12-2009

A Pure Test of Backward Induction

Kelly Padden Hall
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Kelly Padden Hall entitled "A Pure Test of Backward Induction." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Economics.

William S. Neilson, Major Professor

We have read this dissertation and recommend its acceptance:

Robert Bohm, Bruce Tonn, Michael Price

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Kelly Padden Hall entitled "A Pure Test of Backward Induction." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy with a major in Economics.

William S. Neilson, Major Professor

We have read this dissertation
and recommend its acceptance:

Robert Bohm

Bruce Tonn

Michael Price

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

A PURE TEST OF BACKWARD INDUCTION

A Dissertation Presented for
the Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Kelly Padden Hall
December 2009

DEDICATION

This dissertation is dedicated to my best friend, biggest cheerleader, and love of my life, my dear husband Brent Hall. Without your support and encouragement I would not have made it successfully through this program. I definitely consider this accomplishment a team effort, and this degree is just as much yours as it is mine. Thank you for marrying me!

Also to my loving parents, Brian and Jennifer Padden, who taught me from an early age the value of hard work and self-discipline, and who remind me often that nothing worth having comes easily. I am proud to be your daughter.

ACKNOWLEDGEMENTS

I wish to thank Dr. Bill Neilson for his mentorship, unfailing patience, and friendship over that last three and a half years. I would also like to thank Drs. Mike Price and Mike Shor for teaching me how to be a better researcher and for sharing their infectious enthusiasm for experimental methods. Thanks also to Dr. Bob Bohm for serving on my committee and taking a chance on a government employee whose background in economics was woefully inadequate. Finally, thanks to Dr. Bruce Tonn for participating on my committee. Only because of them can I call myself an economist. Someday I hope to be half as smart as they are.

I would be remiss if I failed to acknowledge Dr. Richard Lester and the staff at the Defense Financial Management & Comptroller School, Maxwell AFB, Alabama, for allowing me the incredible opportunity to take three years away from my primary career field to think and study at this great institution. I promise you got your money's worth.

Finally, I need to extend a very special thank you to Brigadier General (Ret.) Sandra Gregory. I have a Ph.D. because of you, and for that I am forever in your debt. My greatest wish is that someday I'll be in a position to do for another young officer what you have done for me. I'm so thankful and I'll never forget it.

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

ABSTRACT

This dissertation proposes a simple computerized game to serve as a pure test of backward induction and then tests the game in the laboratory. One of the fundamental assumptions of neoclassical economic theory is that human beings function as fully rational agents who maximize their utility over multidimensional alternatives under economic constraints. However, numerous studies have shown systematic deviation from rational decision making in a laboratory setting. While no single explanation is obvious for this suboptimal behavior, the literature suggests other motivations (besides maximizing utility) may be at play, including reciprocity, trust, reputation, and welfare. The "Race to 21" game we test renders these other-regarding preferences irrelevant; therefore we call it a "pure" test of backward induction.

Chapter one introduces the game, as well as tests the effect of adding incentive payments in several places along the path of play. Chapter two continues by analyzing how each different intermediate incentive affects the speed of learning in the game. Chapter three concludes with a look at whether individual differences among laboratory subjects explain some of our experimental results. Common to all chapters is the result that incentive payments offered on the subgame perfect equilibrium path near the midpoint of the game particularly enhance the use of backward induction among subjects.

TABLE OF CONTENTS

List of Figures

LIST OF TABLES

CHAPTER ONE: CAN PLACEMENT OF INCENTIVES WITHIN THE STRUCTURE OF A RACE GAME ENHANCE BACKWARD INDUCTION?

1. Introduction

In economics, as in life, we tend to assume agents make decisions only after careful consideration of the choices available, the chances afforded by nature, and the outcomes that are possible as a result. In fact, the fundamental assumption of neoclassical economic analysis is that human beings function as fully rational agents who maximize their utility over multidimensional alternatives under economic constraints. Most real-world decisions require multiple stages of actions, events, and consequences that are inherently complex. As a result, no single explanation is obvious for the deviations from optimality we observe in the experimental laboratory. Plausible explanations include cognitive limitation, incomplete specification, information availability (or lack of), and attitudes toward risk, among others.

