Cognitive Reflection and (the Limits of) Strategic Thinking in a Market Choice Game

By ERIK DUHAIME*

Previous experimental studies have documented quick convergence to equilibrium in N-player market entry games, a result that has been replicated under a wide variety of experimental conditions and that looked "like magic" to Kahneman. However, few studies have examined the individual characteristics or strategies that lead to success in the early rounds of market entry games, before equilibrium is approached and the potential for excess returns vanishes. I introduce a "market choice" game, which can be thought of as a one-shot, forced choice, multiple-market version of a market entry game with known market capacities. This game reflects the fact that one is typically faced not with the decision of whether to enter a particular market (or whether to travel to a particular bar, as in the El Farol bar problem), but rather, with the decision of which market (or bar) to enter. Here I analyze the behavior of 285 participants in a series of market choice games and find that those who score higher on the Cognitive Reflection Task (CRT) exhibit higher level-k thinking, earning them higher profits. However, this group of high-CRT individuals still would have been outperformed both by a pure Nash strategy and a strategy of choosing markets completely randomly with no consideration for market size.

* Sloan School of Management, Massachusetts Institute of Technology, 100 Main St, Cambridge, MA 02139 (e-mail: eduhaime@mit.edu)

I. Introduction

Deviations from the efficient markets hypothesis (EMH) have been well documented for decades (e.g., Lo and MacKinlay, 1988) and the past two decades in particular have seen the emergence of behavioral economics and behavioral finance as a major challenge to this perspective. However, one might still characterize economics and finance as largely adhering to the EMH with a number of asterisks to account for all of the anomalies documented in the behavioral economics and finance literatures. Only recently has an alternative framework emerged, the Adaptive Markets Hypothesis (AMH), which attempts to reconcile the tension between these two perspectives by applying the principles of evolution to financial interactions (Lo, 2004; Lo, 2005).

The strength of the AMH is in recognizing that prices are not simply perfect aggregations of all existing information about an economic environment, but they also reflect the number and nature of the actors – or "species" – in a market (Lo, 2005). The market environment, then, has an inherently social and contextual component, and performance in a market therefore depends to some extent on an actor's ability to accurately decipher how large numbers of unknown others will behave. Consider, for instance, how an investor should behave if they believe that market prices have deviated substantially from their fundamental value, such as a fund manager in 1998 who believes that technology stocks are in a bubble. Shorting technology stocks might have driven the manager bankrupt because – as Keynes is purported to have said – "markets can stay irrational longer than one can stay solvent." In this case, understanding the rationality (or lack thereof) of the other market actors would have been far more important than having accurate information about the future cash flows of technology companies.

Understanding the beliefs and behaviors of a large number of unknown others is similarly important outside of financial markets. Consider, for instance,

W. Brian Arthur's (1994) "El-Farol problem." In this problem a finite population of people all want to go to the El-Farol bar every Thursday night. However, because the bar is small and prone to overcrowding, if over 60% of the population goes to the bar on any given night than each person would rather have just stayed home. Because everyone has to decide at the same time whether or not to go to the bar on a particular night, each person must make assumptions about what all of the others will do, and the accuracy of these assumptions will determine how much each person enjoys their Thursday night. We all face similar decisions throughout life, whether it is what color shirt to wear to a party, what route to take during rush hour, or what career to pursue.

Market Entry Games

Market entry games are one way to study how people navigate these social dilemmas because participants are typically given complete information about the environment's economic conditions (i.e. the market size). The basic structure of the game is that a number of participants each are faced with the decision of whether to enter a market of size X and receive a payment of X/n, where n is the number of participants that enter the market, or to refrain from entering the market and receive c, some fixed amount smaller than X. As in the El-Farol bar problem, then, the only uncertain variable is what the other participants will do. Remarkably, even the earliest plays of such games produce aggregate patterns close to those implied by the Nash equilibrium, where $X/n = c$, a fact that "cannot be given a game-theoretic explanation. Nor can [it] be accounted for by any other theory of which I am aware," wrote Ochs (1999). After the first play in such games participants quickly converge to equilibrium, a result that "to a psychologist" looks "like magic" (Kahneman 1988).

