Know When to Fold 'Em: The Grit Factor*

Larbi Alaoui† Christian Fons-Rosen‡

May 1, 2016

Abstract

This paper investigates the way different sides of grit influence behavior. We show that, in addition to grit's well-known upside in achieving economic success, it has a potential downside too. Specifically, we conduct an experiment using a game of luck and elicit each individual's intended plan of action and compare it to actual choice. We find that grittier individuals have a higher tendency to overplay. We then split grit into two new categories, tenacity and diligence, and obtain that tenacity alone captures the difficulty in respecting ex-ante preferences when this means accepting defeat. Both components of grit correlate with lower self-reported procrastination problems and higher self-esteem, with diligence being the more beneficial trait. Overall, the results indicate that diligence has a clear upside while tenacity has both the upside of not giving up and the downside of not letting go.

Keywords: noncognitive skills; grit; tenacity; diligence

JEL codes: C91; D03; I20

*Alaoui and Fons-Rosen gratefully acknowledge financial support from the Spanish Ministry of Economy and Competitiveness under projects ECO2014-56154-P and ECO2014-55555-P as well as the Barcelona GSE grants. We thank Sule Alan, Jose Apesteguia, Ghazala Azmat, Roy Baumeister, Seda Ertac, Rosa Ferrer, David Laibson, Pablo Lopez-Aguilar, Christian Michel, Helena Perrone, Charles Sprenger, Adam Szeidl, Ana Tur-Prats and Kathleen Vohs for their thoughtful comments.

†Universitat Pompeu Fabra and Barcelona GSE. E-mail: larbi.alaoui@upf.edu

‡Universitat Pompeu Fabra and Barcelona GSE. E-mail: christian.fons-rosen@upf.edu

1 Introduction

The influence of personality traits and noncognitive skills on success has been progressively recognized in recent years (Almlund, Duckworth, Heckman and Kautz (2011)). Among these traits, grit, which is connected to perseverance, has been singled out in the psychology literature and increasingly in economics. Studies consistently show that it is a critical trait in education (see, e.g., Heckman and Rubinstein (2001), Duckworth, Peterson, Matthews and Kelly (2007), Dobbie and Fryer (2015) and Burks et al. (2015)), and that it even outperforms IQ in determining lifetime success (Duckworth and Seligman (2005), Roberts et al. (2007), Kautz et al. (2014)). The upside of grit is clear upon introspection, since perseverance, determination, dedication and resilience are all positive words that are nearly universally accepted as crucial to achieving success. At the same time, related terms like stubbornness, obstinacy and bullheadedness all have a decidedly negative connotation. The same traits which are so critical for success also seem, when cast under a different light, linked to the idea of not knowing when to let go. Little is known about the repercussions of this second aspect, and about whether the effects of the two components, the positive and the potentially negative, can be disentangled using standard measures of grit. Our aim in this paper is to investigate this subject.

The first task is to define the meaning of not letting go at the right moment, as it is an inherently subjective idea. Within economics, a widely accepted viewpoint for evaluating decision-making consists of using the individual's own ex-ante preferences, or plan of action, as the metric. Under this view, the determining factor is internal consistency, which compares the ex-ante plan of action to the actual behavior exhibited ex-post. But making this comparison using traditional datasets poses a challenge, because the plan of action is typically not observed. Suppose, by way of example, that when playing a game of luck (or investing in the stock market), an individual loses a large amount but refuses to stop. This may be fully consistent with an ex-ante preference of pursuing a risky strategy, but it is also consistent with him deviating from his plan and "overplaying," finding it more difficult to accept a loss than he would like. Without more information on his plan of action, we cannot rule in favor of either narrative. For this reason, we design an experiment that separates preferences from overplaying and relates them to the different facets of grit.

Our experiment elicits both the ex-ante plan of action and actual behavior using a simple single-agent game of pure chance. This game has no ability component and no scope for overconfidence, and it induces well-defined temptations to go beyond the established plan. In particular, subjects start with a fixed amount of money, which they can win or lose. They can effectively keep playing for as much or as little as they wish. Their elicited plan of action determines the maximum loss that they are willing to incur before stopping the game. Going beyond the plan, therefore, means incurring the risk of a bigger loss. We measure grit using the Duckworth (2009) questionnaire and ask a set of self-reported

questions on procrastination, temptation and the locus of control. We run the experiment over eight sessions, on 138 total subjects.

The results on the upside of grit all support known findings – for instance, subjects with a higher Grit Index report lower temptation problems, lower procrastination problems and higher self-esteem. Grittier subjects also have an ex-ante preference for playing less. But when we compare actual choice to the plan of action, we find that grittier subjects overplay the *most*, even after controlling for preferences. This result is not driven by a few subjects.[1]

This last finding is indicative of the downside of grit. Grittier agents persevere and are not easily discouraged, which is a clear prerequisite for success, but we hypothesize that their 'stubbornness' also makes it difficult for them to let go and accept failure, thereby leading them to go beyond their plan of action. Said differently, it is harder for a grittier person to quit when he wants to.

To further explore our hypothesis, we split the grit questions into two new categories: 'diligence' and 'tenacity.' Diligence captures the notion of being hard-working, which appears to be an unambiguously positive skill. Tenacity, instead, is more nuanced. We conjecture that it leads to overplaying, because it includes both the positive side, perseverance, and the negative side, stubbornness. Consistent with our hypothesis, only tenacity drives overplaying in the regressions. However, both diligence and to a lesser degree tenacity are predictive of lower self-reported procrastination and temptation problems, as well as higher self-esteem. This is consistent with the idea that while diligence is unambiguously desirable, tenacity instead has both a downside and an upside. We perform robustness checks on these categories by considering variations of the partition. Our results remain unchanged, thereby ensuring that any potential subjectivity in the split is not an issue.

Since the diligence and tenacity categories are drawn directly from the standard grit questions, they are available as regressors in all datasets in which this questionnaire has been administered. As an external validity test of their predictive power with conventional performance measures, we analyze an existing dataset. The sample contains 4199 people, and it includes the grit measure as well as educational outcomes and other demographics. We find that both measures are positive predictors of educational level, and, interestingly, that diligence outperforms tenacity. However, we note that this data does not allow us to identify a notion of not letting go at the right moment, unlike the data obtained from our experiment. The experiment also reveals that grit, which is typically linked with long term effects, is predictive of behavior in a short-term, controlled environment too.

Looking ahead, the finding that grittier individuals have a higher tendency not to let go when they should, even according to their own preferences, raises policy implications and questions for future research. Our results indicate that it is specifically tenacity and

[1] In fact, we find that at least 30% of subjects overplay, irrespective of characteristics such as undergraduate degree and gender. This illustrates that our design can be useful for future experiments which require a domain in which temptation is easily elicited and measured. As discussed later on, this domain of temptation is relatively unexplored in the literature.

not diligence that should be investigated for its potential downside. This aspect may be particularly relevant for high achievers, precisely because grit is strongly connected to success. For instance, it is well-known that investors often have difficulty in accepting their losses (e.g. Odean (1998)), and numerous firm managers and entrepreneurs maintain their strategy despite all indications that they should readjust. To the extent that these investors, managers and entrepreneurs have high grit, they themselves may wish to counter its flip side. Future research can focus on the interplay between grit's different facets as a function of the environment.

The rest of the paper is structured as follows. We first discuss the related literature. Section 2 then presents the experiment, Section 3 provides the main findings, and Section 4 discusses the external validity of our analysis. Section 5 concludes.

Related literature

This paper relates to several strands of research. The first concerns the relevance of noncognitive skills, and particularly grit, to economic outcomes, such as education and employment. The psychology literature is more extensive while the economics literature is small but growing. Here we do not provide an exhaustive survey, but see Borghans, Duckworth, Heckman and Weel (2008) and Almlund et al. (2011) for detailed discussions. Grit is especially important for achieving long-term goals, even in the absence of positive feedback (Duckworth et al. (2007)). Heckman and Rubinstein (2001) explain the importance of these factors in skill formation and success, and highlight the gap in the economics literature in analyzing these crucial traits. Burks et al. (2015) find that within conscientiousness (a "big five" personality trait, which also include neuroticism, openness to experience, extraversion and agreeableness), the 'hard work and persistence' component predicts collegiate success. Dobbie and Fryer (2015) consider the impact of charter schools on outcomes such as academic achievement, and among their measures, administer the short-scale Duckworth Grit Index used in this paper. Cubel, Nuevo-Chiquero, Sanchez-Pages and Vidal-Fernandez (2016) conduct a laboratory experiment on the effect of big five personality traits on performance, and find that conscientious subjects perform better.

Considering the importance of noncognitive skills for success, a natural policy issue concerns the degree to which these skills can be influenced. Cunha, Heckman and Schennach (2010) analyze the production function of both cognitive and noncognitive skills, in a dynamic setting in which skills are determined through parental investment, and estimate optimal targeting of interventions. Recent papers use large-scale policy interventions in elementary schools to investigate the malleability of noncognitive skills. Alan and Ertac (2014) find that groups treated towards being more patient have improved outcomes, even a year after the intervention. Their findings strongly suggest that noncognitive skills can be influenced in a persistent manner. Alan, Boneva and Ertac (2016) focus more explicitly on a notion of grit, and also find that treated subjects are more likely to obtain higher

grades.

We view this paper as complementary to the ones above. By revealing another side of grit, that of not letting go, our paper shows that grit is not only a prerequisite for achievement, it can also shed light on other patterns of behavior. In particular, it can help explain the difficulty in respecting one's own plan of action and accepting a loss. Our paper also shows that the relevance of grit can be elicited in a laboratory environment and in a short timespan. It further provides a way of splitting the existing grit scale into new categories of diligence and tenacity. This division can be used to analyze existing data which uses the grit questionnaire, and it can further serve in considering future policy implications.