One such deviation from rationality we see is the failure of individuals to use backward induction in decision making tasks. Backward induction involves solving first for optimal behavior at the "end" of a game, and then determining what optimal behavior is earlier in the game given the anticipation of this later behavior. Many studies attribute backward induction failures to bounded rationality[1]. Nevertheless, others have shown that some notion of training or experience actually increases the propensity for subjects to backward induct. To that end, Kagel and Levin (1999) showed that "super-experienced" subjects behaved differently than inexperienced. Likewise, in a field experiment whereby baseball cards were auctioned at a trade show, List and Lucking-Reiley (2002) found that dealers, who commonly participate in auctions, bid more strategically than nondealers. Further, Levitt et al. (2008) show that chess grandmasters are able to use backward induction in simple decision tasks[2]. Finally, Johnson et al. (2002) concluded that untrained subjects deviate from equilibrium as a result of limited look ahead (vs. backward

[1] For example, see McKelvey and Palfrey (1992), Busemeyer et al. (2000), Costa-Gomes, Crawford, and Broseta (2001), Johnson, Camerer, Sen, and Rymon (2002), and Johnson and Busemeyer (2001).
[2] Another notable exception is Palacios-Huerta and Volij (2008).

induction), and that training them in backward induction draws them far closer to equilibrium (though, curiously, not all the way).

To attempt an explanation ourselves of the seemingly irrational behavior commonly observed in the laboratory, we implement a zero-sum game called Race to 21, whereby two players take turns choosing numbers that are added in sequence until one reaches the sum 21. This game is not new or unique to our set of experiments; in fact, we know of several experimentalists who have used variations on this game to test various aspects of learning, strategic sophistication, bounded rationality, and backward induction[3]. Given that most previous empirical work shows that people tend not to behave rationally (at least in the lab), the challenge is to structure a game in such a way that we may elicit rational backward induction. Our game is both different from and preferable to those games others have used to test for backward induction in the laboratory setting in that we employ "teaser" payments both on and off the equilibrium path of play to see whether and how incentives impact an individual's ability to backward induct. Additionally, we test two strategy space choice sets that are different from those previously tested[4], and our subjects play against an emotionless computer opponent so that deviations from the optimal path are automatically punished[5]. We find that subjects more often solve the game when offered incentive payments for staying on the equilibrium path (more specifically, when offered an incentive near the midpoint of the game) than when they play the baseline game with no teaser payments.

The remainder of the study is organized as follows. Section 2 outlines previous related literature. Section 3 describes our experimental design, including a detailed description of the game, subject pool, and laboratory procedure. In section 4, we make predictions based on theory and discuss their implications. We present our empirical results in section 5, and section 6 concludes.

[3] See, for example, Levitt, List, and Sadoff (2008); Gneezy, Rustichini, and Vostroknutov (2007), Costa-Gomes and Crawford (2006), and Dufwenberg, Sundaram, and Butler (2008).

[4] Levitt, List, and Sadoff (2008) test strategy spaces (1-9) and (1-10), Dufwenberg, Sundaram, and Butler (2008) test strategy space (1-2). We test strategy spaces (1-3) and (1-4), as do Gneezy, Rustichini, and Vostroknutov (2007).

[5] Similar to Johnson, Camerer, Sen, and Rymon's (2002) "robot" players.

2. Related Literature

Decision theory, a complex body of knowledge that has been studied by economists, mathematicians, and psychologists for over 40 years, has proven useful to economic theorizing in several ways. Certainly, if we can describe what variables affect decisions, we can attempt to prescribe how decisions should be made. The major dichotomy that exists in decision theory is that between normative and positive disciplines. The vast majority of the prior work in decision theory falls under the normative heading, i.e., concerned with how people *should* make decisions in theory (for a comprehensive review of the classic literature, see S. O. Hansson, 2005, and Bell et al., 1988). These theoretically-oriented studies are useful to the extent that they suggest tools, methodologies, and software interfaces to help people make better decisions.

