwick.ac.uk`

§University of Warwick. `a.j.p.tuckwell@warwick.ac.uk`

"No matter how far you've gone down the wrong road, turn back."

- Turkish proverb

1 Introduction

In a famous example by Thaler (1980), an individual who purchased a $40 ticket to a basketball game finds themselves driving for miles through a snowstorm, just because they spent money (and perhaps also effort, time, etc.) to get the ticket. A seminal definition of this "sunk cost effect" or "fallacy" by Arkes and Blumer (1985) characterizes individuals as falling prey to the effect when they continue an endeavor as a result of previously invested resources such as time, money, or effort. Many definitions have been offered, but central to them all is the idea that some actions (sunk) in the past constrain decision-making in the present despite the fact that the actions do not affect the attractiveness of the available options. The effect seems widespread, with work showing that it is even present in mice and rats (Sweis et al., 2018). It is also an important part of the work by Thaler (1999) on mental accounting, which is based on the idea that individuals often fail to treat money as a fully fungible resource. We offer contributions to the identification, understanding, and measurement of the sunk cost effect in humans.

First, we conduct an experiment with pecuniary incentives to detect behavior consistent with the sunk cost effect in a controlled setting. Subjects in the primary treatment completed a real-effort task to earn a lottery. Subjects who earned the lottery ($n = 268$) then had the (unanticipated) choice to switch to a dominant lottery; 23% chose to stick with their earned (and dominated) lottery. We argue that the endowment effect is an intrinsic part of the sunk cost effect. To uncover how large a part, we ran a second treatment in which subjects ($n = 197$) did not face the real-effort task, but were instead endowed with the inferior lottery before having the choice to switch to the superior lottery; 7% chose to stick with their endowed (and dominated) lottery.

Second, we document correlations between sunk cost behavior and different traits that have been argued to be related to the effect, which we preregistered in advance.[2] One type of cognitive ability introduced by Frederick (2005), "cognitive reflection", concerns the ability of an individual to override a heuristically-primed ("System 1") response and engage in further deliberate ("System 2") reflection to figure out the correct answer, and has been found to be predictive of various types of biases (Toplak et al., 2011).[3] We test its capacity to explain our behavioral measure of the sunk cost effect and find it to be a highly significant predictor. Because cognitive reflection is itself likely to be a function of different aspects of intelligence and thinking style (Stanovich and West, 2008), we also test three of these measures and find crystallized intelligence to be a significant driver.

[2]In our pretrial registration (Ronayne et al., 2019), we listed a small number of variables linked to cognitive ability, which have been argued to be related to the sunk cost effect. We did this both to test the associations posited in the literature, and to limit our ability to find and exploit spurious correlations.

[3]Terms "System 1" and "System 2" are those famously used by Kahneman (2011).

Third, we offer a scale consisting of hypothetical scenarios to measure individuals' susceptibility to the sunk cost effect. We propose the scale for use by researchers looking to measure subjects' susceptibility to the effect without the need to elicit their own incentivized behavioral measure. The lack of a validated scale in the extant literature along with the costs of devising and running an experiment suggests some potential benefits of this. From a set of 18 scenarios, we use factor analysis to select 8 items to form a concise scale: the SCE-8.

Fourth, to reflect the wide-ranging nature of the effect, we include multiple types of sunk resources including effort, time, money, and emotional attachment. Effort and time are perhaps the most relevant costs in our experiment, but all four are represented in the SCE-8. Our results indicate that many different sunk resources contribute to a singular, underlying sunk cost effect.

We also provide a definition of the sunk cost effect and contrast it with the endowment effect. Given the structure of our definition it is also possible to distinguish between the sunk cost effect and the sunk cost fallacy, which until now many have used interchangeably.

2 Literature

Many papers report evidence of behavior consistent with the sunk cost effect. These studies typically rely on responses to hypothetical questions, are subject to various confounds, or are field experiments that tend to suffer from real-world factors that complicate the interpretation of the effect. For example, it can be hard to disentangle the sunk cost effect from other possible reasons for behavior in real world situations involving price data (Ashraf et al., 2010; Berry et al., 2020; Cohen et al., 2015), while consumer goods may be subject to heterogeneous mental accounting (Arkes and Blumer, 1985; Just and Wansink, 2011), and penny auctions can suffer from intractable bidding environments (Augenblick, 2016). In contrast, lab and online experiments offer greater control and cleaner measurement, which are attractive merits for our purposes.

Some experiments have attempted to identify the sunk cost effect. Friedman et al. (2007) find limited evidence of the effect, but with a design where subjects' actions are not responsible for generating the sunk cost, an important factor in generating the effect (Arkes and Blumer, 1985; Staw, 1976).[4] The design of Weigel (2018) features such actions but employs a penny auction task, the dynamic nature of which raises the need to control for confounds such as the gambler's fallacy.[5] In contrast, we conduct an experiment with a simple structure in which subjects sink resources then make a one-off decision.

[4]Responsibility may operate by increasing the strength of *positive association* between the sunk resources and choice alternatives (see Section 3.1).

[5]The lab experiment of Haita-Falah (2017) also uses a related design, but the task and instructions are complex and there is evidence subjects may not understand the task (see e.g., Weigel, 2018, footnote 3).

Several papers have tested the role of cognition in accounting for sunk cost behavior.[6] Evidence has been mixed, with often a small or insignificant correlation found between intelligence measures and sunk costs. However, the majority of studies measure sunk cost behavior using unvalidated hypothetical scenarios e.g., Bruine de Bruin et al. (2007); Larrick et al. (1993); Parker and Fischhoff (2005); Stanovich and West (2008); Strough et al. (2008); Toplak et al. (2011). We incentivize subjects to sink resources via a real-effort task and find that cognitive reflection and crystallized intelligence measures significantly explain the variation in our data.

Early in the study of the sunk cost effect, there were discussions of the extent to which non-monetary resources might be fungible. For instance, Thaler (1999) briefly discusses how people appear to allocate time suboptimally, while Soman (2001) argues people treat sunk time and sunk money differently. However, there have been almost no attempts within the literature to explore the commensurability of time, effort, money, or emotional attachment as separate factors in a sunk cost context. This might be considered important especially if these concepts cannot be converted readily into monetary values (Leclerc et al., 1995). We include different resources in our scale (time, money, effort, and emotion) and find them all to be relevant in describing a single underlying factor: susceptibility to the sunk cost effect.

When considering the role of cognition in explaining behavioral biases, it is important to distinguish between different measures. Cognitive reflection is a type of cognitive ability introduced by Frederick (2005) relating to the ability to override a heuristically-primed or knee-jerk response and engage in further reflection to figure out a correct answer, which has been found to predict various types of bias (Toplak et al., 2011). We test its potential to explain the sunk cost effect using our behavioral and scale measures, and find it highly predictive of both. An individual's cognitive reflection, in turn, is likely dependent on other aspects of intelligence and thinking style (Stanovich and West, 2008; Stanovich, 2012). In particular, accumulated stocks of knowledge and experience that might help one to recognize the need to override an instinctive response (crystallized intelligence) and styles of thinking conducive to discovering new perspectives on a problem (open-minded thinking).[7] Interestingly, the literature identifies fluid intelligence (the ability to think logically or algorithmically, as measured by various IQ tests) as having less of a role, because the computational power required to override an impulsive response is often only very slight (Stanovich, 2008); what matters is recognizing the need to override it in the first place. Our results support this hypothesis in the context of sunk costs.

Our paper is the first to validate a scale composed of hypothetical scenarios to measure the sunk cost effect. The effect is not currently measured by any widely accepted scale. It evades, for example, the "Big Five" Personality Inventory (Costa and McCrae, 1989). As one of the most

[6] A more general drive in economics has examined the relationship of cognitive ability to other important characteristics e.g., temporal and risk preferences (Dohmen et al., 2010) and cooperation (Proto et al., 2019).

[7] A related concept to open-minded thinking is that of mindfulness; Hafenbrack et al. (2014) find that both the trait of mindfulness and mindful states induced through meditation increase resistance to the sunk cost effect.

widely known behavioral biases, it seems important to have a reliable yet easy to use scale to measure susceptibility to the effect. To this end, we offer our SCE-8 scale to researchers for any study in which the sunk cost effect may explain outcomes, reducing or eliminating the need to run a full experiment. Collecting SCE-8 data would allow susceptibility to the effect to be measured and used in the same way as other common items including risk tolerance (Blais and Weber, 2006), patience levels (Brockhoff et al., 2015), and personality measures (Costa and McCrae, 1989).