A second strand of the literature concerns dynamic inconsistency in decision making, which occupies a central part in behavioral economics. The most commonly used modeling approach is to use present-biased preferences. The reach of this approach is too vast to document in this paper, but see, among other contributions, Strotz (1956), Laibson (1997), O'Donoghue and Rabin (1999); see also O'Donoghue and Rabin (2015) and Sprenger (2015) for discussion of the literature; for other modeling approaches, see, e.g., Gul and Pesendorfer (2001), and see Ameriks, Caplin, Leahy and Tyler (2007) for a survey-based measure of self-control problems. An important question within this literature concerns separating ex-ante preferences from behavior and measuring the extent of present-bias. Doing so is a challenging task that requires carefully conducted experiments, and the same subjects are followed over time. A recent contribution by Augenblick, Niederle and Sprenger (2015) resolves the main difficulties of this exercise in a laboratory experiment, over a span of several weeks. For papers that focus on commitment, see, for instance, Ashraf, Karlan and Yin (2006) on commitment devices for savings products, Kaur, Kremer and Mullainathan (2015) for a field experiment on self-control and commitment devices in the workplace, and Alan and Ertac (2015) for a field experiment that analyzes patience, self-control and the demand for commitment among school children. Our paper does not focus on present-bias preferences, but the connection is in the dynamic consistency problem itself. Grit is sometimes linked to – though distinct from – self-control ability and self-discipline (see Duckworth and Gross (2014)).

Our setting also has a self-control element, but the temptation is very different in nature. In this sense, our paper brings to light a relatively unexplored domain of temptation, namely that of not letting go and accepting failure. This form of temptation may relate to noncognitive skills in a way that is distinct from conventional domains. This distinction is clearly revealed by our results. Grittier subjects report less self-control problems, in line with known results, but within our game they are also more likely to deviate from their own plan of action. Another contribution of our paper is methodological. Around 35% of subjects overplay, illustrating that our design can be useful for experiments that require a domain in which temptation is easily elicited in a short period of time.

A separate and extensive literature analyzes the effects of overconfidence (see, for

instance, Malmendier and Tate (2005), among many contributions). Our experiment has been designed to avoid overconfidence effects, in that there is no ability component on the subject's part, and the probabilities of failure are known and objective. However, an interesting direction for future research may involve exploring the relation between the grit and tenacity of CEOs and their overconfidence. For instance, Galasso and Simcoe (2011) find that more overconfident CEOs, i.e., those who underestimate the probability of failure, are more likely to pursue innovation. An open question concerns measuring CEO grit, comparing their ex-ante plan to their ex-post choices, and analyzing the tradeoff between the upside and the downside of tenacity. Overconfidence is also linked to self-confidence and self-image (see, for instance, Benabou and Tirole (2002), Compte and Postlewaite (2004), Brunnermeier and Parker (2005) and Mobius, Niederle and Niehaus (2014)). Given this link, another question concerns whether grittier subjects also have more self-worth attached to their successes and failures, which in turn may correlate with difficulty in letting go.

Lastly, there is a large literature on the disposition effect, which refers to investors' tendency to keep assets whose value have dropped (see, e.g., Shefrin and Statman (1985), Odean (1998), and for a discussion see Barberis and Thaler (2003)). This effect, and the related idea of loss aversion, may be present in the way subjects approach losses in our experiment. In light of our results, another avenue of research would explore whether more tenacious subjects also have a higher tendency to hold losing assets than they had planned on ex-ante.

2 Experimental Design

This experiment was conducted at the Behavioral Sciences Laboratory (BESLab) of Pompeu Fabra University, and included 138 subjects. There were eight sessions spread over several weeks, and typically lasted between 75 and 90 minutes, with 15 to 20 subjects per session. Each session is divided into three main stages: the first consists of eliciting the subjects' plan of action for the (single-agent) games in mind, the second consists of subjects actually playing the games, and the third consists of a questionnaire. We explain each stage below, after describing the main game that subjects can play.

2.1 Description of the game

The relevant game for our analysis is a simplification of the familiar (American) roulette. We modified a standard online roulette by removing some betting options (e.g., betting on evens and odds and on first, second or third 12th). For screenshots of the roulette, see Figure 1, and for the actual game used see http://experimentalgames.upf.edu/roulette/.

Each subject has 2000 tokens to start with, where 150 tokens is worth 1 euro. He or she (henceforth he) can place a bet on any of the 38 numbers (0, 00, and 1 to 36) or colors (red or black), or on nothing at all. He can bet anywhere between 0 and 500 tokens per spin,

in any increment of 50. Once he finishes placing his bets, the roulette wheel spins, and the ball lands on a number and its associated color. We use the standard roulette betting rules: if the ball lands on a number on which the subject has placed a bet, he receives the original amount wagered on the number plus 35 tokens for each token wagered. If it lands on a color on which he has bet, then he receives the original amount plus 1 token for each token wagered. All other tokens bet are lost, and the game repeats with the updated amount of tokens. Subjects can play for as long or as little as they wish, and they do not have to play if they do not want to. They can also spin the wheel even when their bet is 0. Subjects can see the tokens they have, the amount bet, and the history of the last five roulette outcomes (color and number). They cannot earn less than 0 tokens or more than 8000. Notice that the expected returns are negative for any strategy that involves betting a positive amount.

We use this game because it has several advantages for our objectives. First, this game is familiar to many, so subjects do not incur a large cognitive cost to understand it. This notion was confirmed during the experiment, in which almost no subject displayed problems understanding our roulette. Second, there is no ability component. This is a game of pure luck, and furthermore each draw is independent of the others. This feature reduces any potential confounds from overconfidence in one's own ability, which may otherwise form before or during the game.[2] Third, this game quickly draws the subjects in, absorbs their attention, and makes them value the outcome. While it has a component of entertainment, it simulates an environment in which losing and winning are important to the subjects and in which deviating from a plan of action may prove costly to them. The amount they receive is itself meaningful, in that their potential earnings (8000 tokens, or 53 euros) are considered substantial in Spain, as is the magnitude of the loss from their initial wealth to 0. This can easily be observed during the experiment from the subjects' reactions to their outcomes. Fourth, the nature of the game is such that the agents, when incurring a loss and reaching the limit they set for themselves (ex-ante), are faced with a clear tradeoff between following their plan of action and stopping, or continuing to play at the risk of losing more. Lastly, even if subjects play less or more in an experimental setting than they would otherwise, this would not introduce any bias in favor of our findings. That is, it should not lead grittier individuals to have a higher tendency to deviate from their plan of action.

2.2 Plan of action

The first stage of the game consists of eliciting the subjects' plan of action.[3] We inform them of the game rules and of their initial endowment of 2000 tokens. We then ask each subject for the range of tokens within which he would like to play. Specifically, we ask

[2]The data shows that subjects do not use gambler's fallacy strategies, so this is not an issue. Moreover, such behavioral biases would not correlate with grit or other critical factors for our analysis.

[3]The experiment was programmed in Qualtrics, aside from the roulette itself which was programmed in Adobe Flash.

for the minimum limit which he is not willing to surpass. (For completeness we also ask for the maximum of the range, but it is not the object of our analysis.) This limit is the subject's intended bound, in that his plan is to stop playing if it is ever reached or even before. He can choose any number in increments of 100 between 2000 tokens and 0 tokens. That is, he can choose within the range of not playing at all and risking his entire wealth. We refer to the elicited lower bound as the "planned minimum bound," or simply planned minimum.

We believe the subjects to be truthful about their plan of action. They have no incentives to be dishonest, and lying is well-known to be psychologically costly (see, for instance, Gneezy (2005)). Moreover, any noise in the responses should not correlate with grit, and would therefore not bias our results. As an additional precaution, however, we check for the reliability of the stated plan of action in the following way. Before introducing the roulette game and asking for the plan of action, we ask the subjects for their plan in another game.[4] We do not expect to observe a high degree of overplaying in this first game, as it is designed to be much less captivating and tempting. Comparing the stated plan of action to actual play in this game then provides a measure of how accurate elicited plans are.

The rules of the first game are the following. Subjects are informed that they will start with 2000 tokens, just as they will later be informed of the same initial endowment for the roulette. At each round of the game, he chooses a number between 1 and 38. The computer chooses a number at random in that range as well, and if they match then he receives 3200 tokens and the game stops automatically. Otherwise, he loses 100 tokens. Here too, the subject can play for as many rounds or as few as he wishes, unless he reaches 0 tokens or unless he wins. We ask for the subject's plan of action, and specifically for the number of rounds that he would like to play, provided he has not won before. Since the game ends if the subject wins, his maximum earnings are therefore 5200 tokens (35 euros), which he would receive if he wins in the first round. Clearly, this game is much more passive and tedious than the roulette, as the only choices available here are whether or not to stop at any stage and to pick one of the 36 numbers. Furthermore, since the game ends as soon as the subject wins once, there can be no excitement or momentum from playing after a win, as there would be with the roulette.

When comparing the stated plan of action to the behavior ex-post for this first game, we see that nearly 84% of subjects play in accordance with their stated plan. This suggests that subjects are indeed truthful in their response and wish to respect their established plan of actions. Moreover, we do not observe any correlation between grit and overplaying in this game, which indicates that there is no systematic bias that may drive our results for the main game. These results demonstrate that the stated plan of action is indeed a reliable measure of their true ex-ante preferences.

[4]We also use an incentive compatible mechanism to elicit the plans of action, although we defer this discussion to the Appendix, as this mechanism is more involved.

