STAFF REPORTS

NO. 1044
DECEMBER 2022

Strategic Sophistication and Trading Profits: An Experiment with Professional Traders

Marco Angrisani | Marco Cipriani | Antonio Guarino

FEDERAL RESERVE BANK *of* NEW YORK

Strategic Sophistication and Trading Profits: An Experiment with Professional Traders
Marco Angrisani, Marco Cipriani, and Antonio Guarino
Federal Reserve Bank of New York Staff Reports, no. 1044
December 2022
JEL classification: C93, G11, G14

Abstract

We run an experiment where professional traders, endowed with private information, trade an asset over multiple periods. After the trading game, we gather information about the professional traders' characteristics by having them carry out a series of tasks. We study which of these characteristics predict profits in the trading game. We find that strategic sophistication, as measured in the Guessing Game (for example, through level-k theory), is the only significant determinant of professional traders' profits. In contrast, profits are not driven by individual characteristics such as cognitive abilities or behavioral traits. Moreover, higher profits are due to the ability to trade at favorable prices rather than to the ability to earn higher dividends. Comparing these results to those of a sample of students, we show that whereas cognitive skills are important for students, they are not for traders, whereas the opposite is the case for strategic sophistication.

Key words: experiments, financial markets, professional traders, strategic sophistication

---

Cipriani: Federal Reserve Bank of New York (email: marco.cipriani@ny.frb.org). Angrisani: University of Southern California, Center for Economic and Social Research (email: angrisan@usc.edu). Guarino: Department of Economics, University College London and CEPR (email: a.guarino@ucl.ac.uk). The authors are grateful to Giorgio Coricelli, Rosemarie Nagel, Julen Ortiz De Zarate Pina, and participants in the ESA Meeting 2022 for helpful comments. They also thank Antonella Buccione, Andrea Giacometti, Seungmoon Park, and James Symons-Hicks for terrific research assistance, and the professional traders and portfolio managers who volunteered for this study.

To view the authors' disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr1044.html.

1 Introduction

In the General Theory, Keynes (1933, Section 12) writes: "[...] Professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees."

In modern Game Theory, this idea has been revisited in the celebrated Guessing Game (Nagel, 1995), in which the strategic sophistication of agents is measured by how they pick a number with the goal of matching a fraction of the group's average pick. Level-k theory classifies as level-0 the agents who chose randomly; as level-1 those who best respond to the belief that all others are level-0; as level-2 those who best respond to the belief that all others are level-1, and so on. This classification mirrors the hierarchy of beliefs indicated by Keynes: the higher an agent's level-k, the higher their strategic sophistication.

In this paper, we study whether strategic sophistication explains how successfully professional traders trade in financial markets, as Keynes himself suggested: "The actual, private object of the most skilled investment to-day is 'to beat the gun', as the Americans so well express it, to outwit the crowd, and to pass the bad, or depreciating, half-crown to the other fellow." (Keynes, 1933, Section 12).

To this purpose, we recruited traders and portfolio managers—financial professionals that actually trade or invest in financial markets—working in London and had them play both a Trading Game and a Guessing Game; we then studied the relationship between their degree of strategic sophistication in the Guessing Game and their profits in the Trading Game.[1] Note that although there is a literature of experimental finance with financial market professional (see, among many papers, Alevy et al., 2007; Cipriani and Guarino, 2009; List and Haig, 2005; List and Haig, 2010), our subjects are all specifically engaged in trading activity, as traders or portfolio managers; for this reason, the experimental outcomes are informative of professional traders' behavior in actual financial markets.

Our experimental design, while borrowing from past studies, presents several novelties. In the Trading Game, subjects trade an asset that earns dividends over multiple periods in a continuous double auction. Our experimental asset market innovates with respect to the standard setup (Smith et al., 1988) by introducing private information about the asset's fundamental value. Since trading occurs over multiple periods in this Trading Game, subjects make a profit both from earning the asset's dividends and from any capital gain they make by buying the asset at a lower price than that at which they sell it. Hence, profits do not depend only on the ability to forecast the fundamental value correctly, but also on the ability to predict what

[1] That is, we run an artifactual field experiment (see Harrison and List, 2004).

other traders will be willing to pay or receive to buy or sell the asset (the ability "to outwit the crowd," as Keynes put it).

In the second task of our experiment, the Guessing Game (Nagel, 1995), each subject chooses one number and the winner is the subject whose number is closest to a fraction (2/3 in our case) of the average. A large literature (e.g., Nagel, 1995; Camerer et al., 2004; Crawford et al., 2013) has shown that student subjects choose numbers far away from the Nash equilibrium (in which everyone chooses 0). We study how professional traders play this game. Finally, after these two games, subjects participate in a series of individual-level tasks, including an "Individual Guessing Game" (aimed to disentangle the reasons for different levels of strategic sophistication), two measures of cognitive ability (Raven IQ Test and Cognitive Reflection Test), tasks aimed to elicit risk preferences and confidence and, additionally, questionnaires to measure non-cognitive abilities.

We find that strategic sophistication does explain traders' trading profits: more sophisticated professional traders earn higher payoffs in the Trading Game. An increase in sophistication, measured by a 10-point decrease in the Guessing Game choice, is associated with a 4% increase in payoffs relative to the mean. When we classify traders according to Nagel (1985)'s criterion, we find that level-(1-3) traders earn 17% more than level-0 traders, and level-4 to level-∞ traders earn 31% more than their level-0 counterparts. Importantly, the higher profits secured by more sophisticated traders are not obtained at the cost of a riskier trading strategy. On the contrary, we find evidence that the volatility of traders' portfolio decreases as their level of sophistication increases.

We then ask how more sophisticated traders earn higher profits than their less sophisticated counterparts. We find that this is due to their ability to obtain trading profits, that is, their ability to sell at higher prices than those at which they buy. For a 10-unit decrease in the Guessing Game choice, per-period trading profits increase by an amount corresponding to about 2.5 times the total per-period profits secured by the average trader. Strategic sophistication is instead unrelated to the dividends traders receive. It seems, therefore, that, as in Keynes's description, strategic sophistication helps because it increases the ability to outwit the crowd, rather than to forecast the fundamentals.

In a recent paper, Corgnet et al. (2018) report the results of a trading experiment with undergraduate students in which they explore which of the students' characteristics explain successful trading behavior. In their experiment, high cognitive abilities (IQ and Cognitive Reflection) increase profits. As a control, we also conduct our experiment with undergraduate students and confirm Corgnet et al. (2018)'s results: with student subjects, higher cognitive abilities are associated with higher profits, with strategic sophistication playing a minor role. In contrast, in our trader sample, whereas strategic sophistication matters, cognitive abilities do not. This suggests that for students, who are not selected to trade in real markets and have no trading experience, differences in cognitive abilities drive differences in trading performance in the laboratory. For professional traders, who are self-selected into trading and have experience in markets, cognitive abilities are less relevant; what matters is the difference in strategic reasoning.

The paper is organized as follows: Section 2 describes the experiment. Section 3 describes the sample of participants. Section 4 presents some preliminary results. Section 5 describes the relation between strategic sophistication and trading performance. Section 6 investigates the mechanism behind this relation. Section 7 contrasts our results for professional traders with those for a sample of undergraduate students. Section 8 concludes. The online Appendix contains the instructions and other results.

2 The Experiment

We ran the experiment in the Experimental Laboratory for Finance and Economics (ELFE) in the Centre for Finance at the Department of Economics at UCL.

In our experiment, subjects participate in the following sequence of tasks: i) the Trading Game; ii) the Guessing Game; iii) the Individual Guessing Game; iv) individual-level tasks aimed to infer their cognitive abilities, risk preferences, and confidence; v) a series of questionnaires aimed to elicit non-cognitive abilities (Big 5, Locus of Control, Grit, Self-Monitoring).

2.1 The Trading Game

We first provide a simple model of the Trading Game and then describe the procedures that we use to implement it in the laboratory.

2.1.1 Setup

We consider a market with a continuum of traders who trade a risky asset. Time is discrete and indexed by $t = 1, 2, .., 10$. The numeraire is cash. Traders are risk neutral and do not discount the future.

At each period t, the asset yields dividends d_t equal to 50 or 150 with equal probability. Dividends are independently distributed over time. The asset has no residual value after period 10.

The realization of the dividends is unknown to traders. In each period t, however, they receive noisy private information about the value of the current dividend d_t in the form of a symmetric binary signal with precision 0.75. In particular, each trader i receives a signal s_t^i distributed as $\Pr(s_t^i = 0|d_t = 50) = \Pr(s_t^i = 1|d_t = 150) = 0.75$. Conditional on the dividend value, signals are independently distributed. Note that the signal is only informative about the dividend in the current period, not about future dividends.

In order to generate gains from trade, at the beginning of the period, half of the traders are selected to pay a fee of 50 for each asset held in their portfolio at the end of the period. Whether a trader has to pay a fee in a given period is independent of whether they had to pay it in previous periods.

In period $t = 1$, each trader has an endowment of 3 assets and 7,000 of cash. After trading in period t and receiving the dividend, the portfolio of cash and assets carries over to period $t + 1$.

2.1.2 Equilibrium Prediction

In the Rational Expectations Equilibrium (REE) of the model, the price is equal to the asset's fundamental value (i.e., the expected value of current and future dividends). At each period t, private signals are perfectly aggregated by the price. Moreover, in each period, all traders who have to pay a fee for holding the asset at the end of the period sell it to those who do not pay the fee, thereby realizing all the gains from trades.

In particular, in period 1, demand and supply clear the market for a price of 950 when $d_1 = 50$ and for a price of 1,050 when $d_1 = 150$ (see Figure B.1 in Appendix): the equilibrium price is equal to the sum of the expected value of dividends in all nine subsequent periods ($100 \times 9 = 900$) and the realized dividend in period 1 (50 or 150). The same logic applies to any subsequent period. Equilibrium prices are reported in Table 1. Note that the equilibrium price of the asset is weakly decreasing over time: between time t and time $t+1$, it decreases by 100 (when the dividend is either 150 or 50 in both periods t and $t+1$), or by 200 (when the dividend is 150 in period t and 50 in period $t+1$), or it remains constant (when the dividend is 50 in period t and 150 in period $t+1$).

Table 1: REE by Period and Dividend Realization

Period: t	1	2	3	4	5	6	7	8	9	10
$d_t = 50$	950	850	750	650	550	450	350	250	150	50
$d_t = 150$	1,050	950	850	750	650	550	450	350	250	150

2.1.3 Trading in the Laboratory

In the experiment, we have 8 subjects acting as traders. At the beginning of period 1, each subject receives an endowment of 3 assets and 7,000 Experimental Currency Units (ECU).

Subjects receive a private signal at the beginning of each period. Specifically, when the dividend is equal to 150, 6 subjects observe a blue ball and 2 subjects a red ball; when the dividend is equal to 50, 6 subjects observe a red ball and 2 subjects a blue ball. This signal structure guarantees that, in each period, private signals jointly reveal the dividend even if the number of subjects is finite. Other signal structures (for instance, i.i.d. signals with precision 0.75), even if informative, may not deliver the same result.[2] At the beginning of each period, subjects also learn whether they are fee-paying or non-fee-paying subjects for that period.

In each period, subjects trade for 150 seconds in a double-auction market. They post offers to sell or buy one asset. To post a sell offer, a subject enters the minimum price they are willing to accept and clicks on a sell button. The offer appears immediately on everyone's screen, in a column labeled "Sell Offers" (the identity of the subject making the offer is not revealed). Similarly, to post a buy offer, a subject enters the maximum price they are willing to pay and clicks on a buy button. A trade is automatically executed whenever the lowest sell offer (ask) is lower than the highest buy offer (bid). In other words, if a subject wanted, for instance, to buy at the prevailing (i.e., the lowest) ask, they could simply enter a price equal to or greater than that price, and the trade would be immediately executed (at the outstanding price). Subjects can also buy or sell by clicking on "BUY" or "SELL" buttons, which automatically accept the best outstanding sell or buy offer.

[2]Note that, in the theoretical model, we consider i.i.d. signals conditional on the dividend value, because we have a continuum of traders and the dividend realization is revealed even with i.i.d. signal. It is easy to show that the equilibrium in the eight-subject economy with the signal structure explained above is the same as the equilibrium described in the theoretical section.

Each subject can post a maximum number of sell offers equal to the number of assets held in their portfolio and the sum of all the outstanding buy offers cannot exceed the cash held in their portfolio. At any time, a subject can withdraw outstanding buy or sell offers that have not already been executed by clicking a button labeled "Cancel." A subject's screen displays their current portfolio of cash and assets, the list of past trades (with their own executed trades highlighted), all the outstanding bid and ask prices, and the time left before the end of the period (see Section B.2 of the Appendix for instructions and decision screen shots).

At the end of each period, subjects are informed about the dividend's realization and their end-of-period portfolio. Changes in the portfolio between adjacent periods are due to the dividends earned, the fee payments, and the trading profits (i.e., profits or losses from trading). The portfolio at the end of a period carries over to the next period. After one period ends, trading in the following period starts, according to the same rules, until all ten periods are completed.

2.2 The Guessing Game

After the Trading Game, subjects play the standard Guessing Game (Nagel, 1995). Each subject chooses one number in $[0, 100]$. Subjects are asked to guess a target number, defined as 2/3 of the average of the 8 numbers entered by all the subjects (2/3 is a standard target, used, e.g., in Nagel's original paper). The subject whose number is closest to the target number earns £5; the others earn nothing. In the case of a tie, the amount is split equally. The decision screen shot for the Guessing Game is in Figure B.3 in the Appendix.

The only Nash equilibrium of the game is all subjects choosing 0. The most common interpretative framework for behavior in the Guessing Game is level-k theory (e.g., Nagel, 1995; Camerer et al., 2004; Crawford et al., 2013). A subject who chooses randomly in $[0, 100]$ is a level-0 subject. A subject who believes that all other subjects are level-0 and best responds on the basis of this belief is a level-1 subject. Since a level-1 subject believes that other subjects choose, on average, 50, their best response is to choose 31.8.[3] If a subject believes that all other subjects are level-1 and best responds accordingly, they choose 21.2 and are said to be a level-2 subject. In general, a level-k subject is one who best responds to the belief that others are level-(k-1) subjects. When k tends to infinity, the best response converges to the Nash equilibrium action. The higher the level-k, the higher the level of strategic sophistication of a subject.

2.3 Individual tasks

After the Trading Game and the Guessing Game, we asked subjects to perform a series of individual tasks to infer their cognitive abilities, their risk aversion, and their non-cognitive skills.

2.3.1 Individual Guessing Game

After the standard Guessing Game, we ask subjects to take part in a modified, individual-level version of it. In this Individual Guessing Game (IGG), each subject chooses 8 numbers in

[3] This number is slightly below 2/3 because the target is computed taking the subject's own guess into account.

[0, 100] and is given a target number of 2/3 of the average of the 8 chosen numbers. Then, one of the subject's 8 chosen numbers is randomly selected for payment and, if it equals the target number, the subject earns £5. The decision screen shot for the IGG is in Figure B.4 in the Appendix.

According to level-k theory, subjects with higher levels of strategic reasoning choose lower numbers in the Guessing Game. However, for any choice in the Guessing Game between 1 and 50, one cannot disentangle a subject's ability to reason from the subject's belief about other subjects' ability to reason. For instance, suppose a subject chooses 31.8 and is, therefore, classified as a level-1 subject. This choice could be because the subject lacks the ability to reason past level 1; or it could be because, although the subject does have such an ability, they believe that other subjects are level-0 subjects.