3 Design

Our design combines a randomized experiment, a scale composed of hypothetical questions, and trait measures that have been linked to susceptibility to the sunk cost effect. We organize and structure our approach around a definition of the sunk cost effect, which we detail now.

3.1 Sunk cost effect definition and the endowment effect

If an individual is more likely to choose an alternative when they have sunk resources *positively associated* with that alternative, we define them as exhibiting the sunk cost effect. More formally, consider a binary choice set $\{X, Y\}$ and a quantity of resources $r \geq 0$ sunk in ways positively associated with X (but not Y),[8] where r does not directly affect the utility garnered from either X or Y, and let $P_i^K(r)$ be the probability with which i chooses $K \in \{X, Y\}$ given r. Under our definition, i exhibits the sunk cost effect when $r > 0$ and $P_i^X(r) > P_i^X(0)$.[9,10,11,12]

Within the definition of the sunk cost effect, the term "positively associated" is necessarily general to reflect the diverse range of contexts across which the effect applies. A simple example may be the money spent on (and hence positively associated with) a theater ticket; a more complex one may be the various resources sunk into (and hence positively associated with)

[8]More generally, one could incorporate the relative contributions of X and Y via a definition in which the resources sunk may be positively associated to both, but *more* with X.

[9]Moreover, if this is true *and* i incorrectly believes X yields higher utility, we say that i exhibits the sunk cost *fallacy*, which is therefore a subset of the sunk cost *effect*. Our definitions imply that within mainstream economics the effect and fallacy are identical as choices of X over Y typically imply (by revealed preference) that i believes $U_i(X) > U_i(Y)$, where U_i is $i's$ utility function. This may explain why the two terms are often used interchangeably within economics. Our definitions otherwise serve to label the probabilistic outcome as the effect and the beliefs supporting that outcome as the fallacy. Others have also drawn distinctions between the terms but without explicit mention of payoffs or beliefs, e.g., Olivola (2018) remarks that the fallacy refers to taking an inferior action due to sunk costs whereas the effect refers only to taking a different action.

[10]Note also that we require that r does not directly affect the utility garnered from X or Y. This is required to distinguish from cases in which a sunk resource *does* raise utility. For example, consider sinking effort into learning to play the violin well. Having sunk this effort, it seems likely that the utility derived from playing the violin would be higher than if the costs were not sunk. In our experimental design, scenarios, and examples throughout this paper, this is not the case; spent resources do not directly affect utility.

[11]The definition given is sufficient to cover the scope of our experiment. However, more generally we envisage P_i^X to be a monotonic function such that for all $(r', r)\colon r' > r > 0$ we have $P_i^X(r') > P_i^X(r) > P_i^X(0)$.

[12]A related and older term is the "Concorde effect" (Dawkins and Carlisle, 1976), named after the sustained investment in the Concorde supersonic jet project, after it was recognized to be an unprofitable venture.

developing previous stages of an ongoing project. In our experiment, we expect subjects to positively associate the resources they sink into a real-effort task with the lottery they earn if they perform well enough in that task. We are also agnostic about what might count as a cost and our hypothetical scenarios cover several: effort, time, money, and emotion.

What distinguishes the sunk cost effect from the endowment effect (Kahneman et al., 1990) is that the former requires some resources to be sunk. In contrast, the latter occurs when an individual receives an endowment costlessly and derives value from the very fact they have it. The sunk cost effect is thereby more nuanced, specifying that the individual must have spent (sunk) some resources to have the "endowment". As such, we view the endowment effect as an essential and necessary part of the sunk cost effect; when an individual exhibits the sunk cost effect they also exhibit the endowment effect, but the reverse is not necessarily true.[13]

To detect the part of the sunk cost effect net of the endowment effect (and hence the existence of the sunk cost effect in its own right), we dedicate a treatment group to measuring the latter. By comparing responses there to those in our primary sunk cost group, we can identify whether sunk costs have an effect over and above the otherwise standard endowment effect.[14]

3.2 Experiment

3.2.1 Subjects

Subjects were recruited via the Amazon Mechanical Turk (MTurk) online platform, commonly used throughout the social sciences including economics and finance (e.g., DellaVigna and Pope, 2017; Kuziemko et al., 2015; Lian et al., 2019). MTurk's population has been shown to be more demographically diverse and to produce data of a comparable quality to some more traditional participant pools (Chandler et al., 2014; Crump et al., 2013; Paolacci and Chandler, 2014), with many studies replicating classic experiments across various domains including cognitive psychology (e.g., Goodman et al., 2013; Paolacci et al., 2010) and economics (e.g., Horton et al., 2011). Snowberg and Yariv (2020) compare many behavioral measures across representative, MTurk, and student samples. They find the levels of some behaviors vary across subject pools but that those from their MTurk sample generally sat between those of their representative

[13]For some instances in which an individual may exhibit the sunk cost effect, there may not be a final or concrete achievement or outcome (e.g., partway through an ongoing R&D process or a human relationship) making it less clear what any endowment is. We note that our view that the sunk cost effect subsumes the endowment effect includes the case in which the latter is small or zero.

[14]A further possible dichotomy is that which examines the interaction between "loss framing" (Kahneman and Tversky, 1979) and the sunk cost effect. In particular, some literature explains that it is possible to define the sunk cost effect as reflecting a rise in risk-seeking behavior after incurring a sunk cost. Similarly, if the risky choice that comes after a sunk cost is framed in terms of a loss we may instead see a rise in risk aversion, which can then be considered a "reverse sunk cost effect" (Zeelenberg and van Dijk, 1997). Heath (1995) also discusses something akin to a reverse of the normal sunk cost effect arguing that, where individuals have a fixed budget allocated to any given activity, they might feel they have already sunk sufficient resources and may then suboptimally reduce investment in the future as a form of compensation.

and student samples (and more often closer to the representative). Perhaps more importantly, they found correlations between behaviors to be similar across all samples. A disadvantage of MTurk they highlight is that the data can contain more noise. As such, researchers using MTurk may wish to collect larger samples to detect behaviors. In line with this, we collected data from 528 subjects, many more than typically invited to lab studies.

We restricted participation to those in the US, with a good track record (at least 95% of MTurk jobs approved), at least some experience (successfully completed at least 50 MTurk jobs), and who had not taken part in any of our pilots. We used the software Qualtrics to implement the experimental design and collect the response data.

3.2.2 Pretrial registration

The experiment was registered in advance in the AEA RCT Registry (Ronayne et al., 2019). There, we provided an experimental design, power calculations, and detailed our intention to study: reflective thinking as captured by the Cognitive Reflection Test (CRT); fluid and crystallized intelligence captured by Raven's progressive matrices and the verbal reasoning item of the International Cognitive Ability Resource (ICAR), respectively; the Openness scale from the Big Five Personality Inventory; and various demographics. All these measures are included in our analysis, and no other measures were preregistered.

3.2.3 Design overview

We conducted one wave of data collection from 528 subjects.[15] Subjects were randomly allocated to one of three groups such that approximately 60% and 35% fell into our primary and secondary treatment groups, with 5% in the final group.

In the primary *sunk cost* group, subjects completed a real-effort task (counting letters in blocks of text composed of Latin words). If they did well, they earned a lottery (termed an "asset" for subjects) paying \$10 with a 10% chance (else \$0), before being given an (unanticipated) choice to switch to a dominant lottery paying \$10 with a 20% chance (else \$0). The \$1 difference in expected payoff corresponds to 25% of the subjects' participation fee, and represents a meaningful amount to the MTurk population, who regularly respond to similar stakes. In addition, there is growing evidence that online and offline subjects make consistent choices when faced with (the same) lotteries (see, e.g., Lian et al., 2019; Snowberg and Yariv, 2020).

We detail the task below (and present a redacted transcript in the Online Appendix), but the key idea is that the resources (time, effort, etc.) exerted in the task to obtain the lottery, of quantity $r > 0$, are positively associated with it, and lead a subject, if susceptible to the sunk cost effect, to be more likely to make a different choice (the inferior lottery) to that if they had spent no such resources (the superior lottery).

[15] Calculations suggested this would detect effects with 80% power and 5% significance (Ronayne et al., 2019).

In the *endowment* group, subjects did not complete the real-effort task to earn the asset. Instead, these subjects were endowed with the inferior (10% chance of $10) lottery before facing the same (unanticipated) choice to switch to the dominant (20% chance of $10) lottery.

The third and final group was run to check that subjects could generally be expected to maximize their expected pecuniary outcome and understood the wording and descriptions of the lotteries. Subjects in this condition did not face any real-effort task and were not endowed with anything. They were simply given a straight choice between the two lotteries.

