Know when to Fold 'em: The Grit Factor

Larbi Alaoui
Christian Fons-Rosen

This version March 2019
(April 2016)

Barcelona GSE Working Paper Series

Working Paper nº 899

Know When to Fold 'Em: The Flip Side of Grit*

Larbi Alaoui† Christian Fons-Rosen‡

March 1, 2019

Abstract

This paper investigates the way different sides of grit influence behavior. In addition to grit's upside in achieving economic success associated with not giving up, it might also have a downside associated with not letting go. We provide a stylized model that illustrates this potential downside. We split grit into two new categories, tenacity and diligence, and hypothesize that tenacity can lead individuals to go beyond their own intended plan of action when making a loss. We test the predictions with an experiment that elicits each individual's plan of action which we compare to actual choice in a game of luck. Consistent with our priors, grittier individuals have a higher tendency to overplay, and tenacity alone captures the difficulty in respecting ex-ante preferences when this means accepting defeat. We then discuss the external validity of our findings.

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, Carlos Alos-Ferrer, Jose Apesteguia, Ghazala Azmat, Roy Baumeister, Roland Benabou, Veronica Benet-Martinez, Laura Doval, Angela Duckworth, Seda Ertac, Rosa Ferrer, Andrea Galeotti, James Heckman, Johannes Horner, Philipp Kircher, David Laibson, Jennifer Lerner, Pablo Lopez-Aguilar, Christian Michel, Magne Mogstad, Ignazio Monzon, Antonio Penta, Helena Perrone, Richard Roberts, Santi Sanchez-Pages, Charles Sprenger, Adam Szeidl, Ana Tur-Prats, Kathleen Vohs, Leeat Yariv, participants of the HCEO Chicago conference, the UPF behavioral science workshop, the SAET conference, and members of the Bogazici University, the EUI, Sussex university and the University of Pennsylvania Character Lab for their thoughtful comments.

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

‡University of California at Merced and CEPR. E-mail: cfons-rosen@ucmerced.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 skills, grit and conscientiousness, which connect to perseverance, have attracted attention in the psychology literature and increasingly in economics. Studies consistently show that such skills are critical 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 may even outperform IQ in determining lifetime success (Duckworth and Seligman (2005), Roberts et al. (2007), Kautz et al. (2014)). The upside of grit and conscientiousness 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 characteristics 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.[1] 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 than expected to accept a loss. 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

[1] A number of studies have found a high correlation between grit and conscientiousness (a "Big Five" personality component, together with neuroticism, openness to experience, extraversion and agreeableness), and both are intimately linked to the idea of perseverance. We focus on the grit survey because it includes questions which are more directly phrased in the language of what we term tenacity, but, as we discuss explicitly below, we expect some of our findings to hold with conscientiousness as well.

Duckworth and Quinn (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.

To organize ideas, we provide a stylized model that illustrates our hypothesis discussed above on the link between grit and not letting go. Agents in this framework have ex-ante preferences which determine their plan of action, and in particular the maximum loss they are willing to incur. They also have a given amount of grit, which is made up of two possibly correlated components, 'diligence' and 'tenacity.' Diligence captures the notion of being hard-working. Tenacity, instead, is more nuanced. It includes both the positive side, perseverance, and the negative side, stubbornness. In the context of our setting, tenacity is the relevant component. In particular, we assume that agents who are more tenacious are more likely to incur a cost of failure, or letting go, when they are losing. In turn, this will make them more likely to overplay and keep going past their planned maximum loss. However, agents with higher diligence and the same level of tenacity are neither more nor less likely to overplay. Moreover, since grit is made up of both tenacity and diligence, agents with more grit are also more likely to overplay, although the likelihood is smaller than it is for tenacity alone. Lastly, we show that tenacity can have an upside in settings for which agents have a tendency to give up earlier than they would like. In that case, the incurred cost of failure pushes agents to keep going further, and potentially closer to their objective. Our model thus serves to illustrate when tenacity would have a downside and when it would have an upside.

The results of our experiment are consistent with our hypothesis. Comparing the actual choices of subjects to their 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. In fact, at least 30% of subjects overplay, irrespective of characteristics such as undergraduate degree and gender.[2] This finding is indicative of the downside of grit. Grittier agents persevere and are not easily discouraged, but their 'stubbornness' also makes it difficult for them to let go and accept failure, thereby leading them to go beyond their plan of action.

To further explore our predictions, we split the grit questions along the two categories discussed above, diligence and tenacity. This split is a new decomposition of grit, and it serves our objectives because it isolates the central aspect of not letting go (tenacity) from the aspect of working harder (diligence). 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 then perform robustness checks on these categories by considering variations of the partition. Our results remain unchanged. In addition, we perform a two-factor confirmatory factor analysis (CFA) of the main decomposition of grit into diligence and tenacity as well as of the variations used in the robustness checks, and find that the results of this analysis are consistent with our categorization.

[2] 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.

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. We analyze existing datasets, as external validity tests of their predictive power with conventional performance measures in domains in which we would conjecture that their positive aspect dominates. We specifically consider data on educational performance. We expect that diligence plays an unambiguously positive role in these contexts, and that tenacity, by making it more likely not to give up, plays a mainly positive role as well. The first sample consists of an ICPSR school survey. It includes the grit measure and data on educational performance for 14,465 students between grades 6 and 9. The second sample consists of a psychological test repository and contains 3,988 people. It also includes the grit measure as well as educational outcomes and other demographics. We find in the two datasets that the tenacity and diligence measures are both positive predictors of educational level. Interestingly, for some of these educational outcomes, 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, since, unlike the data obtained from our experiment, we do not have the ex-ante plan of action. The second dataset also includes a measure of conscientiousness, and so we use it to explore the relation between our categorization of diligence and tenacity to conscientiousness.

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 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. The aspects of conscientiousness most connected to tenacity would also be useful measures in analyzing this subject. Future research can focus on the interplay between the different facets of grit and conscientiousness as functions 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 stylized model, Section 4 provides the main findings, and Section 5 discusses additional results. Section 6 addresses the external validity of decomposing grit into tenacity and diligence. Section 7 concludes.

Related literature

This paper relates to several strands of research. The first concerns the relevance of noncognitive skills, and particularly grit and conscientiousness, 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 noncognitive skills in academic success, and highlight the gap in the economics literature in analyzing them. MacCann, Duckworth and Roberts (2009) find that the different facets of conscientiousness are linked to higher education performance, and Burks et al. (2015) find that within conscientiousness, 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. Gerhards and Gravert (2017) propose a behavioral measure of grit. Cubel, Nuevo-Chiquero, Sanchez-Pages and Vidal-Fernandez (2016) conduct a laboratory experiment on the effect of the Big Five on performance, and find that conscientious subjects perform better. Gill and Prowse (2016) analyze how both cognitive skills and personality, including grit, influence behavior in the beauty contest. Proto, Rustichini and Sofianos (2018) study how intelligence and personality affect group outcomes, and find that conscientiousness and agreeableness have an impact on cooperation. Papageorge, Ronda and Zheng (2017) study how another noncognitive skill, known as externalizing behavior, reduces attainment but is productive in the labor market. Within neureconomics, Rueter et al. (2018) test the hypothesis that a neural network denoted as the "goal priority network" is a neural correlate of conscientiousness. Within the psychology literature, Lucas et al. (2015) find that grittier subjects are less willing to quit when failing. The focus of their study is distinct from ours, however, and since there is no measure of the plan of action, the subjects' intentions cannot be compared to their actual behavior. Boyce, Wood and Brown (2010) use data from a German socio-economic survey and find that the negative correlation between unemployment spells and life satisfaction is stronger for individuals with high levels of conscientiousness. Using another survey, Heineck (2011) finds a non-linear association between conscientiousness and wages.

To some extent our analysis using grit is likely to have implications on conscientiousness as well, since recent studies find that they are intricately related. In particular, Credé, Tynan and Harms (2016) provide a meta-analysis that finds a high correlation between grit and conscientiousness, particularly with respect to the 'perseverance' facet of grit. The focus of our analysis is orthogonal to establishing the relation between grit and conscientiousness, but an interesting subject for future research concerns the extent to which our results carry through using the conscientiousness measure. We discuss this point in Sections 6.2.

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)). We also note that Amador, Werning and Angeletos (2006) consider the optimal tradeoff between commitment and flexibility. Our paper has a different focus and we do not test whether grittier subjects are more or less likely to commit to a goal beforehand or to prefer flexibility. Rather, we analyze whether they are more likely to systematically go beyond their pre-established plan.

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 (but see, within the psychology literature, Gollwitzer (1999) and Achtziger, Bayer and Gollwitzer (2012)). 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. At least 30% 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, Bénabou and Tirole (2002, 2016), 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. It may also be interesting to analyze the degree to which individuals with more grit also invest more in the abilities that will lead to higher positive self-image (see Santos-Pinto and Sobel (2005) for a model in which there are multiple skill dimensions, and agents invest along some dimensions in a manner that contributes to their self-image when assessing their abilities relative to others).

Within the prospect theory literature, Barberis (2012) discusses how it can explain gambling behavior, and particularly the idea of gambling past the plan to follow a particular strategy. Ebert and Strack (2015) show that, strikingly, a naive agent with prospect theory preferences will gamble until the bitter end, while Ebert and Strack (2018) show that a sophisticated one will not start gambling at all. It would be interesting to link tenacity to higher loss aversion through a channel that relates to prospect theory. The focus of these papers is distinct from ours, which focuses on comparative statics on overplaying along the grit (and tenacity) dimension. While using a prospect theory setting would be interesting, in our stylized model, we do not explore whether tenacity links directly to loss aversion (in a prospect theory sense) or probability weighting. This is partly because it would be challenging to explain through such a channel why, in our experiment, tenacious subjects do not have significantly different plans of action and do not play more than others; rather, they *overplay* more. In addition, Ebert and Strack's (2015, 2018) results suggest some caution in the modeling aspect. There is also 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) and Odean (1998)). This effect may be present in the way subjects approach losses in our experiment. For papers exploring the disposition effect combined with a plan of action, but without incorporating grit or tenacity, see Andrade and Iyer (2009), Ploner (2017). Relatedly, Fischbacher, Hoffmann and Schudy (2017) investigate how to mitigate the disposition effect via automatic trading systems. In light of our results, another avenue of research would explore whether more tenacious subjects also have a higher tendency to hold losing assets longer 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 2, 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.[3] 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

[3]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.

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.[4] 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 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. 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 36. 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. This game was designed to be much more passive than the roulette, since the only choices available here are whether or not to stop at any stage and to pick one of the 36 numbers. Furthermore,

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

since the game ends as soon as the subject wins once, there can be no post-victory momentum from playing after a win takes place.