Each subject's final payment is based on the earnings of either the first game or the roulette, chosen at random. We inform the subjects beforehand, explaining that at the end of the experiment a die will be tossed to determine the earnings. If it lands between 1 and 3 then one game will determine their earnings, otherwise the other will. We do so to avoid any wealth effects, as may occur if the final earnings were based on the sum of the earnings in the two games. This ensures that the games are independent and that earnings in one have no financial impact on earnings in the other. Using the physical die serves as a guarantee to the subjects that the decision is truly random and that the likelihood of either game being chosen is the same.

After eliciting the subjects' plan of action in the first game and in the main (roulette) game, we ask additional questions with the aim of de-framing them before they play the actual games. These questions are designed not to impose a heavy cognitive burden or induce fatigue. They consist, for instance, of filling in a caption for a drawing and of checking boxes on the kinds of news that they follow. The subjects are informed that these questions will not impact their earnings in any way. They are free to answer them as they wish, or even randomly if they are inclined to do so, without any consequences.

2.3 Game play

After subjects provide their plan of action and answer the questions that follow, they play the two games. As discussed above, they are free to play as much or as little of each game as they wish. Hence, the length of the experiment depends on the subject's own choice– there are no time constraints of any kind, and each subject's game is private and independent of everyone else's. Each subject plays the games in the same order as they were presented in the 'plan of action' stage: he starts with the first game presented (the simpler 1 to 38 game), and then plays the main one (the roulette).

2.4 Questionnaire

The subjects are asked to complete a 5 minute questionnaire in the last stage for an additional 600 tokens. This survey includes the standard 8-question short grit measure (see Duckworth and Quinn (2009)). These questions are used to construct the Grit Index, which ranges from 1 to 5, where grit is increasing in the number assigned. We also ask questions from a shortened Locus of Control (Rotter (1954)), and construct a Locus Index that will be used as a control variable. The Locus Index ranges between 0 and 1, where a higher score is associated with a higher belief that external factors determine events and outcomes. Subjects are then asked additional questions, particularly on their self-assessed degree of temptation and procrastination problems, as well as self-esteem. We specifically use the single-item measure of self-esteem that consists of the question "I have high self-esteem" which ranges from 1 to 5, where 5 indicates higher self-esteem. This measure is viewed as being highly correlated with the Rosenberg Self-Esteem Scale in adult samples

(Robins, Hendin and Trzesniewski (2001)). Lastly, subjects were asked to indicate their age, gender and field of study.

3 Results

3.1 Descriptive statistics

Since the roulette is a game of chance, a lucky subject may never be close to his minimum limit, and so may never be confronted with the temptation to overplay. Hence, we only capture a lower bound on overplaying. Nevertheless, we still find that 48 of our total of 138 subjects overplay (35%). The high figure illustrates that our roulette is an effective tool for eliciting temptation preferences in a short span of time, in a laboratory setting, and in an easily measured way. Moreover, overplaying is neither gender-specific nor degree-specific; 31% of women and 40% of men overplay, and 34% of those pursuing technical degrees and 36% of those pursuing non-technical degrees overplay. Of the 48 subjects who overplay, 20 of them lose their entire endowment, ending up with 0 tokens.

Table 1 provides further descriptive statistics. The mean for the Grit Index is 3.38, which is in line with other studies. By way of comparison, the mean in Duckworth and Quinn (2009) is 3.4. An average subject plans to stop playing when left with about 900 tokens. In addition, 60% are female and the average age is close to 22.

3.2 Overplaying and grit

Overplaying is defined as a dummy whose value is 1 if the subject goes below his planned minimum bound, and it is 0 otherwise.[5] For instance, if the subject's plan is to stop at 1000 tokens but he goes below that amount at any stage of the game, then this dummy will take a value of 1. Our main explanatory variable of interest is the Grit Index.

Table 2 presents the first main empirical finding of the paper. In column (1), we use a standard OLS econometric specification to regress the dummy for overplaying on the Grit Index. The estimated coefficient is positive, statistically significant at the 1% level, and economically relevant: a one unit increase in the index is associated with a 25.9% increase in the mean predicted probability of overplaying. As individuals with different grit levels also tend to have differing elicited preferences on when to stop playing, our estimated coefficient might be biased. A preference for playing less, which is defined as having a higher planned minimum bound, may lead to a higher likelihood of overplaying, since this bound may be reached with fewer spins. For this reason, we control for preferences (plan of action) in column (2), and find that its estimated coefficient is positive, as expected. Furthermore, the estimated coefficient on the Grit Index is practically unchanged. In columns (3)-(5) we gradually augment the specification with variables controlling for the log of age, a dummy for gender, and a dummy for whether they studied a technical

[5]Recall that we use the terms 'plan of action' and 'planned minimum bound' interchangeably to refer to the lower limit which the subject does not plan on exceeding.

undergraduate degree.[6] The estimated coefficient on the Grit Index remains stable and strongly statistically significant.

In Table 3 we gradually drop subjects with the largest gains to ensure that they do not drive the results. In column (1), we drop the 5 individuals who ended up in the range of 6000-8000 tokens. In column (2), we drop the 40 additional individuals with a final outcome between 2000 and 6000 token. Finally, in column (3), we only keep the 40 subjects with the worst final outcomes (below 1000 tokens). Remarkably, the estimated coefficient remains at approximately 0.25 and is statistically significant at conventional levels throughout these specifications.

We then verify in Table 4 that results are not driven by subjects with little interest in becoming involved in the game. Such players have a high planned minimum bound. For instance, if the planned minimum is 1900 tokens, then the subject does not wish to go beyond 100 tokens below his original wealth of 2000 tokens. In column (1) we drop the 9 subjects whose minimum plan of action was exactly 2000 tokens, the highest value we allowed for. In column (2), we additionally drop the 20 individuals with minimum plan of action in the range 1500–2000, while in column (3) we only keep the 70 subjects with the greatest appetite for risk (minimum plan of action below 1000 tokens). The estimated coefficient on the Grit Index remains strongly statistically significant with values around 0.25 in all three specifications.

Another robustness check consists of dropping subjects who overplayed by a small amount. If our results were solely driven by them, then our claim that grittier people have a higher tendency to overplay would be weakened. In Table 5 we drop the 10 subjects who overplayed by 100 tokens or less.[7] As before, our main coefficient of interest is unaffected. This is true when we perform the simplest regression (column (1)), when controlling for the plan of action (column (2)), and also when we further account for age, gender, and type of degree (column (3)).

As an additional robustness test, we consider alternative econometric specifications in Table 6. In columns (1)-(2) we run a Logit specification. The estimated coefficient on the Grit Index is again statistically significant at the 1% level with the following economic significance: the mean predicted probability of overplaying increases by 17% for a one standard deviation increase in grit, averaging across the sample values of the other regressors (plan of action, age, gender, and degree). The magnitude and statistical significance of the effect is similar when we rather use a Probit (columns (3)-(4)) or a Poisson (columns (5)-(6)) specification.

[6] We define technical degrees as: Economics, Business Administration, Engineering, and Biology. The non-technical degrees are: Political Science, Law, Criminology, Humanities, Sociology, Marketing, Language Translation, Audiovisual Arts, Design, Tourism, Journalism, and Sociology.

[7] Here we do not explore further how the amount by which subjects overplay relates to grit. There is a difficulty in defining a natural benchmark of comparison for degree of overplaying because the amount is highly sensitive to the subject's draw, for instance if the subject's aim is to return to his inital wealth. Nonetheless, subjects who overplay and lose all their wealth have a higher average grit than the others, and grittier subjects seem to have a higher tendency to attempt to regain their initial wealth.

In Table 7, we add the Locus of Control Index as a control variable in the OLS regressions. In column (1), we regress the dummy for overplaying on the Locus Index alone and do not find that the estimated coefficient is statistically significant. Its negative sign may loosely suggest that those who believe more in external control are less likely to try their luck. We then include the Grit Index in addition to the Locus Index (column (2)). Only the former can help to explain overplaying behavior, and its coefficient maintains the same magnitude and level of significance. This is still the case after controlling for the plan of action (column (3)) and even after accounting for the other standard controls we have used so far (column (4)).

Overplaying is a form of dynamic inconsistency, in that there is a difference between ex-ante plan of action and actual behavior. The question then arises as to whether these results suggest that grittier agents are more present-biased or impatient. But since our experiment takes place over a short span of time, temporal preferences should not play a role under any feasible parameters. Moreover, results on temporal preferences and noncognitive skills do not suggest that grittier subjects would be more present-biased (see Alan and Ertac (2015) for a discussion of noncognitive skills and patience).[8] We also do not find any difference in amount of time played, as would have been the case if receiving their final earnings earlier affected subjects in a way that varied with grit. This is in line with the intuition that the form of dynamic inconsistency captured in our environment, which links to difficulty in accepting defeat, is conceptually different from that due to present-biased preferences– in fact, it is possible that they go in opposite directions. As an additional robustness check, however, we use the self-reported measures on self-control problems (procrastination and temptation) as controls in our regression. Specifically, Table 8 follows the structure of Table 7, but using control variables for (self-reported) temptation and procrastination problems instead of the Locus of Control. A similar picture emerges: on their own, neither of the two estimated coefficients is statistically significant (column (1)). When we add the Grit Index in columns (2)-(4) it is always positive and statistically significant, while neither the temptation nor procrastination variables are relevant.

Lastly, in Table 9 we split the sample by technical and non-technical degree in columns (1) and (2) and by gender in columns (3) and (4), while maintaining the remaining usual controls. Remarkably, the estimated coefficient on the Grit Index is positive and statistically significant at least at the 10% level in all four specifications, with coefficient values ranging between 0.2 and 0.3.

To give visual intuition for our findings, we graphically compare the kernel density estimation of grit for subjects who overplayed and those who did not (Figure 2). This comparison reveals a striking difference between the two, in that the distribution of those who overplayed is markedly shifted to the right. In fact, the cumulative distribution of those who overplayed effectively first order stochastically dominates the distribution of

[8]This is further confirmed by the subjects' behavior in the first game (the simpler 1 to 38 game) in which grit does not predict any detectable difference in behavior, which would have been the case if deep temporal parameters had been significantly different.

those who have not (Figure 3).