Our novel Individual Guessing Game disentangles one's own ability from their belief about the ability of other subjects. In particular, in the IGG, the only 8 numbers that guarantees earning £5 are 8 zeros. A subject who enters this answer has the ability to reason to an arbitrarily high level. Therefore, using the example above, if a subject chooses 31.8 in the Guessing Game and correctly answers the Individual Guessing Game, they show the ability to reason in a sophisticated way but a low-level belief about the ability to reason of their peers.

2.3.2 Cognitive Ability and Confidence

Our first test of cognitive ability is the Raven (1941)'s progressive matrices test (perhaps, the most well-known IQ test). We selected 18 of the most difficult Raven's matrices, available in Raven's Advanced Progressive Matrices (Raven, 1990). We did this to avoid a ceiling effect on the scores (which we avoided, as the highest score recorded was 15 out of 18). A subject's IQ score is the number of correct answers within a 10-minute period.

Our second test of cognitive ability is the Cognitive Reflection Test (CRT; Frederick, 2005). The test measures a subject's tendency to override an incorrect impulsive response and engage in further reflection that leads to the correct answer. For instance, one question is the following: "A bat and a ball cost £1.10 in total. The bat costs one pound more than the ball. How much does the ball cost?" An answer of 5 pence is correct while an answer of 10 pence suggests a lack of reflection. We use the extended 7-question version of the CRT (Toplak, West, and Stanovich, 2014). A subject's CRT score is the number of correct answers within a 5-minute period.

After the Raven's and CRT tests, we present subjects with an 8-row table, in which each row reads: "What is the percent chance that X of the 7 other participants had more correct answers than you did?" (for $X = \{0, 1, ...7\}$). In other words, we elicit the entire distribution of beliefs about the subject's ranking for each of these two cognitive ability tests. Given the subject's answers, for each test, we compute the expected number of outperforming subjects. The difference between the actual and expected numbers is our confidence measure for the subject. A positive number indicates that the subject is over-confident (because they expected to be outperformed by fewer subjects than they actually did); a negative number indicates that the subject is under-confident; and a score of 0 means that the subject is not biased. We elicit confidence separately for the Raven's and CRT tests. We average these two measures to get a composite measure of confidence about one's own cognitive performance relative to others, which we use in the empirical analysis. More details about the administered cognitive ability tasks and elicitation of confidence are provided in Figures B.5-B.7 in the Appendix.

2.3.3 Risk preferences - Bomb Risk Elicitation Task (BRET)

We measure risk preferences through the Bomb Risk Elicitation Task (BRET; Crosetto and Filippin, 2013). In the BRET, subjects are shown a screen with 100 boxes and are asked to choose to open a number of boxes (between 1 and 99). Each box contains 20 pence; therefore, earnings increase linearly with the number of boxes chosen. Among the boxes, however, there is one that, if opened, makes the subject lose all money earned in the other opened boxes (in the original version by Crosetto and Filippin (2013) this box was described as a box containing a bomb; we used a more neutral description).[4] Subjects earn 20 pence for each box they open, which are added to any profits made in the Trading Game and earnings in the Guessing Game and Individual Guessing Game.

The decision about the number of boxes is a decision under risk. A risk-neutral subject opens 50 boxes. A risk-averse subject opens less than 50 boxes and a risk-loving one more than 50. The more boxes a subjects opens, the higher their degree of risk-seeking preference. The instructions and decision screen shots for the BRET task are in Figures B.8 and B.9 in the Appendix.

2.3.4 Non-cognitive Skills

Finally, we gathered data on participants' non-cognitive skills using a non-incentivized survey given at the end of the experiment. Specifically, we elicited the Big 5 personality traits—openness (to experience), conscientiousness, extraversion, agreeableness, and neuroticism—through the 10-item Big 5 inventory of Rammstedt and John (2007), the level of Grit through the 8-item scale of Duckworth and Quinn (2009), and the level of Locus of Control through the 7-item questionnaire as in Cobb-Clark and Schurer (2013). For all these non-cognitive skills, participants were asked to state the extent to which they agreed with a specific statement using a 1 (Disagree strongly) to 5 (Agree strongly) scale. The items associated to the Big 5 personality traits, Grit and Locus of Control are in Figure B.10 in the Appendix and indicated with the initial of the corresponding non-cognitive skill. We reverse coded answers so that higher values correspond to more of each skill and took the average across all items associated to each skill.

The level of Self-Monitoring was captured by a series of 18 true/false questions (coded as 0 or 1) proposed by Snyder and Gangestad (1986), which are reported in Figure B.11 in the Appendix. We computed the proportion of times a participant's choice aligns with the higher Self-Monitoring answer. This measure takes a value between 0 and 1; a higher value represents higher Self-Monitoring.

[4]Crosetto and Filippin (2016) provide a comparison of different risk elicitation methods. Angrisani at al. (2020) study the behavior of risk aversion in our sample of professional traders at the onset of the COVID-19 pandemic. List and Haig (2005) test the expected utility theory using professional traders.

3 Experimental Subjects

The subjects are traders and portfolio managers working in London (UK). The experiment consists of 7 sessions; in each session, 8 subjects perform the tasks described above. Overall, our sample includes 56 professional traders. Subjects had no previous experience with this experiment and participated in one session only. Each session lasted approximately 2 hours.

The subject pool consists of: 30 traders, 4 proprietary traders, 4 sales-traders, 12 portfolio managers, and 6 belonging to other categories (e.g., trading strategists or sales with management of virtual portfolios). In our call for subjects, we wrote that "*we need participants who are either traders or portfolio managers or who have had such roles in the past. You are also eligible if you do not have the formal title of trader or portfolio managers, but you perform activities that are closely related to that of a trader or portfolio manager (e.g., sales-trader or sales on the trading floor)*". Subjects work in a variety of markets, such as equity, equity derivatives, FX, fixed income, and commodities. Thirty-two subjects are employed by an investment bank, 12 by an investment fund and the others by other types of institutions (or chose not to report). Traders' ages range between 22 and 48 years, with a mean of 32 years and a standard deviation of 6.17 years. Their average job tenure is 9.25 years, with a range between 1.5 and 21 years (standard deviation: 5.42 years). Three subjects have an MPhil or Ph.D., 30 a Master's degree, 4 an MBA, 19 a Bachelor's degree. Thirty subjects studied economics or finance, 8 mathematics or physics, 8 engineering or computer science, and the remaining have a degree in other disciplines or did not declare it.

At the end of the experiment, subjects are paid based on their choices in the Trading Game, the Guessing Game, the Individual Guessing Game, and the BRET. In the Trading Game, we convert ECU into British Pounds at the exchange rate of £2.50 = ECU 100.

4 Descriptive Statistics

We start by presenting some descriptive statistics on the Trading Game and the Guessing Game. In the Trading Game, traders' average final profit is ECU 9,397, corresponding to £235 (approximately \$305) with a standard deviation of ECU 1,621, corresponding to £41 (approximately \$52; see Figure 1 and Table 2).

We also compute the profits that traders accrued in each of the 10 periods in which they traded. For each period $t = 1, 2, ..., 10$, we compute the value of a trader's portfolio at the end of the period by summing the cash and the value of the assets in the portfolio, valuing the assets at the REE. For instance, the value of the portfolio in period $t = 1$ (i.e., at the end of $t = 1$) is computed summing the cash and the value of the assets in the portfolio, using the REE price for $t = 2$.[5] At the end of the Trading Game, $t = 10$, the value of the portfolio is the cash held by each trader, since the final value of the asset is zero. The per-period profit for period t is the change in the value of the portfolio between t-1 and t (i.e., between the beginning and the end of period t). We obtain 10 per-period profits for each trader, for a total of 560 observations.

It is important to note that since we evaluate assets in the portfolio at their REE value, in equilibrium, the value of the portfolio would never change, only its composition between cash

[5] This is because the dividend for period 1 is already paid and hence already part of the cash in the portfolio.

and assets would change. Therefore, in equilibrium the per-period profits are by construction 0. Recall that in our market there are gains from trade, since half the traders have to pay a fee; in the REE, the allocation is such that no one pays the fee. In the experiment, not all gains from trade are realized, with the result that even the average and median per-period profit (as we compute them) are negative: as can be seen in the second row of Table 2, on average, traders lost ECU 45 in each period, while the median per-period profit was 0. There is substantial heterogeneity in per-period profits (well captured by the right histogram in Figure 1) as the standard deviation is ECU 611.[6]

A trader's final payoff equals the initial endowment plus the sum of their per-period profits. Therefore, the fact that, on average, traders make negative per-period profits does not imply that they make a loss in the overall experiment, since they start with a positive endowment. Indeed, all traders earned a positive final payoff.

Let us now consider the Guessing Game. Figure 2 shows the distributions of the Guessing Game choice (GG choice). Across traders, the average and median GG choices are 22.63 and 22.50, with a standard deviation of 19 (see Table 2).

Figure 1: Histograms of Final Payoffs and Per-Period Profits

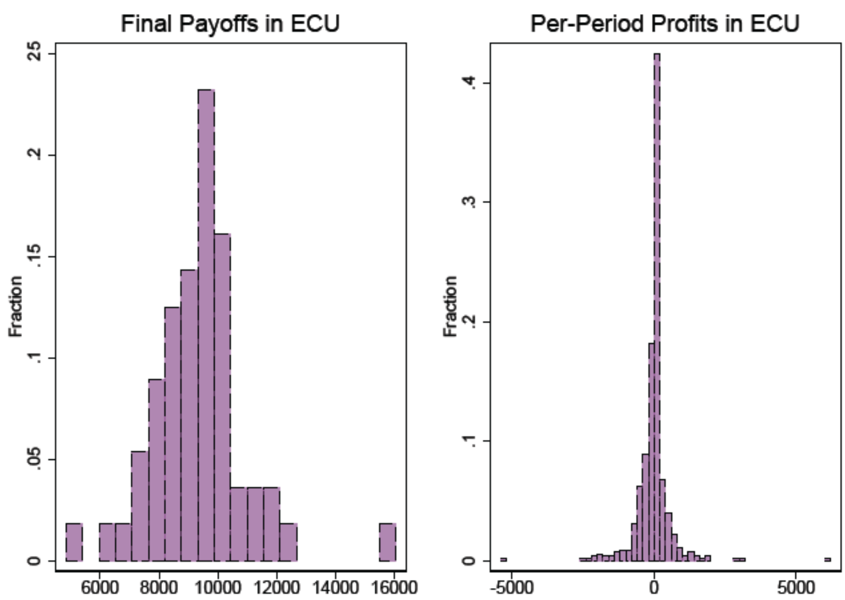


Table 2: Descriptive Statistics for Trading Game and Guessing Game

	Mean	Median	Std.Dev.
Final Payoffs	9,397.32	9,461.50	1,621.17
Per-period Profits	-45.27	00.00	610.87
GG Choice	22.63	22.50	18.75

Following Nagel (1995), we classify a trader as level-k if their choice is between a lower and an upper bound defined as the nearest integers of $\left\{50\left(\frac{2}{3}\right)^{\left(k+\frac{1}{4}\right)}\right\}$ and $\left\{50\left(\frac{2}{3}\right)^{\left(k-\frac{1}{4}\right)}\right\}$, respec-

[6]The high standard deviation reflects the high heterogeneity of performances in each period.

tively. The definition of level-0 is an exception to this rule as the interval is truncated at 50. So, a level-0 trader chooses a value in the [45,50] interval, a level-1 trader chooses a value in the [30,37] interval, a level-2 chooses a value in the [20,25] interval, and so on. A level-∞ trader is one choosing a value in the [0,1] interval. Traders whose choices do not fall into these intervals are not classified. As the left panel of Table 3 shows, we can classify 41 out of 56 (73%) traders. There is a large proportion of level-∞ traders (29.3% of the classified traders). Overall, about 60% of the classified traders are at least level-2. As we will discuss later, these results indicate a higher level of strategic sophistication than what found in the literature with student subjects. The right panel of Table 3 shows a different (broader) categorization of traders: a trader is classified as level-k if their choice is smaller than the level-k upper bound, but bigger than the level-(k+1) upper bound. For instance, a trader whose choice is lower than 25 but bigger than 16 is classified as level-2. This alternative caegorization allows us to classify all 56 traders who participated in the experiment.

Table 3: GG Choice: Level-k Classification

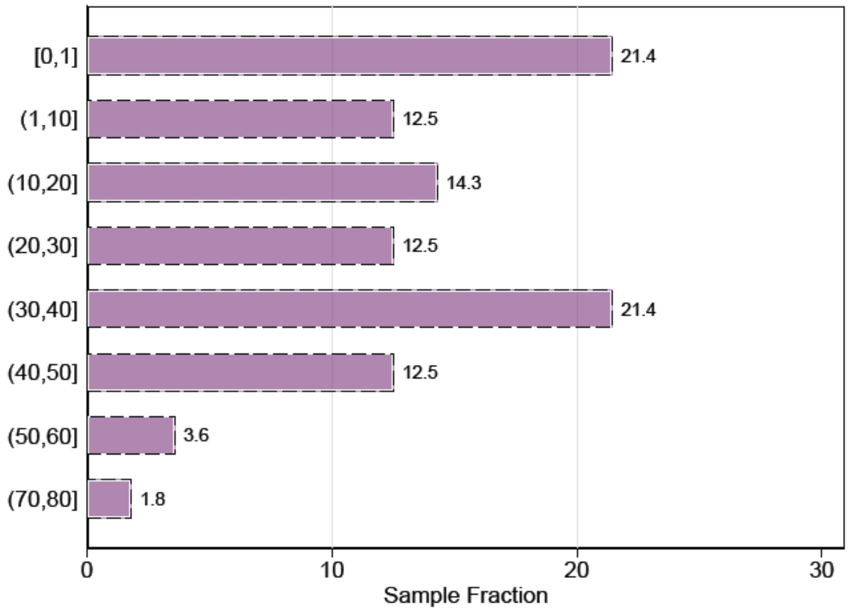
	Nagel's Classification			**Broader Classification**		
	Choice Range	Freq.	Perc.	Choice Range	Freq.	Perc.
Level-0	[45, 50]	5	12.20	(37, 100]	12	21.43
Level-1	[30, 37]	12	29.27	(25, 37]	14	25.00
Level-2	[20, 25]	3	7.32	(16, 25]	6	10.71
Level-3	[13, 16]	2	4.88	(11, 16]	4	7.14
Level-4	[9, 11]	3	7.32	(7, 11]	4	7.14
Level-5	[6, 7]	1	2.44	[6, 7]	1	1.79
Level-6	[4, 5]	0	0.00	[4, 5]	0	0.00
Level-7	[3, 3]	2	4.88	[3, 3]	2	3.57
Level-8	[2, 2]	1	2.44	[2, 2]	1	1.79
Level-∞	[0, 1]	12	29.27	[0, 1]	12	21.43
Total		41	100.00		56	100.00

Descriptive statistics for individual tasks are reported in Table 4. The first individual task we consider is the Individual Guessing Game. As explained above, in the Individual Guessing Game, traders' beliefs about the behavior of the others traders are irrelevant; only a trader who chooses all zeros plays this game correctly. Eleven traders, 20% of the sample, solved the IGG correctly. As for cognitive abilities, traders have an average Raven IQ score of 8.7, and an average CRT score of 5.14. We do not observe overconfidence in these cognitive abilities: the average measure of confidence across traders is positive (0.24) but not statistically different from zero (t-test p-value = 0.292). Finally, the average and median trader is risk neutral, choosing to open 50 boxes in the BRET task, though there is heterogeneity in risk preferences (the standard deviation of the BRET choice is 13).

5 Trading Profits and Strategic Sophistication

In Figure 3, we report the results of a non-parametric regression of final payoffs on the GG choice: the relationship is monotonic (almost linear), with payoffs decreasing as the GG choice

Figure 2: Distribution of GG Choice



increases. The estimated linear regression coefficient is -38 and statistically significant (p-value<0.001): a 10-unit decrease in the GG choice increases final payoffs by ECU 380, a 4% increase relative to the average final payoff.