After their group-specific tasks, all subjects completed a set of psychometric measures. First, 18 hypothetical scenarios (presented in a random order) to each of which they responded via a 6-point Likert scale. These scenarios form the basis of the scale we set out to validate (the numbered list is in the Online Appendix). Table 1 details which scenarios highlight which sunk resources: *effort*, *time*, *money*, *emotion*, and a final category *belief*, shorthand for the resources sunk during the process of belief formation, which is likely to be a subset of effort.[16]

Table 1: Different types of sunk resources highlighted by the hypothetical scenarios

Resource	Scenario ID																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Effort		✓	✓	✓		✓			✓	✓					✓	✓	✓	
Time		✓	✓	✓					✓		✓			✓	✓	✓		
Money	✓				✓		✓	✓				✓	✓					✓
Emotion	✓					✓								✓				
Belief															✓	✓	✓	

The full text of the 18 scenarios is given in the redacted transcript in the Online Appendix. A dark tick indicates the scenario's primary focus, while lighter ticks indicate some secondary resources.

Subjects then completed the CRT, measures of fluid and crystallized intelligence, and openness to experience. The CRT measures reflective capacity, i.e., the ability to think about problems that tend to induce people to fall prey to a behavioral bias (Frederick, 2005). We adopt the extended version of the test (Toplak et al., 2014), which consists of 7 right/wrong questions, giving each subject a score in $\{0, \ldots, 7\}$.

Fluid intelligence was measured using 10 Raven's progressive matrices (John and Raven, 2003), giving scores $\{0, \ldots, 10\}$. We follow the common practice of assessing crystallized intelligence by verbal reasoning ability (e.g., Bruine de Bruin et al., 2007) and use the 16-item verbal reasoning subset of the ICAR (Condon and Revelle, 2014), giving scores $\{0, \ldots, 16\}$. Potential for open-minded thinking was gauged via the openness to experience personality trait, which

[16]The scenarios are taken or adapted from known sources: 1-10 from the University of Michigan's Health and Retirement Study (`https://hrs.isr.umich.edu/about`, also used by Olivola, 2018); 11-12 from Arkes and Blumer (1985); and 13 from Thaler (1980). We added 14-18 to balance and enrich the set of scenarios across resources.

we measured using a 12-item subscale from the NEO Five Factor Inventory (Costa and McCrae, 1989); each item is responded to via a 5-point Likert scale coded $\{0, \ldots, 4\}$, giving scores $\{0, \ldots, 48\}$.

The measures were incentivized such that one of each subject's answers (across the CRT, ICAR, and Raven's tests) was chosen at random, and if correct, \$2 was added to their payment. Last, demographic information was collected. We now outline the sunk cost task in greater detail.

3.2.4 The sunk cost task

Subjects in our primary (sunk cost) group were able to earn a lottery (termed an "asset" in the experimental instructions) by sinking sufficient resources into a task of counting letters (similar to that used by Rosaz and Villeval, 2012). Each subject faced a sequence of five blocks of text. For each block they were asked to count the number of occurrences of two different letters within the time limit of 60 seconds (see Figure 1 for a screenshot). For each letter correctly counted (within a margin of error of one), they got one point. If they got a total of 6 out of 10 points or more, they earned the (inferior) lottery paying \$10 with a 10% chance, \$0 otherwise. To avoid potential emotional primes from valences of familiar words, randomly-selected Latin words (from the Lorem Ipsum corpus) were used.[17]

We consider the cleanest design to be one in which all subjects incur the same level of sunk cost. A challenge to achieving this arises due to the highly subjective nature of sunk cost perceptions, e.g., the same task may be easy for one subject and difficult for another, the utility two subjects derive from \$10 may differ, etc. For our task, providing only one difficulty level to all subjects would lead to substantial variation in the amount of perceived sunk costs across subjects (indeed, pilots revealed substantial heterogeneity in subjects' performance in the task). To combat this, we implemented subject-specific block paths (unknown to subjects). Specifically, based on their performance on the first text block, subjects were branched into five routes (very hard, hard, medium, or easy; some were branched into a very easy route if they also performed poorly in their second block, which may have included anyone not paying attention). It appears we were somewhat successful in this regard; 94% of subjects completing the task followed the easy, medium, hard, or very hard paths achieved almost completely statistically indistinguishable scores on average.[18]

[17] Subjects did not show signs of fatigue due to the task, which otherwise may have impacted their performance in later questions or tasks; those randomly allocated the task (those in the sunk cost group) generated CRT, ICAR, and Raven's scores that were statistically indistinguishable from those not allocated the task (those in other groups).

[18] Average scores for easy, medium, hard and very hard groups were 7.2, 7.6, 8.4 and 7.6, respectively. Pairwise t-tests show these to be statistically indistinguishable save the comparison hard vs very hard; although the difference was only 0.7, it suggests the hard route could have been made slightly easier. Moreover, a more reasonable comparison may remove subjects for which there was no suitable path i.e., those so good that even the hard path was too easy and those so bad that even the easier paths were too hard or bothersome. Those subjects are unlikely to perceive the same sunk costs as those not at the extremes. Removing, for example, the 11.1% who scored 10/10 and the 5.6% who scored between 0-4/10 brings the average scores even closer together, to 7.5, 7.7, 7.9 and 7.4.

Figure 1: Real-effort task

Subjects in the sunk cost group entered letter counts for five blocks of text. A timer showed the amount of time remaining (here, 55 seconds). If subjects did not enter a count, their answer was logged as incorrect. If time expired, the answers present were submitted, and the subject automatically progressed to the next page. All those points were covered in the instructions. The words were uploaded as an image file to prevent "CTRL+F" commands from giving the answer. Only integers were accepted.

We consider the inclusion of endogenous task difficulty a contribution to the literature. Existing designs to measure the sunk cost effect do not attempt to control for variation in the level of sunk costs as subjectively perceived by subjects, and as such may be underpowered. This may be one reason several prior studies find weak or no evidence of its existence.

The path-dependent design also increases the experiment's efficiency by allowing a greater share of subjects to get at least 6/10 in the task (and thus earn a lottery). With a fixed set of text for all subjects, we expected to lose a lot of data because many would fail to reach 6/10 (and thus not earn a lottery). We believe we were successful in this regard; 268 of 306 subjects (88%) scored at least 6/10, making it through to the asset choice, and so can be included in all our analyses.

4 Results

We collected data from 528 subjects over July 22-24, 2019. Average completion time was 27m20s. Subjects received $4.00 for participating, corresponding to an average hourly wage of $8.78 ($11.21 including incentive payments). Subject demographics are given in Table A1 in the Online Appendix.

4.1 The sunk cost effect

Table 2 reports the proportion of subjects in the three different treatment groups who chose the dominated lottery. We first report that 23% (95CI = [18, 28]%) of the 268 subjects in the sunk cost group who earned the dominated lottery from the real-effort task chose to stick with it, which we interpret as evidence of behavior consistent with the sunk cost effect.[19]

In the endowment group, 7% (95CI = [4, 11]%) of the 197 subjects chose the dominated asset. The difference in the proportion of irrational decisions between the sunk cost and endowment groups is significantly different from zero ($d = 0.16$; $p < 0.001$); the sunk cost effect was present and not entirely explained by the endowment effect. In fact, in our data the sunk cost motive appeared to exert a significantly greater influence on decisions than the endowment motive per se, with the latter accounting for approximately only a third of the overall effect.

Last, we point out that none of the 25 subjects who were presented with a straight choice chose the dominated alternative. This suggests it is unlikely an individual would choose that asset due to mathematical deficiency, misunderstanding the text, or (at least some sorts of) experimental demand effect.

Table 2: Behavior consistent with the sunk cost effect

Condition	Choice			Pr(dominated)
	Dominated	Dominant	n	
Earned via sunk costs	62	206	268	0.231
Endowment only	14	183	197	0.071
Straight choice	0	25	25	0.000
Difference in proportions				0.160
P-value				<0.001

4.2 Drivers of sunk cost behavior

Table 3 first reports that the average marginal effect (AME) of cognitive reflection on the probability of sunk cost behavior is significant (both $p < 0.001$) and negative; a one standard deviation increase corresponds to a decrease of approximately 0.11-0.12 (11-12 percentage points).