In summary, the finding that subjects with higher grit are most likely to overplay is highly statistically and economically significant, and it is not driven by outliers. Note that this result would have been missed had we not separated ex-ante preferences from actual behavior. Not only would it be difficult to define overplaying (from the subject's perspective), but grittier subjects also tend to prefer playing less. Hence, the net effect on behavior alone is ambiguous.[9] Our experimental design, therefore, serves to emphasize the importance of separating plan of action from behavior to address our questions of interest.

3.3 Overplaying, tenacity and diligence

Our explanation for the main finding is that grittier subjects, by being more tenacious and stubborn, find it difficult to let go and accept failure, even if it means deviating from their initial plan of action. To test this hypothesis, we partition the grit questions into two new categories, which we refer to as 'diligence' and 'tenacity'. By way of illustration, consider the image of a person with a rowboat that he can both row and steer. Together, these two factors determine grit, and separately they determine diligence and tenacity. A more diligent person rows harder, and a more tenacious one is more reluctant to steer away from his trajectory, regardless of the setbacks he encounters along the way. These two traits are correlated but distinct.[10] The unwillingness to accept defeat and stop playing once the planned minimum bound is reached is a refusal to steer away. Hence, it falls squarely within tenacity, not diligence. Our prediction, then, is that tenacity alone should explain overplaying.

The tenacity category consists of all questions that specifically ask about not letting go, while the diligence category consists of the remaining questions, which ask about being hard-working. For instance, the question "Setbacks don't discourage me" clearly falls within tenacity, while "I am diligent" obviously falls within diligence (see Table 10). The grit questions are cleanly and unambiguously separated into these dimensions. As with the Grit Index, the Tenacity Index and the Diligence Index can take values between 1 and 5, where 5 indicates higher tenacity and higher diligence, respectively.

Table 11 presents the results for these indices. Column (1) regresses overplaying on the Tenacity and Diligence indices, and column (2) controls for plan of action, age, gender and degree (cf. columns (1) and (5) in Table 2). Columns (3) and (4) provide sample splits by degree and columns (5) and (6) by gender, analogously to Table 9. Column (7) checks for robustness to overplaying by more than 100 tokens, analogously to column (3) of Table 5. We find that the coefficient for the Tenacity Index is highly significant for all 7 columns, and its magnitude tends to be larger than the one obtained for the Grit Index in previous tables. The Diligence Index, however, is never significant, and its estimated coefficient

[9] For instance, grit on its own does not correlate with final earnings.

[10] The correlation coefficient is 0.56.

remains close to zero for all columns. These findings jointly support the hypothesis that it is the tenacity component of grit that drives overplaying.[11]

Graphically, the kernel density estimation for tenacity of those who overplayed and of those who have not shows a pronounced shift of the distribution (Figure 4), while the analogous kernel density estimation for diligence shows very little difference (Figure 5). This provides additional visual assurance that the strong positive relationship between grit and overplaying is through tenacity.

Our categorization is distinct from the standard one of 'perseverance' and 'consistency of interest,' both in its interpretation and in the actual partition of the grit questionnaire. Returning to the rowboat metaphor, a more perseverant person rows harder and does not relent when faced with obstacles, and one who has consistency of interest does not change his mind often on his final destination. Neither perseverance nor consistency of interest, therefore, perfectly maps either to our diligence category (being hard-working) or to our tenacity category (refusing to steer when faced with obstacles). It is also not obvious, from this image, whether either of the two is a natural candidate to explain overplaying. The difference in what these categories aim to capture is naturally reflected in the distinct partitioning; for instance, "Setbacks don't discourage me" and "I am diligent," which are grouped separately in our categorization, are instead grouped together under 'perseverance' in the standard classification.

To compare the explanatory power of our categorization to the standard one, Table 12 replicates the same regressions as Table 11, but using the perseverance and consistency of interest split instead of our tenacity and diligence split. The coefficient estimates are clearly less stable. Furthermore, in some columns, perseverance is significant but consistency is not (e.g. column (3)), in some it is the reverse (e.g. column (5)), and in others neither is significant (e.g. columns (6) and (7)).

We consider alternative specifications of tenacity and diligence to ensure that our categorization is robust and that our own subjectivity is not a factor. In particular, we moved and removed questions whose categorization carries any hint of ambiguity. Our results remain unchanged to these alternative specifications. We do not discuss them all for the sake of brevity, but we provide here the most striking such alternative specification. This consists of removing all but the two clearest questions that measure tenacity ("Setback don't discourage me" and "I finish whatever I begin") and the two that measure diligence ("I am diligent" and "I am a hard worker"). Running all previous regressions on these indices leaves the significance levels unchanged and the coefficient in the same range. Moreover, these four questions, which are split equally into two separate categories in our alternate reduced specification, are all grouped together in the standard classification under 'perseverance.' This further illustrates the different rationale behind classifying grit into tenacity and diligence compared to consistency of interest and perseverance.

[11]The results for tenacity and diligence are robust to all the specifications discussed in the previous subsection, although we do not present them all here for the sake of brevity.

3.4 Additional outcomes: procrastination, temptation and self-esteem

We close this section by exploring the link between grit (and its decompositions into tenacity and diligence) and self-reported procrastination, temptation and self-esteem. While the regressions discussed here are not the main object of our analysis, they serve to confirm that there is an upside of grit present even in this setting. In particular, higher grit is typically associated with reporting lower procrastination and temptation problems and higher self-esteem. The regressions serve the added purpose of checking whether our new categories of tenacity and diligence, when taken separately, predict less procrastination and temptation problems and higher self-esteem as well.[12]

In Table 13 we regress procrastination and temptation on grit (columns (1) and (5), respectively), and find that both estimated coefficients are highly significant and negative, in line with our expectations. When regressing these outcome variables on the split of tenacity (columns (2) and (6), respectively) and diligence (columns (3) and (7)) separately, we still obtain that the coefficients are highly significant and negative for both variables. Interestingly, when we regress procrastination on both tenacity and diligence together (columns (4) and (8)), we find that only the coefficient of diligence is highly statistically significant, which again confirms the upside of diligence. The coefficient for tenacity is also negative, but not significant.

We follow the same sequence in Table 14, and regress self-esteem on grit (column (1)), tenacity (column (2)), diligence (column (3)), and both tenacity and diligence (column (4)). The estimated coefficient is positive for all specifications, and it is highly significant for grit, tenacity and diligence separately. When the regression includes both tenacity and diligence, we find instead that only the coefficient for diligence is significant. This exactly mirrors the pattern from Table 13, with the sign reversed, which confirms our expectations that these traits are linked with higher self-esteem. As a robustness check, for all the tables above we consider once more alternative specifications to the tenacity and diligence split, and the results remain essentially unaltered.

Notice that the different facets of grit explain different outcomes. In the regressions that include both tenacity and diligence, the following pattern emerges. Tenacity alone explains overplaying, while diligence has more explanatory power for lower temptation and procrastination and higher self-esteem.

4 External validity: grit components and education

Partitioning grit into tenacity and diligence is natural for our objective of understanding overplaying behavior. But as this is a new way of splitting grit, we consider how well these categories explain conventional outcome variables, and specifically education. To address this point, we use a sample of 4199 individuals which includes data on grit, education as

[12] These questions were asked to 118 of the 138 subjects.

well as other demographics.[13] Respondents are informed before taking the test that the data may be used for research purposes. The survey can be taken online by anyone who chooses to do so and it is anonymous. While all the answers are self-reported, including educational outcomes, we do not expect any bias in our estimation. This dataset is particularly informative because the respondents have a wide heterogeneity in educational level, age, ethnic background and geography.

The mean of the Grit Index for this dataset is 3.25 (3.31 for the US), which is in line with the mean in our experiment (3.38). This similarity also holds when splitting the questions into tenacity and diligence: the mean for diligence (tenacity) is 3.62 (3) in this dataset compared to 3.65 (3.22) in our experiment.

In Table 15, we regress education on the Diligence and Tenacity indices in column (1). We include country fixed effects in column (2) and further control for age, gender, racial background, and urbanization in column (3). The estimated coefficient of the Diligence Index is positive and highly statistically significant for all specifications. It is also economically significant, as a 1 point increase in diligence is associated with a 0.1 to 0.2 point increase in the educational index. It is also noteworthy that adding the control variables increases our R-squared from 0.04 to 0.4, but does not affect the statistical significance of diligence. The coefficient for tenacity is positive and statistically significant in columns (1) and (2), albeit with a lower magnitude. This result confirms the upside of diligence discussed in our analysis, and shows that tenacity, while more nuanced, can also have a positive aspect.

In column (4) we restrict the sample to the 2112 US respondents, and include the same controls in column (5). A similar pattern emerges. Diligence is the main component of grit explaining educational levels, even after a large increase in the R-squared from 0.06 to 0.4 once control variables are added.

5 Conclusions

Our results indicate that grittier subjects have a higher tendency to play past the point at which they would have liked to stop. We have further shown that when grit is split into tenacity and diligence, tenacity alone explains this tendency to overplay. Diligence, instead, explains lower procrastination and temptation problems within our experiment, and higher educational level when applied to an existing survey.

The upside of grit seems beyond dispute, but our analysis reveals that it has an important flip side too. Individuals with higher grit also have more difficulty in stopping and accepting failure, even when they would have liked to. This tendency is contained within the tenacity facet of grit, which itself has both the positive aspect associated with not giving up and the negative aspect associated with not letting go. Diligence, instead,

[13]This data is drawn from the online psychology survey repository available at personality-testing.info/_rawdata (March 17, 2016). This archive of psychological tests has been used in several articles in the psychology literature.

appears unambiguously positive in our setting. Returning to the image of the rowboat introduced in the paper, a propensity for rowing hard, which describes diligence, is clearly necessary for getting to the destination. Resistance to steer away from the current route and onto another, which describes tenacity, intertwines stubbornness with steadfastness.