While this convergence phenomenon has been widely studied, relatively few have examined the individual characteristics or strategies that lead to success in early rounds of a market entry game, before equilibrium is approached and excess returns vanish. Camerer and Lovallo (1999) create an experimental setting where in some conditions the payment to market entrants is not split equally but instead depends on a participant's relative skill on logic puzzles or trivia. They find that a number of participants are overconfident about their relative abilities, which leads to excess market entry and, consequently, lower returns for those participants who were overconfident about their abilities. This means that at least some people were better able to understand their own skill and choice in relation to the skills and choices of others (i.e. the social aspects of the market environment) and made relatively higher earnings as a result. Presumably these participants are relatively less likely to be deluded about their chances of success in an overly saturated market such as the restaurant industry.

To the extent that there are traits that can be positively associated with returns in a market entry game, success in a market entry game can be thought of as a measure of a form of social intelligence. Related games (e.g., the prisoner's dilemma) have examined similar forms of social intelligence, but few studies have considered whether some individuals are reliably better at predicting the knowledge and behaviors of large numbers of unknown others. This is the "skill" that performance in a market entry game tests: not the ability to decipher the knowledge and decisions of another person per se, but the ability to determine the systematic biases of hundreds, thousands, or even millions of unknown others.

The Market Choice Game

To examine whether some individuals are in fact better able to recognize and act on the biases of large groups of unknown others I introduce here the "market choice" game, which can be thought of as a one-shot, forced choice, multiple-market version of a market entry game with known market capacities. In contrast to prior experiments on market entry games, which typically have around a dozen participants, I utilize the availability of large samples online to examine a more complex market environment. Whereas a participant's decision in a market entry game can be thought of as the decision of whether or not to enter a specific industry (e.g. the restaurant industry), the decision in this game can be thought of as one of which market – each with their own market capacity and potential entrants – to enter. Some industries may be more salient to people and lure a disproportionate number of entrants (e.g. an airplane pilot, a rock star, etc.) while others may be less salient than their market capacity (e.g. an engineer, an accountant, etc.), and surely there are some accountants who were more enticed by the prospect of being a rock star but realized that there would be too many entrants in the music industry. Were these accountants just lucky for realizing this, or did they possess some higher degree of social intelligence than their starving-artist peers? Or, put another way, the question is: are some people better at knowing where the herd will run?

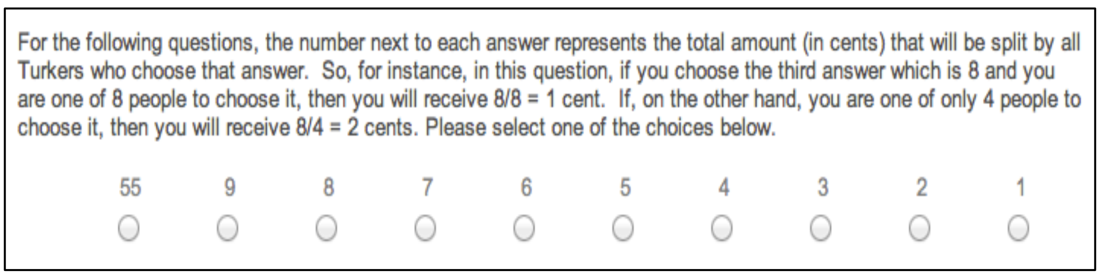
II. Methods

Participants

Two hundred and eighty-five U.S. residents were recruited for a "10 minute decision-making survey" on Amazon Mechanical Turk and participated in return for a small payment. Participants were 42% male and 58% female, with an average age of ~36 years. The subject pool was relatively well educated: 14% reported that they had an advanced degree, 42% reported having a college degree, and 34% reported that they had at least some college education.

Materials and Procedure

After agreeing to a consent form, participants were asked a series of eight hypothetical questions designed to gauge risk aversion (e.g., "Imagine that you have just won $5,000 in a game show. Would you risk it all for a 50% chance of walking away with $50,000?”). Participants were next administered the Cognitive Reflection Test (CRT), a well-validated, three-item test designed to assess an individual’s ability to suppress an intuitive “system 1” wrong answer in favor of a reflective “system 2” right answer (Frederick, 2005). In previous work, performance on the CRT has proven to be negatively correlated with risk aversion and positively related to patience, as well as measures related to intelligence such as IQ and SAT scores. Participants were then informed that in the next section of the survey they would be asked “REAL financial questions” and that their answers to these questions would influence how much they would receive for a bonus payment. Participants were then shown the prompt shown in Figure 1, below:


For the following questions, the number next to each answer represents the total amount (in cents) that will be split by all Turkers who choose that answer. So, for instance, in this question, if you choose the third answer which is 8 and you are one of 8 people to choose it, then you will receive 8/8 = 1 cent. If, on the other hand, you are one of only 4 people to choose it, then you will receive 8/4 = 2 cents. Please select one of the choices below.