Table 4: Descriptive Statistics for Individual Tasks

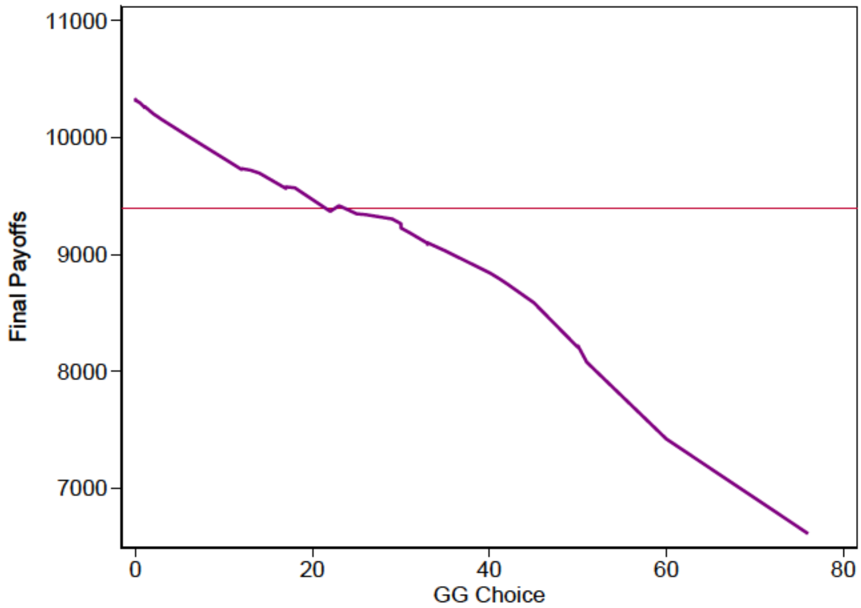
	Mean	Median	Std.Dev.
IQ Score	8.70	9.00	2.22
CRT Score	5.14	6.00	1.72
Confidence	0.24	0.07	1.69
BRET Choice	49.68	50.00	12.57

To explore the relationship between the performance in the Trading Game and strategic sophistication further, we estimate random-effects (RE) panel regressions of traders' per-period profits on their strategic sophistication. We use progressively richer specifications to control for traders' gender and age, correct answers to the Individual Guessing Game ("IGG Correct"), cognitive skills as measured by individuals' performances in the IQ and CRT tests, confidence in own cognitive ability, risk aversion, and non-cognitive skills.[7] We include period fixed-effects in all specifications. To account for unobserved correlations among traders within each session, we cluster standard errors at the session level and, following Cameron and Miller (2015), compute bootstrap standard errors using 500 replications. The results of these RE regressions are reported in Table 5.

The panel-data estimates confirm the strong association between strategic sophistication and trading performance. The RE coefficients of the GG choice are of similar size across

[7]We standardize IQ and CRT. One trader did not reveal their gender; we impute this missing value with the sample mean to maintain the same sample size across specifications; excluding this trader does not affect our results.

Figure 3: Locally Weighted Regression of Final Payoffs on GG Choice



The red line indicates average final payoff, corresponding to ECU 9,397.

all specifications and always statistically significant (p-value<0.001 in specifications (i)-(v), p-value<0.05 in specification (vi)). Taken all together, the estimated coefficients indicate that a 10-unit decrease in GG choice is associated with about ECU 35 more per-period profits, a sizeable increase leading a trader to earn a payoff 4% higher at the end of the trading game. None of the controls are significantly associated with trading performance. In particular, column (ii) shows that there is no significant association with demographics; column (iii) shows that there is no evidence of association with IGG, IQ or CRT scores. The joint test that cognitive skills (and IGG) have no explanatory power for trading performance never rejects the null at conventional significance levels (p-values range between 0.487 and 0.714 across specifications). In column (iv), we include as control our measure of confidence, which is also not significant. In column (v), we add traders' risk aversion. Since risk neutral subjects maximize expected payoff, the relationship between returns and risk attitude is likely to be non linear: as traders become more risk loving or more risk averse, their average payoff should decrease; for this reason, we use the distance from risk neutrality, separately for risk averse and risk loving traders;[8] if deviations from risk neutrality decrease profits, the coefficients of these two variables should be negative and statistically different from zero. As column (v) shows, there is no significant relation between trading performance and traders' risk preferences. Importantly, the effect of strategic sophistication on the per-period profits remains quantitatively similar across all five specifications.

Finally, in column (vi), we also condition on traders' non-cognitive skills, which include the Big 5 personality traits, Locus of Control, Grit, and Self-Monitoring. None of these non-cognitive skills significantly correlates with trading performance (the p-value of a joint test that all the non-cognitive measures have a coefficient of zero is 0.749), nor does their inclusion affect

[8] That is, we construct two regressors, $\max(BRET - 50, 0)$ and $-\min(BRET - 50, 0)$.

Table 5: Panel Regressions of Per-Period Profits on GG Choice

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	-3.787***	-3.777***	-3.284***	-3.329***	-3.218***	-3.796**
	(0.612)	(0.737)	(0.762)	(0.777)	(1.215)	(1.553)
Male		-10.178	-4.045	-4.347	-1.249	-41.945
		(60.418)	(64.669)	(63.321)	(64.300)	(76.740)
Age		-3.793	-1.091	-1.062	-0.969	-0.984
		(2.812)	(3.871)	(3.784)	(4.099)	(5.050)
IGG Correct			68.975	71.245	77.759	73.139
			(78.832)	(78.519)	(75.974)	(120.803)
IQ Z-Score			55.602	47.473	45.913	45.681
			(36.522)	(41.075)	(45.570)	(55.980)
CRT Z-Score			-30.093	-38.432	-31.967	-45.332
			(30.836)	(39.956)	(43.034)	(48.004)
IQ/CRT Confidence				-10.352	-10.191	-13.081
				(15.399)	(19.641)	(20.167)
$\|BRET-50\|_{BRET<50}$					1.371	1.362
					(1.995)	(2.464)
$\|BRET-50\|_{BRET>50}$					-1.411	-0.013
					(2.677)	(2.735)
Constant	1.796	133.208	15.680	18.079	8.240	53.607
	(62.412)	(128.044)	(175.757)	(172.949)	(184.813)	(192.552)
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.487	0.542	0.605	0.714
Test Non-Cog=0 (p-val)						0.749
N	560	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$. Estimated coefficients for non-cognitive skills are in Table A.1 in the Appendix.

the estimated association between the GG choice and per-period profits.[9]

Overall, the pattern in Figure 3 and the regression results in Table 5 reveal a strong and significant relationship between the GG choice and trading performance. Such relationship is remarkably stable after accounting for several other potential determinants of traders' profits. Moreover, the importance of other possible determinants of traders' profits is small compared to strategic sophistication, and almost never statistically significant.

An interesting question is whether the impact of strategic sophistication on traders' profits is stronger in the earlier or later rounds of the experiment. We answer this question by constructing a dummy variable for the beginning, the middle and the end of the experiment—Phase 1: Periods 1-3; Phase 2: Periods 4-7; Phase3: Periods 8-10—and estimating a regression with the interaction between GG choice and each of these dummies, in addition to time fixed effects. Results are shown in Table 6.

[9] The estimated coefficients for non-cognitive skills are reported in Table A.1 in the Appendix. Besides estimating the model in column (vi) of Table 5, where all non-cognitive skills are entered simultaneously, we also consider specifications where the Big 5 personality traits, Locus of Control, Grit and Self-Monitoring entered separately; the coefficients of non-cognitive skills are never statistically significant.

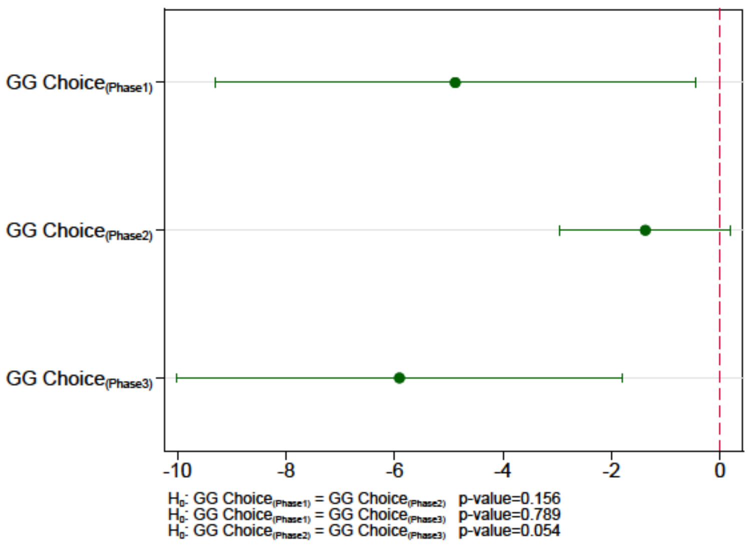
Table 6: Panel Regressions: Differential Effect of GG Choice on Per-Period Profits across Experiment Phases

Phase1 × GG Choice	-4.881**
	(2.256)
Phase2 × GG Choice	-1.376*
	(0.804)
Phase3 × GG Choice	-5.910***
	(2.093)
Constant	26.540
	(65.040)
N	560

Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Phase1 is a dummy for periods 1-3; Phase2 is a dummy for periods 4-7; Phase3 is a dummy for periods 8-10. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

As can be seen, the effect of strategic sophistication is relatively more apparent at the beginning (Phase1) and at the end (Phase3) of the experiment. Specifically, a 10-unit decrease in GG choice is associated with about ECU 50 and ECU 60 increases in per-period profits in Phases 1 and 3, respectively, but only ECU 14 in Phase 2. Whereas the difference between the effect in Phase 1 and Phase 2 is not significant, the p-value for the null that the coefficient of GG choice is the same in Phase 2 and Phase 3 is 0.054. A graphical illustration of these patterns, alongside statistical tests of pairwise differences in the effect of GG choice between experiment phases, is provided in Figure 4.

Figure 4: Differential Effect of GG Choice on Per-Period Profits across Experiment Phases



In their experiment with student subjects, Corgnet et al. (2018) find that Theory of Mind (ToM; Baron-Cohen et al., 2001) is positively related to subjects' payoffs. We did not collect data on ToM at the time of the experiment, but, similarly to Corgnet et al. (2018), contacted the traders afterwards. Thirty six traders completed a new survey in which we measured ToM using the Eye Gaze test (Baron-Cohen et al., 2001).[10] Within this sub-sample of traders, the ToM score and the GG choice are negatively correlated (the correlation is -0.46 and significant at 1% level), that is, those with higher ToM score also exhibit a higher level of strategic sophistication.[11] However, the ToM score has only a weak, positive correlation with traders' final payoffs, with a non-statistically significant correlation of 0.18 (p-value=0.290).[12] Consistent with the unconditional correlations, if we regress per-period profit on both ToM and the GG measure, we find that while the ToM score is not predictive of per-period profits, the impact of GG measure on profits remains significant and of a similar magnitude to that of our baseline regressions (column (i) of Table 5) even in this subsample.

5.1 Alternative specifications of strategic sophistication

Figure 5 shows average final payoffs, average portfolio value over time, and average per-period profits separately: i) for traders who are level-0; ii) for traders who are level-1 to level-3 (level-(1-3)); and for traders who are level-4 to level-∞ (level-(4-∞)). In the top panel, traders are classified on the basis of Nagel's classification; in the bottom panel, on the basis of our broader classification, which allows all traders to be assigned a level of strategic sophistication. As the figure shows, traders with higher level-k earn more. The top panel shows that level-(4-∞) traders maintain the value of their portfolio constant throughout the experiment, the equilibrium outcome (see the previous section). In contrast, the value of the average portfolio of level-(1-3) traders decreases slightly, while the one of level-0 traders exhibits a sharp decline especially towards the end of the experiment. Hence, level-(4-∞) traders obtain the highest level of final payoffs, and level-0 traders obtain the lowest.

In Table 7 we show the results of RE panel regressions of per-period profits on an indicator variable for level-(1-3) and on an indicator variable for level-(4-∞). The top panel of the table shows the results when we use Nagel's classification, while the bottom panel reports the results based on our broader classification. As before, we take into account unobserved correlation among traders within each session, by computing bootstrap standard errors clustered at the session level.

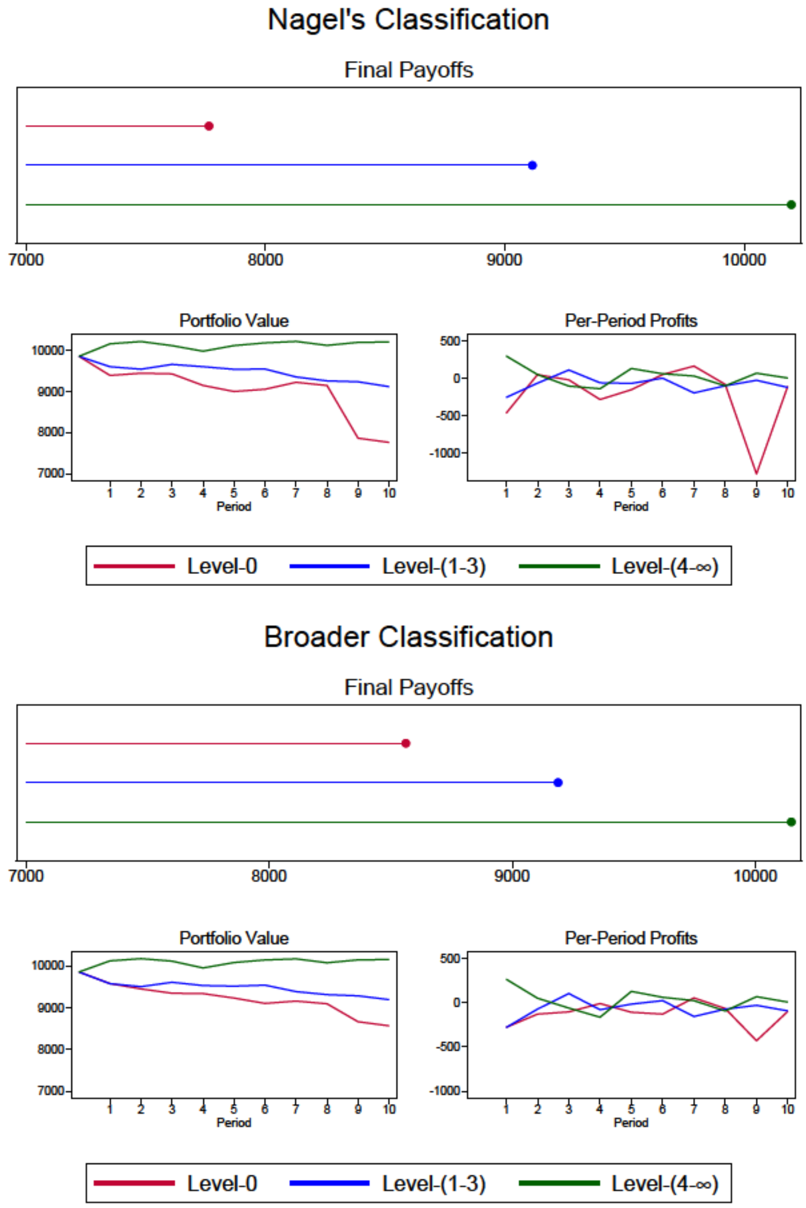
On average, level-(1-3) traders secure ECU 135 more in per-period profits than level-0 traders, while level-(4-∞) traders earn, on average, ECU 243 more in per-period profits than level-0 traders. Both coefficients on the level-(1-3) and level-(4-∞) indicator variables are significant at the 1% level (p-values<0.001). Moreover, we reject the null that the per-period profits realized by level-(1-3) and level-(4-∞) traders are the same (p-value=0.013). These estimates imply that, compared to level-0 traders, level-(1-3) traders make 17% more in final payoffs (ECU 9,115 versus ECU 7,764), and level-(4-∞) traders make 31% more in final payoffs (ECU 10,197

[10] We refer the reader to the Appendix for the details.