Our findings raise new questions. They first support that diligence has clear value in achieving success. The upside of tenacity may well outweigh its downside in many contexts, but in others it may prove costly. This tension should be investigated further to shed light on addressing the flip side of tenacity without diminishing the positive side.

Table 1: Descriptive statistics

Note: Overplaying is a dummy variable taking value 1 if the subject played beyond his plan of action. Plan of Action is the planned minimum bound, which ranged between 0 and 2000 tokens. The Grit, Tenacity, and Diligence Indices range from 1 to 5. The Procrastination, Temptation and Self-esteem measures are self-reported measures that also range from 1 to 5.

Variable	Observations	Mean	Std. Dev.	Min	Max
Overplaying	138	.35	.48	0	1
Grit Index	138	3.38	.54	2	4.5
Plan of Action	138	877.54	578.19	0	2000
Tenacity Index	138	3.22	.57	2	4.6
Diligence Index	138	3.65	.68	2	5
Locus Index	138	.47	.17	.06	.88
Age	138	21.72	3.73	18	47
Female	138	.58	.49	0	1
Procrastination	118	3.29	.80	1	5
Temptation	118	3.31	.85	1	5
Self-esteem	138	3.14	.99	1	5

Table 2: Baseline regressions

Note: An Ordinary Least Squares specification is used. The dependent variable is a dummy variable taking the value 1 in case the subject overplayed. Grit Index takes values between 1 and 5 and increases with the level of grit. Plan of Action is the log of the planned minimum bound. Standard errors are robust to heteroskedasticity. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Overplaying	(1)	(2)	(3)	(4)	(5)
Grit Index	0.259***	0.248***	0.237***	0.248***	0.260***
	(0.072)	(0.070)	(0.071)	(0.070)	(0.071)
Plan of Action		0.067***	0.068***	0.072***	0.074***
		(0.013)	(0.013)	(0.014)	(0.014)
ln(Age)			0.236	0.154	0.076
			(0.274)	(0.267)	(0.262)
D(Female)				-0.124	-0.156*
				(0.080)	(0.082)
Technical Degree					-0.095
					(0.080)
Constant	-0.529**	-0.902***	-1.602*	-1.335	-1.081
	(0.239)	(0.234)	(0.844)	(0.826)	(0.811)
Observations	138	138	138	138	138
R-squared	0.087	0.145	0.150	0.165	0.174

Table 3: Robust to level of final gains

Note: This table reports regressions similar to those in column (5) of Table 2. Column (1) excludes individuals with final gains above 6000 tokens. Column (2) additionally excludes individuals with final gains between 2000 and 6000 tokens. Finally, column (3) only keeps individuals with final gains below 1000 tokens. See also notes to Table 2. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)
Dep. Var.: Overplaying	<6000	<2000	<1000
Grit Index	0.264***	0.299***	0.239**
	(0.075)	(0.098)	(0.094)
Plan of Action	0.074***	0.079***	0.140***
	(0.014)	(0.016)	(0.020)
ln(Age)	0.045	-0.091	0.428
	(0.266)	(0.376)	(0.394)
D(Female)	-0.161*	-0.160	-0.133
	(0.083)	(0.100)	(0.106)
Technical Degree	-0.081	-0.011	0.094
	(0.082)	(0.100)	(0.123)
Constant	-1.001	-0.687	-2.169
	(0.829)	(1.150)	(1.294)
Observations	133	93	40
R-squared	0.173	0.199	0.525

Table 4: Robust to different plans of action

Note: This table reports regressions similar to those in column (5) of Table 2. Column (1) excludes individuals with a plan of action of 2000 tokens. Column (2) additionally excludes individuals with plans of action between 1500 and 2000 tokens. Finally, column (3) only keeps individuals with plans of action below 1000 tokens. See also notes to Table 2. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.:	(1)	(2)	(3)
Overplaying	<2000	<1500	<1000
Grit Index	0.240***	0.233***	0.264***
	(0.074)	(0.081)	(0.086)
Plan of Action	0.057***	0.049***	0.039**
	(0.014)	(0.014)	(0.015)
ln(Age)	0.063	0.119	-0.145
	(0.274)	(0.284)	(0.306)
D(Female)	-0.166*	-0.148	-0.117
	(0.084)	(0.092)	(0.108)
Technical Degree	-0.062	-0.010	-0.046
	(0.082)	(0.085)	(0.100)
Constant	-0.919	-1.068	-0.326
	(0.837)	(0.865)	(0.920)
Observations	129	109	70
R-squared	0.141	0.132	0.155

Table 5: Robust to overplaying >100 tokens

Note: This table uses some of the specifications of Table 2, but drops the 10 subjects that overplayed by less than 100 tokens. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)
Dep. Var.: Overplaying			
Grit Index	0.200***	0.187***	0.203***
	(0.074)	(0.071)	(0.073)
Plan of Action		0.070***	0.078***
		(0.013)	(0.014)
ln(Age)			0.110
			(0.257)
D(Female)			-0.184**
			(0.083)
Technical Degree			-0.114
			(0.081)
Constant	-0.371	-0.760***	-1.038
	(0.244)	(0.241)	(0.793)
Observations	128	128	128
R-squared	0.057	0.129	0.171

Table 6: Robust to alternative econometric specifications

Note: This tables uses alternative econometric specifications. Columns (1) and (2) uses a Logit specification, where the latter column is otherwise equivalent to column (5) of Table 2. Columns (3) and (4) proceed similarly with a Probit estimator. Finally, columns (5) and (6) use a Poisson specification. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep.Var.: Overplaying	Logit	Logit	Probit	Probit	Poisson	Poisson
Grit Index	1.281***	1.322***	0.751***	0.785***	0.823***	0.764***
	(0.412)	(0.454)	(0.239)	(0.256)	(0.252)	(0.252)
Plan of Action		0.631***		0.360***		0.405***
		(0.244)		(0.126)		(0.137)
ln(Age)		0.506		0.293		0.311
		(1.327)		(0.812)		(0.530)
D(Female)		-0.768*		-0.467*		-0.385*
		(0.428)		(0.259)		(0.221)
Technical Degree		-0.560		-0.311		-0.324
		(0.428)		(0.255)		(0.228)
Constant	-5.031***	-10.088**	-2.962***	-5.872**	-3.929***	-6.967***
	(1.459)	(4.654)	(0.837)	(2.749)	(0.937)	(2.097)
Observations	138	138	138	138	138	138

Table 7: Adding Locus Index as a control variable

Note: This table resembles the baseline Table 2, but now also adds the Locus Index. This index is defined in the range between 0 and 1, where higher values mean that the subject believes to a larger extent that outcomes in life are driven by external factors. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
Dep. Var.: Overplaying				
Grit Index		0.246***	0.236***	0.254***
		(0.073)	(0.071)	(0.072)
Locus Index	-0.373	-0.227	-0.196	-0.125
	(0.226)	(0.229)	(0.223)	(0.221)
Plan of Action			0.066***	0.073***
			(0.013)	(0.014)
ln(Age)				0.047
				(0.261)
D(Female)				-0.150*
				(0.083)
Technical Degree				-0.098
				(0.081)
Constant	0.522***	-0.378	-0.767***	-0.909
	(0.116)	(0.288)	(0.284)	(0.837)
Observations	138	138	138	138
R-squared	0.018	0.093	0.149	0.175

Table 8: Adding temptation and procrastination as control variables

Note: This table resembles the baseline Table 2, but now also adds temptation and procrastination as control variables. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. See also notes to Table 2.

	(1)	(2)	(3)	(4)
Dep. Var.: Overplaying				
Grit Index		0.308***	0.317***	0.348***
		(0.087)	(0.084)	(0.079)
Procrastination	-0.046	0.029	0.048	0.038
	(0.062)	(0.061)	(0.059)	(0.060)
Temptation	-0.013	0.030	0.043	0.063
	(0.060)	(0.055)	(0.054)	(0.057)
Plan of Action			0.071***	0.084***
			(0.016)	(0.015)
ln(Age)				0.056
				(0.269)
D(Female)				-0.257***
				(0.084)
Technical Degree				-0.146*
				(0.084)
Constant	0.542**	-0.876*	-1.454***	-1.631*
	(0.226)	(0.447)	(0.462)	(0.958)
Observations	118	118	118	118
R-squared	0.008	0.099	0.158	0.226

Table 9: Sample split by degree and gender

Note: This table splits the full sample based on the type of undergraduate degree studied or the gender of the subject. Column (1) uses the subset of individuals who studied a technical degree, while column (2) uses the remaining subjects. Column (3) uses only data on women, while column (4) only men. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
Dep. Var.: Overplaying	Technical	Non-technical	Female	Male
Grit Index	0.206*	0.287***	0.312***	0.201*
	(0.106)	(0.093)	(0.094)	(0.104)
Plan of Action	0.064***	0.086***	0.063**	0.073***
	(0.014)	(0.026)	(0.028)	(0.015)
ln(Age)	-0.468	0.583**	-0.369	0.573**
	(0.402)	(0.270)	(0.372)	(0.259)
Constant	0.660	-2.922***	-0.033	-2.488***
	(1.204)	(0.814)	(1.113)	(0.815)
Observations	71	67	80	58
R-squared	0.117	0.242	0.157	0.205

Table 10: Decomposition of Grit Index into tenacity and diligence

Note: The table below decomposes the 8 questions of the Grit Index into the Tenacity Index and the Diligence Index. Furthermore, the sentences in italics belong to 'Consistency of Interest', while the remaining ones rather correspond to 'Perseverance'. The alternative specification used as a robustness check consists of only the non-italic questions (i.e., the first two under tenacity and the first two under diligence).