Since it is obvious that human beings do not always behave optimally, the positive, or descriptive, discipline consists of tests of actual behavior against the predictions of the aforementioned theoretical models. Most of this work, though smaller in volume than its normative counterpart, has exploded since the mid-20th century as experiments became infinitely easier to administer with the proliferation of computer technology. There is a growing body of experimental literature that studies the principles that govern both strategic behavior and individual decision making, surveyed in both Kagel and Roth (1995) and Crawford et al. (1997). Historically, most applications of individual decision and strategic game theories assumed equilibrium strategies selected by backward induction in their predictions. Backward induction, one of the most important solution concepts in game theory, relies on a set of commonly accepted economic assumptions; namely, that individuals are rational and have common knowledge that all other individuals are rational as well (Aumann, 1995). Despite its widespread theoretical application, empirical evidence suggests that economic agents may engage in backward induction less frequently than the theory would predict. In fact, nearly all laboratory experiment results indicate people are not able to backward induct successfully; see, for example, McKelvey and Palfrey (1992), Busemeyer et al. (2000), Costa-Gomes et al. (2001), Johnson et al. (2002), and Johnson and Busemeyer (2001). These deviations from equilibrium are often explained by social preferences (G. E. Bolton and A. Ockenfels, 2000; E. Fehr and K. M. Schmidt, 1999), risk attitudes (R. P. Cubitt et al., 1998; D. Kahneman and A. Tversky, 1979), or failures of rationality. By far the most often cited and generally accepted explanation for bounded rationality is the practical reality that humans have finite computational resources available for decision making; in fact, Rapoport (1975) shows that human beings may be capable of planning only two or three stages ahead. Seminal bodies of work by Tversky and Kahneman (1974) and Fischhoff, Slovic, and Lichtenstein (1978) introduced useful rules of thumb, or heuristics, to overcome the strict rigidity of optimization required by rational agent models. More recently, Costa-Gomes and Crawford (2004) and others have suggested that allowing structured boundedly rational decision rules in some of these applications can resolve the apparent contradiction between theory and observation.

If perfect backward induction yields the optimal solution, and the literature suggests people don't actually use backward induction unless trained, might we be able to improve outcomes by using incentives to help guide individuals to the subgame perfect equilibrium path identified by backward induction? This study is our contribution to the larger literature on the propensity of people to actually use backward induction.

3. Experimental Design

The "Race to 21" game

While much of the experimental literature tests whether subjects use backward induction in strategic settings such as bilateral bargaining, centipede, and prisoner's dilemma games, these games actually may be testing a host of motivations other than backward induction (e.g. reciprocity, trust, welfare, etc.). The "Race to 21" game we studied (a variation on Levitt, List, and Sadoff's (2008) Race to 100 game), by contrast, serves as a pure test of backward induction. In the baseline Race to 21 game, a human subject plays opposite a computer opponent. The human and computer alternately choose numbers within a given range (strategy space 1-3 or 1-4, inclusive) which are then added in sequence. In the 1-3 version, perfect backward induction yields a subgame perfect equilibrium strategy of choosing "1", then on subsequent turns selecting whatever number sums to 5, 9, 13, 17, and 21. In the 1-4 game, the subgame perfect equilibrium path is 1, 6, 11, 16, and 21. The computer is programmed to make random selections within the given action space, unless the human subject chooses a number that takes him off the equilibrium path of play. In that case, the computer will move to the equilibrium path for the remainder of the game, rendering a victory for the human subject impossible. Our setup guarantees a first-mover advantage; therefore the human player always makes his selection first. Play continues until either the human or the computer chooses a number that makes the sequence sum to 21. This player is the winner and receives a predetermined payoff of 100 points while the loser receives nothing. Race to 21 is a zero-sum game; therefore it allows us to isolate the test for backward induction while disregarding any assumptions on social preferences or beliefs about other players.