[11] The positive relation between the ToM score and measures of strategic sophistication is in line with existing results in the literature, e.g., Choi et al. 2022.

[12] The ToM score is positively and significantly (at the 1% level) correlated with whether a trader plays the IGG correctly (0.28) and with the CRT score (0.20). Its correlation with the IQ score is 0.05 and not statistically significant. The results described in the text and reported in Table A.3 in the Appendix remain unchanged when we control for IGG, IQ and CRT.

Figure 5: Trading Game Performance by Level of Strategic Sophistication



versus ECU 7,764). Results are similar when we adopt our broader classification of strategic sophistication (see lower panel of Table 7). These patterns are also robust to alternative ways of aggregating levels of strategic sophistication (e.g., level-0, level-(1-8), level-∞; or level-0, level-1, level-(2-∞)).

Finally, in Table 8, we use a different measure of strategic sophistication, the distance from the target, instead of the simple GG choice. The estimated RE coefficient is -0.101, which is statistically significant at the 1% level (p-value$<$0.001), indicating that as the distance from the target increases, subjects make lower per-period profits. The estimated coefficient implies that a 10-unit decrease in the distance from the target is associated with an ECU 32 increase in the per-period profit. Over the course of the experiment, this would result in a increase of ECU 320 in final payoffs, corresponding to 3.4% of average final payoff in the sample.

Table 7: Panel Regressions of Per-Period Profits on Level of Strategic Sophistication

Level-K: Nagel's Classification	
Level-(1-3)	135.153***
	(23.779)
Level-(4-∞)	243.353***
	(30.714)
Constant	-188.934***
	(54.111)
Test Level-(1-3) = (4-∞) p-val	0.013
N	410
Level-K: Broader Classification	
Level-(1-3)	62.629
	(51.647)
Level-(4-∞)	158.618***
	(31.265)
Constant	-167.419***
	(62.559)
Test Level-(1-3) = (4-∞) p-val	0.021
N	560

Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Level-(1-3) = (4-∞) is a one-tail t-test; the null that the coefficient of Level-(1-3) is equal to the coefficient of Level-(4-∞) is tested against the alternative that the former is smaller than the latter. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

Table 8: Panel Regressions of Per-Period Profits on Distance from Target

(GG Choice-Target)^2	-0.101***
	(0.024)
Constant	-46.003
	(49.225)
N	560

Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

5.2 Profit Volatility and Strategic Sophistication

An important question is whether traders with higher levels of strategic sophistication secure higher profits at the cost of higher profit volatility. We measure a trader's profit volatility as the standard deviation of the percent difference in the value of their portfolio between two consecutive periods. We then regress this volatility measure on the GG choice using progressively richer specifications. The results of these regressions are reported in Table 9.

The coefficient of the GG choice across all specifications is positive and relatively stable at about 0.1. This indicates that the volatility of a trader's portfolio *decreases* as the level of strategic sophistication increases (GG choice decreases). This estimated relation between profit volatility and GG choice is significant at the 5% level in columns (i)-(ii) and at the 10% level in columns (iii)-(iv), while it is not statistically distinguishable from 0 in columns (v)-(vi), with

Table 9: Regressions of Profit Volatility on GG Choice

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	0.094**	0.094**	0.100*	0.103*	0.091	0.086
	(0.042)	(0.038)	(0.055)	(0.058)	(0.058)	(0.066)
Male		0.605	0.578	0.593	0.493	-0.123
		(1.419)	(1.629)	(2.024)	(2.009)	(3.359)
Age		0.159**	0.161	0.160	0.159	0.225
		(0.073)	(0.102)	(0.103)	(0.103)	(0.156)
IGG Correct			-0.333	-0.447	-0.253	-0.259
			(1.513)	(1.687)	(1.864)	(1.616)
IQ Z-Score			0.043	0.450	0.298	-0.236
			(0.666)	(1.081)	(1.168)	(1.625)
CRT Z-Score			0.288	0.705	0.720	1.901
			(1.135)	(1.439)	(1.545)	(2.420)
IQ/CRT Confidence				0.518	0.370	0.091
				(0.756)	(0.797)	(1.285)
$\|BRET-50\|_{BRET<50}$					0.075	0.141
					(0.115)	(0.173)
$\|BRET-50\|_{BRET>50}$					0.087	0.120
					(0.069)	(0.113)
Constant	2.982***	-2.681	-2.814	-2.934	-3.180	-5.031
	(1.078)	(2.451)	(3.485)	(3.612)	(3.618)	(5.173)
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.988	0.952	0.962	0.874
Test Non-Cog=0 (p-val)						0.825
N	56	56	56	56	56	56

OLS estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

p-values of 0.117 and 0.193, respectively. None of the other regressors correlates significantly with profit volatility with the exception of age in one specification only (column (ii)).

6 Strategic Sophistication and Trading Skills

There are three ways in which a trader can earn profits: i) by holding the asset and earning its dividends; ii) by avoiding paying the fees; iii) by buying and selling the asset at different prices ("trading profits"). Trading profits are the most related to Keynes (1933)'s informal description of the stock market as a beauty contest and are the reason why strategic sophistication as measured in the Guessing Game may be important for trading activity: as in the Guessing Game, a subject is asked to guess what others will guess, similarly, profits from trading are due to one's ability to guess correctly other traders' willingness to buy or sell the asset in future periods.

In this section, we study the relationship between a trader's performance in the Guessing Game and each of these sources of profits (or loss); the dividends received, the fees paid, and the realized trading profits.

To study the relationship between dividends earned and strategic sophistication, for each

Table 10: Panel Regressions of Dividends on GG Choice

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	5.681	5.645*	4.438	4.559	4.658	4.673
	(3.537)	(3.344)	(3.606)	(3.649)	(3.702)	(4.221)
Male		-45.926	-36.705	-35.899	-36.433	-43.414
		(132.520)	(144.707)	(154.189)	(164.660)	(248.106)
Age		8.233	9.564	9.487	9.450	11.928
		(6.715)	(7.487)	(7.632)	(8.026)	(11.434)
IGG Correct			101.876	95.811	88.875	120.252
			(179.867)	(186.952)	(184.590)	(139.233)
IQ Z-Score			39.129	60.844	64.023	27.619
			(29.462)	(55.858)	(49.013)	(56.904)
CRT Z-Score			-91.990**	-69.715	-74.028	-4.603
			(42.366)	(63.661)	(70.594)	(96.968)
IQ/CRT Confidence				27.652	29.686	34.771
				(51.376)	(47.068)	(54.730)
$\|BRET-50\|_{BRET<50}$					-1.950	-0.054
					(7.586)	(11.148)
$\|BRET-50\|_{BRET>50}$					-0.359	1.011
					(5.162)	(10.240)
Constant	21.428	-204.159	-247.979	-254.387	-244.593	-339.594
	(85.816)	(286.254)	(321.507)	(327.486)	(335.248)	(441.474)
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.091	0.167	0.126	0.782
Test Non-Cog=0 (p-val)						0.147
N	560	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

trader, we compute the amount of dividends received at the end of each period. In Table 10, we show the results of a RE regression of per-period dividends on GG choice; each column contains the results of a progressively richer specification. There is no evidence that more sophisticated traders receive higher dividends: while the coefficient for the GG choice is positive across specifications, suggesting that less sophisticated traders receive higher dividends, it is not statistically significant except in column (ii) where it is significant at the 10% level.

In Table 11, we show the results of a RE regression of fees on strategic sophistication. Recall that, at the beginning of each period, half of the traders are randomly selected to pay a fee of ECU 50 for each asset held in their portfolio at the end of the same period (this random selection is independently performed for each period). We regress the total fees paid by each trader at the end of each period on GG choice and other traders' characteristics, with each column of Table 11 representing a richer specification. The estimated coefficient for the GG choice is positive and statistically significant at the 10% in three out of six specifications; that is, there is weak evidence that less sophisticated traders pay more fees.

Finally, to study the relation between strategic sophistication and realized trading profits,

Table 11: Panel Regressions of Fees on GG Choice

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	1.735*	1.728*	1.396	1.420	1.570*	1.398
	(0.983)	(0.973)	(0.934)	(0.957)	(0.883)	(0.870)
Male		-16.541	-14.000	-13.841	-13.345	-17.239
		(24.881)	(27.890)	(29.250)	(31.499)	(49.641)
Age		1.200	1.607	1.592	1.583	1.749
		(1.555)	(2.012)	(2.069)	(2.078)	(2.810)
IGG Correct			27.556	26.356	20.938	14.610
			(43.288)	(44.799)	(45.351)	(34.518)
IQ Z-Score			12.247	16.542	19.526	17.931
			(8.647)	(15.239)	(14.731)	(16.256)
CRT Z-Score			-26.057***	-21.651	-24.171	-7.500
			(6.347)	(14.735)	(16.104)	(20.238)
IQ/CRT Confidence				5.469	7.787	11.655
				(13.720)	(12.190)	(12.967)
$\|BRET-50\|_{BRET<50}$					-1.685	-1.388
					(1.091)	(1.915)
$\|BRET-50\|_{BRET>50}$					-0.895	-0.152
					(1.353)	(2.406)
Constant	23.233	-1.021	-14.313	-15.580	-8.155	-10.024
	(26.283)	(68.966)	(88.602)	(89.352)	(88.459)	(102.199)
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.001	0.028	0.051	0.628
Test Non-Cog=0 (p-val)						0.357
N	560	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

for each period we compute the difference between the total amount that subjects receive from selling assets and the total amount they pay to purchase them.

The RE regression analysis in Table 12 shows that sophisticated traders gain substantially more from trading activity than their less sophisticated counterparts. Across all specifications, a 10-unit decrease in the GG choice is associated with approximately a ECU 70 increase in realized trading profits in each period. Over the course of the entire trading game, this would translate into an increase of final payoffs corresponding to 7.4% of the average final payoffs. In other words, sophisticated traders' abilities to make higher profits are mainly driven by their ability to time the market and trade at favorable prices. This is similar to the ability to assess market sentiment, which Keynes described as a feature of a successful trader.

7 A Comparison with a Sample of Student Subjects

To contrast traders' behavior to that of students, the traditional subjects in laboratory experiments, we also repeated our experiment with undergraduate students from all disciplines enrolled in the subject pool of ELFE at UCL. We know from Corgnet et al. (2018) that students

Table 12: Panel Regressions of Realized Trading Profits on GG Choice

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	-7.733***	-7.694***	-6.326**	-6.468**	-6.306*	-7.071*
	(2.590)	(2.299)	(3.202)	(3.244)	(3.408)	(3.709)
Male		19.208	18.660	17.711	21.838	-15.770
		(68.224)	(85.599)	(100.468)	(101.178)	(194.615)
Age		-10.826*	-9.048	-8.957	-8.836	-11.163
		(6.096)	(5.577)	(5.488)	(5.890)	(9.465)
IGG Correct			-5.346	1.790	9.822	-32.503
			(107.483)	(114.351)	(115.080)	(129.713)
IQ Z-Score			28.720	3.170	1.416	35.993
			(46.153)	(59.941)	(51.499)	(68.560)
CRT Z-Score			35.841	9.631	17.890	-48.229
			(58.972)	(67.482)	(76.094)	(117.011)
IQ/CRT Confidence				-32.535	-32.090	-36.198
				(37.456)	(30.922)	(55.957)
$\|BRET - 50\|_{BRET<50}$					1.636	0.027
					(5.959)	(10.009)
$\|BRET - 50\|_{BRET>50}$					-1.947	-1.177
					(4.985)	(9.122)
Constant	175.030**	507.775***	420.775*	428.314*	416.106*	554.606*
	(69.058)	(191.859)	(216.126)	(218.682)	(227.401)	(332.390)
Period Fixed Effects	No	No	No	No	No	No
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.283	0.999	0.996	0.930
Test Non-Cog=0 (p-val)						0.142
N	560	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Period fixed-effects are omitted since the average of realized capital gains is zero by construction in each period. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p - value < 0.1$, **: $p - value < 0.05$, ***: $p - value < 0.01$.

with higher cognitive abilities earn higher profits in a trading game. However, their design of the trading game is different from ours and they do not run a Guessing Game in their experiment.[13] It is, therefore, worth considering how student subjects behave in our own experiment.

We recruited the students so that their gender composition would match that of the traders: 48 traders (86% of the sample) and 44 students are male (79% of the sample). Table 13 shows descriptive statistics for students' final and per-period profits, GG choice, IGG choice, cognitive skills, confidence, and risk aversion. The last column of the table reports the p-value of a Mann-Whitney rank-sum test comparing students' and traders' outcomes.[14] Students' average final payoff was ECU 9,340, similar to that of traders and not statistically different from it. Students were paid at the exchange rate of £0.25 = ECU 100; therefore, their final average payoff corresponds to £23 (approximately equal to $30.50). There is substantial volatility in students' final payoffs with a standard deviation of ECU 2,424, which is significantly larger than

[13]In particular, in their experiment, each market only lasts for one period.

[14]The tests are run at the individual level, except for final and per-period profits in the Trading Game, and GG choice where the tests are run at the session level to take into account interactions among participants.

Table 13: Students' Experiment Outcomes and Comparison with Traders

	Students Mean	Traders Mean	Students vs. Traders MW test (p-value)
Final Profits	9,340.18	9,397.32	0.655
Per-period Profits	-68.12	-45.27	0.503
GG Choice	37.58	22.63	0.002
IQ Score (0-15)	9.82	8.70	0.013
CRT Score (0-7)	4.59	5.14	0.101
Confidence	-0.64	0.24	0.153
BRET Choice	39.68	49.68	0.000

the one observed among traders.[15]

Table 14: Students' GG Choice: Level-k Classification

	Nagel's Classification			**Broader Classification**		
	Choice Range	Freq.	Perc.	Choice Range	Freq.	Perc.
Level-0	[45, 50]	7	20.00	(37, 100]	24	42.86
Level-1	[30, 37]	15	42.86	(25, 37]	15	26.79
Level-2	[20, 25]	7	20.20	(16, 25]	9	16.07
Level-3	[13, 16]	3	8.57	(11, 16]	3	5.36
Level-4	[9, 11]	0	0.00	(7, 11]	2	3.57
Level-5	[6, 7]	0	0.00	[6, 7]	0	0.00
Level-6	[4, 5]	1	2.86	[4, 5]	1	1.79
Level-7	[3, 3]	0	0.00	[3, 3]	0	0.00
Level-8	[2, 2]	0	0.00	[2, 2]	0	0.00
Level-∞	[0, 1]	2	5.71	[0, 1]	2	3.57
Total		35	100.00		56	100.00

Compared to traders, students guess higher numbers in the GG (average and median of 37.58 and 35.00 versus 22.63 and 22.5 for traders), thereby exhibiting a significantly lower level of strategic sophistication (p-value=0.002). As can be seen from Table 14, there is a small proportion of level-∞ students (about 6% using the Nagel's classification, and 4% using our broader classification), whereas roughly two-thirds of students are classified as level-1 or level-0.