Tenacity Index

Setbacks don't discourage me.

I finish whatever I begin.

New ideas and projects sometimes distract me from previous ones.

I have been obsessed with a certain idea or project for a short time but later lost interest.

I often set a goal but later choose to pursue a different one.

Diligence Index

I am diligent.

I am a hard worker.

I have difficulty maintaining my focus on projects that take more than a few months to complete.

Table 11: Splitting Grit Index into tenacity and diligence

Note: This table builds on Table 2, but now splits Grit Index into Tenacity Index and Diligence Index. Column (1) uses the full sample of observations with no controls, similar to column (1) on Table 2. Column (2) adds the full battery of control variables, as in column (5) on Table 2. Columns (3) to (6) split the data by educational degree and gender, similar to Table 9. Finally, column (7) drops the 10 subjects who overplayed by less than 100 tokens. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Overplaying			Technical	Non-technical	Female	Male	>100
Tenacity Index	0.298***	0.278***	0.237**	0.371***	0.280***	0.313***	0.227***
	(0.074)	(0.068)	(0.091)	(0.096)	(0.095)	(0.096)	(0.073)
Diligence Index	-0.028	-0.011	-0.033	-0.058	0.042	-0.107	-0.017
	(0.073)	(0.071)	(0.108)	(0.098)	(0.092)	(0.104)	(0.075)
Plan of Action		0.071***	0.057***	0.093***	0.068**	0.063***	0.077***
		(0.013)	(0.015)	(0.030)	(0.030)	(0.018)	(0.013)
ln(Age)		0.086	-0.494	0.637**	-0.357	0.519*	0.104
		(0.274)	(0.437)	(0.261)	(0.390)	(0.277)	(0.265)
D(Female)		-0.150*					-0.180**
		(0.081)					(0.082)
Technical Degree		-0.068					-0.088
		(0.083)					(0.084)
Constant	-0.509**	-1.088	0.846	-3.166***	-0.099	-2.197**	-1.011
	(0.237)	(0.856)	(1.350)	(0.807)	(1.180)	(0.910)	(0.827)
Observations	138	138	71	67	80	58	128
R-squared	0.112	0.192	0.136	0.280	0.167	0.250	0.186

Table 12: Splitting Grit Index to consistency and perseverance

Note: This table builds on Table 2, but now splits Grit Index into Consistency Index and Perseverance Index. Column (1) uses the full sample of observations with no controls, similar to column (1) on Table 2. Column (2) adds the full battery of control variables, as in column (5) on Table 2. Columns (3) to (6) split the data by educational degree and gender, similar to Table 8. Finally, column (7) drops the 10 subjects who overplayed by less than 100 tokens. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Overplaying			Technical	Non-technical	Female	Male	>100
Consistency Index	0.136**	0.104	-0.003	0.185**	0.107	0.107	0.092
	(0.066)	(0.063)	(0.093)	(0.080)	(0.083)	(0.102)	(0.064)
Perseverence Index	0.122*	0.162**	0.242**	0.096	0.215**	0.093	0.114
	(0.072)	(0.070)	(0.109)	(0.093)	(0.100)	(0.093)	(0.075)
Plan of Action		0.076***	0.078***	0.086***	0.062**	0.073***	0.079***
		(0.015)	(0.020)	(0.028)	(0.027)	(0.017)	(0.015)
ln(Age)		0.072	-0.510	0.588**	-0.396	0.569**	0.110
		(0.267)	(0.395)	(0.271)	(0.378)	(0.270)	(0.260)
D(Female)		-0.157*					-0.185**
		(0.083)					(0.084)
Technical Degree		-0.105					-0.119
		(0.082)					(0.085)
Constant	-0.525**	-1.094	0.561	-2.928***	0.030	-2.466***	-1.051
	(0.241)	(0.820)	(1.174)	(0.825)	(1.126)	(0.851)	(0.801)
Observations	138	138	71	67	80	58	128
R-squared	0.087	0.176	0.144	0.247	0.163	0.205	0.172

Table 13: Post-Questions on procrastination and temptation

Note: In columns (1) to (4), the dependent variable is the degree of procrastination problems, in an increasing range from 1 to 5. Columns (5) to (8) rather use the degree of temptation problems (also in an increasing range from 1 to 5) as dependent variable. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep.Var.:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Overplaying	Procrastination				Temptation			
Grit Index	-0.660***				-0.570***			
	(0.141)				(0.130)			
Tenacity Index		-0.497***		-0.217		-0.454***		-0.249
		(0.150)		(0.184)		(0.131)		(0.162)
Diligence Index			-0.525***	-0.427***			-0.425***	-0.311**
			(0.096)	(0.128)			(0.110)	(0.136)
Constant	5.498***	4.874***	5.185***	5.520***	5.212***	4.753***	4.839***	5.224***
	(0.460)	(0.474)	(0.332)	(0.443)	(0.430)	(0.420)	(0.387)	(0.429)
Observations	118	118	118	118	118	118	118	118
R-squared	0.198	0.125	0.201	0.218	0.129	0.091	0.115	0.134

Table 14: Post-Questions on self-esteem

Note: In this table the dependent variable is the degree of self-esteem, in an increasing range from 1 to 5. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Selfesteem	(1)	(2)	(3)	(4)
Grit Index	0.576***			
	(0.163)			
Tenacity Index		0.425***		0.150
		(0.156)		(0.177)
Diligence Index			0.480***	0.409***
			(0.128)	(0.150)
Constant	1.193**	1.772***	1.386***	1.163**
	(0.560)	(0.513)	(0.470)	(0.554)
Observations	138	138	138	138
R-squared	0.099	0.060	0.108	0.113

Table 15: Online education survey

Note: In this table the dependent variable is the level of education attained, ranging from 1 (lowest) to 5 (highest). Column (1) includes subjects from all over the world in a specification that only has the Diligence and Tenacity Indices. Column (2) adds country fixed effects, and column (3) also controls for gender, race, and whether the subjects lives in an urban area. Column (4) and (5) limit the data to US respondents, without and with control variables, respectively. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Education	(1)	(2)	(3)	(4) US only	(5) US only
Diligence Index	0.148***	0.214***	0.094***	0.258***	0.126***
	(0.021)	(0.021)	(0.018)	(0.032)	(0.025)
Tenacity Index	0.088***	0.052**	0.009	0.028	-0.009
	(0.023)	(0.023)	(0.019)	(0.032)	(0.026)
ln(Age)			1.313***		1.337***
			(0.044)		(0.060)
Gender Dummies	N	N	Y	N	Y
Race Dummies	N	N	Y	N	Y
Urban Dummies	N	N	Y	N	Y
Country FE	N	Y	Y	N	N
Observations	4,199	4,162	4,162	2,112	2,112
R-squared	0.040	0.134	0.399	0.062	0.394

Figure 1: Roulette Pictures

Note: Snapshots of the roulette game. In the first picture, the subject places a bet (at the start of the game) of 50 tokens on the number 4, 100 tokens on the number 10 and 50 tokens on the color red. The second picture takes place later in the game, and the subject sees the history of the last five outcomes (with the last being 12 red), how much he has won and bet following the previous spin, and his remaining tokens. He has not yet placed the new bet. He is always free to quit, and he is allowed to spin even without placing any bet.

Figure 2: Kernel Density - Grit Index

Note: This figure separately traces the kernel density distributions for the Grit Index for subjects who did not overplay (blue dashed line) and for those who overplayed (red solid line).

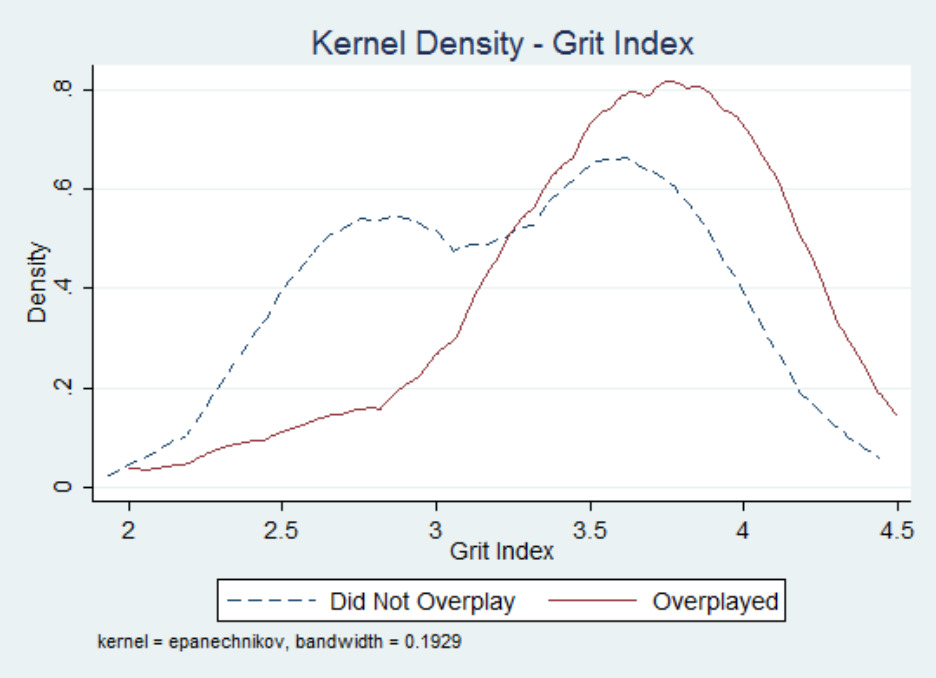


Figure 3: Cumulative Density Function - Grit Index

Note: Building on Figure 1, this figure separately traces the cumulated density functions for subjects who did not overplay (blue dashed line) and for those who overplayed (red solid line).