As discussed above, we do not expect the subjects to lie about their plan of action. But to reinforce their incentives to tell the truth, we use a variation of the Becker-DeGroot-Marschak (BDM) method to make this stated plan of action incentive-compatible. We leave the details of this point to the Appendix, as they are involved and not central to our analysis. In brief, however, at the start of the experiment we explain to the subjects the BDM mechanism and its incentive compatibility. After asking for their stated plan of action, we ask for their willingness to sell their preferred plan of action in exchange for having to play until they either lose all their tokens or win the maximum amount possible. We explain that we may compare their chosen amount against an amount chosen by the computer, according to the BDM (second-price auction) way. While the complexity of the multiple-stage lottery of these games makes it difficult to take the willingness to sell at face value, notice that we are not using this amount in our analysis, but only the stated plan of action, which is the focus of our study.

The results from the first game suggest that the subjects are indeed truthful in their response and wish to respect their established plan of actions. Specifically, 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. 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.

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 attenuating any anchoring effect from the recently elicited plans of action before they play the actual games. These questions, which take approximately 10 minutes to answer, 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 Theoretical Framework

This section shows how higher grit can be linked to behavior through a stylized model. Formally, suppose that an individual's grit G is the weighted sum of two components, tenacity T and diligence D, i.e., $G = \alpha_T T + (1 - \alpha_T)D$, where $\alpha_T \in (0, 1)$. 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 encountered along the way. In Section 4.3 we will discuss how the standard grit questionnaire can be decomposed into these two measures. Here instead we focus on the theoretical notion of these terms, and we will specifically assume that higher tenacity makes it more likely to incur a psychological cost to stop performing a task when failing. Our mechanism relates to tenacity alone and not diligence. The actual game used in our experiment is more complex, as are other relevant tasks in which tenacity may be a factor, such as tasks with elements of procrastination or temptation. But the model below captures the main features of interest. While our central assumptions on tenacity can be adapted to fit within a prospect theory setting instead of a von Neumann-Morgenstern Expected utility setting, we abstract away from doing so, for reasons explained below.

3.1 General setting

Description of strategies: plan of action and behavior

Let the agent's initial wealth be $z_b > 0$, and suppose that the agent has τ periods in which to act, where $\tau \geq 2$, and possibly infinite. Let $t = \{1, .., \tau\}$ denote a period. At period t, the agent has the option to stop a task ($a_t = N$) or to continue ($a_t = Y$). If the agent chooses N then the task stops, and the agent cannot return to it. As a tie-breaker, we assume that when indifferent between continuing and not, the agent chooses N. If the agent chooses Y then, at the end of that period, he wins $z_w > 0$ with probability p_w and loses $z_l > 0$ with probability $1 - p_w$. If his earnings reach or drop below 0, or if he was at the last period $t = \tau$, then he must stop. Otherwise, he chooses again between Y and N.

There are therefore two kinds of histories, *terminal* and *non-terminal* histories. Let y^t denote the terminal history of actions and outcomes reached if the agent chooses to stop at period t. Terminal history for $t = 1$ is denoted $y^1 = \{N\}$, and for $t \geq 2$ the history is denoted $y^t = ((Y, x)_1, ...(Y, x)_{t-1}, N)$, where $x \in \{z_l, z_w\}$ is the (new) earnings made at the end of the corresponding period. Let h^t denote a non-terminal history, which has the same form but without the N, and represents histories up to the end of period $t - 1$ for which the agent has not stopped playing. Specifically, $h^1 = \emptyset$, and for $t \geq 2$, $h^t = ((Y, x)_1, ...(Y, x)_{t-1})$.

A strategy maps from non-terminal histories to actions. But since the agent may not be dynamically consistent, we specify the strategy s_t, for $t \geq 1$, which is the strategy at period t for all possible subsequent histories. In particular, a strategy s_t will prescribe an action $a_{t'}(s_t|h^{t'}) \in \{Y, N\}$ for every $t' \in \{t, .., \tau\}$ and every (non-terminal) history $h^{t'}$. Note that there is no action to take for terminal histories y^t.

Then a chosen strategy s_1, in period 1, corresponds to the ex-ante plan of action for every future possible history reached. In particular, the ex-ante action that the agent would take at period t (if it is reached), given history h_t, is $a_t(s_1|h^t)$. Instead, if at history h^t the agent chooses strategy s_t, then the action actually taken in that period is $a_t(s_t|h^t)$. That is, $a_t(s_t|h^t)$ represents the actual behavior at period t if that period is reached, given history h^t. Notice that if, at period t and history h^t, the action chosen, $a_t(s_t|h^t)$, is not the same as the ex-ante planned action for that node, $a_t(s_1|h^t)$, then the strategies s_1 and s_t are inconsistent. That is, actual behavior deviates from the plan of action.

Description of preferences

The agent's value function for period t reached by history h^t is denoted $U^t(s_t, h^t)$, and is a function from strategies and histories to $\mathbb{R}$. The reason that the superscript t is specified in U^t is that the agent may not be temporally consistent, and the value function when a decision node is reached may not be the same as it was ex-ante. Throughout, we make the assumption that temporally inconsistent agents are naive about their inconsistency, as we will make precise. Below, we will consider different types of underlying preferences for different contexts, and explain what higher tenacity entails for each. We will maintain the same assumptions on tenacity, which we now discuss.

Assumptions on tenacity and failure

Our central assumption is that if the agent is losing, then with some probability, he immediately incurs an arbitrarily large cost of failing if he decides to stop. Moreover, this probability increases in tenacity. The modeling choice of having a probability of incurring a cost is made for simplicity. The choice that it is arbitrarily large serves to avoid the need for conditions on the minimal bound of the cost, and could be weakened for our results below.

To avoid making ad-hoc assumptions on the meaning of losing, we take the initial endowment z_b to correspond to the baseline, and take earnings $z < z_b$ to correspond to losing. The agent is naive about his future probabilistic cost of failing. Formally, let $m_w(h^t)$, $m_l(h^t)$ be the count of the number of wins and losses, respectively, received in all periods up to non-terminal history h^t, and let $z_e(h^t) \equiv z_b + m_w(h^t)z_w - m_l(h^t)z_l$ be the total earnings at that history. We also write $z_e(y^t)$ for total earnings at terminal history y^t.

Assumption 1. *Suppose that $z_e(h^t) < z_b$ at history h^t. Then, if the agent's chosen strategy s_t prescribes that he stops (i.e., $a_t(s_t|h^t) = N$), he incurs an immediate and arbitrarily large cost of failure $c_f > 0$ with probability $q_f(T) \in [0, 1)$, where $q_f'(T) > 0$.*

The assumption that the agent does not anticipate his cost of failure beforehand is the source of his naivete about his possible temporal inconsistency. Without this assumption, an agent who anticipates this cost of failure may choose not to play at all. Introducing some degree of sophistication would complicate our analysis, and would also require a maximum bound on the cost of failure to obtain our results.

Relationship with loss aversion and other effects

Assumption 1 could in principle serve as a factor towards the disposition effect, and may be an additional force towards loss aversion. Interestingly, it may also relate to the distinction between realized versus paper (non-realized) losses (Imas (2016)). In this setting, by stopping, a losing agent effectively converts his paper losses into realized losses, and the cost of failure c_f can be viewed as a cost of making this conversion. Note, however, that our focus is distinct from that of this literature, in that we analyze the comparative statics as tenacity increases. We view this paper as complementary to these strands of research. In particular, introducing our assumptions on tenacity within a dynamic prospect theory setting would add complexity (see, for instance, Barberis (2012), where a closed-form analytical solution is not found for finite periods). It would therefore distract from the main objective of conducting comparative statics across tenacity in a simple way. Furthermore, according to Assumption 1, even two agents with identical ex-ante preferences, including in their loss-aversion parameters, may still differ in actual behavior if their tenacity T is different. That is, in a prospect theory setting with a utility function that is concave over the gains, convex over the losses and a (cumulative) probability weighting function, if an increase in tenacity mapped to either an ex-ante change in the shape of the utility function or that of the weighting function, then it would possibly affect not only overplaying, but also the plan of action itself, and the two could vary simultaneously. Within our model, instead, comparative statics on tenacity T lead to a change in overplaying alone while maintaining the plan of action fixed, and therefore result in clear comparative statics on

overplaying as well (maintaining an assumption on naivete in both settings).[5]

Moreover, accounting for the results by Ebert and Strack (2015), with a prospect theory representation with commonly-used parameters, effectively all the subjects, regardless of tenacity, who make a loss would then play until the 'bitter end' of reaching zero, but this is not what we find. A number of subjects (30%) make losses and stop before losing everything, or even before overplaying.[6] Our model provides a clean and simple way of capturing this phenomenon, and removes orthogonal considerations.

In brief, we will not explore a prospect theory representation of preferences, because we view our approach to be more effective way in explaining our mechanism in a clean and simple way. We note, however, that our assumptions can be implemented into a prospect theory setting (specifically Barberis (2012)), although this is beyond the scope of our paper.

In closing this discussion, we note too that if the objective were to separate our assumptions from one in which loss aversion (in a prospect theory sense) increases, the following experiment could be conducted. Suppose that the subjects, after a certain number of losses, did not have the possibility of recovering their initial endowment (which could be occur, for instance, with finite periods). In a prospect theory setting, such agents would still overplay, under common assumptions concerning the utility and weighting functions. Under our assumptions above, however, an increase in tenacity would not have any effect. This is because without the possibility of recovering the loss, the cost of failure is effectively sunk. In other words, an increase in tenacity, as defined in our model, would not in itself be predictive of an increase in the probability of overplaying in this case. Conducting such an experiment would be interesting, but is not the objective of our paper.

3.2 Predictions for the experiment

Let $E_{y^{t'}|h^t,s}$ denote the expectations operator given strategy s and reached history h^t for every possible subsequent terminal history $y^{t'}$, for $t' \geq t$. Assume that the ex-ante (i.e., period $t = 1$) value function is of the von Neumann-Morgenstern Expected-Utility form $U^1(s_1, h^1) = E_{y^t|h^1=\emptyset,s_1} u(z_e(y^t))$, in which the agent's preferences depend on the expected utility of final earnings. This assumption can be relaxed, and we also abstract away from describing the curvature of u. Note that in our experimental setting, the expected returns from playing are negative, and so agents who play at least once are risk-loving over at least some interval. In practice, while this is not the subject of our analysis, we expect the risk attitude to be in part context-dependent, and influenced by the roulette wheel framing. Likewise, we may expect q_f to be context-dependent, which is why we do not perform comparative statics across different

[5]In particular, if there are cumulative probability functions $w^+ : [0,1] \to [0,1]$ for the gains, $w^- : [0,1] \to [0,1]$ for the losses, and a utility function u, then the plan of action would rely on all these functions, as would overplaying. If a change in tenacity were purely about the curvature of any of these functions, it would change both the plan of action and the propensity to overplay, even for the naive agent.