The number of students able to solve the Individual Guessing Game correctly is almost the same as that of traders (10 versus 11, corresponding to about 18% of the sample). Students perform better than traders in the IQ test by about 1 point on average (9.8 versus 8.7), a statistically significant difference (p-value=0.013). They score slightly lower than traders in the CRT, with an average of 4.59 versus 5.14 (p-value=0.101). Students tend to be underconfident, with an average confidence index of -0.64 (statistically different from 0 at the 10% level – p-value=0.089). They are more risk averse than traders: the average number of boxes opened by

[15] We use the Brown and Forsythe (1974)'s test for the equality of variances, which is robust under non-normality and and when distributions are skewed.

students in the BRET is 40, versus 50 opened by traders, a statistically significant difference (p-value < 0.001). This result aligns with previous research findings on professional traders (e.g., Grinblatt et al., 2012).

Table 15: Panel Regressions of Per-Period Profits on GG Choice (students)

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
GG Choice	-2.451***	-1.675	-2.211	-2.599	-2.939**	-2.451
	(0.845)	(1.203)	(1.574)	(1.795)	(1.437)	(2.716)
Male		121.052	91.395	84.592	43.305	4.565
		(127.246)	(147.037)	(140.295)	(156.130)	(170.225)
Age		-4.230	-17.472	-18.904	-5.425	-7.746
		(13.782)	(11.832)	(11.926)	(19.855)	(25.033)
IGG Correct			-6.170	-1.861	-14.465	2.256
			(62.009)	(55.603)	(66.666)	(81.088)
IQ Z-Score			83.620***	79.249***	71.902**	83.739*
			(27.727)	(27.334)	(28.221)	(45.380)
CRT Z-Score			17.755	14.383	13.360	44.591
			(44.759)	(47.094)	(38.470)	(47.426)
IQ/CRT Confidence				-10.372	-7.702	-12.730
				(15.051)	(14.334)	(22.026)
$\lvert BRET-50\rvert_{BRET<50}$					-3.701	-5.271
					(2.985)	(3.614)
$\lvert BRET-50\rvert_{BRET>50}$					-14.575	-14.988
					(12.732)	(11.421)
Constant	-23.075	-63.021	252.752	294.701	135.296	208.574
	(79.702)	(316.557)	(310.057)	(310.056)	(367.517)	(511.312)
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-Cognitive Skills	No	No	No	No	No	Yes
Test Cog=0 (p-val)			0.011	0.024	0.003	0.053
Test Non-Cog=0 (p-val)						0.371
N	560	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Cog=0 is a joint test that the coefficients of IGG Correct, IQ Z-Score and CRT Z-Score are all equal to 0. Test Non-Cog=0 is a joint test that the coefficients of the non-cognitive skills are all equal to 0. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

Table 15 shows the results of RE regressions of per-period profits in the Trading Game on GG choice. Following the same approach as that adopted for the traders' pool, we use progressively richer specifications to control for individuals' demographic characteristics and skills. The estimates for the GG choice coefficients are significant only for some specifications. When significant, a 10-unit decrease in GG choice is associated with an increase in per-period profits of ECU 24, in the baseline specification (column(i)), and ECU 29 in the specification with all the controls except non-cognitive skills (column (v)); these effects translate into increases in final payoffs ranging from 2.6% to 3.1% of the average final payoff. As was the case for traders, we also observe that the effect of GG choice on per-period profits in the students' sample is stronger at the beginning and at the end of the experiment relative to the middle phase (see Table A.2 in the Appendix).

In Table 16, we explore differences in per-period profits by level of strategic sophistication.

Table 16: Panel Regressions of Per-Period Profits on Level of Strategic Sophistication (students)

Level-K: Nagel's Classification	
Level-(1-3)	-116.069
	(82.643)
Level-(4-∞)	-224.095***
	(77.087)
Constant	-134.600
	(83.502)
Test Level-(1-3) = (4-∞) p-val	0.000
N	350
Level-K: Broader Classification	
Level-(1-3)	79.431
	(52.806)
Level-(4-∞)	39.092
	(52.886)
Constant	-156.966***
	(40.223)
Test Level-(1-3) = (4-∞) p-val	0.280
N	560

Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Test Level-(1-3) = (4-∞) is a one-tail t-test; the null that the coefficient of Level-(1-3) is equal to the coefficient of Level-(4-∞) is tested against the alternative that the former is smaller than the latter. *: $p - value < 0.1$, **: $p - value < 0.05$, ***: $p - value < 0.01$.

Using both Nagel (1985)'s and our broader classifications, we find no evidence that level-(1-3) or level-(4-∞) students exhibit a better trading outcome throughout the trading game compared to their level-0 counterparts.[16] In Table 17, we find a negative and significant relationship between per-period profits and distance between GG choice and target. Specifically, we estimate that a 10-unit decrease in the distance from the target is associated with an ECU 23 increase in per-period profits, an effect somewhat smaller in magnitude than what we found for traders (ECU 32). Overall, these results suggest that, within the students' pool, the relationship between trading performance and GG choice is weaker and less robust than that within the traders' pool.

In contrast with what we have observed for professional traders, an important predictor of trading performance among students is IQ level, confirming Corgnet et al. (2018)'s result on the importance of cognitive skills. A one standard deviation increase in the IQ score is associated with an increase in per-period profits ranging between ECU 72 and ECU 83. As can be seen in Table 15, the effect of IQ on per-period profits is highly significant across specifications, except in column (vi) where it is only significant at the 10% level. These estimates imply that a one standard deviation increase in IQ score is associated with an increase in final payoffs approximately between 8% and 9% of students' average final payoff. Similarly to our findings for professional traders, no other demographic variable or skill correlates significantly with trading

[16]Note that there are only two level-∞ students who make *less* per-period profits than everybody else.

Table 17: Panel Regressions of Per-Period Profits on Distance from Target (Students)

(GG Choice-Target)^2	-0.063**
	(0.026)
Constant	-78.612
	(60.993)
N	560

Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

performance among students.

In general, the analyses performed in this section show that, in a sample of undergraduate students which is typically used in laboratory experiments, cognitive abilities, as measured by the IQ score, appear to be the main predictor of trading performance. Although the impact of strategic sophistication is positive, it is relatively small in magnitude and not always statistically significant. The results for student subjects contrast with the patterns observed for professional traders, for whom differences in IQ do not matter, whereas the driver of trading performance is the level of strategic sophistication.

8 Conclusions

We have shown that that strategic sophistication, as measured in the Guessing Game (e.g., through level-k theory), is the only significant determinant of professional traders' profits in a trading game. In contrast, profits are not driven by individual characteristics such as cognitive abilities or behavioral traits. Moreover, higher profits are due to the ability to trade at favorable prices rather than to the ability to forecast the fundamental value, that is, to earn higher dividends. As seen from a theoretical viewpoint, our experimental results suggest that depth of reasoning, as captured by the level-k theory, has a predictive power in how well professional traders trade in financial market. From a different angle, the results indicate that, in the selection of traders, financial companies may want to look at the candidate's strategic sophistication, rather than at their cognitive or non-cognitive abilities or other individual traits.

9 References

References

[1] Alevy, Jonathan E., Michael S. Haigh, and John A. List. (2007). "Information cascades: Evidence from a field experiment with financial market professionals." The Journal of Finance, 62(1): 151-180.

[2] Angrisani, Marco, Marco Cipriani, Antonio Guarino, Ryan Kendall, and Julen Ortiz de Zarate Pina. (2020). "Risk Preferences at the Time of COVID-19: An Experiment with Professional Traders and Students." Staff Reports 927, Federal Reserve Bank of New York.

[3] Baron-Cohen, Simon, Sally Wheelwright, Jacqueline Hill, Yogini Raste, and Ian Plumb. (2001). "The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism." The Journal of Child Psychology and Psychiatry and Allied Disciplines 42(2): 241-251.

[4] Brown, Morton and Forsythe, Alan. (1974). "Robust Tests for the Equality of Variances." Journal of the American Statistical Association, 69(346): 364-367.

[5] Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. (2004). "A Cognitive Hierarchy Model of Games." The Quarterly Journal of Economics, 119, 861-898.

[6] Cameron, Colin A., and Douglas L. Miller. (2015). "A Practitioner's guide to cluster-robust inference." Journal of Human Resources, 50.2: 317-372.

[7] Cipriani, M., and Antonio Guarino. (2009). "Herd behavior in financial markets: an experiment with financial market professionals." Journal of the European Economic Association, 7(1): 206-233.

[8] Choi, Syngjoo, Seonghoon Kim, and Wooyoung Lim. (2022). "Strategic Thinking Skills: A Key to Collective Economic Success". Working paper.

[9] Cobb-Clark, Deborah A., and Stefanie Schurer. (2013)."Two economists' musings on the stability of locus of control." The Economic Journal 123(570): F358-F400.

[10] Corgnet, Brice, Mark DeSantis, and David Porter. (2018). "What makes a good trader? On the role of intuition and reflection on trader performance." The Journal of Finance, 73(3): 1113-1137.

[11] Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. (2013). "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." Journal of Economic Literature, 51, 5-62.

[12] Crosetto, Paolo, and Antonio Filippin. (2013). "The "bomb" risk elicitation task." Journal of Risk and Uncertainty, 47(1): 31-65.

[13] Crosetto, Paolo, and Antonio Filippin. (2016). "A theoretical and experimental appraisal of four risk elicitation methods." Experimental Economics, 19(3): 613-641.

[14] Duckworth, Angela Lee, and Patrick D. Quinn. (2009). "Development and validation of the Short Grit Scale (GRIT–S)." Journal of Personality Assessment 91(2): 166-174.

[15] Frederick, Shane. (2005). "Cognitive reflection and decision making." Journal of Economic perspectives, 19(4): 25-42.

[16] Grinblatt, Mark, Matti Keloharju, and Juhani T. Linnainmaa. (2012). "IQ, trading behavior, and performance." Journal of Financial Economics, 104(2): 339-362.

[17] Haigh, Michael S., and John A. List. (2005). "Do professional traders exhibit myopic loss aversion? An experimental analysis." The Journal of Finance, 60(1): 523-534.

[18] Harrison, Glenn W., and John A. List. (2004). "Field experiments." Journal of Economic literature, 42(4): 1009-1055.

[19] Keynes, John Maynard. (2018). "The general theory of employment, interest, and money." Springer. (Original work published 1937).

[20] List, John A., and Michael S. Haigh. (2010). "A simple test of expected utility theory using professional traders." Proceedings of the National Academy of Sciences, 102(3): 945-948.

[21] List, John A., and Michael S. Haigh. (2010). "Investment under uncertainty: Testing the options model with professional traders." The Review of Economics and Statistics, 92(4): 974-984.

[22] Nagel, Rosemarie. (1995). "Unraveling in guessing games: An experimental study." The American Economic Review, 85(5): 1313-1326.

[23] Rammstedt, Beatrice, and Oliver P. John. (2007). "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German." Journal of Research in Personality 41(1): 203-212.

[24] Raven, John C. (1941). "Standardization of progressive matrices." 1938. British Journal of Medical Psychology, 19(1): 137-150.

[25] Raven, John C. (1990). "Advanced progressive matrices sets 1 and 2." prepared by J. C. Raven. Oxford: Oxford Psychologists Press.

[26] Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams. (1988). "Bubbles, crashes, and endogenous expectations in experimental spot asset markets." Econometrica: Journal of the Econometric Society, 1119-1151.

[27] Snyder, Mark, and Steve Gangestad. (1986). "On the nature of self-monitoring: Matters of assessment, matters of validity." Journal of Personality and Social Psychology 51(1): 125-139.

[28] Toplak, Maggie E., Richard F. West, and Keith E. Stanovich. (2014). "Assessing miserly information processing: An expansion of the Cognitive Reflection Test." Thinking & Reasoning, 20(2): 147-168.

Appendix

A Additional Results

Table A.1: Panel Regressions of Per-Period Profit on GG Choice
Focus on Non-Cognitive Skills

	(i)	(ii)	(iii)	(iv)	(v)
GG Choice	-3.573***	-3.261***	-3.223**	-3.189**	-3.796**
	(1.198)	(1.244)	(1.294)	(1.260)	(1.553)
Male	-7.084	-4.170	-4.260	1.350	-41.945
	(78.275)	(65.063)	(69.271)	(54.857)	(76.740)
Age	-0.968	-1.076	-1.032	-0.992	-0.984
	(4.295)	(4.114)	(4.424)	(4.257)	(5.050)
IGG Correct	74.736	91.063	73.541	78.287	73.139
	(89.589)	(93.171)	(69.553)	(83.141)	(120.803)
IQ Z-Score	48.416	45.662	46.317	46.080	45.681
	(51.858)	(46.810)	(46.747)	(46.435)	(55.980)
CRT Z-Score	-31.692	-38.203	-32.053	-32.063	-45.332
	(40.965)	(45.712)	(45.560)	(47.274)	(48.004)
IQ/CRT Confidence	-12.482	-11.602	-9.613	-10.192	-13.081
	(17.422)	(19.922)	(20.867)	(20.266)	(20.167)
$\|BRET-50\|_{BRET<50}$	0.986	1.474	1.388	1.291	1.362
	(1.905)	(2.291)	(2.020)	(2.149)	(2.464)
$\|BRET-50\|_{BRET>50}$	-0.247	-1.576	-1.386	-1.454	-0.013
	(3.157)	(2.885)	(2.947)	(2.323)	(2.735)
Openness Z-Score	-6.246				-15.114
	(34.917)				(43.386)
Conscientiousness Z-Score	-3.872				6.083
	(28.351)				(47.147)
Extraversion Z-Score	-6.974				-18.734
	(29.970)				(27.578)
Agreeableness Z-Score	32.586*				35.574*
	(16.958)				(21.604)
Neuroticism Z-Score	14.854				22.071
	(17.539)				(21.473)
Locus of Control Z-Score		13.047			21.598
		(19.019)			(27.687)
Grit Z-Score			-3.535		-27.678
			(19.607)		(44.431)
Self Monitoring Z-Score				-3.064	13.604
				(31.215)	(38.719)
Constant	19.634	13.158	13.560	6.468	53.607
	(171.018)	(183.879)	(199.786)	(169.821)	(192.552)
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes
N	560	560	560	560	560

Random-effects estimates. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

Table A.2: Panel Regressions: Differential Effect of GG Choice on Per-Period Profits across Experiment Phases (students)

Phase1 × GG Choice	-3.119**
	(1.399)
Phase2 × GG Choice	-0.804
	(1.285)
Phase3 × GG Choice	-3.979***
	(1.098)
Constant	2.031
	(85.696)
N	560