55	9	8	7	6	5	4	3	2	1
○	○	○	○	○	○	○	○	○	○

Figure 1: A sample prompt for one round in a market choice game

Participants were shown a series of seven questions in the manner shown in Figure 1 with the following “market choices” available as shown in Table 1. For each question, the total value of all the market choices equals exactly 100. After completing these questions participants were asked to report basic demographic information.

Market Choice Game Round #	Choices Available
#1	55, 9, 8, 7, 6, 5, 4, 3, 2, 1
#2	30, 25, 20, 15, 4, 3, 2, 1
#3	60, 20, 5, 8, 7
#4	30, 25, 20, 15, 10
#5	50, 11, 10, 9, 7, 5, 4, 3, 1
#6	25, 16, 14, 11, 10, 8, 7, 5, 3, 1
#7	24, 21, 14, 11, 8, 7, 6, 5, 3, 1

Table 1: The set of "market choices" available for participants to choose

III. Results

Degree of Tacit Market Coordination

In the Appendix is a table for each round that includes the 1) market choices available, 2) number of participants choosing each market choice, 3) percent of participants choosing each market choice, and 4) the payout for each participant choosing that market. Given the one-shot nature of the market choice task and the fact that each round had at least five market choices and as many as ten, the degree of tacit market coordination achieved is striking. Consider, for instance, that in all there were fifty-seven market choices spread out over seven rounds, and in only one instance was the difference between the percentage of participants choosing a single market and that market's "share of the total economy" greater than twenty.

Average Earnings

The average bonus amount earned by participants was 2.456 cents, which follows from the fact that there were seven rounds with aggregate market size of one hundred cents each and two hundred and eighty five total participants (7 * 100 / 285 = 2.456).

Distribution of Earnings

As can be seen in Figure 2 below, earnings were roughly normally distributed around the mean with a standard deviation of .538. Importantly, despite the high level of tacit coordination achieved there was still plenty of variation in earnings, with some participants making almost four times as much as others.

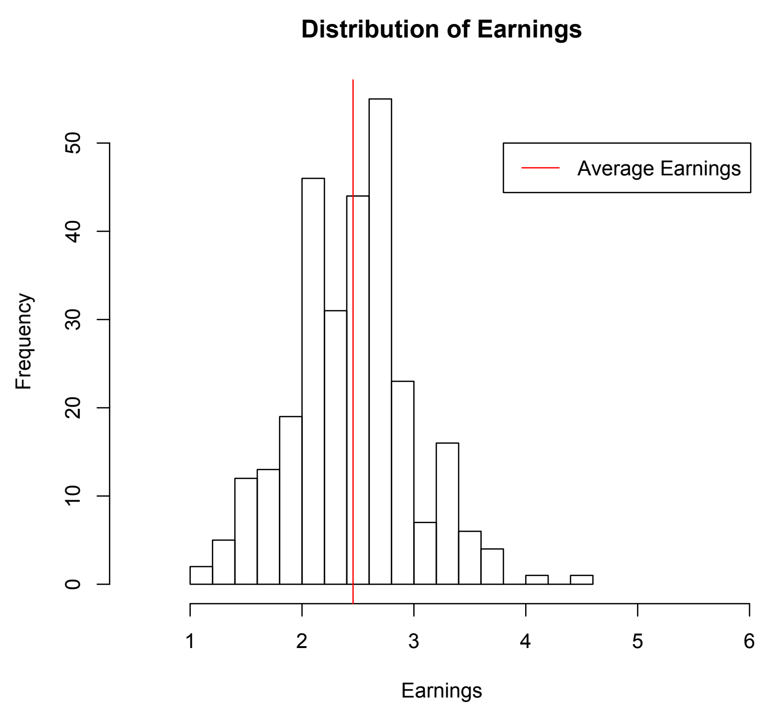


Figure 2: The distribution of total earnings in the market choice games

Strategic Determinants of Earnings

Figure 3 below shows the distribution of earnings with vertical lines indicating what a participant would have earned if they implemented a strategy of choosing either the largest, second largest, third largest, smallest, second smallest, or third smallest market in each round. Notably, while 37 participants did in fact implement the relatively successful "pick the largest market each time" strategy, only 4 participants implemented the relatively unsuccessful "second largest market each time" strategy and no other participants implemented any of the other easily decipherable strategies considered here.