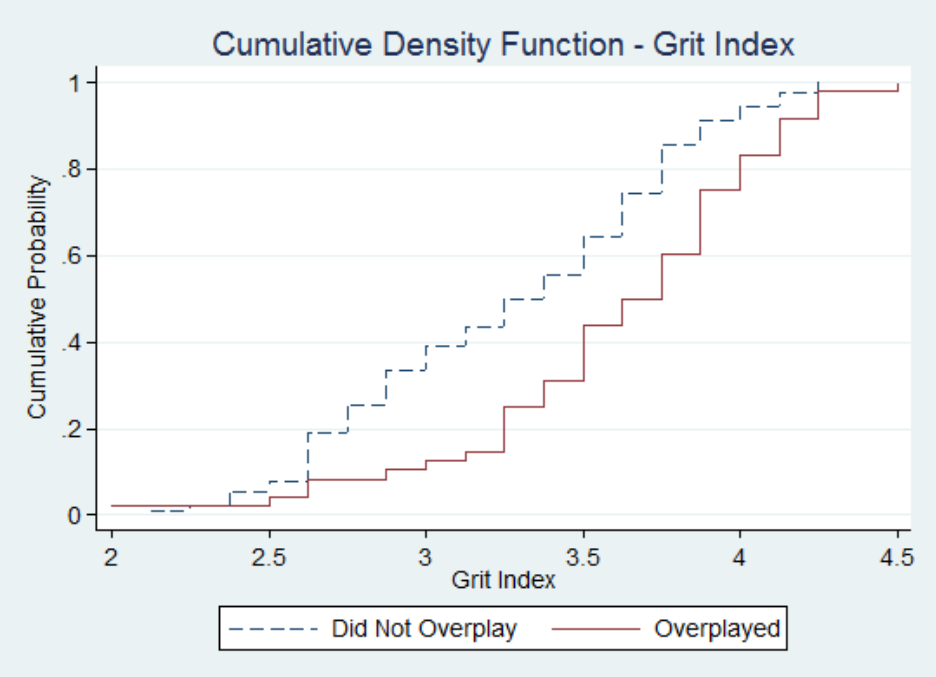


Figure 4: Kernel Density - Tenacity Index

Note: This figure separately traces the kernel density distributions for the Tenacity Index for subjects who did not overplay (blue dashed line) and for those who overplayed (red solid line).

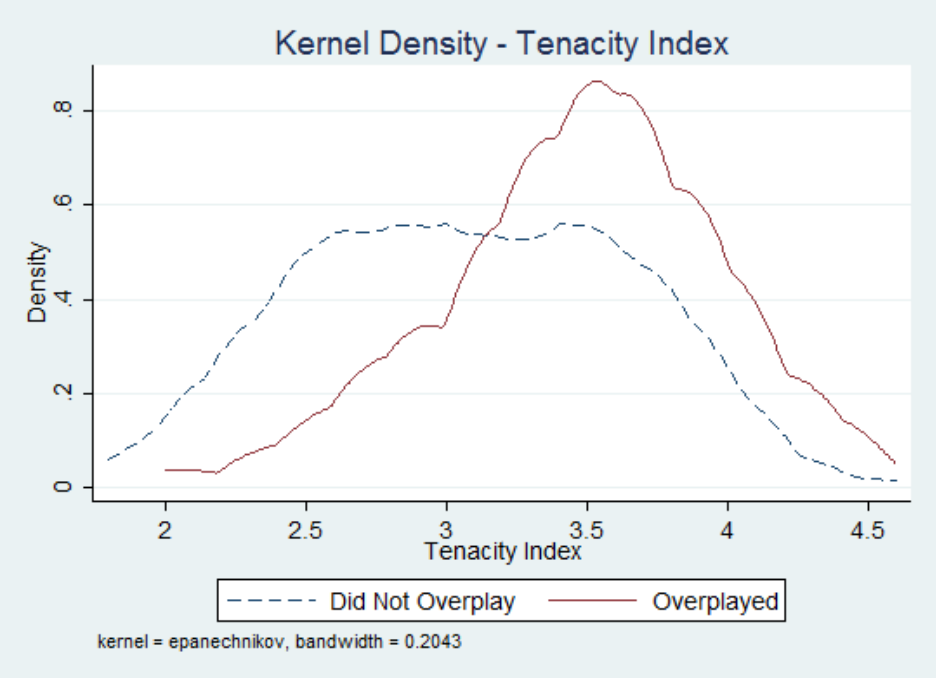
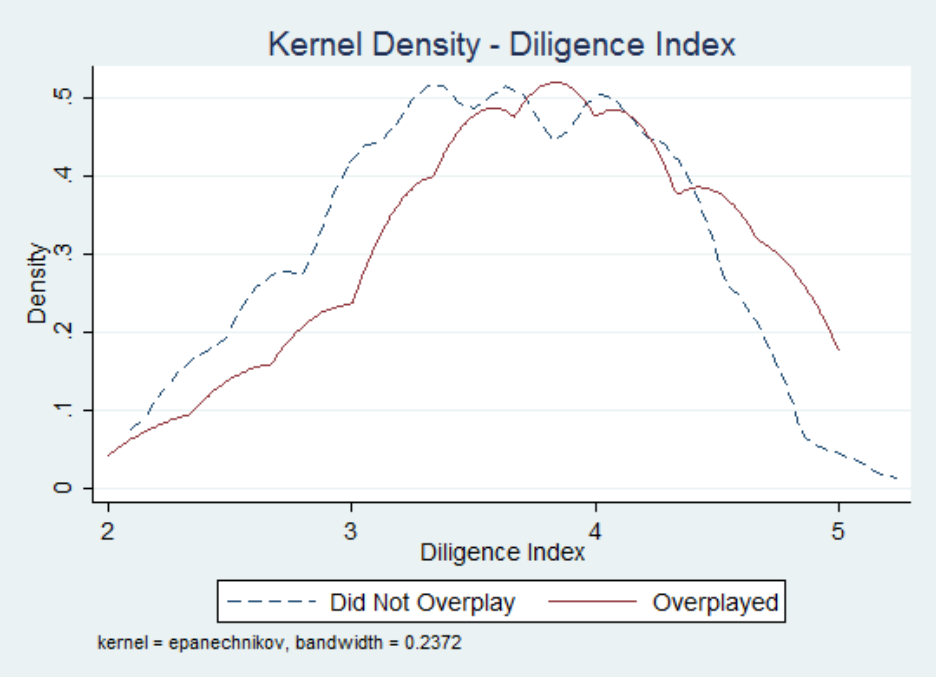


Figure 5: Kernel Density - Diligence Index

Note: This figure separately traces the kernel density distributions for the Diligence Index for subjects who did not overplay (blue dashed line) and for those who overplayed (red solid line).



References

[1] Alan, S., Boneva, T., and Ertac, S. Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *Working Paper* (2016).

[2] Alan, S., and Ertac, S. Good things come to those who (are taught how to) wait: Results from a randomized educational intervention on time preference. *Working Paper* (2014).

[3] Alan, S., and Ertac, S. Patience, self-control and the demand for commitment: Evidence from a large-scale field experiment. *Journal of Economic Behavior and Organization 115* (2015), 111–122.

[4] Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. Personality psychology and economics. *Handbook of the Economics of Education 4 (Chapter 1)* (2011).

[5] Ameriks, J., Caplin, A., Leahy, J., and Tyler, T. Measuring self-control problems. *American Economic Review 97*, 3 (2007), 966–972.

[6] Ashraf, N., Karlan, D., and Yin, W. Tying Odysseus to the mast: Evidence from a commitment savings product in the philippines. *Quarterly Journal of Economics 121*, 2 (2006), 635–672.

[7] Augenblick, N., Niederle, M., and Sprenger, C. Working over time: dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics* (2015), 1067–1115.

[8] Barberis, N., and Thaler, R. A survey of behavioral finance. *Handbook of the Economics of Finance 1B (Chapter 18)* (2003), 1053–1128.

[9] Benabou, R., and Tirole, J. Self-confidence and personal motivation. *Quartely Journal of Economics 117*, 3 (2002), 871–915.

[10] Borghans, L., Duckworth, A. L., Heckman, J. J., and ter Weel, B. The economics and psychology of personality traits. *Journal of Human Resources 43*, 4 (2008), 972–1059s.

[11] Brunnermeier, M. K., and Parker, J. A. Optimal expectations. *American Economic Review 95*, 4 (2005), 1092–1118.

[12] Burks, S. V., Lewis, C., Kivi, P. A., Wiener, A., Anderson, J. E., Gotte, L., DeYoung, C. G., and Rustichini, A. Cognitive skills, personality, and economic preferences in collegiate success. *Journal of Economic Behavior & Organization 115* (2015), 30–44.

[13] Compte, O., and Postlewaite, A. Confidence-enhanced performance. *American Economic Review 94*, 5 (2004), 1536–1557.

[14] Cubel, M., Nuevo-Chiquero, A., Sanchez-Pages, S., and Vidal-Fernandez, M. Do personality traits affect productivity? Evidence from the lab. *Economic Journal forthcoming* (2016).

[15] Cunha, F., Heckman, J. J., and Schennach, S. M. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*, 3 (2010), 883–931.

[16] Dobbie, W., and Fryer Jr., R. G. The medium-term impacts of high-achieving charter schools. *Journal of Political Economy 123*, 5 (2015), 985–1037.

[17] Duckworth, A. L., and Gross, J. J. Self-control and grit: related but separable determinants of success. *Association for Psychological Science 23*, 5 (2014), 319–325.