Since the point of this experiment is to test the effect of incentives on backward induction, we complicate the baseline treatment by adding what we call "teaser" payments at various points along the equilibrium path of play. Subjects are offered an additional 50 points for choosing a number that yields a sum of, for example, 13 during the course of play. If the action space is 1-3, the teaser payment for landing on 13 serves as an incentive to stay on the equilibrium path, and the subject earns 150 points. When the action space is 1-4, a teaser payment at 13 could serve as a distraction from the equilibrium path; if he chases the teaser in this case, the subject sacrifices the larger payoff of 100 for the intermediate teaser payoff of 50. To isolate whether the subject is actually chasing the teaser, we test the teaser treatments on and off the equilibrium path, as well as early in the game and nearer the middle of the game, against a control treatment with no teaser. The computer is programmed such that if the human subject is on the equilibrium path the computer plays randomly, unless it can grab the teaser, in which case it does.

Subject Pool

We recruited approximately 200 undergraduate students during the summer of 2009 at the University of Tennessee-Knoxville (UT). The UT economics department recruits generically for experiments through the existence of a database called the Online Recruiting System for Economic Experiments (ORSEE). Currently, there are 1,543 students registered in ORSEE; of those, we invited 1,512 to register for the Race to 21 experiment. Invitees received an e-mail describing the task as "a market experiment … [in which] earnings are determined by the decisions you make." We had 212 students register, and 199 actually showed up to participate. Academic majors represented were quite diverse, including liberal and fine arts, business, social and hard sciences, education, and nursing.

Experimental Procedure

The experiment sessions took place in the UT Experimental Economics Laboratory. The lab is set up with 27 individual client workstations (Dell PCs) networked to an intranet server. Study carrel walls separate each workstation to maintain subjects' privacy. Subjects remained anonymous to each other and their decisions remained private throughout the experiment.

The Race to 21 game is programmed in Perl[6], a highly flexible general-purpose dynamic programming language. Perl allows the subject in the lab to play the game via an html interface, and records the subject's and computer's decisions to a text file for analysis. The flexibility inherent in the program allowed us to make quick parameter changes between lab sessions to accommodate all 10 teaser treatments.

At the start of each lab session, experimenters gave subjects a copy of written instructions for the game to complement the same instructions visible on the participants' computer screens. Subjects were asked to follow along as experimenters read the instructions aloud. During each lab session, subjects

[6] Special thanks to Dr. Mike Shor at Vanderbilt University for programming our game.

played two Race to 21 games; game one consisted of 30 rounds of either action space 1-3 or 1-4, while game two consisted of 15 rounds of the action space that was not played in the first game[7]. The decision task was framed as one of maximizing earnings by removing between 1 and 3 (or 1 and 4) numbered stones, alternating with the computer opponent, until all 21 stones were removed from the screen. Figure 1 shows a screenshot of the graphic interface the subjects faced[8].

Subjects were told they earned points when they, and not the computer, removed the green stones. The boldface numbers on the green stones represent the points they earned for removing those stones. Players removed stones by mousing over the stones they wanted to remove (which highlighted them on the screen), and then clicking on the last stone in the series they wanted to remove. Participants were given the opportunity to practice with the interface and ask questions before actual play began. All navigation through the pages of the experiment occurred by clicking a button at the bottom of each screen labeled "Proceed". At the completion of the final round of game two, we asked subjects to complete a short questionnaire. Immediately following the experiment, we paid subjects their earnings privately in cash. Payments were calculated at the rate of $1 per 150 points. Please see Appendix B for a copy of the written instructions, screen shots for all phases of the experiment, and the questionnaire.

Each lab session represented one of 10 different treatment parameters. Table 1 summarizes the 10 treatments we tested in game one.

4. Hypotheses

Levitt, List, and Sadoff (2008) found in their Race to 100 game that simply changing the action space from (1-9) to (1-10) dramatically reduced the share of subjects who fully backward induct from nearly sixty percent to less than fifteen percent—a striking result given their subject pool consisted of chess

7 Only the results of game one are reported in this essay.
8 All figures and tables may be found in Appendix A.

grandmasters with extensive experience in backward induction. To explore whether the same phenomenon occurs with less experienced student subjects, we test each teaser against both action spaces (1-3) and (1-4). Theoretically, it's unclear which of these should make backward induction easier. Asking a subject to choose a number between one and three in a race to 21 leaves him with five decision nodes to contend with, and three options at each node. Conversely, selecting from one to four leaves him with only four decision nodes, but four options to choose from at each node. Whether backward induction is easier with fewer decision nodes to the end of the game but more options at each node, or fewer options to choose from at each node but more total nodes to analyze, remains an empirical question.