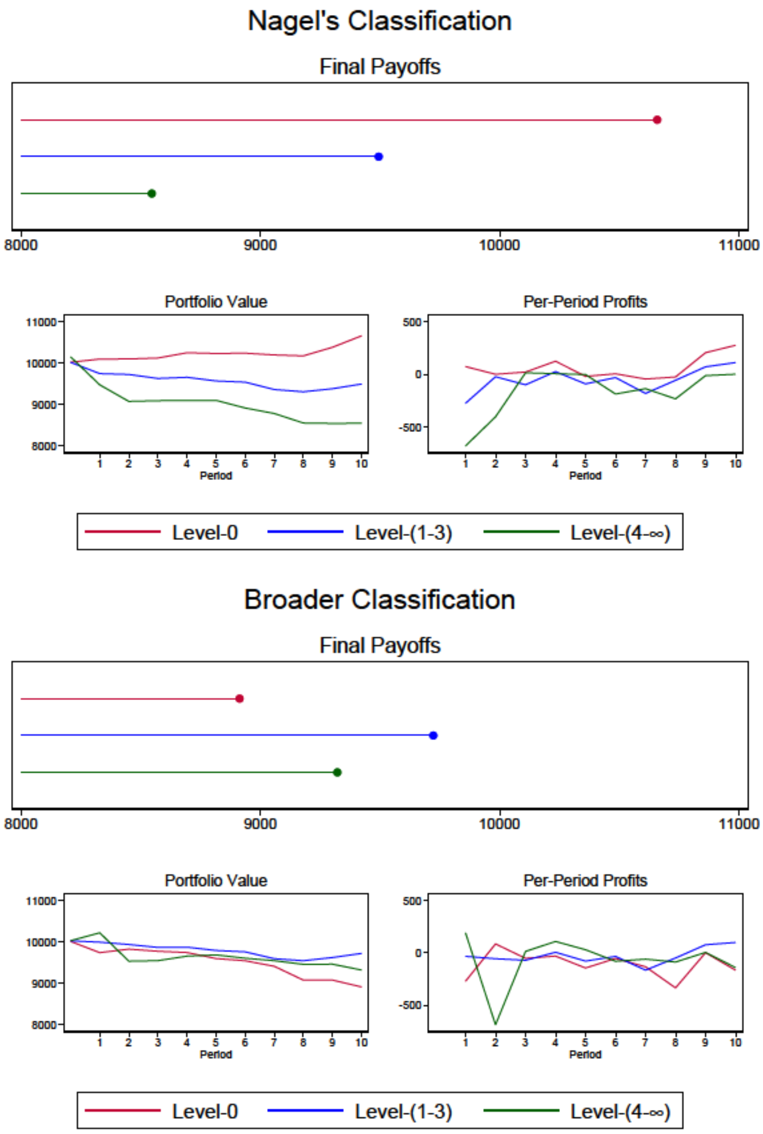
Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. Phase1 is a dummy for periods 1-3; Phase2 is a dummy for periods 4-7; Phase3 is a dummy for periods 8-10. H_0 : GG Choice$_{(Phase1)}$=GG Choice$_{(Phase2)}$, p-value=0.290. H_0 :GG Choice$_{(Phase1)}$=GG Choice$_{(Phase3)}$, p-value=0.580. H_0 :GG Choice$_{(Phase2)}$=GG Choice$_{(Phase3)}$, p-value=0.001. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

Table A.3: Panel Regressions: Effect of GG Choice and ToM on Per-Period Profits

	(i)	(ii)	(iii)
GG Choice	-3.303***		-3.053***
	(0.733)		(0.756)
ToM Score		8.305	2.335
		(7.257)	(7.437)
N	360	360	360

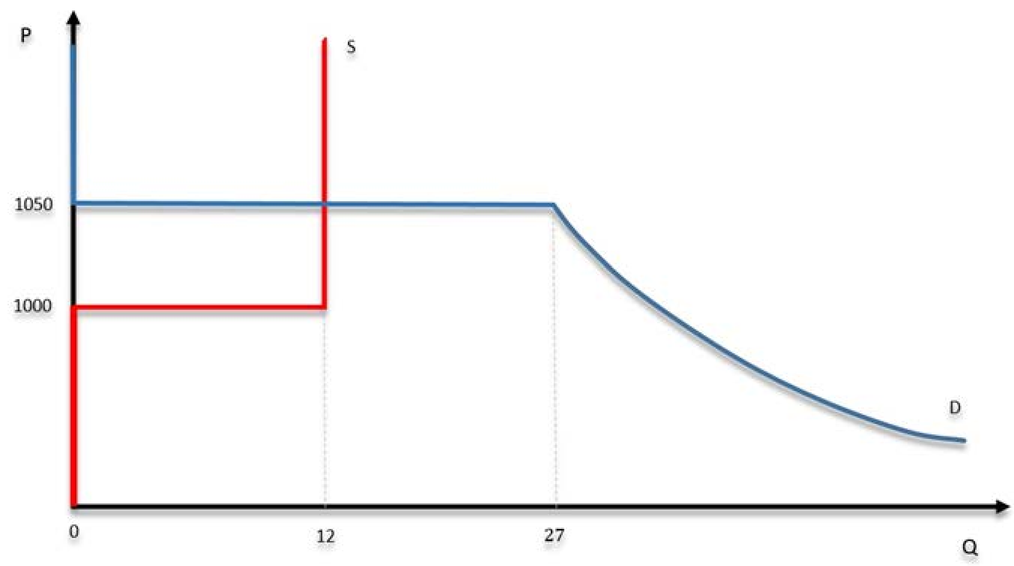
Random-effects estimates. Period fixed-effects are included. Bootstrap standard errors (500 replications) clustered at the session level in parentheses. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

Figure A.1: Trading Performance by Level of Strategic Sophistication (students)



B Additional Experiment Details and Instructions

Figure B.1: REE for $d_t = 150$



B.1 Experiment timeline

Each session was organized in the following way:

- Upon arrival to the laboratory, subjects received documentation about the general nature of the study and the ethical protocol (e.g., no deception, data confidentiality). We took a scan of their right hand, for the purpose of measuring the ratio between the lengths of the 2nd and 4th fingers. We recorded this measure based on scans of the subjects' right hands, since previous research shows that this is more reliable than using left hand measurements (Hönekopp et al., 2010). The digit ratio is taken as a measure of exposure to prenatal androgens, a genetic feature that plays a role in brain development and future behavior.

- We then started the Trading Game. Subjects read written instructions on their screens (see section B.2). The instructions included questions, so that subjects could check their understanding of the rules of the game. Subjects could also ask us clarifying questions, which we answered privately.

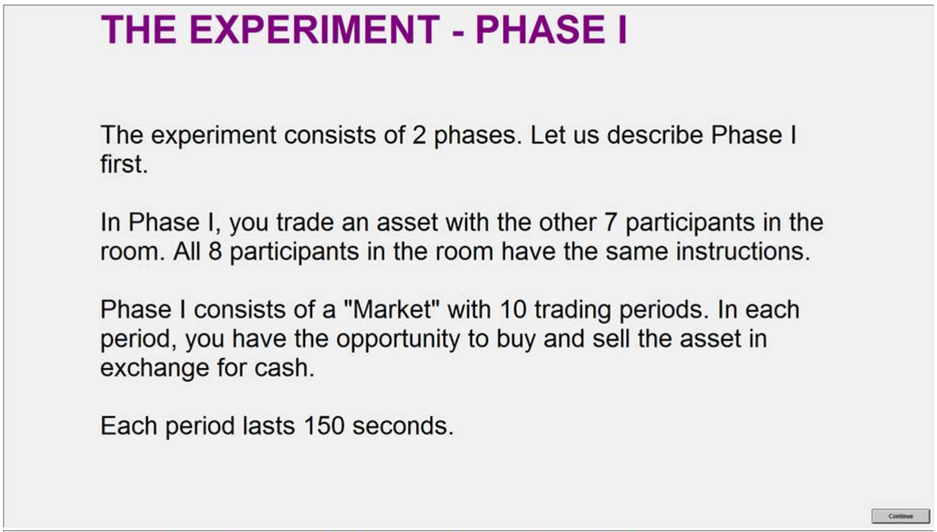
- After reading the instructions, subjects traded in a two-period "practice market" to familiarize themselves with the trading platform (see Figure B.2).

- After the practice market, a 10-period trading experiment started.

- When the Trading Game was over, subjects participated in the Guessing Game, in the Individual Guessing Game, and in a series of individual-level tasks to measure cognitive and non-cognitive skills. The Guessing Game, the Individual Guessing Game, and the test for risk attitudes were computerized, whereas the other tests were conducted using pen and paper (see section B.3).

- Finally, subjects completed a form with demographic and professional information and received their payments in private.

B.2 Trading Game instructions



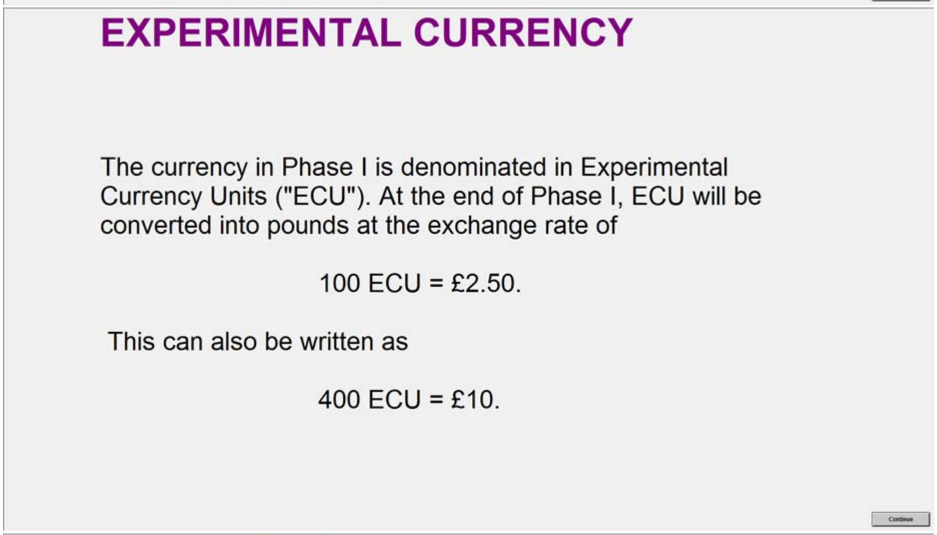
THE EXPERIMENT - PHASE I

The experiment consists of 2 phases. Let us describe Phase I first.

In Phase I, you trade an asset with the other 7 participants in the room. All 8 participants in the room have the same instructions.

Phase I consists of a "Market" with 10 trading periods. In each period, you have the opportunity to buy and sell the asset in exchange for cash.

Each period lasts 150 seconds.



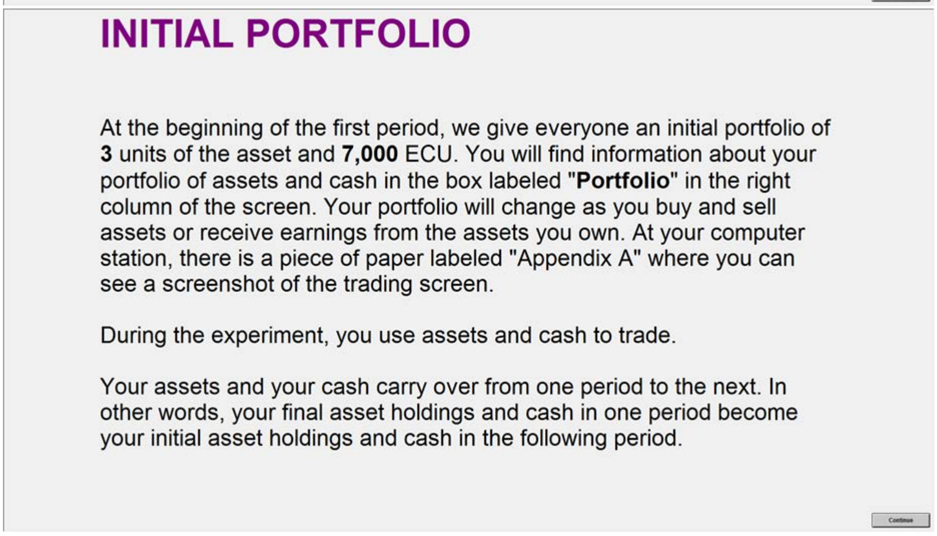
EXPERIMENTAL CURRENCY

The currency in Phase I is denominated in Experimental Currency Units ("ECU"). At the end of Phase I, ECU will be converted into pounds at the exchange rate of

100 ECU = £2.50.

This can also be written as

400 ECU = £10.



INITIAL PORTFOLIO

At the beginning of the first period, we give everyone an initial portfolio of **3** units of the asset and **7,000** ECU. You will find information about your portfolio of assets and cash in the box labeled "**Portfolio**" in the right column of the screen. Your portfolio will change as you buy and sell assets or receive earnings from the assets you own. At your computer station, there is a piece of paper labeled "Appendix A" where you can see a screenshot of the trading screen.

During the experiment, you use assets and cash to trade.

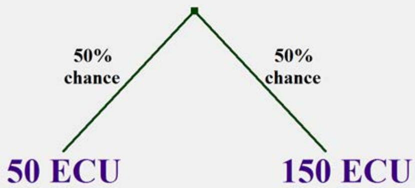
Your assets and your cash carry over from one period to the next. In other words, your final asset holdings and cash in one period become your initial asset holdings and cash in the following period.

ASSET EARNINGS

At the end of each period, the asset produces earnings. Each unit of the asset produces either **50** ECU or **150** ECU.

The computer determines whether the earnings are 50 or 150 with a mechanism simulating the tossing of a fair coin. That is, each period there is an equal chance that the earnings per unit of the asset are 50 or 150.

The diagram below summarises how the asset earnings are determined in each period.



ASSET EARNINGS IN FUTURE PERIODS

The coin toss is repeated each period. Whether the earnings are 50 or 150 in one period does not depend on whether they were 50 or 150 in previous periods.

A unit pays you earnings at the end of each period in which you hold it in your portfolio. For instance, if you buy a unit of the asset in period 3 and you choose to keep it in your portfolio until period 10, you will receive earnings from period 3 until period 10, that is for 8 periods.

In any future period, you know that there is a 50/50 chance that earnings will be 50 or 150. Therefore, if you hold one unit of the asset in a future period, in expectation, it will pay you 50%*50 + 50%*150= **100** ECU.

CALCULATING THE EXPECTED EARNINGS

This table shows how much earning per unit of the asset you should expect in future periods (column 3) and in future and current periods combined (columns 4 and 5) according to whether you expect the earnings to be 50 or 150 in the current period.

Current period	Periods remaining	Future earnings in expectation	Current and future earnings in expectation if current earnings are 50	Current and future earnings in expectation if current earnings are 150
1	9	100*9 = **900**	900 + 50 = **950**	900 + 150 = **1050**
2	8	100*8 = **800**	800 + 50 = **850**	800 + 150 = **950**
3	7	100*7 = **700**	700 + 50 = **750**	700 + 150 = **850**
4	6	100*6 = **600**	600 + 50 = **650**	600 + 150 = **750**
5	5	100*5 = **500**	500 + 50 = **550**	500 + 150 = **650**
6	4	100*4 = **400**	400 + 50 = **450**	400 + 150 = **550**
7	3	100*3 = **300**	300 + 50 = **350**	300 + 150 = **450**
8	2	100*2 = **200**	200 + 50 = **250**	200 + 150 = **350**
9	1	100*1 = **100**	100 + 50 = **150**	100 + 150 = **250**
10	0	100*0 = **0**	0 + 50 = **50**	0 + 150 = **150**

EXAMPLES

In period 1, the asset pays 150.
What is the chance that earnings will be 150 in period 2?

- 0%
- 50%
- 100%

Assume you buy an asset and keep it until the end of period 10.
Do you expect the overall earnings to be higher if you buy it in period 5 or period 8?

- Period 5
- Period 8

In period 1, suppose that you think the current earnings are 150. What do you expect the current and future earnings will be if you hold the asset until the end of period 10?

- 900 ECU
- 950 ECU
- 1000 ECU
- 1050 ECU

In period 3, suppose that you buy an asset and that you think earnings are 150. Suppose you hold the asset until the end of period 6. How much do you expect to earn in those 4 periods?

- 400 ECU
- 450 ECU
- 500 ECU
- 550 ECU

Continue

INFORMATION

You will learn whether asset earnings are 50 or 150 only at the end of each period.

However, at the beginning of each period you will receive some information about earnings.

If earnings are 50, the computer will distribute blue and red balls across participants, as if it were using an urn containing 2 blue balls and 6 red balls. The computer will give each participant one ball. This means that 2 participants receive blue balls and 6 participants receive red balls.

If earnings are 150, the computer will distribute blue and red balls across participants, as if it were using an urn containing 6 blue balls and 2 red balls. The computer will give each participant one ball. This means that 6 participants receive blue balls and 2 participants receive red balls.

Therefore, if earnings are 50 ECU per unit, there is a 2/8 (25%) chance that you receive a blue ball and a 6/8 (75%) chance that you receive a red ball. If earnings are 150 ECU per unit, there is a 6/8 (75%) chance that you receive a blue ball and a 2/8 (25%) chance that you receive a red ball.