[6]This result would be difficult to explain with a notion of fatigue without making post-hoc assumptions, since those agents play on average less, and not more, than the rest- hence, all else being equal, their fatigue should not be higher. Instead, this is consistent with their playing with the plan.

games.[7] We also assume that u is a continuous and strictly increasing function.[8] Let s_1^* denote the optimal ex-ante strategy. Assume that $z_w = z_l$, so that for any given number of losses, the same number of wins will bring the agent back to the initial amount z_b.

Concerning value functions U^t for histories h^t where $z_e(h^t) \geq z_b$, the value function is fully in line with U^1. Hence the optimal strategy $s_t^*(h^t)$ is fully consistent with the ex-ante strategy $s_1^*(h^t)$, and the action taken, $a_t(s_t^*|h^t)$, remains the same as the planned action $a_t(s_1^*|h^t)$. For histories h^t where $z_e(h^t) < z_b$ instead, the value function may now be different. If the agent had intended to continue at that node (i.e., if $a_t(s_1^*|h^t) = Y$), then the value function U^t for the optimal strategy remains the same, and the strategies remain consistent. If, instead, the planned action was to stop (i.e., $a_t(s_1^*|h^t) = N$), then with probability q_f the value function for that strategy is now $U^t(s_1^*, h^t) = u(z_e(h^t)) - c_f$, which is *not* consistent with the ex-ante value function. Instead, action $a_t = N$ is now strictly dominated by a strategy for which $a_t = Y$, so as to avoid incurring the (large) cost c_f. Hence, for the revised optimal action s_t^*, it must be that $a_t(s_t^*|h^t) = Y$, meaning that the agent is inconsistent (see Figure 1 for an illustration with two periods). This contingency occurs with higher probability for higher T, by Assumption 1.

Within the context of our experiment, this argument leads to the following predictions, which we formally prove in Appendix C. Assume that the Grit Index measures grit G, and that its partitions into the Tenacity and Diligence Indices, which we construct formally in Section 4.3, measure tenacity T and diligence D. We also assume that these indices are not negatively correlated, which is consistent with the data. Overplaying refers to playing past the planned minimum bound, which is the threshold of $z \in [0, z_b]$ at which the agent had decided to stop, according to strategy s_1^*. Note that there is only one such threshold, under our assumptions.

Prediction 1: The probability of overplaying (weakly) increases in the Tenacity Index, and does not change in the Diligence Index, except through its possible correlation with the Tenacity Index.

This first prediction will be tested in Section 4.3.

Prediction 2: The probability of overplaying also increases in the Grit Index.

This prediction follows from Prediction 1 and from attribute G consisting of both T and D, using the assumption that the two are not negatively correlated. It will be tested in Section 4.2. Notice that for both predictions, the comparisons are made between two individuals who have the same plan of action but different levels of tenacity. Since different individuals may have different utility functions u, it is important to control for the plan of action, and specifically the planned minimum bound, in the analysis. In particular, our empirical analysis (Section 4) will estimate how the likelihood of overplaying varies with the Tenacity, Diligence and Grit Indices, controlling for the minimum bound. Our prediction, driven by the assumption $q_f'(T) > 0$, is

[7] Formally, q_f would be enriched to be a function $q_f(T, C)$, where C would be the context (or environment), and may be influenced by factors such as arousal and entertainment. In the case of the first game, we would expect q_f to be very low. But we maintain all our comparative statics across T, holding C fixed.

[8] As mentioned previously, in our stylized model we do not explore whether tenacity links directly to loss aversion in prospect theory or probability weighting, partly because it would be challenging to explain through such a channel why, in our experiment, tenacious subjects do not have significantly different plans of action and do not play more than others, but that instead they overplay more.

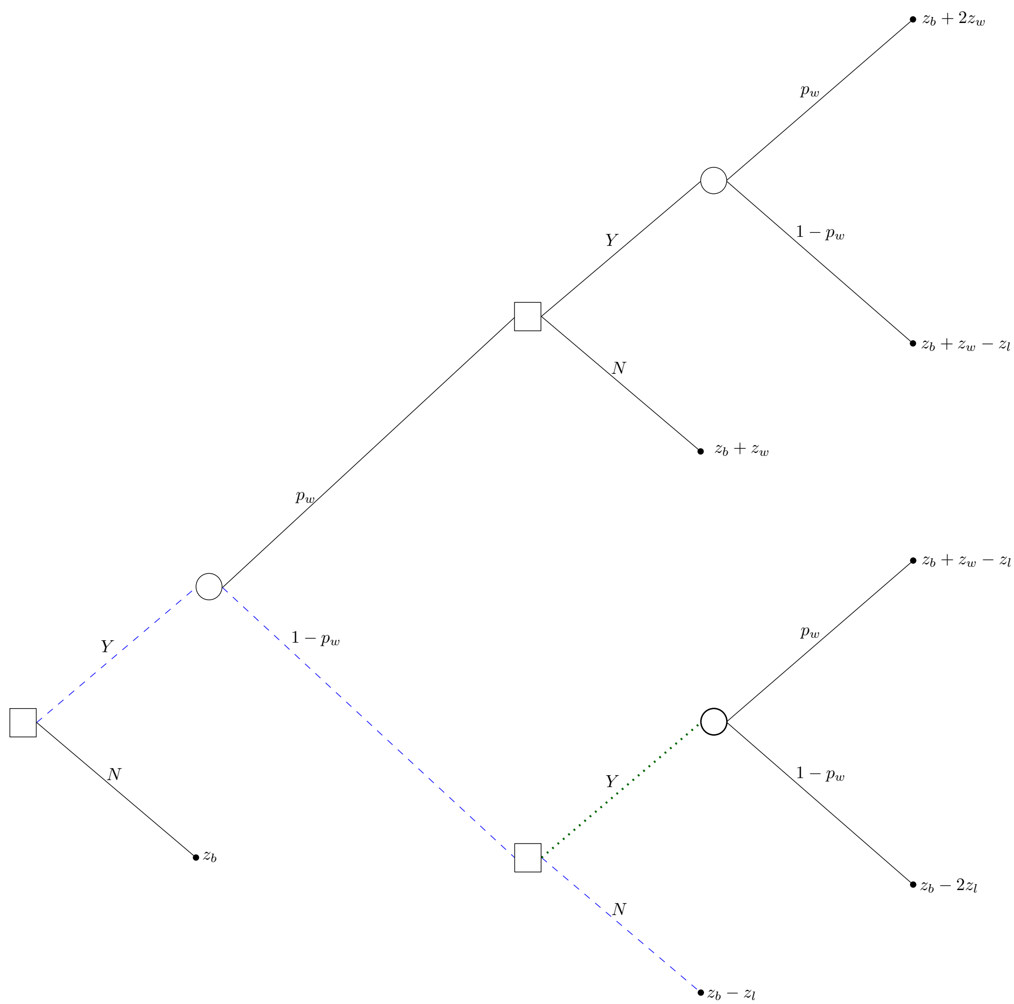


Figure 1: The figure describes the behavior of an agent whose plan of action (s_1^*) is to play once and stop if he loses ($a_1 = Y$ and $a_2(s_1^*|(Y, z_l)) = N$), represented by the blue dashed lines, but who instead deviates from his plan of action to strategy s_2 by playing a second time if he loses ($a_2(s_2|(Y, z_l)) = Y$)), represented by the green dotted line. The squares represent the decision nodes (Y to continue or N to stop) and the circles represent the chance nodes (p_w of winning z_w and $1 - p_w$ of losing z_l). Here, $\tau = 2$.

that the sign of the coefficient is positive for the Tenacity and Grit Indices. The sign of the coefficient for the Diligence Index should be 0 once we control for the Tenacity Index.[9]

Prediction 3: Assume that $z_b - Mz_l = 0$ for some integer M and that $z_w = z_l$. Among agents who overplay, the ones with a higher Tenacity Index are more likely to end with final earnings of either 0 or an amount greater or equal to z_b.

Here as well, the same link would hold for the Grit Index, accounting for the planned minimum bound. This prediction is more difficult to test, however, due to the highly stochastic

[9] In the following sections we use the terms tenacity, diligence and grit interchangeably with the Tenacity, Diligence and Grit Indices, respectively, when it is clear that it is a reference to the empirical measures.

nature of final earnings.

3.3 Upside of grit

The subsections above presented cases in which tenacity would have a clear downside, from the perspective of the decision-maker's own ex-ante preferences. Here, we consider a case in which tenacity may have an upside. This upside takes place when the agent has a force influencing him towards being discouraged when the task gets harder, and quitting earlier than he would have liked to. Higher tenacity then serves as a countervailing effect, encouraging him to pursue the activity. This can be seen, for instance, in the case of present-biased preferences. Given our setting, this can be thought of as an agent who must decide whether or not to pursue a risky strategy for which he may win (z_w) or not (z_l).

We consider the case $\tau = 3$, which suffices for our purposes. Suppose, moreover, that if the agent reaches the last period $t = 3$, it must be that $a_t = N$, so that he must stop the activity if he has not already. Hence, the agent has at most 2 periods of decision making. Suppose that the agent has $\beta-\delta$ preferences (see, e.g., O'Donoghue and Rabin (1999)), setting $\delta = 1$ for simplicity and $\beta \in (0, 1)$. The agent is naive about his present bias. Formally, the agent's value function is $U^1(s_1, h^1) = E_{y^t|h^1,s_1}(u_1 + \beta(u_2 + u_3))$ in period 1 and $U^2(s_2, h^2) = E_{y^t|h^2,s_2}(u_2 + \beta u_3)$ in period 2. At time $t \in \{1, 2\}$, the agent maximizes his value function of period t without accounting for his future inconsistency. The agent's utility u_t is decomposed into the (strictly increasing) value of the earnings v and a cost c, for each period t. The earnings value v depends on his final earnings $z_e(y^t)$, and he obtains utility $v(z_e(y^t))$ in the period following the one in which he stops playing, and 0 beforehand. If he keeps conducting the activity in period t ($a_t = Y$), then he incurs an immediate cost $c_e^t > 0$, in addition to the potential unanticipated cost of failure, c_f, under the maintained assumptions discussed above. In other words, this setting is one in which the cost is immediate, but the reward is delayed. Note that the agent is naive about two components: his present bias β, and the possibility of incurring a cost of failure c_f.