Guided by the tripartite theory of the mind of Stanovich (2012), we now assess the strength of inputs of the reflective mind in explaining sunk cost behavior.[20] We find crystallized intelligence to be significant (both $p < 0.001$), but not fluid intelligence or openness to experience, as reported in specifications (3) and (4) of Table 3.[21] Interpreting the estimated AMEs, a one

[19] A different explanation is that performance, rather than sunk cost, matters. If so, subjects scoring higher would be more likely to stick with the dominated lottery, but we find a negative correlation ($r = -0.16$; $p = 0.009$).

[20] Auxiliary regressions show both fluid and crystallized intelligence are associated with CRT scores (see Table A3).

[21] The two intelligence measures are of course correlated ($r = 0.60$; $p < 0.001$) and either measure alone explains

Table 3: Determinants of susceptibility to the sunk cost effect

Average marginal effects	Dependent variable: Behavior			
	(1)	(2)	(3)	(4)
Cognitive reflection	-0.112***	-0.116***		
	(0.023)	(0.023)		
Fluid intelligence			-0.013	-0.009
			(0.030)	(0.028)
Crystallized intelligence			-0.173***[a]	-0.175***[a]
			(0.028)	(0.030)
Openness			-0.000	-0.000
			(0.022)	(0.022)
Demographics		X		X
Observations	265	265	265	265

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. [a] Bonferroni-adjusted $p < 0.001$. Average marginal effects with robust standard errors in parentheses. Estimates from the underlying probit regressions are in Table A2 in the Online Appendix. Regressors are standardized. Behavior is a binary variable = 1 if the subject displayed sunk cost behavior. Demographics include sex, age, race, household income, education, and conservatism. Of the 268 subjects in the sunk cost treatment group, 3 chose not to specify their sex, leaving 265 for analysis.

standard deviation increase in crystallized intelligence decreases the probability of sunk cost behavior by approximately 17-18 percentage points on average.

4.2.1 Interpretation

Previous work (Stanovich, 2012; Stanovich and West, 2008, discussed in Section 2) suggests the key to being able to override a heuristically-primed response is recognizing the need to override it in the first place (while the capacity required to avoid the bias once recognized is relatively slight). As such, stocks of knowledge or experience (crystallized intelligence) and openness of mind are likely to matter more than computational power (fluid intelligence). Our results concerning intelligence measures support this hypothesis in the context of sunk costs. A natural interpretation is that a bigger stock of experience is helpful in enabling individuals to recognize instances of the sunk cost effect and thereby avoid them.

Stanovich and West (2008) also argue conventional measures lack the scope to adequately capture the types of knowledge required for situations that invoke behavioral biases, but our evidence suggests they do not. Contrary to their hypothesis we do not find a relationship between sunk cost behavior and openness to experience. This could be because that personality trait is not a major driver of this mode of thinking, because it is not a major driver of the sunk cost effect, or something particular to our sample: a question for future research.

significant variation in sunk cost behavior. However, when both are included in regressions, only crystallized intelligence retains explanatory power. See Table A4 for the supporting regressions.

4.3 SCE-8: A scale to measure susceptibility to the sunk cost effect

We first analyze the latent factor structure of our 528 subjects' responses to the 18 scenarios to identify any underlying factors causing them to covary. Informed by that analysis, we select the scenarios to be included in our scale, which we then relate back to behavior.

4.3.1 Factor analysis

We find that a single principal factor is the most reasonable and explains 90% of the variation in our subjects' response data. Following standard procedures (details can be found in the Online Appendix) we retain (and confirm) the scenarios with sufficiently strong factor loadings (reported in Table A6), resulting in the selection of 8 of them. We name the scale formed of these scenarios the "SCE-8". This scale is provided in the Online Appendix, along with the corresponding scores of our subjects (presented in Figure A1).

4.3.2 Interpretation

Across scenarios, different types of costs are sunk to different degrees. We interpret the emergence of a single factor as the best representation of the data as reflecting both the idea that the sunk cost effect applies across resources and the highly interdependent nature of the resources involved. Moreover, among the scenarios with the highest loadings, there is at least one scenario for which each of the main resources covered is the predominant resource (see Tables 1 and A6), further suggesting the factor is relevant for various kinds of sunk cost.

4.3.3 Validating a sunk cost scale with real decisions

Sunk cost behavior and our SCE-8 scale have a significant pairwise correlation ($r = 0.26$; $p < 0.001$).[22] Furthermore, as Table 4 shows, the scale is significantly associated to cognitive reflection (both $p < 0.001$), just as our behavioral measure is (as seen in Table 3). Interpreting the coefficients, a one standard deviation increase in cognitive reflection is associated with an approximate decrease of 0.4 standard deviations in susceptibility to the sunk cost effect as measured by the SCE-8. Moreover, unpacking cognitive reflection into three components, we find that crystallized intelligence is significantly associated with SCE-8, just as our behavioral measure is. These findings and consistencies lead us to conclude that the SCE-8 scale is an appropriate substitute for a behavioral measure.

[22]We report the pairwise correlations between sunk cost behavior and susceptibility in each scenario in Table A5. Despite the interdependent nature of resources and no scenario perfectly reflecting our experiment's specific real effort task, it is notable that of the 8 scenarios that pairwise correlate with sunk cost behavior, we identified 5 to feature effort as a salient sunk resource (Table 1).

Table 4: Determinants of the SCE-8

Average marginal effects	Dependent variable: SCE-8			
	(1)	(2)	(3)	(4)
Cognitive reflection	-0.421***	-0.417***		
	(0.055)	(0.058)		
Fluid intelligence			-0.121	-0.125
			(0.068)	(0.068)
Crystallized intelligence			-0.339***[a]	-0.330***[a]
			(0.081)	(0.082)
Openness			-0.152*	-0.116
			(0.066)	(0.070)
Demographics		X		X
Observations	265	265	265	265

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. [a] Bonferroni-adjusted $p < 0.01$. OLS regressions. All variables are standardized. Robust standard errors are in parentheses. Specification testing prompted the inclusion of squared and interaction terms of the three trait measures in (3)-(4), and so average marginal effects are reported. Demographics include sex, age, race, household income, education, and conservatism. Of the 268 subjects in the sunk cost treatment group, 3 chose not to specify their sex, leaving 265 for analysis.

5 Conclusion

The sunk cost effect is one of the most well-known biases in decision making. Our work advances the identification, understanding, and measurement of the effect.

In contrast to existing research, we provide significant evidence of the sunk cost effect through an incentivized experiment with human subjects. In addition, and to aid our design, we offer a formal choice-based definition. In our sample, we also find that the endowment effect, far from accounting for all of it, is approximately only a third as strong as the sunk cost effect.

Second, we find strong evidence that capacity for cognitive reflection is negatively related to sunk cost behavior: the ability to override one's instinctive response matters for overcoming the effect. Moreover, our results support the intuitive hypothesis that one's stock of knowledge and experience is predictive of susceptibility to the sunk cost effect, rather than computational ability. This carries an important and subtle point: we find some measures of intelligence to be highly correlated with the sunk cost effect and others not. This could explain the mixed results in the literature and provides a warning: depending upon which measure is used, it is possible to miss the association between the sunk cost effect and cognitive ability. We encourage future research to further investigate the cognitive, and other, underpinnings of the effect.

Third, we offer a scale – the SCE-8 – for researchers to measure susceptibility to the sunk cost effect, without needing to conduct an experiment. The SCE-8 covers a range of costs, capturing the generality of the effect, and appears a good substitute for a behavioral measure. The SCE-8

can be incorporated easily into applied work and can either serve as a measure of interest per se, or as a control just as other measures have been for decades.

References

Arkes, Hal R and Catherine Blumer (1985), "The psychology of sunk cost." *Organizational Behavior and Human Decision Processes*, 35, 124–140.

Ashraf, Nava, James Berry, and Jesse M Shapiro (2010), "Can higher prices stimulate product use? evidence from a field experiment in Zambia." *American Economic Review*, 100, 2383–2413.

Augenblick, Ned (2016), "The sunk-cost fallacy in penny auctions." *The Review of Economic Studies*, 83, 58–86.

Berry, James, Greg Fischer, and Raymond P Guiteras (2020), "Eliciting and utilizing willingness to pay: Evidence from field trials in northern Ghana." *Journal of Political Economy* - forthcoming.

Blais, Ann-Renée and Elke U Weber (2006), "A domain-specific risk-taking (DOSPERT) scale for adult populations." *Judgment and Decision Making*, 1.

Bollen, Kenneth A and J Scott Long (1993), *Testing Structural Equation Models*. SAGE Publications, Newbury Park, CA, USA.

Brockhoff, Klaus, Maximilian Margolin, and Juergen Weber (2015), "Towards empirically measuring patience." *Universal Journal of Management*, 3, 169–178.