[18] Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology 92*, 6 (2007), 1087–1101.

[19] Duckworth, A. L., and Quinn, P. D. Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment 91*, 2 (2009), 166–174.

[20] Duckworth, A. L., and Seligman, M. E. Self-discipline outdoes IQ in predicting academic performance for adolescents. *Psychological Science 16*, 12 (2005), 939–944.

[21] Galasso, A., and Simcoe, T. CEO overconfidence and innovation. *Management Science 57* (2011), 1469–1484.

[22] Gneezy, U. Deception: the role of consequences. *American Economic Review 95*, 1 (2005), 384–394.

[23] Gul, F., and Pesendorfer, W. Temptation and self-control. *Econometrica 69*, 6 (2001), 1403–1435.

[24] Heckman, J. J., and Rubinstein, Y. The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review 91*, 2 (2001), 145–149.

[25] Kaur, S., Kremer, M., and Mullainathan, S. Self-control at work. *Journal of Political Economy 123*, 6 (2015), 1227–1277.

[26] Kautz, T., Heckman, J., Diris, R., ter Weel, B., and Borghans, L. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. *NBER Working Paper 20749* (2014).

[27] Laibson, D. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*, 2 (1997), 443–477.

[28] Malmendier, U., and Tate, G. CEO overconfidence and corporate investment. *Journal of Finance 60*, 6 (2005), 2661–2700.

[29] Mobius, M. M., Niederle, M., and Niehaus, P. Managing self-confidence. *Working Paper* (2014).

[30] Odean, T. Are investors reluctant to realize their losses? *Journal of Finance 53* (1998), 1775–1798.

[31] O'Donoghue, T., and Rabin, M. Doing it now or later. *American Economic Review 89*, 1 (1999), 103–124.

[32] O'Donoghue, T., and Rabin, M. Present bias: Lessons learned and to be learned. *American Economic Review 105*, 5 (2015), 273–279.

[33] Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., and Goldberg, L. R. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Sciences 2* (2007), 313–345.

[34] Robins, R. W., Hendin, H. M., and Trzesniewski, K. H. Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale. *Personality and Social Psychology Bulletin 27*, 2 (2001), 151–151.

[35] Rotter, J. B. *Social learning and clinical psychology.* Prentice-Hall, Englewood Cliffs, NJ, 1954.

[36] Shefrin, H., and Statman, M. The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance 40*, 3 (1985), 777–790.

[37] Sprenger, C. Judging experimental evidence on dynamic inconsistency. *American Economic Review: Papers & Proceedings 105*, 5 (2015), 280–285.

[38] Strotz, R. H. Myopia and inconsistency in dynamic utility maximization. *Review of Economics Studies 23*, 3 (1956).

Appendix

A Logistics of the Experiment

The experiment was conducted at the Behavioral Sciences Laboratory (BESLab) of Pompeu Fabra University (UPF), over 8 sessions, and included 138 subjects. Subjects were students of UPF, recruited using the BESLab system. No subject took part in more than one session. Subjects were paid 3 euros for showing up. They were also paid based on their earnings in the game (chosen at random) plus 4 euros for the end of the experiment survey. Earnings ranged from 7 euros to 60 euros. All earnings in the game were in tokens, where 150 tokens are worth 1 euro. Experimental sessions typically ranged between 60 and 75 minutes, with some subjects playing for longer (no time constraint was imposed).

A.1 Instructions of the Experiment

The sequence of the experiment is as follows. The subjects were informed that they would be paid at random between the games, that all the games were single-agent, and that there were no right or wrong answers. In particular their first (Qualtrics) screens had the following information (translated from Spanish):

***
Welcome to the experiment. Please, do not use mobile phones or any other electronic devices. Talk to other participants is not allowed. Please raise your hand if you have any questions.
Important: Do not close the web browser at any time!
During the experiment, your earnings in each game will be counted in tokens. A euro is worth 150 tokens. At the end of the experiment, you will be paid in euros.
Your earnings in each game will be independent of each other. At the end you will randomly be paid the based on the earnings of one of these games. For example, if in a game you end up with 100 tokens and in another one you end up with 200 tokens, then your payment will be either 100 or 200 tokens, chosen at random.
In this experiment there are no right or wrong questions. All answers will only depend on your preferences, and different people will answer the questions in different ways. Simply answer the questions based on your preferences.
***

The first phase of the game consists of eliciting the subjects' plan of actions. We ask them to state their preferences for the first game, which we use as a robustness check for the plan of action. This is done as follows:

***
Consider the next game:
In each round of the game, you will choose a number between 1 and 38, and the computer will then also randomly choose a number between 1 and 38. Each number has the same likelihood of being chosen. If your chosen number is the the same as the computer's, then

you win 3200 tokens. If not, then you lose 100 tokens. If you win a round, then you cannot continue playing. That is, at most you can only win once.
If you start with 2000 tokens, after how many rounds would you like to stop playing?
***

They can choose any number between "0 (and you will end up with 2000 tokens)" and "20 (and if you don't win you will end up with 0 tokens)."
We then asked them for their plan of action of the roulette game. We explained the rules, but subjects were familiar with this game.

***
Consider this second game. The computer will choose a number between 1 and 36, as well as 2 additional numbers that we will call 0 and 00. Each number has the same likelihood of being chosen. In addition, half of the numbers are red and the other half are black.
You can choose a number between 1 and 36. You can also choose a color (red/black).
If your chosen number is the the same as the computer's, then you get back your bet and you also receive 35 extra tokens for each token you bet. If your number is not selected, then you lose your bet and do not receive anything.
Similarly, if you have chosen a color and it is selected, then you get back your bet and additionally receive 1 extra token for each token you bet. If your color is not selected, then you lose your bet and do not receive anything.
You can play for as long as you want, provided you have enough tokens. There is a maximum of 8000 tokens that you can win. You can stop playing at any time. For this game, the game does not end if your chosen number is the same as the computer's.
Example: Suppose that you bet 100 tokens on number 15.
If the computer also chooses number 15, then you receive 3600 tokens (the 100 tokens of your initial bet and 3500 additional tokens).
If the computer chooses either 0, 00, or any other number between 1 and 36 that is not number 15, then you lose the 100 tokens you bet.
Now, suppose you bet 200 tokens on the color red. If the computer also chooses red, then you receive 400 tokens (the 200 tokens of your initial bet and 200 additional tokens). If not, then you lose the 200 tokens you bet.
If you start with 2000 tokens, what is the range of tokens inside which you would like to keep playing?
***

They are given the choice between 0 and 2000 tokens, in increments of 100. We also used a simplified Becker-Degroot-Marshak (BDM) mechanism, although here they have no incentive to lie. In particular, we first explain in detail the mechanism at the beginning of the experiment. Then, this simplified mechanism is used only to ensure that they have incentives to truthfully report the stated preferences in the main game. (We rely on the stated preferences because the complexity of this setting would make the provided willingness to trade amount overly noisy, and these amounts are not useful for our objectives.)
As noted previously, in the first game, designed as a robustness check, a high percentage of subjects play consistently with their plan of action, confirming that the stated preferences are reliable.

Before moving on to the second phase of the experiment (the actual games), we ask questions to deframe the subjects. The subjects are informed that these questions will not matter for their earnings (this is done so that subjects do not incur much cognitive strain for these questions):

***

Before proceeding to play the games on which your earnings will be based, please answer the following questions.
Your earnings will not depend on the answers given to these questions.
***

The questions themselves are designed not to induce cognitive fatigue. For instance:

***

How much time per day do you spend reading the news?
Which topics do you spend more time reading about?
Which of these is the closest to your favorite color?

There are two different shapes in this image. Which ones are they? (They could leave this question blank if they choose to do so.)

***

The second phase of the experiment consists of the actual games, in the same order in which the plans of action were elicited. They play the first (simpler) game, which they will see until they choose to quit (by choosing 39), or until they win, or until they lose all their tokens. Before playing either game, we inform them again that their earnings will be determined at random. In particular, the instructions explain that at the end of the experiment a die will be tossed to determine the earnings. If it lands between 1 and 3 then one game will determine their earnings, otherwise the other will.

After they are done with the first game, they can play the roulette game. We provide them with a sheet with the rules of the roulette (same as explained above), and they are free to raise their hands if they have any doubts, which almost never occurred. The subjects could quit at any moment, and they could also bet 0 tokens if they so wishes. The game can be found at: http://experimentalgames.upf.edu/roulette/ ; we include snapshots above in

Figure 1. The roulette was coded in Adobe Flash.
The third phase of the experiment consists of the questionnaire. We ask them the 8 short grit questions, followed by the shortened 17 locus of control questions. We also ask them self-esteem, temptation and procrastination questions:

***

I have high self-esteem.

- 1 - Completely disagree.
- 2 - Somewhat disagree.
- 3 - Somewhat agree.
- 4 - Very much agree.
- 5 - Completely agree.

There may be tasks that you have to perform but that are not fun to do. For example, this could include studying courses that you dislike, waking up early, etc. Do you find yourself postponing these tasks or performing them less often than you should?

- 1 - Never
- 2 - Almost never.
- 3 - Sometimes.
- 4 - Often.
- 5 - A lot.

Similarly, there may be activities that you should not do too often, but which you enjoy doing. For example, this could include spending the day watching episodes of TV series, eating excessive amounts of chocolate, etc. Does it happen to you often that prioritize or end up spending too much time doing these activities? Do you find yourself putting these activities ahead of more important ones or spending too much time performing them?

- 1 - Never
- 2 - Almost never.
- 3 - Sometimes.
- 4 - Often.
- 5 - A lot.

***

The last questions ask for their age, gender and field of study. Once a subject finishes the survey, he or she enters the control room (individually, to avoid social effects) and can choose to toss the die to determine which game will matter for the earnings.