Continue

INFORMATION

To recap,
- If, in a period, earnings are 50 ECU per unit, there are more RED balls in the urn.
- If, in a period, earnings are 150 ECU per unit, there are more BLUE balls in the urn.

Therefore, the colour of the ball gives you some information about whether earnings are 50 ECU or 150 ECU.
You will see your ball colour in the box labeled **Earnings Information** on the right side of the screen.

In a period, earnings are 150 ECU.
How many participants will receive a BLUE ball?

- 0
- 2
- 6
- 8

In a period, earnings are 50 ECU.
Will you get a RED ball for sure?

- No
- Yes

In a period, earnings are 50 ECU.
How many participants will receive a BLUE ball?

- 0
- 2
- 6
- 8

In a period, earnings are 150 ECU.
Are you more likely to receive a BLUE ball or a RED ball?

- BLUE
- RED

Continue

TRADERS REQUIRED TO PAY A FEE

At the beginning of each period, the computer randomly selects 4 participants. These participants have to pay a "fee" of **50** ECU for each unit of the asset they have in their portfolio **at the end of the period**. The other 4 participants will not pay a fee.

The fee is automatically subtracted from the earnings at the end of each period. This means that if the asset pays 50, a participant who must pay the fee receives 50-50=0 ECU for each unit they hold at the end of the period. If the asset pays 150, they receive 150-50=100 ECU for each unit they hold at the end of the period.

You will be informed about whether you have to pay the fee at the beginning of the period. This will also be displayed in the box labelled **Fee information** on the right of the screen. Whether you have to pay the fee in one period does not depend on whether you have had to pay the fee in previous periods.

In period 1, you have to pay the 50 ECU fee.
What is the chance that you will have to pay the fee in period 2?

- 0%
- 25%
- 50%
- 100%

Continue

DIFFERENT VALUATIONS

Since every period has 4 participants who have to pay a fee, it will always be the case that other participants value the asset differently than you do.

In other words,
- If you have to pay a fee in a period, you know that there are 4 other participants who earn *more* from the asset than you do in that period.
- If you **do not** have to pay a fee in a period, you know that there are 4 other participants who earn *less* from the asset than you do in that period.

Continue

TRADING

- In each period, trading lasts for 150 seconds.

- During that time, participants trade by submitting Buy and Sell Offers.

Continue

BUY OFFERS

- You can submit a Buy Offer by entering a price in the box titled "Buy Offer (Bid)" and pressing OK.

- You can submit as many Buy Offers as you want, but the sum of the prices of all the offers you submit cannot exceed the amount of cash in your portfolio.

Continue

SELL OFFERS

- You can submit a Sell Offer by entering a price in the box titled "Sell Offer (Ask)" and pressing OK.

- You can submit as many Sell Offers as you want, but the total number of assets you offer cannot exceed the number of assets in your portfolio.

Continue

BUY AND SELL OFFERS

On the top of the screen, you will see 2 boxes: "Buy Offers" and "Sell Offers".

Each box lists all outstanding buy and sell offers. You can see all the offers by scrolling through them. Your own offers are marked by an asterisk in the left column.

You can cancel any offer you submitted that has not been executed. To do so, select one of your outstanding offers and then click the "Cancel" button below the table.

Continue

TRADES

There are 2 ways to trade.

(1) There are "BUY" and "SELL" buttons. Pushing "BUY" will buy one unit of the asset at the lowest (best) submitted Sell Offer. Pushing "SELL" will sell one unit of the asset at the highest (best) submitted Buy Offer.

(2) A trade automatically occurs if the lowest price among all Sell Offers is lower than the highest price among all Buyer Offers.

In this situation, a participant is willing to pay more for the asset than another participant asks for it.

The computer recognizes this situation and the trade occurs automatically at the price of the offer submitted first.

Continue

Example

Suppose that the lowest Sell Offer is 300 and the highest Buy Offer is 200. Then, no trade is possible.

If, instead, you are willing to pay up to 350 for a unit, the only thing you need to do is to submit a Buy Offer of 300 (or higher).

The system recognizes that a trade is possible and immediately executes it. That is, the participant who originally submitted the Sell Offer receives 300 from you and you receive one unit of the asset from them.

Note that it does not matter whether you submit a Buy Offer of 300 or higher; the trade will always be executed at the price of 300 since the Sell Offer was submitted first.

Continue

TRADES (continued)

Therefore,
- To accept an existing Sell Offer and buy the asset, submit a Buy Offer at a price at least equal to the best (lowest) Sell Offer.

- To accept an existing Buy Offer and sell the asset, submit a Sell Offer at a price at most equal to the best (highest) Buy Offer.

The computer immediately executes the trade at the outstanding price.

Continue

TRADES (continued)

When a participant sells a unit of the asset, their cash is increased by the price of the asset, and the number of assets in their portfolio is decreased by one.

When a participant buys a unit of the asset, their cash is decreased by the price of the asset, and the number of assets in their portfolio is increased by one.

Suppose you have 200 in cash and you buy a unit of the asset at a price of 150.
How much cash will you have after the trade is executed?

- 50 ECU
- 150 ECU
- 200 ECU
- 350 ECU

Suppose you have 200 in cash and another participant buys a unit of the asset from you at a price of 200.
How much cash will you have after the trade is executed?

- 0 ECU
- 200 ECU
- 400 ECU
- 600 ECU

Continue

PAST TRADES

You can see information about past trades in the period on the bottom box in the right column of the screen. This box also displays whether you participated in the trade as a Buyer or a Seller or whether the trade was between two other participants ("Others").

Continue

THE END OF THE PERIOD

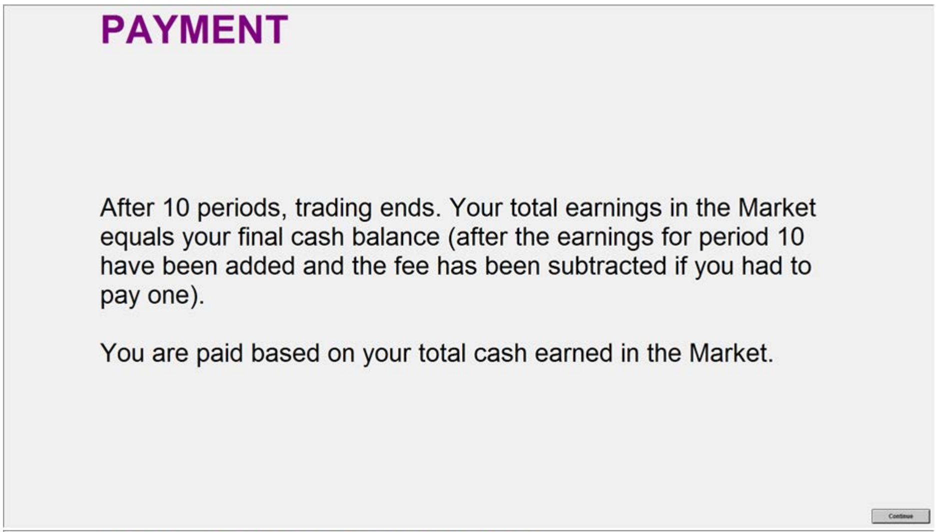
After the 150 second have elapsed, the period ends.

You are told whether earnings in the period were 50 ECU or 150 ECU.

The earnings from the assets you held at the end of the period are credited to your cash balance.

Your final asset holdings and cash (including the earnings from the assets and after deducting the fee if you had to pay it) become your initial cash and asset holdings in the next period.

Continue



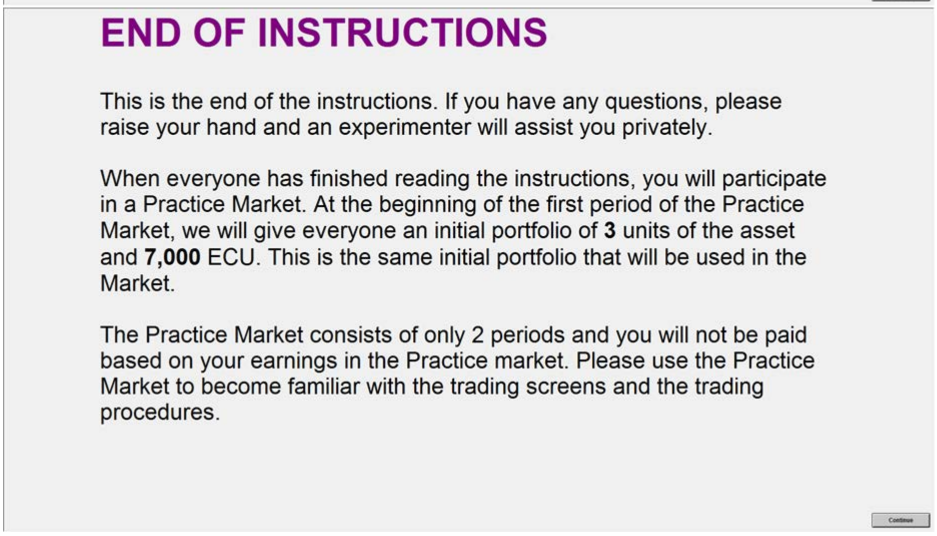
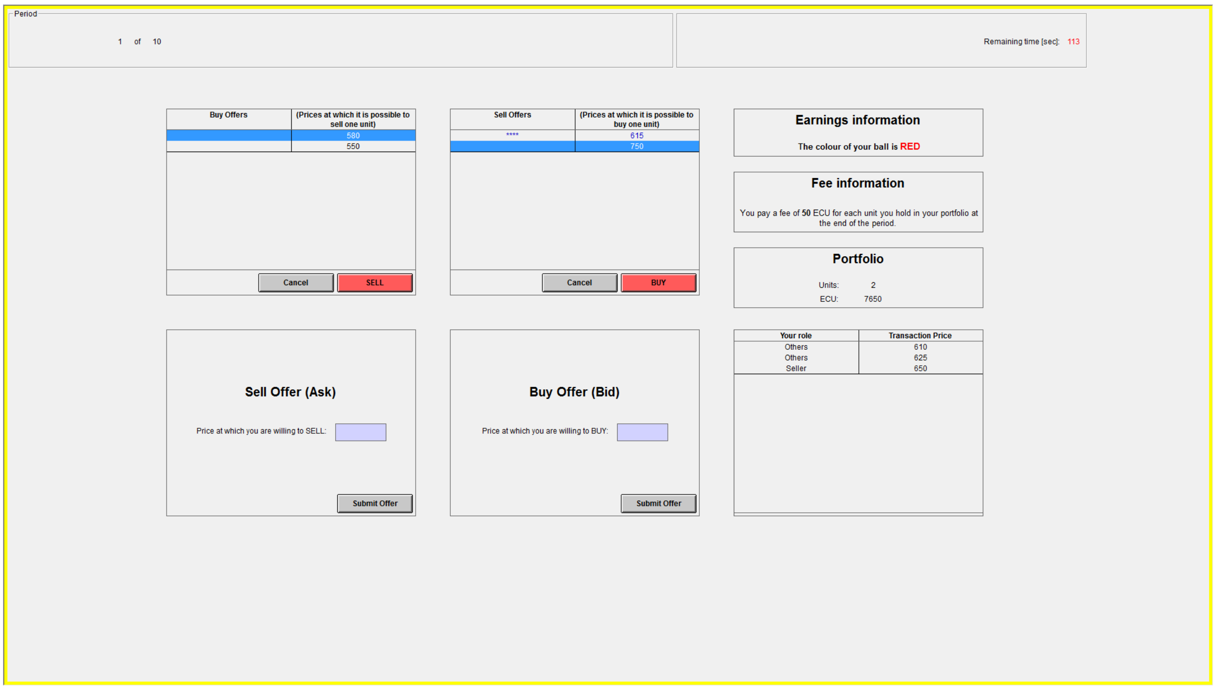


Figure B.2: Trading Platform



B.3 GG, IGG

Figure B.3: Guessing Game Screen

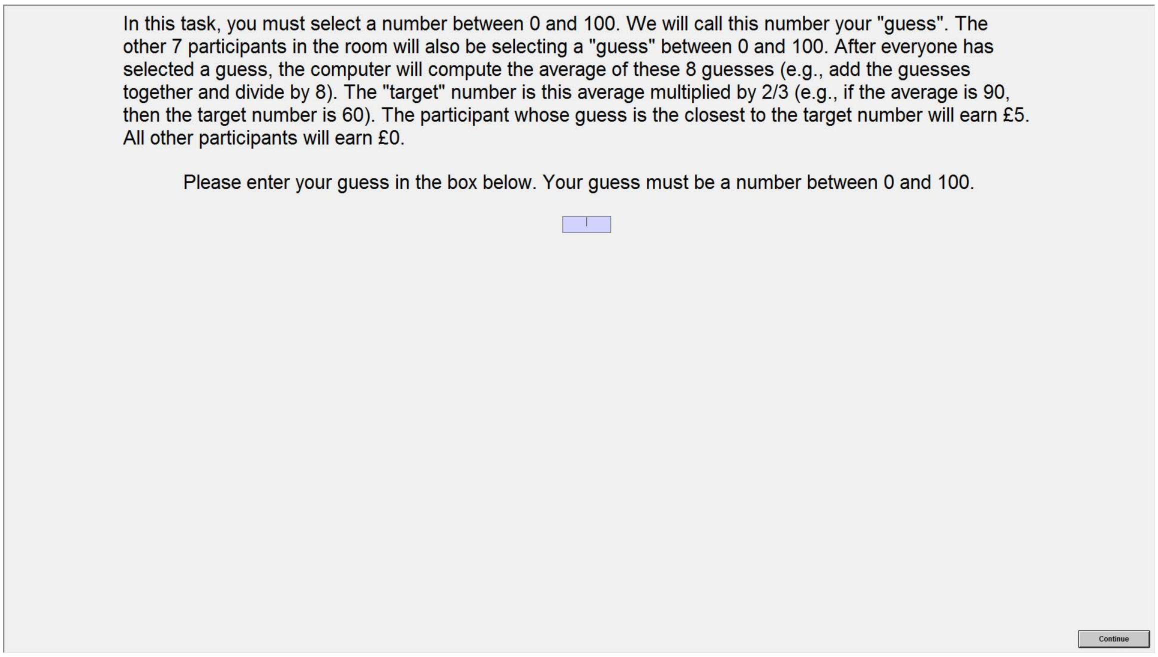
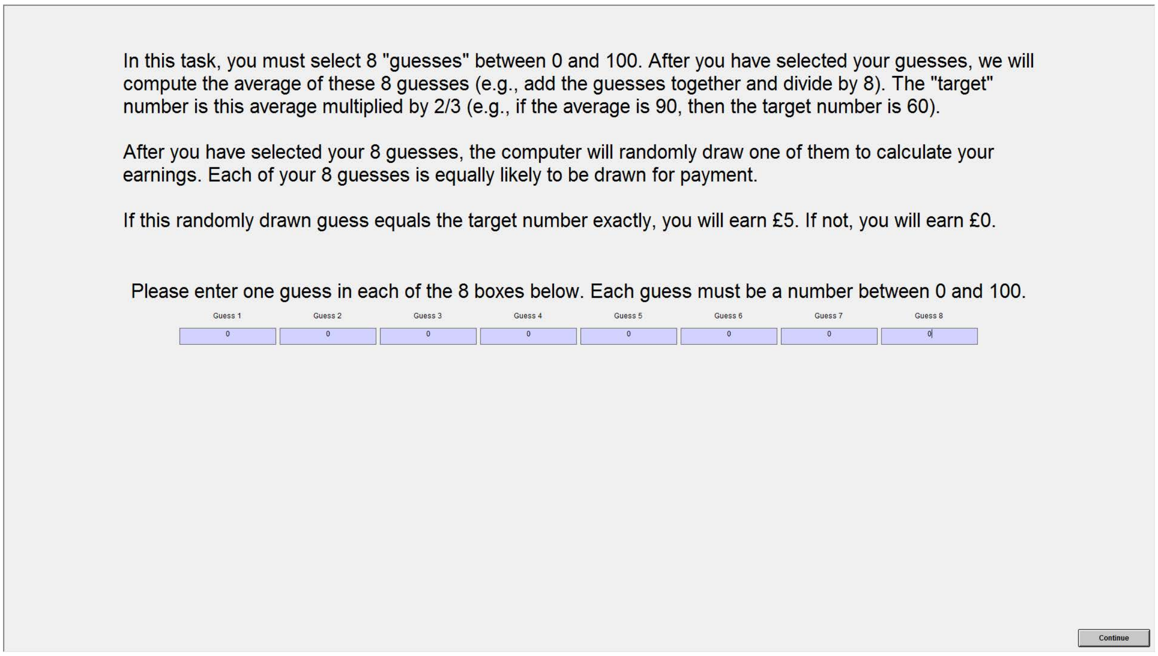