Nonetheless, considering the would-be outcomes of these strategies shines light on relevant decision-making processes and offers insights into the nature of herding behavior. First notice that selecting the largest market each time – the most "obvious" choice – would have earned a participant significantly more than the mean in the sample (t (284) = -9.55, $p < .001$). If a participant realizes that this "obvious" choice may attract too many participants and tries to think one step ahead of the other participants, the next two salient options might be to choose either the second largest market or the smallest market. The fact that a strategy of choosing the largest market each time led to significantly higher earnings while a strategy of choosing the second largest market each time (t (284) = 9.94, $p < .001$) or the smallest market each time (t (284) = 41.98, $p < .001$) would have led to significantly lower earnings suggests that too many people similarly thought along these lines. Thinking yet another step ahead may have led someone to choose either the third largest or the second smallest market, each of which would have earned participants even more than a strategy of choosing the largest market each time. Thinking yet another step ahead may have motivated someone to choose the third smallest market each time, which would have earned a participant the third highest total of all two hundred and eighty-five participants.

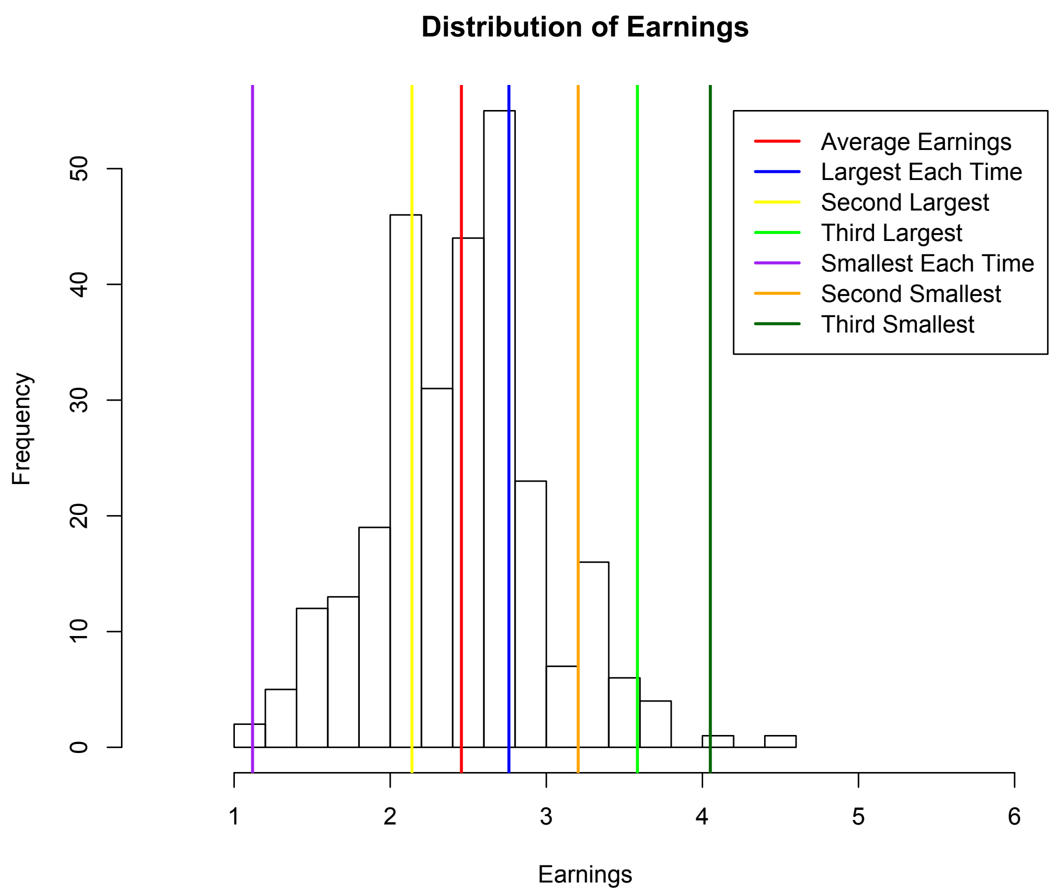


Figure 3: The distribution of total earnings in the market choice games with the would-be returns of simplistic, order-based strategies shown

Individual Determinants of Earnings

1) **Demographics**. None of the demographic factors examined – which included age, gender, income, education and political views – were found to have a significant relationship with earnings in the market choice game.
2) **Risk preferences**. Risk aversion – as measured by our eight-item scale – was found to be positively associated with earnings in the market choice

game ($\beta = .056$, $SE = .023$, $t\ (283) = 2.42$, $p < .05$). Further analysis reveals that this effect is driven primarily by the fact that people high in risk aversion were significantly less likely to choose the smallest possible market ($\beta = -.271$, $SE = .096$, $t\ (283) = -2.80$, $p < .01$), which was consistently one of the lowest returning market choices.