If, according to the optimal period $t = 1$ strategy s_1^*, $a_1(s_1^*) = N$, then the agent stops and is fully consistent. Suppose instead that the optimal strategy is $a_1(s_1^*) = Y$ and $a_2(s_1^*) = Y$, so that the agent's plan is to play twice. In period 2, however, if the agent has not incurred the cost of failure c_f, then it may be that he is inconsistent and chooses to stop, as can occur with present-biased preferences. This is because, from the perspective of period 1, both the value and the period 2 cost are discounted by β, while in period 2, the cost is not discounted.[10] But if $z_e(h^2) < z_b$, then he will incur cost c_f with probability q_f if he does not pursue the activity for another period, in which case he will continue. Since q_f is increasing in T, the following observation therefore follows trivially: the probability of pursuing the activity when $z_e(h^2) < z_b$ increases in T, and is unaffected when $z_e(h^2) \geq z_b$. This observation points to the upside of tenacity. The force (β) that encourages an agent not to respect his ex-ante (period 1) preferences and quit in period 2 is counteracted by the force leading the agent to avoid the cost of failure (c_f).

[10] For instance, let $v(z_e) = z_e$, $c_1 = 0$, $c_2 = 3$, $z_b = 20$, $z_w = z_l = 8$, $p_w = 3/4$ and $\beta = 1/2$. Then, according to s_1^* the agent will keep pursuing the activity in both periods, but in period $t = 2$ he will instead stop.

In brief, therefore, our model points to both the upside and the downside of grit, and specifically of its tenacity component. In essence, tenacity leads the agent to continue pursuing a task for longer when he is losing. This may mean that he continues the task for too long and put himself at risk of losing more, as in our experiment. But in cases in which he may quit too early, then tenacity may push the agent closer to his ex-ante goal.

We note that for clarity we have maintained this setting rather than a more classical procrastination or temptation setting, but our analysis can easily be adapted to such settings as well. Notice that in such settings, however, it is crucial to specify which task the agent is conducting first. If the agent with higher tenacity has already embarked the tempting activity, for example, then he would be more likely to continue. If instead he has embarked on a the non-tempting activity, then he is more likely to continue that task instead. In some cases, therefore, our predictions are distinct from those of present-biased models that do not account for tenacity.

Lastly, we remain agnostic over the relation between tenacity and educational performance, since it is unclear whether the positive or the negative effect will dominate. Concerning diligence, an extension of our model would likely link it to a tendency towards higher educational outcomes, but it is not the object of our analysis and Assumptions 1 and 2 are silent on this point. Hence, even though we conjecture that diligence has a clear upside concerning higher educational outcomes, we do not formally make these theoretical predictions, and defer this analysis to future research.

4 Results

4.1 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.[11]

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.

Table 2 takes the same variables already present in Table 1 and measures their correlation with the Grit Index and its decomposition into the Tenacity and the Diligence Indices. Overplaying, our main dependent variable of interest, correlates positively with the Grit Index (0.29)

[11] The likelihood of overplaying in the second game (35%) is statistically larger than the one of overplaying in the first game (16%).

and especially so with the Tenacity Index (0.33). Also consistent with prior beliefs, the Locus of Control index correlates negatively with the Grit Index (-0.19). Finally, self-reported measures of Procrastination and Temptation correlate strongly and negatively both with the Grit Index and with its facets of tenacity and diligence.

4.2 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.[12] 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.

Table 3 presents the first test of the model's predictions. We begin with the hypothesis that overplaying increases with the aggregate Grit Index (Prediction 2). 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 undergraduate degree.[13] The estimated coefficient on the Grit Index remains stable and strongly statistically significant.[14]

In Table 4 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 5 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,

[12] 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.

[13] 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.

[14] In the Appendix (Table 22) we perform a robustness check by using an expanded measure of overplaying, which we refer to as 'potential overplaying'. In addition to the 48 individuals to actually overplay, we also include the 8 additional ones who placed a bet that would have brought them below their desired minimum, had they lost. For instance, if an individual's desired minimum is 1000 tokens and he places a 200 token bet when he has 1100 tokens remaining, then he has the potential to go to 900, which is below his desired minimum. We therefore would include him in the potential overplaying measure whether or not he subsequently loses. Results are practically unaffected for this expanded measure.

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.

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 3). 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 those who have not (Figure 4).

In summary, the finding that subjects with higher grit are more 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. Our experimental design, therefore, serves to emphasize the importance of separating plan of action from behavior to address our questions of interest.

Difference in present bias or impatience?

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).[15] 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 a robustness check, however, we use the self-reported measures on self-control problems (procrastination and temptation) as controls in Table 6. In column (1), we regress the dummy for overplaying on these two variables alone and do not find that the estimated coefficient is statistically significant. When we add the Grit Index in columns (2) it is positive and statistically significant, while neither the temptation nor procrastination variables are relevant. 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)).

Additional robustness on main specification

Finally, we produce four robustness tests that confirm that our main finding is not driven by

[15] 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.

a small subset of observations or outliers. First, we drop 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 7 we drop the 10 subjects who overplayed by 100 tokens or less.[16] 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)).

Second, we consider alternative econometric specifications in Table 8. 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.[17]

Third, in Table 9 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)).

Fourth, in Table 10 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.

4.3 Decomposing grit into tenacity and diligence

Our explanation for the main finding is that grittier subjects, by being more tenacious, 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, 'diligence' and 'tenacity'. These two components are correlated but distinct.[18] They map to the variables D

[16]We will discuss in Section 5 how the amount by which subjects overplay relates to grit, while noting that it is challenging to define a natural benchmark of comparison for degree of overplaying, because it is highly sensitive to the subject's draw and to parametric assumptions. Despite this difficulty, we still find that subjects who overplay have a higher tendency to either lose all their initial wealth or regain it, and analogously for tenacity. These results are consistent with our hypotheses (Prediction 3). Moreover, we find that agents with higher tenacity overplay more intensely both for number of rounds and for amount played after reaching the minimum bound, , which is also consistent with our hypotheses.

[17]In the Appendix (Table 23) we perform an additional robustness test where we rerun the regressions of Table 8 with 'potential overplaying', the expanded measure of overplaying defined previously (Footnote 14), and results are again practically unaffected.

[18]The correlation coefficient is 0.56 among our 138 participants.

and T, respectively, in our model. According to Prediction 1, tenacity alone should explain overplaying. Intuitively, recalling the image of the rowboat, the unwillingness to accept defeat and stop playing once the planned minimum bound is reached is a refusal to steer away. It therefore falls squarely within tenacity, not diligence.

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 11). The grit questions are cleanly 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 12 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 3). Columns (3) and (4) provide sample splits by degree and columns (5) and (6) by gender, analogously to Table 10. Column (7) checks for robustness to overplaying by more than 100 tokens, analogously to column (3) of Table 7. 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 remains close to zero for all columns. These findings jointly support the hypothesis that it is the tenacity component of grit that drives overplaying.[19]

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 5), while the analogous kernel density estimation for diligence shows very little difference (Figure 6). This provides additional visual assurance that the strong positive relationship between grit and overplaying is through tenacity.

We then ensure that our categorization is robust and not subjective in two different way. The first consists of performing a confirmatory factor analysis (CFA) on the split of grit into tenacity and diligence, and the second consists of considering alternative specifications of this split.

The CFA is provided in the Appendix (Table 24). Its indices provide empirical evidence of fit between the data and the model, thus serving as a validation of measurements of latent construct. Since our split of the Grit Index builds on a specific theory, the factorial validity is tested via a theory-driven CFA rather than a data-driven explanatory factor analysis. Following the standard procedure, we allow for the within-factor statistically significant error covariances (specifically, "I finish whatever I begin" with "New ideas and projects sometimes distract me from previous ones," "I finish whatever I begin" with "I have been obsessed with a certain idea or project for a short time but later lost interest," and "I am a hard worker" with "I am diligent").

To test alternative specifications, we remove questions whose categorization carries any hint

[19]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.

of ambiguity. Our results remain unchanged to these 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 ("Setbacks 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 (Table 25 in the Appendix). 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 perseverance and consistency of interest.

Alternative decomposition into perseverance and consistency of interest

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 or tenacity category. Diligence only concerns working hard and not the reaction to obstacles faced, and therefore neither fully fits within perseverance nor consistency of interest. Tenacity focuses on refusal to steer away, not the strength of the rowing. It is also not obvious, from this image, whether either perseverance or consistency are natural candidate to explain overplaying. The difference in what these categories aim to capture is naturally reflected in the distinct partitioning. For instance, "I am diligent" and "Setbacks don't discourage me" both fall within perseverance in the standard classification. But the first falls within diligence, as it concerns the strength of rowing, while the second falls within tenacity, as it concerns the refusal to steer away.[20]

To compare the explanatory power of our categorization to the standard one, Table 13 replicates the same regressions as Table 12, 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 (4)), and in others neither is significant (e.g. columns (6) and (7)).

5 Further Results

5.1 Playing behavior after hitting the minimum bound

In Table 14, the dependent variable is the number of rounds played after reaching the minimum bound. Column (1) clearly shows that a higher value in the Tenacity Index is associated with

[20]To get a sense of the partial correlations across these variables, we first regress diligence on consistency and perseverance; we obtain an estimated coefficient of 0.37 for the former and 0.71 for the latter. We then set tenacity as the dependent variable, and obtain an estimated coefficient of 0.57 for consistency and 0.37 for perseverance. All four estimated coefficients are highly statistically significant.

overplaying more rounds. In other words, the intensity with which a player deviates from his plan of action increases with tenacity, even after controlling for one's plan of action and the Diligence Index. This baseline result holds with a slightly lower estimated coefficient after adding our usual controls (age, gender, education). Interestingly, this overplaying behavior is not driven by a general tendency to play more throughout the game, as we see from column (3). There, we additionally control for the number of rounds played *before* reaching the plan of action, which is not statistically significant with an estimated coefficient close to zero. In sum, tenacity predicts a larger degree of overplaying even accounting for actual gaming patterns before reaching the minimum bound. Column (4) shows that the result is robust to dropping the individual's plan of action; a one unit increase in the Tenacity Index leads to a 6 additional rounds of overplaying.

Table 15 finds similar results to the previous table after substituting the dependent variable to the (log of) amount played after reaching the minimum bound. We again find that it is tenacity, and not diligence, that strongly correlates with the amount overplayed, and that this variation is still present even after accounting for either the plan of action or the (log of) amount played before reaching the minimum bound.

5.2 Procrastination, temptation and self-esteem

In this subsection, we explore 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 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.[21]

In Table 16 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. The coefficient for tenacity is also negative, but not significant. These results are consistent with the notion that diligence has a clear upside, while the effect of tenacity is more ambiguous. This again is suggestive of the tension present between the upside and the downside of tenacity. For the case of temptation, for instance, there may well be an upside in keep focused on the task if it has already begun, but there is a temptation of sticking with the tempting activity if it is in motion as well. Exploring this tradeoff, and the net effect, is left for future research.

We follow the same sequence in Table 17, and regress self-esteem on grit (column (1)), tenacity (column (2)), diligence (column (3)), and both tenacity and diligence (column (4)).