Figure B.4: Individual Guessing Game Screen



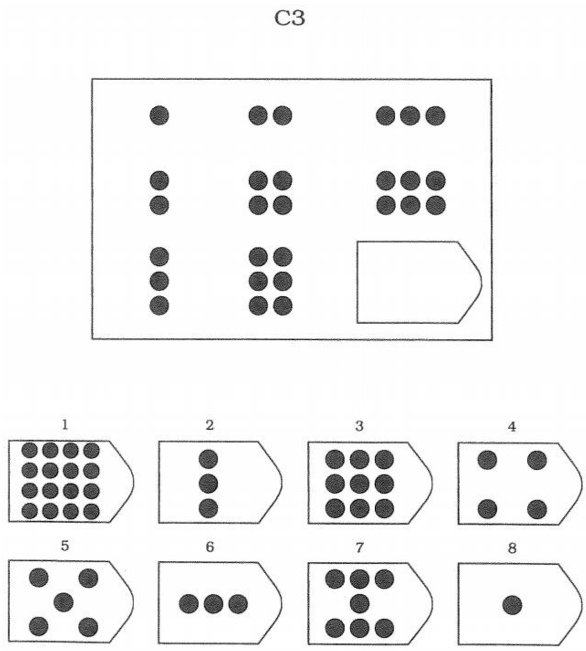
B.4 Eliciting cognitive skills

Figure B.5: Raven's Test Instructions

Station #________________

Please do not turn the page until you are instructed to do so

This is a test of observation and clear thinking It will involve answering problems such as C3 below The top part of problem C3 shows a pattern with a piece cut out of it Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns must be like Then find the right piece out of the 8 pieces shown below Number 3 is the right piece, isn't it? To submit that your answer is Number 3, you will need to circle the Number 3 piece



On every page of this booklet there is a pattern with a piece missing You have to choose which piece is the right one to complete the pattern When you think you have found the right piece, circle it on your booklet and move onto the next page If you make a mistake, or want to change your answer, put a cross through the incorrect answer, and circle your correct answer You will have 10 minutes to complete all 18 pages of this booklet When everyone is ready, the experimenter will let you begin and start the timer The experimenter will announce when 10 minutes is finished and, at that time, you will need to stop working and put down your pen

Figure B.6: CRT Instructions

Station #_________________

Please do not turn the page until you are instructed to do so

On the back side of this page there are seven problems that vary in difficulty You will have 5 minutes to answer as many as you can When everyone is ready, the experimenter will let you begin and start the timer The experimenter will announce when 5 minutes is finished and, at that time, you will need to stop working and put down your pen

(1) In a lake, there is a patch of lily pads Every day, the patch doubles in size If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ____ days

(2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ____ minutes

(3) A bat and a ball cost £1 10 in total The bat costs one pound more than the ball How much does the ball cost? ____ pence

(4) If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? _____ days

(5) Jerry received both the 15th highest and the 15th lowest mark in the class How many students are in the class? ______ students

(6) A man buys a pig for £60, sells it for £70, buys it back for £80, and sells it finally for £90 How much has he made? _____

(7) Simon decided to invest $8,000 in the stock market one day early in 2008 Six months after he invested, on July 17, the stocks he had purchased were down 50% Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75% At this point, which statement is true? (Circle the correct answer)

a Simon has broken even in the stock market

b Simon is ahead of where he began

c Simon has lost money

Figure B.7: Confidence Screen

Station #________________ **R / C**

There are 7 other participants in the room who took the exact same task you just completed. Please answer the following 8 questions.
(The sum of all 8 numbers should equal 100)

What is the percent chance that…

0 of the 7 other participants had more correct answers than you did? _______%

1 of the 7 other participants had more correct answers than you did? _______%

2 of the 7 other participants had more correct answers than you did? _______%

3 of the 7 other participants had more correct answers than you did? _______%

4 of the 7 other participants had more correct answers than you did? _______%

5 of the 7 other participants had more correct answers than you did? _______%

6 of the 7 other participants had more correct answers than you did? _______%

7 of the 7 other participants had more correct answers than you did? _______%

B.5 Eliciting risk preferences (BRET)

Figure B.8: BRET Screens (I)

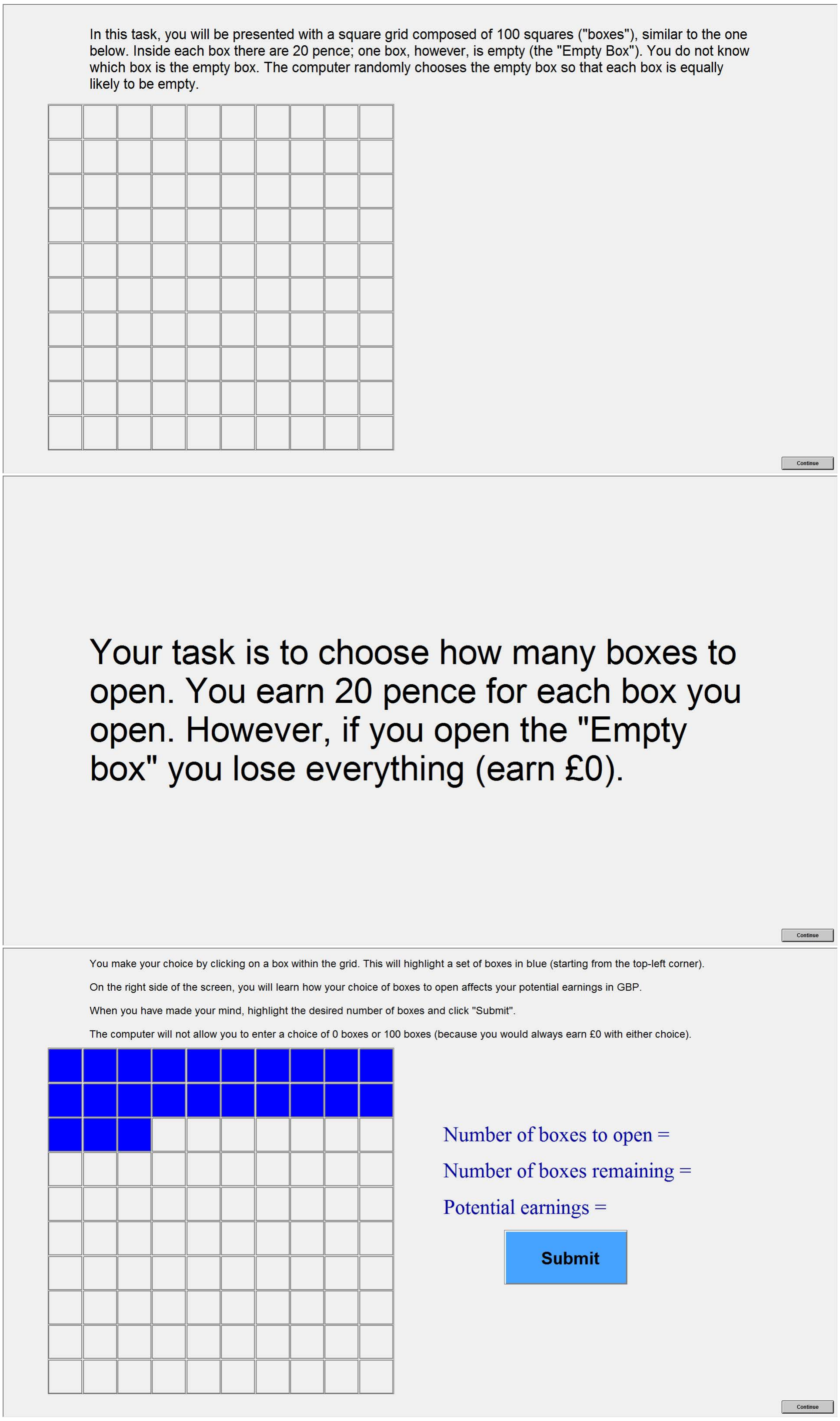
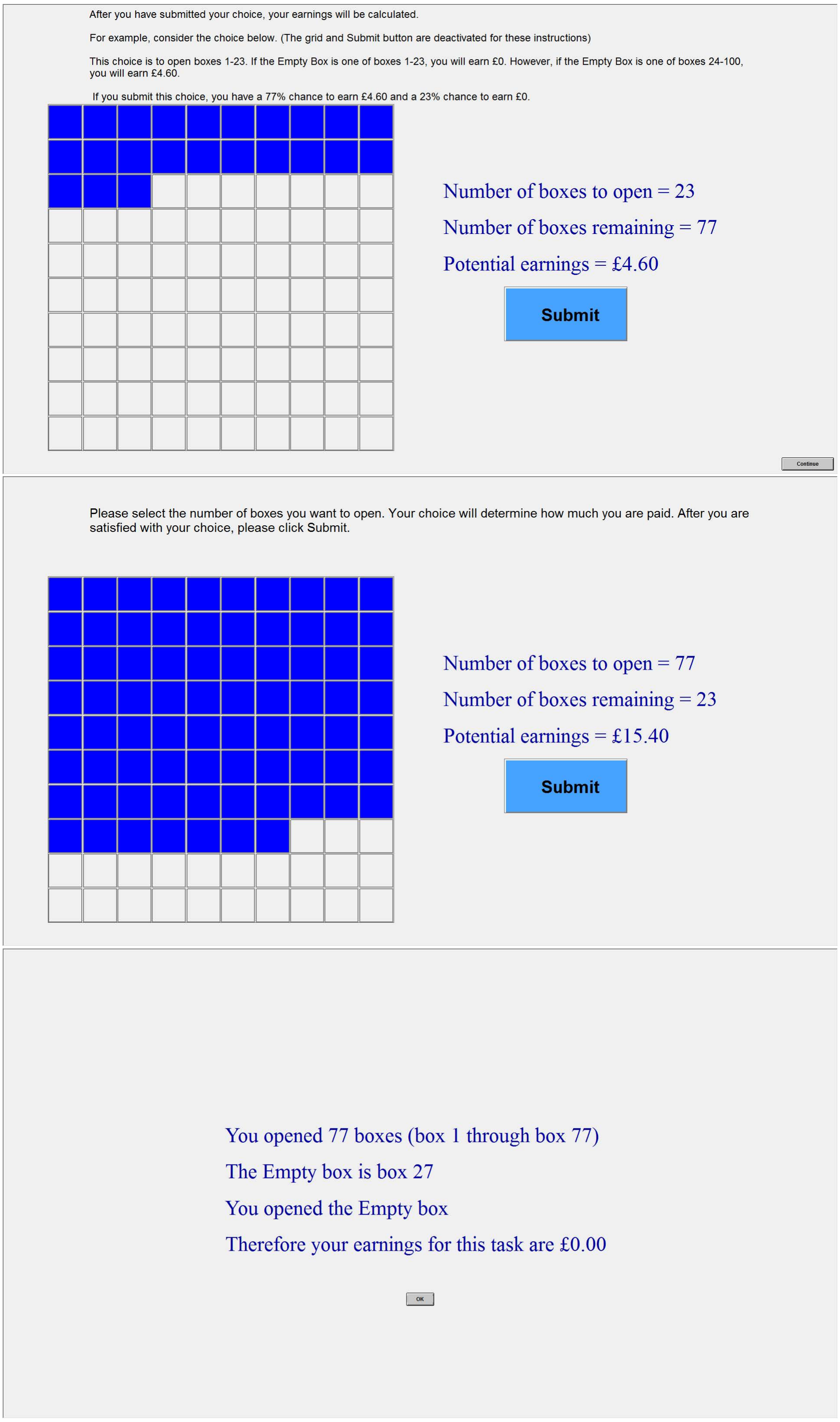


Figure B.9: BRET Screens (II)



B.6 Eliciting non-cognitive skills

Figure B.10: Big5, Grit, and Locus of Control

Station #_________________

How I am in general

Listed below are a number of statements that may or may not apply to you For example, do you agree that you are someone who *likes to spend time with others*? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement

1 Disagree Strongly	**2** Disagree a little	**3** Neither agree nor disagree	**4** Agree a little	**5** Agree strongly

(1) E_____ I am reserved

(2) A_____ I am generally trusting

(3) G_____ Setbacks don't discourage me

(4) C_____ I tend to be lazy

(5) N_____ I am relaxed, I handle stress well

(6) G_____ I finish whatever I begin

(7) O_____ I have few artistic interests

(8) O_____ I have an active imagination

(9) C_____ I do a thorough job

(10) N_____ I get nervous easily

(11) G_____ I am diligent

(12) A_____ I tend to find fault with others

(13) E_____ I am outgoing, sociable

(14) G_____ I am a hard worker

(15) L_____ I have little control over the things that happen to me

(16) G_____ I often set a goal but later choose to pursue a different one

(17) L_____ There is really no way I can solve some of the problems I have

(18) G_____ New ideas and projects sometimes distract me from previous ones

(19) L_____ I often feel helpless in dealing with the problems of life

(20) L_____ Sometimes I feel that I'm being pushed around in life

(21) G_____ I have been obsessed with a certain idea or project for a short time but later lost interest

(22) L_____ What happens to me in the future mostly depends on me

(23) L_____ There is little I can do to change many of the important things in my life

(24) G_____ I have difficulty keeping my focus on projects that take more than a few months to complete

(25) L_____ I can do just about anything I really set my mind to do

A: agreeableness; C: conscientiousness; E: extraversion; G: grit; L: Locus of Control; N: neuroticism; O: openness.

Figure B.11: Self-Monitoring

How I am in general (continued)

As on the previous page, this page lists a number of statements that may or may not apply to you If a statement is true or mostly true as applied to you, make a mark in the "True" column as your answer If a statement is false or not usually true as applied to you, make a mark in the "False" column as your answer Please record your answers in the spaces provided below

	True	False
I find it hard to imitate the behaviour of other people		
At parties and social gatherings, I do not attempt to do or say things that others will like		
I can only argue for ideas which I already believe		
I can make impromptu speeches even on topics about which I have almost no information		
I guess I put on a show to impress or entertain others		
I would probably make a good actor		
In a group of people I am rarely the centre of attention		
In different situations and with different people, I often act like very different persons		
I am not particularly good at making other people like me		
I'm not always the person I appear to be		
I would not change my opinions (or the way I do things) in order to please someone or win their favour		
I have considered being an entertainer		
I have never been good at games like charades or improvisations		
I have trouble changing my behaviour to suit different people and different situations		
At a party I let others keep the jokes and stories going		
I feel a bit awkward in public and do not show up quite as well as I should		
I can look anyone in the eyes and tell a lie with a straight face		
I may deceive people by being friendly when I really dislike them		