Bruine de Bruin, Wändi, Andrew M Parker, and Baruch Fischhoff (2007), "Individual differences in adult decision-making competence." *Journal of Personality and Social Psychology*, 92, 938–956.

Cattell, R B (1966), "The scree test for the number of factors." *Multivariate Behavioral Research*, 1, 245–276.

Chandler, Jesse, Pam Mueller, and Gabriele Paolacci (2014), "Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers." *Behavior Research Methods*, 46, 112–130.

Cohen, Jessica, Pascaline Dupas, and Simone Schaner (2015), "Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial." *American Economic Review*, 105, 609–645.

Condon, David M and William Revelle (2014), "The international cognitive ability resource: Development and initial validation of a public-domain measure." *Intelligence*, 43, 52–64.

Costa, P T and R R McCrae (1989), "NEO five-factor inventory (NEO-FFI)." *Odessa, FL: Psychological Assessment Resources*.

Costello, Anna B and Jason Osborne (2005), "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis." *Practical Assessment, Research, and Evaluation*, 10, 7.

Crump, Matthew JC, John V McDonnell, and Todd M Gureckis (2013), "Evaluating amazon's mechanical turk as a tool for experimental behavioral research." *PloS one*, 8, e57410.

Dawkins, R and T R Carlisle (1976), "Parental investment, mate desertion and a fallacy." *Nature*, 262, 131–133.

DellaVigna, Stefano and Devin Pope (2017), "What motivates effort? Evidence and expert forecasts." *The Review of Economic Studies*, 85, 1029–1069.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2010), "Are risk aversion and impatience related to cognitive ability?" *American Economic Review*, 100, 1238–60.

Frederick, Shane (2005), "Cognitive reflection and decision making." *Journal of Economic Perspectives*, 19, 25–42.

Friedman, Daniel, Kai Pommerenke, Rajan Lukose, Garrett Milam, and Bernardo A Huberman (2007), "Searching for the sunk cost fallacy." *Experimental Economics*, 10, 79–104.

Goodman, Joseph K, Cynthia E Cryder, and Amar Cheema (2013), "Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples." *Journal of Behavioral Decision Making*, 26, 213–224.

Hafenbrack, Andrew C, Zoe Kinias, and Sigal G Barsade (2014), "Debiasing the mind through meditation: Mindfulness and the sunk-cost bias." *Psychological science*, 25, 369–376.

Haita-Falah, Corina (2017), "Sunk-cost fallacy and cognitive ability in individual decision-making." *Journal of Economic Psychology*, 58, 44–59.

Heath, C (1995), "Escalation and de-escalation of commitment in response to sunk costs: The role of budgeting in mental accounting." *Organizational Behavior and Human Decision Processes*, 62, 38–54.

Horton, John J, David G Rand, and Richard J Zeckhauser (2011), "The online laboratory: Conducting experiments in a real labor market." *Experimental Economics*, 14, 399–425.

John and Jean Raven (2003), "Raven progressive matrices." In *Handbook of Nonverbal Assessment* (R Steve McCallum, ed.), 223–237, Springer US, Boston, MA.

Just, David R and Brian Wansink (2011), "The flat-rate pricing paradox: Conflicting effects of "all-you-can-eat" buffet pricing." *The Review of Economics and Statistics*, 93, 193–200.

Kahneman, Daniel (2011), *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York, NY, USA.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1990), "Experimental tests of the endowment effect and the coase theorem." *Journal of Political Economy*, 98, 1325–1348.

Kahneman, Daniel and Amos Tversky (1979), "Prospect theory: an analysis of decision under risk." *Econometrica*, 47, 263–291.

Kaiser, Henry F (1960), "The application of electronic computers to factor analysis." *Educational and Psychological Measurement*, 20, 141–151.

Kaiser, Henry F and John Rice (1974), "Little jiffy, mark iv." *Educational and Psychological Measurement*, 34, 111–117.

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva (2015), "How elastic are preferences for redistribution? Evidence from randomized survey experiments." *American Economic Review*, 105, 1478–1508.

Larrick, Richard P, Richard E Nisbett, and James N Morgan (1993), "Who uses the cost-benefit rules of choice? implications for the normative status of microeconomic theory." *Organizational Behavior and Human Decision Processes*, 56, 331–347.

Leclerc, France, Bernd H Schmitt, and Laurette Dube (1995), "Waiting time and decision making: Is time like money?" *Journal of Consumer Research*, 22, 110–119.

Lian, Chen, Yueran Ma, and Carmen Wang (2019), "Low interest rates and risk-taking: Evidence from individual investment decisions." *The Review of Financial Studies*, 32, 2107–2148.

MacCallum, Robert C, Michael W Browne, and Hazuki M Sugawara (1996), "Power analysis and determination of sample size for covariance structure modeling." *Psychological Methods*, 1, 130–149.

Olivola, Christopher Y (2018), "The interpersonal sunk-cost effect." *Psychological Science*, 29, 1072–1083.

Paolacci, Gabriele and Jesse Chandler (2014), "Inside the Turk: Understanding Mechanical Turk as a participant pool." *Current Directions in Psychological Science*, 23, 184–188.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G Ipeirotis (2010), "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making*, 5, 411–419.

Parker, Andrew M and Baruch Fischhoff (2005), "Decision-making competence: External validation through an individual-differences approach." *Journal of Behavioral Decision Making*, 18, 1–27.

Proto, Eugenio, Aldo Rustichini, and Andis Sofianos (2019), "Intelligence, personality, and gains from cooperation in repeated interactions." *Journal of Political Economy*, 127, 1351–1390.

Ronayne, David, Daniel Sgroi, and Anthony Tuckwell (2019), "Evaluating the sunk cost effect." *AEA RCT Registry*, URL `https://doi.org/10.1257/rct.4083-2.0`.

Rosaz, Julie and Marie Claire Villeval (2012), "Lies and biased evaluation: A real-effort experiment." *Journal of Economic Behavior and Organization*, 84, 537–549.

Snowberg, Erik and Leeat Yariv (2020), "Testing the waters: Behavior across participant pools." *American Economic Review* - forthcoming.

Soman, Dilip (2001), "The mental accounting of sunk time costs: Why time is not like money." *Journal of Behavioral Decision Making*, 14, 169–185.

Stanovich, Keith E (2008), "Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory?" In *Two minds: Dual processes and beyond* (J Evans and K Frankish, eds.), Oxford University Press, Oxford, United Kingdom.

Stanovich, Keith E (2012), "On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning." In *The Oxford Handbook of Thinking and Reasoning* (Keith J Holyoak and Robert G Morrison, eds.), Oxford University Press, Oxford, United Kingdom.

Stanovich, Keith E and Richard F West (2008), "On the relative independence of thinking biases and cognitive ability." *Journal of Personality and Social Psychology*, 94, 672–695.

Staw, Barry M (1976), "Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action." *Organizational Behavior and Human Performance*, 16, 27–44.

Strough, Jonell, Clare M Mehta, Joseph P McFall, and Kelly L Schuller (2008), "Are older adults less subject to the sunk-cost fallacy than younger adults?" *Psychological Science*, 19, 650–652.

Sweis, Brian M, Samantha V Abram, Brandy J Schmidt, Kelsey D Seeland, Angus W MacDonald, Mark J Thomas, and A David Redish (2018), "Sensitivity to "sunk costs" in mice, rats, and humans." *Science*, 361, 178–181.

Thaler, Richard (1980), "Toward a positive theory of consumer choice." *Journal of Economic Behavior & Organization*, 1, 39–60.

Thaler, Richard H (1999), "Mental accounting matters." *Journal of Behavioral Decision Making*, 12, 183–206.

Toplak, Maggie E, Richard F West, and Keith E Stanovich (2011), "The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks." *Memory & Cognition*, 39, 1275–1289.

Toplak, Maggie E, Richard F West, and Keith E Stanovich (2014), "Assessing miserly information processing: An expansion of the cognitive reflection test." *Thinking & Reasoning*, 20, 147–168.

Weigel, Collin (2018), "Don't stop now: The sunk cost effect in an incentivized lab experiment." *Working Paper*.

Zeelenberg, Marcel and Eric van Dijk (1997), "A reverse sunk cost effect in risky decision making: Sometimes we have too much invested to gamble." *Journal of Economic Psychology*, 6, 677–691.