To determine the optimal placement of an incentive within the structure of a race game, we also test teasers placed early in the game and near the midpoint of the game against games with no teasers. A midpoint teaser might enhance backward induction by shortening the game for the subject. In the extreme case, an on-equilibrium teaser at the midpoint of the game, e.g. on 11 in the (1-4) game, may effectively turn the four-node race to 21 game into two separate subgames of two nodes each. Then the subject need only backward induct two moves at a time from the end of the each subgame, lessening his cognitive burden relative to one longer game. Alternatively, encountering a teaser on one of the first several stones in the game could function to entice a forward-looking subject to get on the equilibrium path of play earlier in the game, increasing the probability that a subject will win the game. We will look at the data from several different angles in an attempt to parse these effects.

For obvious reasons, placing a teaser on the equilibrium path should make backward induction easier than in games with no teaser. By offering an extra fifty points for taking the teaser along the way, subjects have an even greater incentive to make the "right" decision at every node. Predicting the effect of an off-equilibrium teaser payment is a more difficult exercise. On one hand, the off-equilibrium teaser effectively shortens the game in the same manner as described above, which should make the game easier to solve by backward induction. On the other hand, this teaser takes the subject off the equilibrium path of play. Since the computer program punishes any deviation from equilibrium in the context of our Race to 21 game, we might expect that backward induction is made more difficult by this effect relative to control treatments with no teaser. Again, we'll use emprics to help us determine which of these effects dominates.

To sum up, relative to a game with no teasers, we predict on-equilibrium teasers to enhance backward induction. For the reasons outlined above, off-equilibrium teasers, early and mid-game teasers, and differences in action space may either enhance or detract from learning via backward induction. We will use the results of our laboratory experiment to draw conclusions regarding the empirical questions.

5. Results

Due to capacity constraints in the lab, technical glitches resulting from data transfer, and the failure of a few subjects to follow directions, we report the results from 179 individual subjects (out of the 199 who showed up to participate). Subjects earned \$18.91 on average, with a median payment of \$18.08, and mode of \$5 (the minimum payoff regardless of performance on the task). We conducted a total of 12 sessions over June and July 2009.

Before it is possible to analyze the effects of teasers on backward induction, first it is necessary to establish that subjects are actually able to backward induct. We begin with the question of whether or not subjects use backward induction in the simple (no teaser) Race to 21 game. Table 2 presents results on the probabilities of remaining on the equilibrium path of play at each node, conditional upon reaching that node. Rows correspond to the action space tested, and columns correspond to each different decision node, working backward from the end of the game. Just beneath each node's description in the first row are the relevant "key numbers" on the equilibrium path of play for each action space version (in parentheses). Equilibrium play dictates that exactly five decision nodes will be reached in the (1-3) game, while only four nodes are required in equilibrium for the (1-4) game.

The first column of Table 2 tells us that of the 360 observations on the (1-3) version of the Race to

21 game, 96% successfully solved the game if they arrived on the final equilibrium node; i.e., of the subjects who chose a number on their penultimate move that summed to seventeen, 96% selected the number that summed to 21 on their next move. Likewise, conditional on reaching the key number sixteen on their penultimate move in the (1-4) version, 97% of subjects will win the game. An alternative way of interpreting these results is this: 46% of subjects made the "right" first move to remain on the equilibrium path in the (1-3) game. Of those, 58% chose a number that added up to 5 on their next move. Of those subjects who still remained on the equilibrium path at node "Final-3", 81% continued to the following equilibrium node, and so on. These results indicate that at least some subjects are able to backward induct in the Race to 21 game absent distracting teaser payments. We will compare all subsequent treatments to this baseline in order to draw conclusions regarding teaser effects on backward induction.