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

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 16, with the sign reversed, which again confirms our view on the upside of diligence and the ambiguity of tenacity. 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.

5.3 Admitting having overplayed after the game

We close this section by discussing an additional survey question given to the subjects, near the end of the experiment. This question asks whether they have overplayed in the roulette or not, where the dummy for saying they have overplayed takes value of 1 if they answered yes, and 0 otherwise. We then compare this answer to whether the agents did in fact overplay, using the measure used throughout this paper of comparing the ex-ante plan of action to actual behavior. In other words, we analyze whether agents 'admit' to having overplayed, by their own measure (Table 18). Our theory and hypotheses do not, in themselves, make predictions on how this question is answered. A reasonable conjecture, however, is that while less stubborn (or tenacious) subjects would admit having overplayed, more stubborn would not. This is consistent with a view that tenacity also involves being stubborn in accepting having gone too far.

We first see in column (1) that the coefficient for overplaying is positive and highly significant, which confirms that overall, subjects do admit having overplayed. Interestingly, and consistent with the conjecture above, column (2) reveals that when tenacity is included, the coefficient for overplaying remains positive and highly significant, while the coefficient for tenacity is negative and significant. That is, more tenacious subjects refuse to admit right after the game that they did overplay. In columns (3) and (4) we add diligence and the plan of action to the regression, and find that neither of them is significant, while the coefficients for overplaying remains positive and highly significant, and the coefficient for tenacity remains negative and significant. Moreover, the same finding holds in column (5) when we condition the sample on overplayers. These results suggest that in addition to tenacity involving refusal to let go through incurring a cost of failure, it may also involve stubbornness in accepting having gone too far, even when the metric used is the subject's own ex-ante preference. Exploring this channel is left to future research.

6 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 in two different datasets. The first dataset, discussed in

Section 6.1 consists of a survey on educational outcomes provided by the Interuniversity Consortium for Political and Social Research (ICPSR). The second, discussed in Section 6.2, consists of an online survey on grit, conscientiousness, and educational attainment.

6.1 ICPSR dataset on educational outcomes

This subsection uses large-scale data from the Measures of Effective Teaching (MET) Project (Bill and Melinda Gates Foundation 2014) to assess how our new categorizations of grit into tenacity and diligence relate to educational outcomes. The MET study provides information between 2009 and 2011 across six districts: Charlotte-Mecklenburg (North Carolina) Schools, Dallas (Texas) Independent School District, Denver (Colorado) Public Schools, Hillsborough County (Florida) Public Schools, Memphis (Tennessee) City Schools, and the New York City (New York) Department of Education.

This dataset is the result of an expansive effort to capture and analyze a variety of measures that could be used to evaluate teachers and provide them with feedback about their professional practice. Standard administrative and achievement data was collected, and students also completed a questionnaire that includes precisely the same eight questions on grit used in our experiment.

The scale of the MET Project affords a unique opportunity to examine variation across grades, subgroups and classrooms. Although the data collected by the MET Project are not a random sample of the national population, they are broadly representative of the population of urban public school students across much of the United States. For our sample of students in Grades 6 to 9, we focus on scores on the standardized ACT test (N=5,733), the standardized SAT9 test (N=5,875), and a standardized Math test (N=18,088) as outcome variables. Together, these tests span a broad range of topics, from more mathematical and technical ones to more reading-oriented ones.

Table 19 summarizes the results on these three dependent variables. Column (1) presents results with the ACT Quality Core test score as the dependent variable and the Grit Index as the main variable of interest. The estimated coefficient of 1.641 is statistically significant and in economic terms it means that a one unit increase in the Grit Index is associated with an increase in the dependent variable of one fifth of a standard deviation. In the next column we decompose the index into its Diligence and Tenacity components. Both indices are statistically significant and positively correlate with the ACT test score. Similar statistical and economic results are found when we use instead the SAT9 test score as the dependent variable (columns (4)-(6)) or the average of a student's math scores in the period 2009-2011 (columns (7)-(9)). These results show that both tenacity and diligence are indicative of higher educational performance in various domains.

6.2 Cross-country online survey of educational attainment

This dataset consists of a sample of 3988 individuals which includes data on grit, conscientiousness, and education as well as other demographics.[22] 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.

While we focus on grit because it incorporates the tenacity category which is central to our hypothesis, grit and conscientiousness have been found to be highly related, which raises the point of whether our results hold with conscientiousness as well. For this reason, in Table 20 we use the online survey to study the connection between the different categories of grit and conscientiousness. Using conscientiousness as the dependent variable, we first confirm it is highly correlated with grit, in both an economic and statistical sense. When decomposing grit into diligence and tenacity, we again find that both are highly positively correlated, but that tenacity does not entirely capture conscientiousness. This suggests that while we do expect that our main findings on the flip side of tenacity may be present in conscientiousness as well, the two are sufficiently distant that understanding more precisely the degree to which our results extend to conscientiousness requires further exploration.

We now document how educational outcomes are explained by the Conscientiousness Index and Grit Index, and the split into Diligence and Tenacity Indices (Table 21). Columns (1)-(5) run the basic specifications, while columns (6)-(10) include country fixed effects and control for age, gender, racial background, and urbanization. In general, the Conscientiousness and Grit Indices are positive and statistically significant on their own and together (columns (1)-(3) and (6)-(8)).

When regressing education on conscientiousness and diligence and tenacity, all three are significant without controls (column (5)), and diligence remains significant with controls as well. Conscientiousness is close to being significant, while the coefficient for tenacity effectively drops to zero (column (10)). The same pattern remains for diligence and tenacity without conscientiousness: they are both significant without controls, and only diligence is significant with controls.

These results confirm the upside of diligence discussed in our analysis, and shows that tenacity is more nuanced. The interplay between the upside and the downside of tenacity is likely to be intricate in a setting as complex and multidimensional as educational attainment. But disentangling the two sides would be of interest in future research. Our results are also in line with

[22] 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.

the findings in the literature that conscientiousness is predictive of educational performance, and at the same time show that our new categories are independently informative of performance as well.

7 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, 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. The upside of tenacity may well outweigh its downside in many contexts, but in others it may prove costly. There may also be other potential downsides of grit; Duckworth and Eskreis-Winkler (2015), for instance, speculate that those who are more likely to stay the course may also ignore new opportunities. This tension should be investigated further to shed light on addressing the flip side of grit and 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 log of 1 plus the planned minimum bound, where the planned minimum bound ranges 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	6.21	1.73	0	7.6
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
Technical Degree	138	.51	.50	0	1
Self-esteem	138	3.14	.99	1	5
Procrastination	118	3.29	.80	1	5
Temptation	118	3.31	.85	1	5

Table 2: Descriptive statistics - Correlations with Grit, Tenacity, and Diligence

Note: Here we take the same variables previously defined in Table 1 and show the correlation coefficients with respect to: (i) Grit Index; (ii) Tenacity Index; (iii) Diligence Index.

Variable	Observations	Correlation with:		
		Grit Index	Tenacity Index	Diligence Index
Overplaying	138	.29	.33	.16
Plan of Action	138	.06	.07	.02
Locus Index	138	-.19	-.19	-.14
Age	138	.16	.16	.11
Female	138	.07	.07	.04
Technical Degree	138	.08	.00	.17
Self-esteem	138	.32	.24	.33
Procrastination	118	-.45	-.35	-.45
Temptation	118	-.36	-.30	-.34

Table 3: 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 1 plus the planned minimum bound. Standard errors are robust to heteroskedasticity. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep.Var.: 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)
Observations	138	138	138	138	138
R-squared	0.087	0.145	0.150	0.165	0.174

Table 4: Different levels of final gains

Note: This table reports regressions similar to those in column (5) of Table 3. 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 3. 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)
Observations	133	93	40
R-squared	0.173	0.199	0.525

Table 5: Different plans of action

Note: This table reports regressions similar to those in column (5) of Table 3. 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 3. 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)
Observations	129	109	70
R-squared	0.141	0.132	0.155

Table 6: Adding temptation and procrastination as control variables

Note: This table resembles the baseline Table 3, 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 3.

	(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)
Observations	118	118	118	118
R-squared	0.008	0.099	0.158	0.226

Table 7: Overplaying >100 tokens

Note: This table uses some of the specifications of Table 3, 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)
Observations	128	128	128
R-squared	0.057	0.129	0.171

Table 8: 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 3. Columns (3) and (4) proceed similarly with a Probit estimator. Finally, columns (5) and (6) use a Poisson specification. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep.Var.: Overplaying	(1) Logit	(2) Logit	(3) Probit	(4) Probit	(5) Poisson	(6) 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)
Observations	138	138	138	138	138	138

Table 9: Adding Locus Index as a control variable

Note: This table resembles the baseline Table 3, 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: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Overplaying	(1)	(2)	(3)	(4)
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)
Observations	138	138	138	138
R-squared	0.018	0.093	0.149	0.175

Table 10: 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)
Observations	71	67	80	58
R-squared	0.117	0.242	0.157	0.205

Table 11: Decomposition of Grit Index into tenacity and diligence

Note: The table below decomposes the 8 questions of the Grit Index and allocates each question into a cell. The rows are defined by our proposed split of the index: Tenacity and Diligence. The columns are defined by the split suggested by Duckworth and Quinn (2009): Perseverance and Consistency. Furthermore, the four questions in italics are the ones kept in our main robustness indices of tenacity and diligence.