Online Appendix to

Evaluating the Sunk Cost Effect

Ronayne, Sgroi & Tuckwell

[To companion our paper as online Supplementary Material]

SCE-8: A scale to measure susceptibility to the sunk cost effect

You will be presented with 8 hypothetical scenarios, each of which lead to a choice. For each one, tell us what you would do. [For each item subjects have a 6-point scale for which the two alternatives are written over the left-most and right-most points. The alternatives are provided after each scenario below.]

A. You have been looking forward to this year's Halloween party. You have the right cape, the right wig, and the right hat. All week, you have been trying to perfect the outfit by cutting out a large number of tiny stars to glue to the cape and the hat, and you still need to glue them on. On the day of Halloween, you decide that the outfit looks better without all these stars you have worked so hard on.
[Wear stars; Go without.]

B. You have been asked to give a toast at your friend's wedding. You have worked for hours on this one story about you and your friend taking drivers' education, but you still have some work to do on it. Then you realize that you could finish writing the speech faster if you start over and tell the funnier story about the dance lessons you took together. [Finish the toast about driving; Rewrite the toast about dancing.]

C. You are painting your bedroom with a sponge pattern in your favorite color. It takes a long time to do. After you finish two of the four walls, you realize you would have preferred the solid color instead of the sponge pattern. You have enough paint left over to redo the entire room in the solid color. It would take you the same amount of time as finishing the sponge pattern on the two walls you have left.
[Finish the sponge pattern; Redo the room in a solid color.]

D. You have invested a good deal of your time into a project and it is failing. You have the option to start on something different that you now know is more likely to be successful but you know you cannot get the time back that you spent on the project. [Keep going with the project; Start something different.]

E. You have an investment strategy that you have developed over several months. It is not working and you are losing money, but there is no way for you to recover the lost effort put into developing the strategy. [Start afresh; Keep going.]

F. Your relationship with your partner is not going well. You have reasoned it out and you have realized that if you knew how it would go when you started the relationship you would not have gone through with it. You now have the opportunity to break up, but you have been together for many months.
[Keep going; Break up.]

G. You have been thinking about how to vote in an election and have invested a good deal of your time to try and make the right decisions including reading newspapers and comment pieces online and thinking hard about the issues. You discover that much of the information you were using is false and a more trustworthy source suggests your initial view was wrong. [Keep beliefs; Change beliefs.]

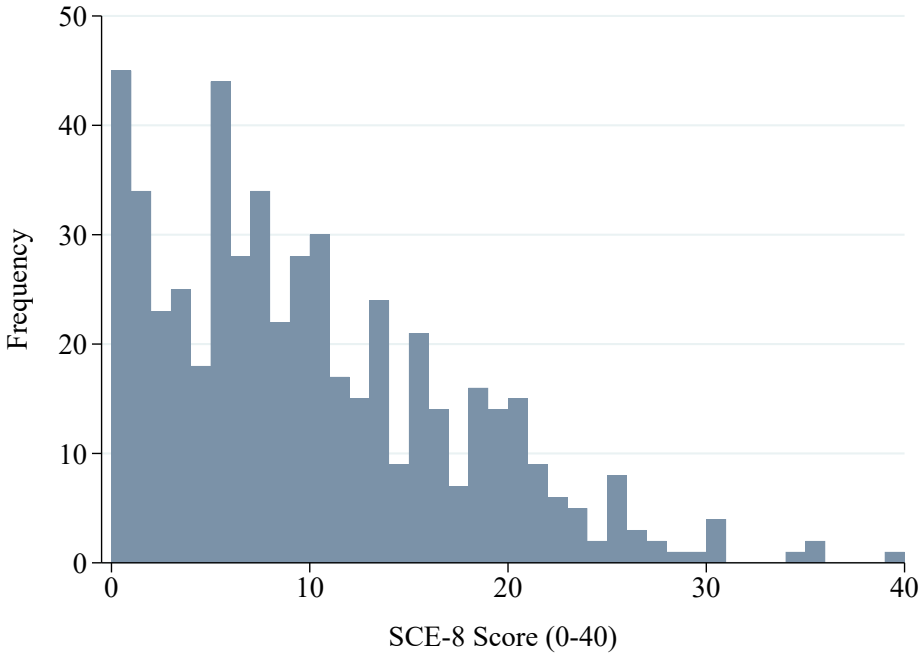
H. You have been thinking hard about the best route to get to somewhere you haven't been to before. Unfortunately, your internet connection isn't working so you have to base your decision on your beliefs about the town's layout. You come to a conclusion on the best possible route but then suddenly the internet is back online. [Look up route online; Stick to planned route.]

Table A1: Subject Demographics

Characteristic	
Sex	
Male	315 (60)
Female	209 (40)
Other / Prefer not to say	4 (0)
Age, mean years [sd]	34.8 [10.3]
18-25	75 (14)
26-30	151 (29)
31-40	176 (33)
41-50	76 (14)
51+	50 (9)
Race	
White	359 (68)
Black or African American	70 (13)
Hispanic or Latino	40 (8)
American Indian or Alaska Native	4 (1)
Asian American	41 (8)
Native Hawaiian or Pacific Islander	0 (0)
Other	14 (3)
Income[a]	
0 – 9.999	16 (3)
10 – 19.999	48 (9)
20 – 29.999	79 (15)
30 – 39.999	79 (15)
40 – 49.999	71 (13)
50 – 59.999	63 (12)
60 – 69.999	38 (7)
70 – 79.999	32 (6)
80 – 89.999	19 (4)
90 – 99.999	30 (6)
100 – 124.999	24 (5)
125 – 149.999	14 (3)
150+	15 (3)
Education	
High school (grades 9-12, no degree)	5 (1)
High school graduate (or equivalent)	66 (13)
Some college (1-4 years, no degree)	175 (33)
Bachelor's degree (BA, BS, AB, etc)	228 (43)
Master's degree (MA, MS, MENG, MSW, etc)	47 (9)
Professional school degree (MD, DDC, JD, etc)	5 (1)
Doctorate degree (PhD, EdD, etc)	2 (0)
Political Affiliation[b] [sd]	36.6 [31.0]
N	528

Frequencies; (% within characteristic); [standard deviation]
[a] Household annual pretax income in '000 USD
[b] 0 = "Entirely Liberal"; 100 = "Entirely Conservative"

Figure A1: SCE-8



Subjects' scores (n = 528) on the SCE-8 scale. Each scenario is responded to on a 6-point Likert scale and is coded 0-5 such that the higher the score, the higher the susceptibility, hence the range is 0-40. Summary statistics for our sample: min 0, max 40, average 9.5, median 8, inter-quartile range [4,15], and standard deviation 7.5.

Table A2: Probit coefficients underlying the average marginal effects of Table 3

	Dependent variable: Behavior			
	(1)	(2)	(3)	(4)
Cognitive reflection	-0.399***	-0.430***		
	(0.091)	(0.093)		
Fluid intelligence			-0.052	-0.038
			(0.120)	(0.117)
Crystallized intelligence			-0.696***	-0.728***
			(0.131)	(0.141)
Openness			-0.002	0.002
			(0.088)	(0.092)
Female		-0.288		-0.210
		(0.190)		(0.200)
White		-0.317		-0.139
		(0.199)		(0.220)
Age		0.035		0.051
		(0.055)		(0.057)
Age^2		-0.000		-0.001
		(0.001)		(0.001)
College		0.278		0.289
		(0.198)		(0.201)
Income		-0.007		-0.000
		(0.024)		(0.028)
Conservatism		-0.000		-0.002
		(0.003)		(0.003)
Constant		-1.283		-1.818
		(1.047)		(1.084)
Observations	265	265	265	265

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Probit regressions. Regressors are standardized. Behavior is a binary variable = 1 if the subject displayed sunk cost behavior. Estimated coefficients shown with robust standard errors in parentheses. Of the 268 subjects in the sunk cost treatment group, 3 chose not to specify their sex, leaving 265 for analysis.

Table A3: Inputs of the reflective mind

Average marginal effects	Dependent variable: Cognitive reflection	
Fluid intelligence	0.196^{***a}	0.193^{***a}
	(0.044)	(0.045)
Crystallized intelligence	0.539^{***a}	0.527^{***a}
	(0.036)	(0.037)
Openness	-0.020	-0.006
	(0.029)	(0.030)
Demographics		X
Observations	524	524

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. ^{a} Bonferroni-adjusted $p < 0.001$. OLS regressions. Robust standard errors in parentheses. Regressors are standardized. Specification testing prompted the inclusion of squared, cubed, and interaction terms of the three trait measures, and so average marginal effects are reported. Demographics include sex, age, race, household income, education, and conservatism. Of the total 528 subjects, 4 chose not to specify their sex, leaving 524 for analysis.