3) **Cognitive reflection ability**. As shown in Figure 4 below, scores on the Cognitive Reflection Test (CRT) were significantly positively associated with earnings in the market choice game ($\beta = .074$, $SE = .030$, $t\ (283) = 2.43$, $p < .05$). Notably, risk aversion and scores on the CRT were negatively correlated, which is consistent with prior research (Frederick, 2005) and suggests that the qualities associated with these tests conferred different advantages in the market choice game. Indeed, regressing earnings on both of these variables only increased the size of the estimates and drove the significance associated with each variable to the $p < .01$ level. Further analysis reveals that participants scoring high on the CRT were somewhat less likely to choose the largest ($\beta = -.03$, $SE = .025$, $t\ (283) = -1.07$, $p = .29$) or smallest ($\beta = -.07$, $SE = .075$, $t(283) = -.92$, $p = .36$) markets and somewhat more likely to choose the second largest ($\beta = .06$, $SE = .039$, $t(283) = 1.54$, $p = .12$), third largest ($\beta = .11$, $SE = .064$, $t(283) = 1.76$, $p = .08$), second smallest ($\beta = .16$, $SE = .111$, $t(283) = 1.45$, $p = .15$), or third smallest market ($\beta = .13$, $SE = .099$, $t(283) = 1.28$, $p = .20$), which suggests that the key to high CRT individuals' success was that they engaged in higher order "level-k" thinking (Crawford and Iriberri, 2007).

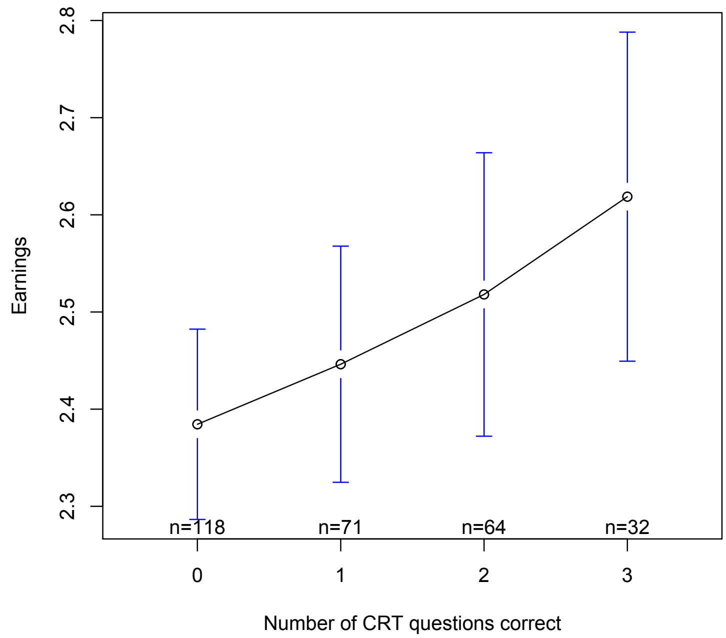


Figure 4: The relationship between total earnings in the market choice game and number of answers correct on the CRT.

Comparison to Nash and Random Strategies

The Nash strategy in the market choice game would be to select each market in the same proportion of its market capacity. So, in Round 1 depicted in Table 1 above, the Nash strategy would be to select the 55 option 55% of the time, the 9 option 9% of the time, and so on and so forth. In reality people do not have a random number generator in their head, and so even if a person tried to implement a Nash strategy their responses would likely reveal systematic biases over time (e.g., the under or overweighting of small probabilities), but here we are able to simulate what a Nash strategy would have returned on average in our environment of two hundred and eighty-five participants. I find that the Nash strategy would have outperformed the average earnings of participants and that this strategy even would have earned significantly higher than the average earnings of those who got all three CRT questions correct (t (31) = 5.20, p < .001). So long as an environment is not in equilibrium it is possible for *some* participants to do better than the Nash strategy on average, and we do indeed see that a number of participants outperformed the Nash strategy here.