	Perseverance	**Consistency**
Tenacity	*Setbacks don't discourage me.* *I finish whatever I begin.*	New ideas and projects sometimes distract me from previous ones. I often set a goal but later choose to pursue a different one. I have been obsessed with a certain idea or project for a short time but later lost interest.
Diligence	*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 12: Splitting Grit Index into tenacity and diligence

Note: This table builds on Table 3, 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 3. Column (2) adds the full battery of control variables, as in column (5) on Table 3. Columns (3) to (6) split the data by educational degree and gender, similar to Table 10. 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.: Overplaying	(1)	(2)	(3) Technical	(4) Non-technical	(5) Female	(6) Male	(7) >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)
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 13: Splitting Grit Index to consistency and perseverance

Note: This table builds on Table 3, 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 3. Column (2) adds the full battery of control variables, as in column (5) on Table 3. Columns (3) to (6) split the data by educational degree and gender, similar to Table 10. 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.: Overplaying	(1)	(2)	(3) Technical	(4) Non-technical	(5) Female	(6) Male	(7) >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)
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 14: Number of Rounds Played After Reaching Minimum Bound

Note: The dependent variable is the number of rounds played after reaching the minimum bound and we have split grit into the Tenacity Index and the Diligence Index. All four specifications use the full sample, and we gradually add control variables. In addition to the usual control variables seen in previous tables, in the last two columns we also include the number of rounds played before reaching the minimum bound. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Rounds after Minimum Bound	(1)	(2)	(3)	(4)
Tenacity Index	6.580**	5.735**	5.648*	6.276**
	(2.729)	(2.774)	(2.914)	(3.075)
Diligence Index	-3.696	-2.754	-2.872	-3.200
	(2.487)	(2.446)	(2.475)	(2.629)
Plan of Action	2.251***	2.344***	2.301***	
	(0.692)	(0.746)	(0.723)	
ln(Age)		3.836	3.743	2.529
		(10.425)	(10.413)	(10.630)
D(Female)		0.835	0.743	1.772
		(2.559)	(2.574)	(2.780)
Technical Degree		-5.416**	-5.270**	-4.494*
		(2.546)	(2.447)	(2.418)
Rounds Before Minimum Bound			-0.029	-0.039
			(0.043)	(0.045)
Constant	-15.005	-25.760	-23.900	-7.495
	(9.361)	(32.122)	(32.116)	(31.614)
Observations	138	138	138	138
R-squared	0.088	0.116	0.119	0.068

Table 15: Amount Played After Reaching Minimum Bound

Note: The dependent variable is the log of the amount played after reaching the minimum bound and we have split grit into the Tenacity Index and the Diligence Index. All four specifications use the full sample, and we gradually add control variables. In addition to the usual control variables seen in previous tables, in the last two columns we also include the log of the amount played before reaching the minimum bound. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Log of Amount after Minimum Bound	(1)	(2)	(3)	(4)
Tenacity Index	2.016***	1.961***	1.668***	1.961***
	(0.521)	(0.524)	(0.572)	(0.524)
Diligence Index	-0.366	-0.243	-0.201	-0.243
	(0.498)	(0.525)	(0.517)	(0.525)
Plan of Action	0.536***	0.585***	0.471***	0.585***
	(0.101)	(0.105)	(0.099)	(0.105)
ln(Age)		0.637	0.999	0.637
		(2.042)	(2.005)	(2.042)
D(Female)		-1.014*	-1.074*	-1.014*
		(0.598)	(0.590)	(0.598)
Technical Degree		-0.675	-0.504	-0.675
		(0.614)	(0.599)	(0.614)
Log of Amount Before Minimum Bound			-0.407***	
			(0.131)	
Constant	-5.978***	-7.578	-4.332	-7.578
	(1.815)	(6.312)	(6.184)	(6.312)
Observations	138	138	138	138
R-squared	0.165	0.189	0.238	0.189

Table 16: 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)
	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)
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 17: 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)
Observations	138	138	138	138
R-squared	0.099	0.060	0.108	0.113

Table 18: The Effect of Stubbornness on Admitting Overplaying

Note: In this table the dependent variable is based on a question posed to the students after the games ended. It is a dummy taking value 1 if the student claims to have overplayed during the Roulette game. Column (1) includes the actual Overplaying variable already used in previous tables. Column (2) adds the Tenacity Index and column (3) also includes the Diligence Index. Column (4) also adds the Plan of Action and, finally, column (5) limits the sample to the 56 students who actually overplayed in the second game. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Admit Overplaying	(1)	(2)	(3)	(4)	(5)
Overplaying	0.331***	0.373***	0.372***	0.401***	
	(0.082)	(0.079)	(0.080)	(0.082)	
Tenacity Index		-0.152**	-0.196**	-0.193**	-0.359***
		(0.069)	(0.086)	(0.086)	(0.121)
Diligence Index			0.065	0.062	0.128
			(0.069)	(0.069)	(0.098)
Plan of Action				-0.033	0.022
				(0.025)	(0.073)
Constant	0.366***	0.838***	0.741***	0.933***	1.274**
	(0.054)	(0.225)	(0.240)	(0.292)	(0.623)
Observations	138	138	138	138	56
R-squared	0.105	0.134	0.139	0.151	0.141

Table 19: ICPSR - Effect of Grit on Education

Note: The three dependent variables correspond to test scores obtained in the ACT Quality Core, the SAT9 exam, and the maths exams between 2009 and 2011. Columns (1)-(3) have the ACT Quality Core test score as the dependent variable. In the first column, the main variable of interest is the Grit Index. In the next two column, the index is decomposed into the Diligence Index and the Tenacity Index, the former including district fixed effects and the later rather including school fixed effects. Columns (4)-(6) replicate the same structure for the SAT9 exam, and columns (7)-(9) do it for the maths exam grade. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Test Scores	ACT	ACT	ACT	SAT9	SAT9	SAT9	Maths	Maths	Maths
Grit Index	1.641***			7.993***			0.166***		
	(0.175)			(0.774)			(0.009)		
Diligence Index		1.079***	1.058***		3.746***	4.133***		0.066***	0.068***
		(0.167)	(0.157)		(0.718)	(0.681)		(0.008)	(0.008)
Tenacity Index		0.540***	0.403**		4.244***	3.781***		0.099***	0.091***
		(0.193)	(0.176)		(0.795)	(0.764)		(0.009)	(0.009)
District FE	Y	Y	N	Y	Y	N	Y	Y	N
School FE	N	N	Y	N	N	Y	N	N	Y
Gender Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age Control	Y	Y	Y	Y	Y	Y	Y	Y	Y
Race Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Additional Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	4,551	4,551	4,551	4,625	4,625	4,625	14,465	14,465	14,465
R-squared	0.228	0.230	0.339	0.319	0.322	0.413	0.333	0.333	0.407

Table 20: Online Survey - Effect of Grit on Conscientiousness

Note: In this table the dependent variable is the level of conscientiousness, ranging from 1 (lowest) to 5 (highest). Column (1) includes subjects from all over the world in a specification that only has the Grit Index and column (2) repeats the same specification with the Diligence and Tenacity Indices. Column (3) adds country fixed effects, and column (4) also controls for gender, race, and whether the subjects lives in an urban area. Column (5) and (6) limit the data to US respondents, without and with control variables, respectively. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Conscientiousness	(1)	(2)	(3)	(4)	(5) US only	(6) US only
Grit Index	0.617***					
	(0.012)					
Diligence Index		0.382***	0.388***	0.377***	0.404***	0.386***
		(0.014)	(0.014)	(0.015)	(0.021)	(0.021)
Tenacity Index		0.235***	0.231***	0.228***	0.221***	0.222***
		(0.015)	(0.016)	(0.016)	(0.022)	(0.022)
ln(Age)				0.126***		0.119***
				(0.025)		(0.032)
Gender Dummies	N	N	N	Y	N	Y
Race Dummies	N	N	N	Y	N	Y
Urban Dummies	N	N	N	Y	N	Y
Country FE	N	N	Y	Y	N	N
Observations	3,988	3,988	3,951	3,951	2,014	2,014
R-squared	0.410	0.430	0.453	0.458	0.417	0.424

Table 21: Online Survey - Effect of Conscientiousness and Grit on Education

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 Conscientiousness Index. Column (2) only includes the Grit Index, while Column (3) includes both items. Column (4) splits the Grit Index into the Diligence and Tenacity Indices, and Column (5) additionally incorporates the Conscientiousness Index. Columns (6)-(10) replicate the same regressions after having added controls for age, gender, race, and whether the subjects live in an urban area, in addition to adding country fixed effects. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep.Var.: Education	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Conscientiousness Index	0.227***		0.114***		0.102***	0.089***		0.039*		0.027
	(0.019)		(0.025)		(0.025)	(0.016)		(0.021)		(0.021)
Grit Index		0.240***	0.170***				0.101***	0.078***		
		(0.018)	(0.024)				(0.015)	(0.019)		
Diligence Index				0.158***	0.118***				0.100***	0.089***
				(0.021)	(0.023)				(0.017)	(0.019)
Tenacity Index				0.083***	0.059**				0.004	-0.002
				(0.023)	(0.024)				(0.019)	(0.019)
Gender, Race, Urban FE	N	N	N	N	N	Y	Y	Y	Y	Y
Country FE	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	3,988	3,988	3,988	3,988	3,988	3,951	3,951	3,951	3,951	3,951
R-squared	0.035	0.042	0.047	0.044	0.048	0.435	0.437	0.437	0.439	0.439

Figure 2: 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 3: 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).

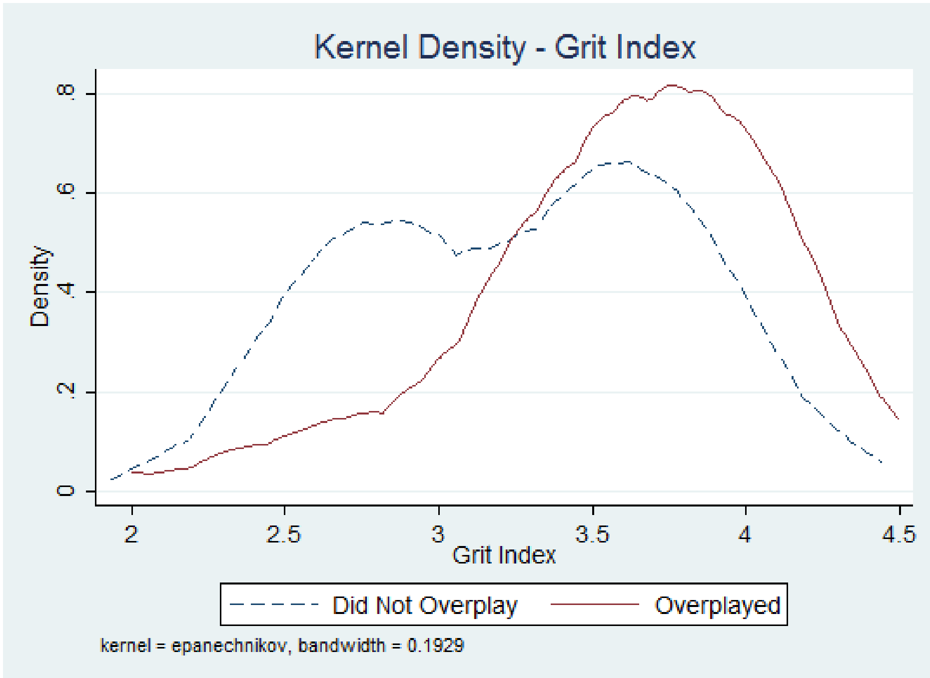


Figure 4: Cumulative Density Function - Grit Index

Note: Building on Figure 3, 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).