Table A4: Auxiliary results to Table 3

AMEs	Dependent variable: Behavior							
Fluid int.	-0.112***	-0.107***					-0.013	-0.009
	(0.025)	(0.025)					(0.030)	(0.028)
Cryst. int.			-0.181***	-0.181***			-0.173***[a]	-0.175***[a]
			(0.023)	(0.026)			(0.028)	(0.030)
Openness					-0.033	-0.024		
					(0.026)	(0.026)		
Demos.		X		X		X		X
Obs.	265	265	265	265	265	265	265	265

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. [a] Bonferroni-adjusted $p < 0.001$. Probit regressions. Average marginal effects with robust standard errors in parentheses. Regressors are standardized. Behavior is a binary variable = 1 if behavior was consistent with the sunk cost effect. Demographics include sex, age, race, household income, education, and conservatism. Of the 268 subjects in the sunk cost treatment group, 3 chose not to specify their sex, leaving 265 for analysis.

Table A5: Pairwise correlations of each hypothetical scenario with sunk cost behavior

Scenario	Pairwise correlation	Scenario	Pairwise correlation
1	0.120*	10	0.142*
2	-0.009	11	0.135*
3	0.141*	12	0.116
4	-0.031	13	0.032
5	-0.088	14	0.139*
6	0.151*	15	0.240**
7	-0.027	16	0.250**
8	0.059	17	0.006
9	0.060	18	-0.046

n = 268 (those in the sunk cost treatment); $^{**}p < 0.01$; $^{*}p < 0.05$. The full text of the 18 scenarios is given in the redacted transcript.

Factor analysis. Below we document in detail how we selected the SCE-8 from the initial 18 scenarios.

Exploratory factor analysis. Various checks support the factorability of the data: multicollinearity between the 18 items is low (mean variance inflation factor: 1.22); Bartlett's sphericity test is significant ($\chi^2(153) = 1160.29$; $p < 0.001$); and the Kaiser-Meyer-Olkin measure of sampling adequacy (0.83) surpasses the advised threshold of 0.6 (Kaiser and Rice, 1974). Extracting factors with eigenvalues > 1 (the Kaiser criterion; Kaiser, 1960), we find one principal factor that explains 90% of the variance (with an eigenvalue of 2.81). A scree test (Cattell, 1966) also supports a one-factor solution, dropping off substantially after the first factor. Investigating the fit of that solution, we extract one factor and find that the majority (11 of 18) of scale items load well onto it (with a loading > 0.32; Costello and Osborne, 2005), as reported in the first row of Table A6.

Table A6: Factor loadings by scenario ID

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
25	18	50	28	24	46	35	39	39	57	49	59	20	43	53	44	30	3
-	-	48	-	-	48	-	-	-	56	48	62	-	44	55	50	-	-

Loadings are multiplied by 100. The first and second rows report the loadings from the factor analysis with all 18 and the selected 8 (SCE-8 scale) scenarios, respectively.

We follow recommended practice and drop items with weak loadings (< 0.32). We also drop items that fail to meet this threshold when excluding subjects ($n = 40$) who completed the 18 scenarios in less than 90 seconds (items 7 and 8). Finally, we drop item 9 because its loading of 0.34 (after dropping 7 and 8) is marginal and because time (the predominant resource it relates to) is well represented by several other items with strong loadings. The loadings generated from the remaining 8 scenarios are given in the second row of Table A6. A reliability analysis of those items demonstrates internal consistency (Cronbach's alpha = 0.75).[23]

Confirmatory factor analysis. To confirm the suitability of a one-factor representation of these 8 items, we estimate a structural equation model that links subjects' responses to them with one latent variable. We find the standardized factor loadings of each of the 8 items to be significant ($p < 0.001$) and above the recommended value (0.32). Goodness-of-fit measures are also satisfactory; the standardized root mean squared error is 0.02, falling in the "good" to "excellent" range (0.01-0.05; MacCallum et al., 1996), while the χ^2-to-degrees-of-freedom ratio is 1.02 (ratios between 1-3 are acceptable, with values closer to 1 indicating a better fit; Bollen and Scott Long, 1993).

[23] A scale with all 11 items with initial loadings > 0.32 yields a negligibly higher Cronbach's alpha of 0.749 (cf. 0.747 with 8 items), owing to the fact that the additional three items have notably weaker loadings.

Redacted transcript: online only. [Bold text in square brackets was not seen by subjects.]

Participation Agreement

You have been invited to take part in a research study run by researchers at the University of Warwick. Please read the following statements carefully and answer the question below.

Our commitments and privacy policy
We never deceive participants. For example, if we inform you that another participant is making a choice on which you can then react, this is indeed the case. We keep our promises made to participants. For example, if we promise a certain payment, participants will indeed receive it. In the event that we are responsible for a mistake that is to the disadvantage of participants, we will inform and compensate the respective participants. We design, conduct and report our research in accordance with recognized scientific standards and ethical principles.

We adhere to the terms of our privacy policy as stated below:
The data in the participants' database will only be used for the purpose of the study. There is no link between the personal data in the participants' database and the data collected during a study. The generated anonymous data will be used for analysis. The end product will be publicly available. Your participation in this study is purely voluntary, and you may withdraw your participation or your data at any time without any penalty to you. Please note that the software (Qualtrics) automatically notes the time you spent on each question and this data will be made available to researchers for analysis.

If you would like to make a complaint about the way you have been dealt with during the study or any possible harm you might have suffered please address your complaint to the person below, who is a senior University of Warwick official entirely independent of this study:

Head of Research Governance
Research & Impact Services
University House
University of Warwick
Coventry
United Kingdom CV4 8UW

Tel: +44 24 76 522746
Email: researchgovernance@warwick.ac.uk

If you are happy to proceed, please press the "I agree" button below to continue to the task.

○ I agree

[EITHER Treatment Group I: Sunk Cost]

Letter-counting task

Your task is to count the number of times specific letters appear in blocks of text (composed of Latin words).

You will be shown five blocks of text, one at a time, and for each you will have to count the number of occurrences of two different letters. The letters to be counted will appear below the block, and next to them will be boxes to input your count for that letter. The blocks will vary in size but the time limit to count both letters in each block is always the same (1 minute). The blocks will appear in no particular order, i.e. they may start difficult and get easy, vice versa, or be in a random order.

For each letter you count correctly you will receive one point. You will be allowed a margin of error of one, so your letter count will be considered correct if it corresponds to the true value plus or minus one. If you are happy with your counts and wish to move on to the next block before the minute is up, simply press the ">>" button at the bottom of the screen. If the time runs out, the letter counts you have entered will be submitted and you will move on automatically to the next block. Once you have moved on from a block, you cannot go back to it.

If you score at least 6 points out of 10 we will reward your effort by giving you an asset which pays $10 with a 10% chance. We will tell you at the end of the blocks whether you scored enough points to earn the asset.

Before the task begins, you first have a practice block. This will allow you to get a feel for the format and time limit. It does not count for your score out of 10. The practice block will **begin immediately** on the next page.

○ I understand these instructions

[IF the subject scores less than 6/10:]

You scored X/10 points. Unfortunately, you needed 6/10 to acquire the asset.

[IF the subject scores at least 6/10:]

You scored at least 6/10, so you earned the asset.

The asset you now own pays $10 with a 10% chance.

Your asset pays $10 with a 10% chance.

Before we cash your asset, you can switch to a new asset that pays $10 with a 20% chance.

What would you like to do?

- ○ Keep the asset I earned
- ○ Switch to the new asset

[OR Treatment Group II: Endowment Group]

You own an asset that pays $10 with a 10% chance.

Your asset pays $10 with a 10% chance.

Before we cash your asset, you can switch to a new asset that pays $10 with a 20% chance.

What would you like to do?

- ○ Keep my asset
- ○ Switch to the new asset

[OR Treatment Group III: Straight Choice Group]

You have a choice between two assets:

Asset A pays $10 with a 10% chance.
Asset B pays $10 with a 20% chance.

Which do you choose?

- ○ Asset A
- ○ Asset B

[Susceptibility to the Sunk Cost Effect]

Task 1: Choices in hypothetical scenarios

You will be presented with 18 hypothetical scenarios, each of which lead to a choice. For each one, tell us what you would do.

[Scenario IDs are shown in braces below. Items were presented in a random order.

Responses were recorded via a 6-point Likert scale with the two alternative actions written above the left-most and right-most radio buttons. For each scenario below, those words are provided. In each case, exactly one alternative corresponds to behavior consistent with the sunk cost effect.]