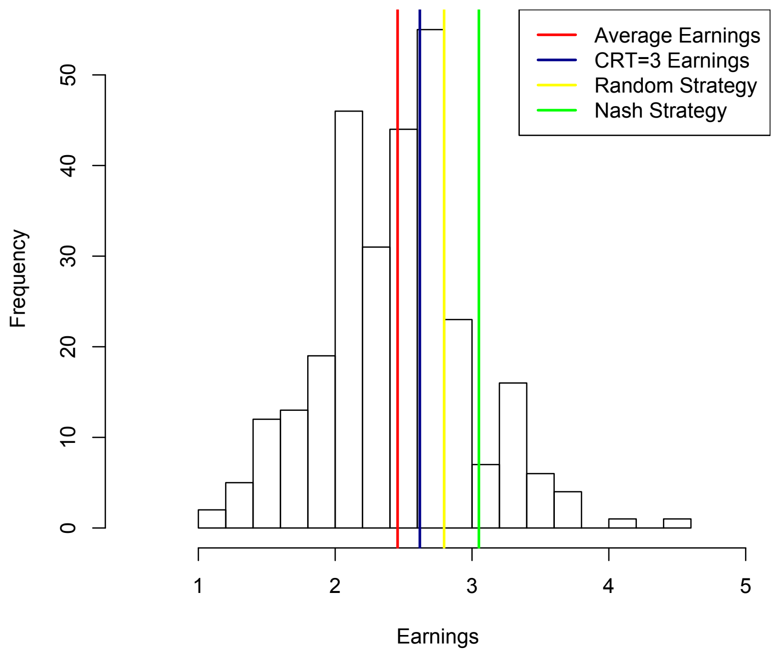


Figure 5: A comparison of average earnings, earnings of high-CRT individuals, the would-be earnings of a random strategy, and the average earnings of a Nash strategy.

What is more surprising is that participants earned, on average, significantly less than if they had implemented a strategy of picking markets *completely randomly.* So, in Round 1 for instance, instead of picking the 55 option 55% percent of the time and the 9 option 9% of the time and so on, a completely random strategy is one in which a participant would pick each option in Round 1 on average 10% of the time (because there were 10 total markets in Round 1). In fact, a strategy of guessing completely randomly would have earned on average even significantly higher than the average earnings of those who got all three CRT questions correct (t (31) = 2.13, $p < .05$). This is particularly striking because compared to the Nash strategy a completely random strategy would have under-weighted the largest markets and over-weighted the smallest markets, which we saw above were particularly high and low returning choices, respectively.

IV. Discussion and Future Directions

The results of this experiment show that some people – in this case, people with relatively high cognitive reflection abilities – are in fact better able to decipher the strategic behavior of a large number of unknown others. While one potential criticism of this study could be that participants did not think hard about their decision because of the relatively low stakes, participants who did not think *at all* about their decision and chose completely randomly still would have done better than others on average, and also better on average than high-CRT individuals. This suggests that the low stakes if anything may have dampened the relative advantage of high-CRT individuals and reveals an important insight: while high-CRT individuals were better able to sense what markets would have excess entry, they too were prone to herding. Consider, for instance, that high-CRT individuals made above average profits by being somewhat more likely to choose the 3rd highest and 3rd smallest markets, but they also erred by being less likely to choose the largest market in favor of the second largest market. Thus, while high-CRT individuals may have been more likely on average to think one step ahead of the majority of participants, many others thought similarly, and –

more often than not – their returns suffered from picking alternative markets that still had excess entry.

Future work will attempt to correlate performance in a number of different market choice game environments with other reliable metrics such as IQ and Myers-Briggs personality types, as well as with real-world behaviors such as trading behavior in real financial markets. For instance, participants who perform well in market choice games may also be more likely to identify social fads or movements, may be better poker players, or may even be better able to identify when "irrational exuberance" has taken hold of equities markets. It may also be the case that poor performance in market choice games is actually a stronger predictive signal in real-world environments. That is, people who perform consistently poorly (perhaps, more specifically, people who are more likely to think that they will be the only unique person to choose the smallest market, or make some other consistent mistake) may be more prone to buy into an irrational social fad or fall victim to other mental biases. If true, an astute observer may improve her own predictions about the social environment simply by discounting the predictions of these poor performers.