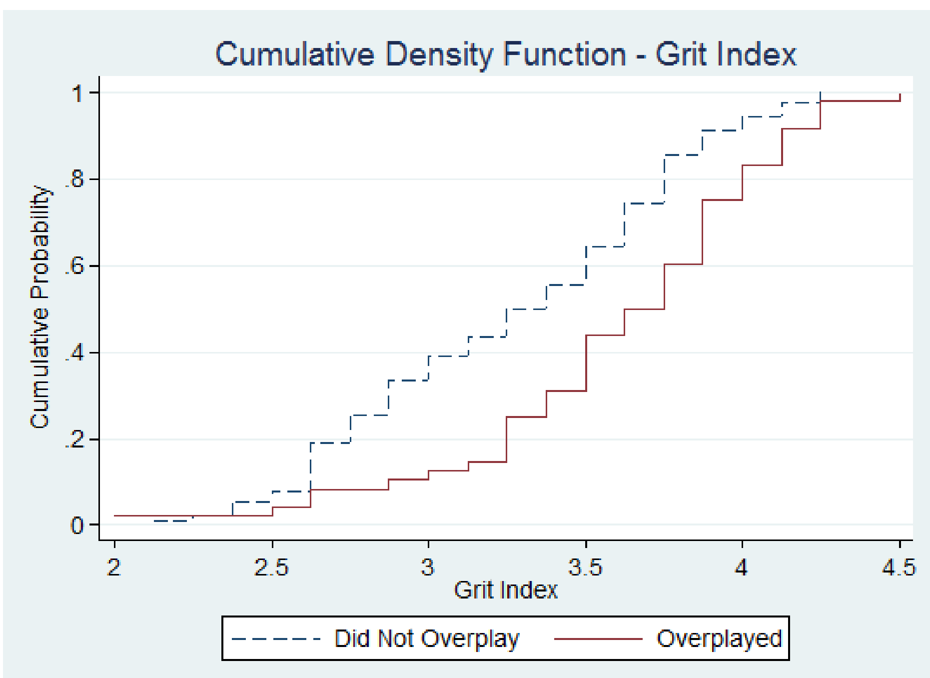


Figure 5: Kernel Density - Tenacity Index

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).

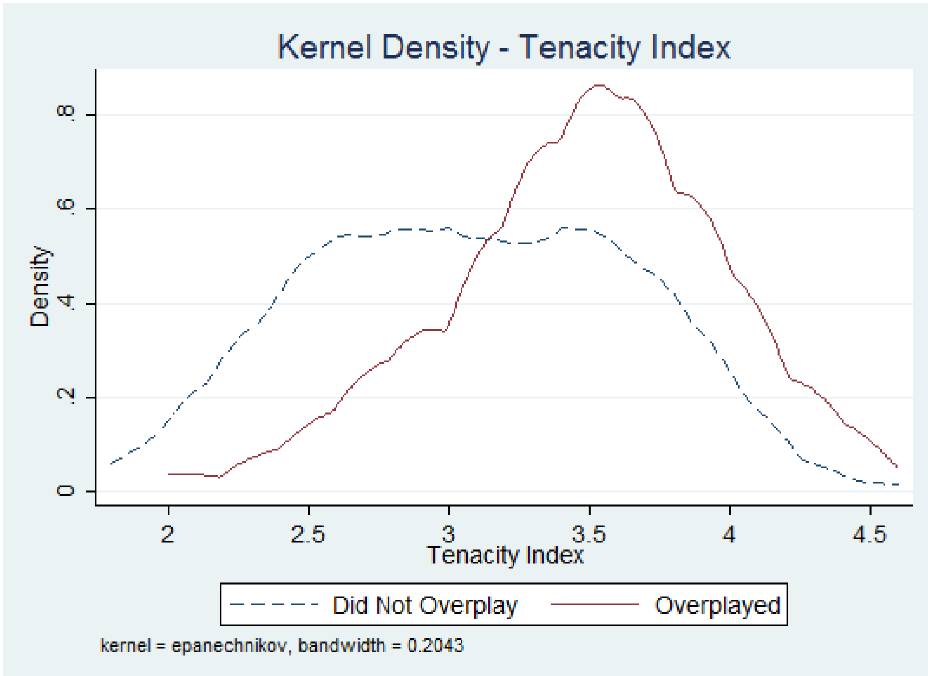
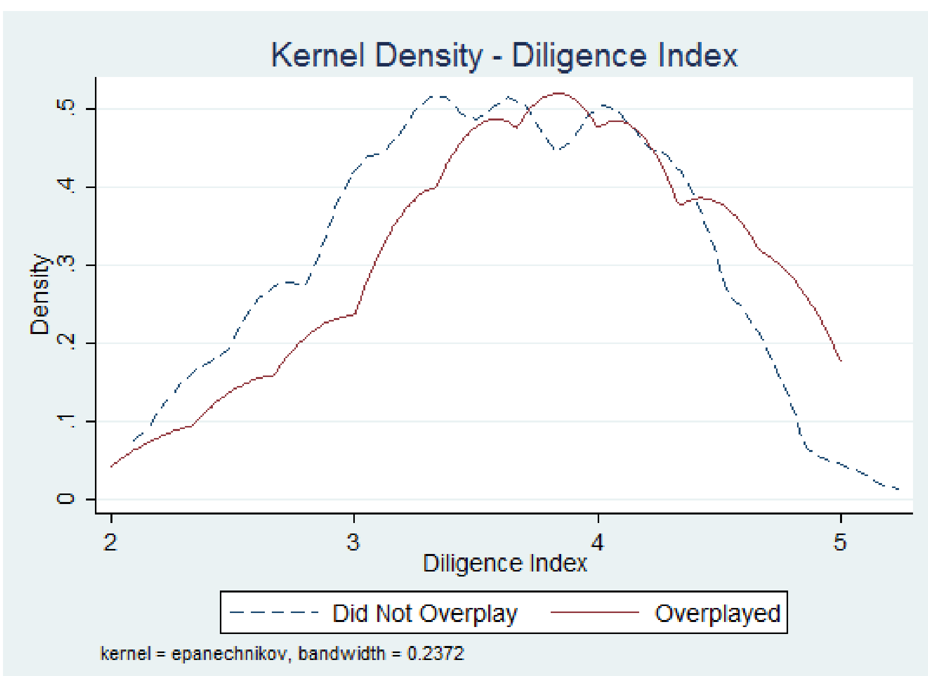


Figure 6: Kernel Density - Diligence Index

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] Achtziger, A., Bayer, U. C., and Gollwitzer, P. M. Committing to implementation intentions: Attention and memory effects for selected situational cues. *Motivation and Emotion 36* (2012), 287–300.

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

[3] 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).

[4] 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.

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

[6] Amador, M., Werning, I., and Angeletos, G.-M. Commitment vs. flexibility. *Econometrica 74*, 2 (2006), 365–396.

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

[8] Andrade, E. B., and Iyer, G. Planned versus actual betting in sequential gambles. *Journal of Marketing Research 46*, 3 (2009), 372–383.

[9] 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.

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

[11] Barberis, N. A model of casino gambling. *Management Science 58*, 1 (2012), 35–51.

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

[13] Benabou, R., and Tirole, J. Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives 30*, 3 (2016), 141–164.

[14] 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.

[15] Boyce, C. J., Wood, A. M., and Brown, G. D. The dark side of conscientiousness: Consciencious people experience greater drops in life satisfaction following unemployment. *Journal of Research in Personality 44* (2010), 535–539.

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

[17] 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.

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

[19] Crede, M., Tynan, M., and Harms, P. Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology* (2016).

[20] 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).

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

[22] 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.

[23] Duckworth, A., and Eskreis-Winkler, L. Grit. In *International Encyclopedia of the Social and Behavioral Sciences. 2nd edition*, J. D. Wright, Ed. Elsevier, 2015.

[24] 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.

[25] 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.

[26] 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.

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

[28] Ebert, S., and Strack, P. Until the bitter end: On prospect theory in a dynamic context. *American Economic Review 105*, 4 (2015), 1618–1633.

[29] Ebert, S., and Strack, P. Never, ever getting started: On prospect theory without commitment. *Mimeo* (2018).

[30] Fischbacher, U., Hoffmann, G., and Schudy, S. The causal effect of stop-loss and take-gain orders on the disposition effect. *Review of Financial Studies 30*, 6 (2017), 2010–2129.

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

[32] Gerhards, L., and Gravert, C. Grit trumps talent? An experimental approach. *Mimeo* (2017).

[33] Gill, D., and Prowse, V. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy 126*, 4 (2016), 1619–1676.

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

[35] Gollwitzer, P. M. Implementation intentions: Strong effects of simple plans. *American Psychologist 54* (1999), 493–503.

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

[37] 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.

[38] Heineck, G. Does it pay to be nice? personality and earnings in the uk. *ILR Review 64*, 5 (2011), 1020–1038.

[39] Imas, A. The realization effect: Risk-taking after realized versus paper losses. *American Economic Review 106*, 8 (2016), 2086–2109.

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

[41] 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).

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

[43] Lucas, G., Gratch, J., Cheng, L., and Marsella, S. When the going gets tough: Grit predicts costly perseverance. *Journal of Research in Personality 59* (2015), 15–22.

[44] MaCann, C., Duckworth, A., and Roberts, R. Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences 19* (2009), 451–458.

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

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

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

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

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

[50] Papageorge, N., Ronda, V., and Zheng, Y. The economic value of breaking bad: Misbehavior, schooling and the labor market. *Mimeo* (2017).

[51] Ploner, M. Hold on to it? An experimental analysis of the disposition effect. *Judgment and Decision Making 12*, 2 (2017), 118–127.

[52] Proto, E., Rustichini, A., and Sofianos, A. Intelligence, personality and gains from cooperation in repeated interactions. *Journal of Political Economy Forthcoming* (2018).

[53] 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.

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

[55] Rueter, A., Abram, S., III, A. M., and Rustichini, A. The goal priority network as a neural substrate of conscientiousness,. *Human Brain Mapping Forthcoming* (2018).

[56] Santos-Pinto, L., and Sobel, J. A model of positive self-image in subjective assessments. *American Economic Review 95*, 5 (2005), 1386–1402.

[57] 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.

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

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

Appendix (For Online Publication)

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 one of the games (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 main 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.
Talking 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 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 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 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 enter the minimum and maximum tokens.]

***

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.
The mechanism used is as follows. After game 1, once the subjects have chosen their plan of action in game 1, they are then asked the following. "Consider the case in which you are given one of the following two options. Option 1: You are given 2000 tokens and you would have to play exactly 20 rounds (unless you win before). Option 2: You are given 2000 tokens and you would play according to the preferences you indicated, namely [quantity chosen by the subject] (unless you win before). How many tokens would you need to receive to prefer option 1? Remember that the trading mechanism is always the one that you have seen at the start of the experiment, and that the best is to choose the amount that corresponds to your preferences." After game 2, the analogous question is asked for the setting: "Consider the case in which you are given one of the following two options. Option 1: You are given 2000 tokens and you would have to play until you lose all your tokens or win the maximum amount (8000 tokens). Option 2: You are given 2000 tokens and you will play according to the range that you have just indicated. How many tokens would you need to receive to prefer option 1? Remember that the trading mechanism is always the one that you have seen at the start of the experiment, and that the best is to choose the amount that corresponds to your preferences." Notice that subjects should anticipate the question when providing their plan of action since it follows the question of game 1, and that it is incentive compatible for them to give the correct plan beforehand. More precisely, it is incentive compatible both to give the preferred plan of action, and then the correct willingness to accept (for option 1). As mentioned above, however, from the complexity of the question we do not think that the willingness to trade is reliable, but it is not relevant for our purposes.