[1.] You are buying a gold ring on layaway for someone special. It costs $200 and you have already put down $100 for it, so you owe another $100. One day, you see in the paper that a new jewelry store is selling the same ring for only $90 as a special sale, and you can pay for it using layaway. The new store is across the street from the old one. If you decide to get the ring from the new store, you will not be able to get your money back from the old store, but you would save $10 overall.
[Continue paying at the old store; Buy from the new store.]

[2.] You enjoy playing tennis, but you really love bowling. You just became a member of a tennis club and a bowling club at the same time. The membership to your tennis club costs $200 per year and the membership to your bowling club $50 per year. During the first week of both memberships, you develop an elbow injury. It is painful to either play tennis or bowl. Your doctor tells you that the pain will continue for about a year.
[Play tennis; Bowl.]

[3.] You have been looking forward to this year's Halloween party. You have the right cape, the right wig, and the right hat. All week, you have been trying to perfect the outfit by cutting out a large number of tiny stars to glue to the cape and the hat, and you still need to glue them on. On the day of Halloween, you decide that the outfit looks better without all these stars you have worked so hard on.
[Wear stars; Go without.]

[4.] After a large meal at a restaurant, you order a big dessert with chocolate and ice cream. After a few bites you find you are full and you would rather not eat any more of it.
[Eat more; stop eating.]

[5.] You are staying in a hotel room, and you have just paid $6.95 to watch a non-refundable movie on pay TV. You then discover that there is a movie you would much rather see on one of the free cable TV channels. You only have time to watch one of the two movies.
[Watch free cable; Watch paid-for movie.]

[6.] You have been asked to give a toast at your friend's wedding. You have worked for hours on this one story about you and your friend taking drivers' education, but you still have some work to do on it. Then you realize that you could finish writing the speech faster if you start over and tell the funnier story about the dance lessons you took together.
[Finish the toast about driving; Rewrite the toast about dancing.]

[7.] You decide to learn to play a musical instrument. After you buy an expensive cello, you find you are no longer interested. Your neighbor is moving and you are excited that she is leaving you her old guitar, for free. You'd like to learn how to play it.
[Practice the cello; Practice the guitar.]

[8.] You and your friend are at a movie theater together. Both of you are getting bored with the storyline. You'd hate to waste the money spent on the ticket, but you both feel that you would have a better time at the coffee shop next door. You could sneak out without other people noticing.
[Finish the movie; Leave for the coffee shop.]

[9.] You and your friend have driven halfway to a resort. You both feel sick and think that you would have a much better weekend at home. Your friend says it is “too bad” you already drove halfway, because you both would much rather spend the time at home. You agree.
[Turn back; Drive on.]

[10.] You are painting your bedroom with a sponge pattern in your favorite color. It takes a long time to do. After you finish two of the four walls, you realize you would have preferred the solid color instead of the sponge pattern. You have enough paint left over to redo the entire room in the solid color. It would take you the same amount of time as finishing the sponge pattern on the two walls you have left.
[Finish the sponge pattern; Redo the room in a solid color.]

[11.] You have invested a good deal of your time into a project and it is failing. You have the option to start on something different that you now know is more likely to be successful but you know you cannot get the time back that you spent on the project.
[Keep going with the project; Start something different.]

[12.] You have an investment strategy that you have developed over several months. It is not working and you are losing money, but there is no way for you to recover the lost effort put into developing the strategy.
[Start afresh; Keep going.]

[13.] Imagine that you have spent $20 on a ticket to a concert. The day of the concert comes but unfortunately it is snowing heavily and you feel tired after a tough day. You know you would not have decided to go to the concert if you hadn't already bought the ticket, but you also know that you cannot get a refund.
[Go to the concert; Stay at home.]

[14.] Your relationship with your partner is not going well. You have reasoned it out and you have realized that if you knew how it would go when you started the relationship you would not have gone through with it. You now have the opportunity to break up, but you have been together for many months.
[Keep going; Break up.]

[15.] You have been thinking about how to vote in an election and have invested a good deal of your time to try and make the right decisions including reading newspapers and comment pieces online and thinking hard about the issues. You discover that much of the information you were using is false and a more trustworthy source suggests your initial view was wrong.
[Keep beliefs; change beliefs.]

[16.] You have been thinking hard about the best route to get to somewhere you haven't been to before. Unfortunately, your internet connection isn't working so you have to base your decision on your beliefs about the town's layout. You come to a conclusion on the best possible route but then suddenly the internet is back online.
[Look up route online; Stick to planned route.]

[17.] You are working on a difficult logic problem. Below the problem is a list of possible answers labelled a to e. Although you are not very confident about your answer you decide to go for option a. A friend you know is usually better at this sort of problem suggests that you should change your answer to option b.
[Answer a; Answer b]

[18.] You have been living in a town where it rains a lot and decide to go and buy a high-quality umbrella that you can carry with you every time you go out. Soon after buying a very expensive umbrella you move to a town where it rains much less often.
[Take umbrella with me; Leave umbrella at home.]

[Fluid Intelligence]

Task 2: Logic puzzles

You will face ten timed multiple choice questions about logic. Each question shows a sequence of nine patterns with one missing. Your task is to select the missing pattern from the drop-down list. There is only one correct answer for each question.

You will have 30 seconds to answer each question. If you are happy with your answer and wish to move on to the next question before the 30 seconds are up, simply press the ">>" button at the bottom of the screen. If the time runs out and you have selected an answer, that answer will be submitted and you will move on automatically to the next question. If the time runs out and you have not selected any answer, that question will be marked as incorrect and you will move on automatically to the next question. Once you have moved on from a question, you cannot go back to it.

Before the task begins, you first have a practice question. This will allow you to get a feel for the format and time limit. It will not count for the bonus payment. The practice question will **begin immediately** on the next page.

○ I understand these instructions

[Crystallized Intelligence and the Cognitive Reflection Test]

Task 3: Knowledge questions

You will now face 23 timed questions testing various aspects of knowledge, with a page after question 12 where you can take a break of up to 1 minute.

You have 30 seconds to answer each question. If you are happy with your answer and wish to move on to the next question before the 30 seconds are up, simply press the ">>" button at the bottom of the screen. If the time runs out and you have selected an answer, that answer will be submitted and you will move on automatically to the next question. If the time runs out and you have not selected any answer, that question will be marked as incorrect and you will move on automatically to the next question. Once you have moved on from a question, you cannot go back to it.

The first question will **begin immediately** on the next page.

○ I understand these instructions

[Openness]

Task 4: Personality questions

For each statement below, please indicate how accurately it describes you. **[Each scale was a 5-point Likert scale with "Strongly disagree", "Moderately disagree", "Neither agree nor disagree", "Moderately agree" and "Strongly agree".]**

[12 questions from the NEO Five Factor Inventory (Costa & McCrae; 1989) followed.]

Final Questions: Demography

What is your sex?

- ○ Male
- ○ Female
- ○ Other
- ○ Prefer not to say

What is your age?

What is your race?

- ○ White
- ○ Black or African American
- ○ Hispanic or Latino
- ○ American Indian or Alaska Native
- ○ Asian American
- ○ Native Hawaiian or Pacific Islander
- ○ Other

What is your household's annual income? (US dollars, before tax)

- ○ 0-9,999
- ○ 10,000 - 19,999

- ◯ 20,000 - 29,999
- ◯ 30,000 - 39,999
- ◯ 40,000 - 49,999
- ◯ 50,000 - 59,999
- ◯ 60,000 - 69,999
- ◯ 70,000 - 79,999
- ◯ 80,000 - 89,999
- ◯ 90,000 - 99,999
- ◯ 100,000 - 124,999
- ◯ 125,000 - 149,999
- ◯ 150,000 +

What is the highest grade of school you have completed, or the highest degree you have received?

- ◯ No schooling (or less than 1 year)
- ◯ Nursery, kindergarten, and elementary (grades 1-8)
- ◯ High school (grades 9-12, no degree)
- ◯ High school graduate (or equivalent)
- ◯ Some college (1-4 years, no degree)
- ◯ Bachelor's degree (BA, BS, AB, etc)
- ◯ Master's degree (MA, MS, MENG, MSW, etc)
- ◯ Professional school degree (MD, DDC, JD, etc)
- ◯ Doctorate degree (PhD, EdD, etc)

Generally speaking, which point on this scale best describes your political affiliation?
[A slider was presented with range [0,100] with "Entirely Liberal" over 0 and "Entirely Conservative" over 100.]