Future work will also explore group-level characteristics in market choice games. For instance, it remains to be seen whether the typical MIT student would outperform the typical Turker if the social environment is composed primarily of Turkers or vice versa. This will depend in part on the extent to which skill in market choice games is based on understanding fundamental human biases compared to the extent to which skill in the game is based on understanding the biases of the other actors in the current environment. It may be, for instance, that a Turker would perform relatively *better* on average in an environment of all MIT students if MIT students tend to think similarly and therefore are more prone to herding, much like we saw amongst high-CRT individuals in this study. If in fact certain "types" of people perform relatively worse when in environments saturated with the same "type" and relatively better when in dynamic, diverse environments, market choice games may also be useful for measuring group-level intelligence. That is, the level of tacit coordination achieved in market choice games may be correlated with the diversity of "types" or perspectives held by group members, and therefore groups that achieve higher levels of tacit coordination may be less likely to fall victim to groupthink or may be less likely to neglect alternative perspectives when trying to decipher the nature of real-world social environments.

Thus, much remains to be explored in future work, but the present study demonstrates clearly that there are in fact certain characteristics that are associated with an individual's ability to decipher the behaviors of large numbers of unknown others. Considering the importance of this ability in many real world settings, market choice games offer a promising experimental paradigm to further examine such large-scale social dynamics.

V. References

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review* 84.2: 406-411

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 306-318.

Crawford, Vincent P., and Nagore Iriberri (2007). Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions. *Econometrica* 75.6: 1721-1770.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.

Kahneman, D. (1988). Experimental Economics: A Psychological Perspective. In *Bounded Rational Behavior in Experimental Games and Markets* (R. Tietz, W. Albers, and R. Selten, Eds.). Berlin: Springer-Verlag.

Lo, A. W. (2004). The adaptive markets hypothesis. *The Journal of Portfolio Management*, *30*(5), 15-29.

Lo, A. W. (2005). Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis. *Journal of Investment Consulting*, *7*(2), 21-44.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies*, *1*(1), 41-66.

Ochs, J. (1999). Coordination in market entry games. *Games and human behavior: Essays in honor of Amnon Rapoport*, 143-172.

VI. Appendix

Round #1

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
55	101	35.44	0.54	2
9	38	13.33	0.24	7
8	23	8.07	0.35	4
7	28	9.82	0.25	6
6	27	9.47	0.22	8
5	38	13.33	0.13	9
4	15	5.26	0.27	5
3	3	1.05	1.00	1
2	4	1.40	0.50	3
1	8	2.81	0.13	10

Round #2

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
30	65	22.81	0.46	3
25	49	17.19	0.51	2
20	60	21.05	0.33	5
15	58	20.35	0.26	6
4	37	12.98	0.11	8
3	8	2.81	0.38	4
2	3	1.05	0.67	1
1	5	1.75	0.20	7

Round #3

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
60	110	38.60	0.55	1
20	70	24.56	0.29	3
8	45	15.79	0.18	4
7	15	5.26	0.47	2
5	45	15.79	0.11	5

Round #4

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
30	122	42.81	0.25	5
25	66	23.16	0.38	3
20	29	10.18	0.69	1
15	29	10.18	0.52	2
10	39	13.68	0.26	4

Round #5

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
50	99	34.74	0.51	3
11	60	21.05	0.18	8
10	41	14.39	0.24	7
9	30	10.53	0.30	5
7	9	3.16	0.78	2
5	12	4.21	0.42	4
4	5	1.75	0.80	1
3	12	4.21	0.25	6
1	17	5.96	0.06	9

Round #6

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
25	107	37.54	0.23	8
16	74	25.96	0.22	9
14	11	3.86	1.27	1
11	24	8.42	0.46	5
10	15	5.26	0.67	3
8	10	3.51	0.80	2
7	13	4.56	0.54	4
5	17	5.96	0.29	7
3	8	2.81	0.38	6
1	6	2.11	0.17	10

Round #7

Market Choice	Number Choosing	Percent Choosing	Payout Per Person	Payout Rank
24	107	37.54	0.22	9
21	64	22.46	0.33	7
14	27	9.47	0.52	4
11	24	8.42	0.46	5
8	12	4.21	0.67	2
7	13	4.56	0.54	3
6	19	6.67	0.32	8
5	7	2.46	0.71	1
3	7	2.46	0.43	6
1	5	1.75	0.20	10