Before moving on to the second phase of the experiment (the actual games), we ask a variety of additional questions to reduce the salience of the previous ones. 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? (*The subjects had the choice of leaving this question blank.*)

***

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 wished. The game can be found at: http://experimentalgames.upf.edu/roulette/ ; we include snapshots above in Figure 2. 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 you 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.

B Further Robustness Estimations

Table 22: Appendix - Robustness to Baseline Regression

Note: An Ordinary Least Squares specification is used. The dependent variable is a dummy variable taking the value 1 in case the subject overplayed or could potentially have overplayed (see Footnote 14). Grit Index takes values between 1 and 5 and increases with the level of grit. Plan of Action is the log of 1 plus the planned minimum bound. Standard errors are robust to heteroskedasticity. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dep. Var.: Potential Overplaying	(1)	(2)	(3)	(4)	(5)
Grit Index	0.209***	0.197***	0.185**	0.202***	0.212***
	(0.074)	(0.072)	(0.073)	(0.071)	(0.072)
Plan of Action		0.070***	0.072***	0.078***	0.080***
		(0.012)	(0.012)	(0.014)	(0.014)
ln(Age)			0.278	0.143	0.072
			(0.269)	(0.256)	(0.254)
D(Female)				-0.202**	-0.231***
				(0.083)	(0.084)
Technical Degree					-0.086
					(0.083)
Constant	-0.301	-0.696***	-1.519*	-1.083	-0.852
	(0.252)	(0.248)	(0.836)	(0.802)	(0.794)
Observations	138	138	138	138	138
R-squared	0.053	0.114	0.121	0.160	0.166

Table 23: Appendix - Robustness to alternative econometric specifications

Note: This tables uses alternative econometric specifications. The dependent variable is a dummy variable taking the value 1 in case the subject overplayed or could potentially have overplayed (see Footnote 14). Columns (1) and (2) uses a Logit specification, where the latter column is otherwise equivalent to column (5) of Table 3. 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.: Potential Overplaying	Logit	Logit	Probit	Probit	Poisson	Poisson
Grit Index	0.920**	0.997**	0.559***	0.611***	0.548***	0.504**
	(0.359)	(0.394)	(0.215)	(0.230)	(0.211)	(0.202)
Plan of Action		0.537***		0.321***		0.330***
		(0.165)		(0.088)		(0.096)
ln(Age)		0.451		0.272		0.237
		(1.244)		(0.775)		(0.460)
D(Female)		-1.089***		-0.672***		-0.521***
		(0.420)		(0.254)		(0.200)
Technical Degree		-0.458		-0.270		-0.248
		(0.409)		(0.246)		(0.204)
Observations	138	138	138	138	138	138

Table 24: Appendix - Confirmatory Factor Analysis (CFA) - split of the grit index

Note: F1 includes the questions assigned to the Tenacity Index; F2 includes the questions assigned to the Diligence Index.

Item Content	F1	F2	Mean	Std. Dev.
Setbacks don't discourage me	0.099		2.898	0.961
I finish whatever I begin	0.578		3.275	1.016
New ideas and projects sometimes distract me from previous ones	0.592		2.956	0.988
I have been obsessed with a certain idea or project for a short time but later lost interest	0.653		3.203	1.033
I often set a goal but later choose to pursue a different one	0.603		3.746	0.793
I am a hard worker		0.362	3.696	0.859
I am diligent		0.527	3.609	0.899
I have difficulty maintaining my focus on projects that take more than a few months to complete		0.755	3.645	0.911

Table 25: Appendix - Robustness to splitting Grit Index into tenacity and diligence

Note: This table builds on Table 12, but now splits Grit Index into the reduced measure of Tenacity Index and Diligence Index. Column (1) uses the full sample of observations with no controls, similar to column (1) on Table 3. Column (2) adds the full battery of control variables, as in column (5) on Table 3. Columns (3) to (6) split the data by educational degree and gender, similar to Table 10. 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.: Overplaying	(1)	(2)	(3) Technical	(4) Non-technical	(5) Female	(6) Male	(7) >100
Tenacity Index	0.159***	0.175***	0.216***	0.147*	0.210**	0.144*	0.140**
	(0.059)	(0.056)	(0.070)	(0.087)	(0.084)	(0.086)	(0.059)
Diligence Index	0.015	0.026	0.018	0.001	0.050	-0.019	0.015
	(0.061)	(0.059)	(0.085)	(0.091)	(0.077)	(0.093)	(0.062)
Plan of Action		0.084***	0.078***	0.098***	0.073**	0.079***	0.086***
		(0.014)	(0.019)	(0.026)	(0.029)	(0.017)	(0.015)
ln(Age)		0.122	-0.578	0.691***	-0.350	0.623**	0.162
		(0.287)	(0.436)	(0.258)	(0.400)	(0.266)	(0.281)
D(Female)		-0.143*					-0.177**
		(0.084)					(0.086)
Technical Degree		-0.107					-0.123
		(0.085)					(0.087)
Constant	-0.198	-1.045	0.855	-2.818***	0.078	-2.381**	-1.046
	(0.227)	(0.888)	(1.347)	(0.803)	(1.167)	(0.898)	(0.874)
Observations	138	138	71	67	80	58	128
R-squared	0.068	0.173	0.176	0.197	0.167	0.199	0.168

C Proofs

Since the proofs are straightforward, we keep them informal for notational simplicity.

Proof of Prediction 1: Considering any optimal ex-ante strategy s_1^* chosen by the agent with tenacity T in period 1. Consider first any non-terminal history h^t reachable by strategy s_1^*. Clearly, if $z(h^t) \geq z_b$, then the agent incurs no cost of failure, and so he does not revise his strategy. That is, the optimal s_t^* is consistent with s_1^*. If $z(h^t) \in (0, z_b)$, then with probability $1-q_f(T)$ he does not incur the cost of failure and so does not revise his strategy. With probability $q_f(T)$ he does incur the cost of failure. Since he had chosen ex-ante not to stop, we know that:

$$u(z_e(h^t)) < E_{\{y^{t'}|h^t,s_1^*\}}u(z_e(y^{t'})), \tag{1}$$

where $t' > t$, and recalling that $E_{\{y^{t'}|h^t,s_1^*\}}$ is the expectations operator for every possible subsequent terminal history $y^{t'}$ given that history h^t has been reached and given that optimal strategy s_1^* is being followed. With the cost of failure, the agent chooses not to stop if:

$$u(z_e(h^t)) - c_f < E_{\{y^{t'}|h^t,s_1^*\}}u(z_e(y^{t'})), \tag{2}$$

which clearly still holds. The reason that the right hand side does not change is due to the assumption that he is naive about his future cost of failure, and so he expects his future behavior to be in accordance with s_1^*. In both cases, therefore, the agent is dynamically consistent.
Consider now any terminal history $y^t = ((Y,x)_1, ...(Y,x)_{t-1}, N)$, for $t \geq 2$ that is reachable by strategy s_1^* and for which $z_e(y^t) \in (0, z_b)$. We do not consider terminal histories for $t = 1$ since the agent clearly could not overplay if he stops at the beginning. History $h^t = ((Y,x)_1, ...(Y,x)_{t-1})$ is the non-terminal history up to the end of the previous period. Then, since the agent had chosen ex-ante to stop in period t, it must be that:

$$u(z_e(y^t)) \geq \max_s E_{\{y^{t'}|h^t,s\}}u(z_e(y^{t'})), \tag{3}$$

where $t' > t$, and the maximization is over all possible strategies s. When this node is actually reached, then with probability $1 - q_f(T)$ the agent does not incur a cost of failure, and the inequality still holds. However, with probability $q_f(T)$ he does incur the cost of failure c_f, and, since c_f is arbitrarily large, the inequality has changed signs, i.e.,

$$u(z_e(y^t)) - c_f < \max_s E_{\{y^{t'}|h^t,s\}}u(z_e(y^{t'})). \tag{4}$$

In this case, the agent chooses *not* to stop but to continue in period t. In other words, he would overplay (and his strategy s_t^* is not equal to s_1^*). It then follows that the agent is more likely to overplay if the probability of overplaying increases, which we have assumed for higher tenacity ($q_f'(T)$). Since diligence D does not affect q_f aside from its possible correlation through tenacity, it has no effect on tenacity, which completes the proof. □

Proof of Prediction 2: Prediction 2 follows trivially from Prediction 1, from grit G being the weighted sum of tenacity T and diligence D, and from the assumption that D is not negatively correlated with T.

Proof of Prediction 3: The proof of Prediction 3 follows from similar logic to the last step of Prediction 1. Specifically, as shown before, for terminal history $y^t = ((Y,x)_1, ...(Y,x)_{t-1}, N)$, for $t \geq 2$ that is reachable by strategy s_1^* and for which $z_e(y^t) \in (0, z_b)$, with probability $q_f(T)$ the agent will deviate from his ex-ante strategy and play another period. If at the next stage, his total earnings are less than z_b and greater than zero, then his cost of failure is still present, and so he continues to play another period, since for arbitrarily large,

$$u(z_e(y^{t+1})) - c_f < \max_s E_{\{y^{t'}|h^t,s\}} u(z_e(y^{t'})), \tag{5}$$

where now $y^{t+1} = ((Y,x)_1, ...(Y,x)_{t-1}, (Y,x)_t, N)$ for his new earnings, and $t' > t+1$. If his earnings are exactly 0, then he must stop. Following the same reasoning, he will continue to play as long as his earnings are less than z_b and greater than 0, and stops if they reach 0. If instead, his earnings are back to z_b (notice that since $z_w = z_l$, they cannot go from being below z_b to strictly above z_b), then his cost of failure is no longer present, and he is back to his initial wealth. If he plays and loses then the same logic starts anew, since again with probability $q_f(T)$ he pays the cost of failure. Hence, since $q_f'(T) > 0$, the agent with higher tenacity is more likely to go through this process not only the first time, but also possible subsequent times if he returns to initial wealth z_b. An agent with higher tenacity is therefore more likely to end up with wealth either of zero or an amount greater or equal to z_b, which completes the proof. □

The model can easily be extended to allow for an increasing opportunity cost of time, as would be realistic in an experimental setting. Our results would still hold in this case, using similar logic to the proofs above.