Table 3 shows the percentage of all rounds played that subjects won for each experimental condition. Recall that each of 179 subjects played thirty rounds of the same game, four different teaser locations were tested (on-early, on-mid, off-early, and off-mid), and two different action space conditions were tested (1-3 and 1-4) for each teaser location, for a total of 5,370 unique observations. In order to win the game, a subject must never deviate from the equilibrium path; therefore, we can equate winning with perfect backward induction[9].

The intersection of the first row and first column of Table 3 indicates that subjects win the game in which the teaser falls on the equilibrium path and early in the game 45% of the time when the action space is (1-3), and 41% of the time when the action space is (1-4). For ease of exposition, we first aggregate the above results for whether the teaser payment falls on the equilibrium path, early or near the middle of the game, and whether the action space presented is (1-3) or (1-4). Then we break the aggregate results down further to support our conclusions. Our analysis leads to the following insight:

RESULT 1: *Teaser payments that fall on the equilibrium path of play make backward induction easier*

[9] The probability of winning just by chance is 0.0041 in the (1-3) game and 0.0039 in the (1-4) game.

relative to games with no teasers.

In Table 4, we pool the data from Table 3 to compare the on-equilibrium teaser condition to the no-teaser condition. Subjects win games when a teaser payment falls on the equilibrium path 52% of the time, more than doubling the probability of winning 25% of the time when no teaser payment exists. This relationship is true regardless of whether the teaser occurs early in the game or nearer the middle.

To complement this analysis, we test the linear probability specification[10] of a binary regression model of the form

$$P(win_{it}) = \beta_0 + \beta_1(on) + \beta_2(off) + \beta_3(on * early) + \beta_4(off * early) + \varepsilon_{it}$$

where $P(win_{it})$ equals one if subject i wins the game in round t and zero otherwise, and the regressors represent interactions between categorical dummy variables. The regressors should be interpreted as follows: "on" is really the "on*mid" interaction, and "off" is really "off*mid". To glean the marginal effect of the on-early teaser condition it is necessary to sum the coefficients on the "on" and "on*early" regressors, and likewise for the "off" and "off*early" interaction coefficients. Empirical results are reported in Table 5, which presents the marginal effects associated with a change in each of the regressors relative to the control (no-teaser) condition.

This model explains 11% of the variation in the data set. The coefficients on the independent variables tell us that relative to the observed probability of winning over all treatments (36%), presenting the subject with an on-mid teaser increases the probability of winning by 34%, holding all other variables in the model constant. Likewise, putting the teaser on the equilibrium path but early in the game increases the probability of winning by 18% (.34-.16=.18) relative to the overall probability of winning, *ceteris paribus*. It should be noted that the default treatment for the regression, the no-teaser case, corresponds to

[10] A potential drawback of this model is that the estimated coefficients can imply probabilities outside the unit interval [0,1]. For this reason, the marginal effects probit model is often used instead. After running the regressions using both linear probability and marginal effects probit using Stata, the coefficients we found were almost identical in each and predicted probabilities never fell outside the unit interval. This is probably attributed to having a very large data set.

the constant term. These results corroborate our earlier conclusion that on-equilibrium teasers are better than no teasers at all.

Finally, as a test of robustness we can look at subjects' average earnings across treatments. On-equilibrium treatments result in average earnings per participant of 1,763 points ($11.75), after backing out the teaser payment of 50 points for each round won. This is indeed much greater than the 717.86 points ($4.79) earned on average in the no-teaser setting, and the difference is statistically significant at the 99% confidence level.

Now that we've established that on-equilibrium teasers outperform no-teaser treatments, further exploration into the data leads to the following result:

RESULT 2: *Teaser payments presented near the midpoint of the game yield better results than those offered near the beginning of the game—but only if those teasers are simultaneously on the equilibrium path.*

In Table 6, again we aggregate the data on mid- and early-game teasers from Table 3 and compare to the aggregated no-teaser treatment results. Subjects win games 39% of the time when a teaser occurs near the midpoint of the game, and 35% of the time when they find teasers early in the game. While both aggregate measures are significantly better than the no-teaser treatment, it is obvious from Table 3 that this result is driven by the teaser falling simultaneously on the equilibrium path. This merely reinforces the conclusions drawn in Result 1 above.

Probing this result further, we re-examine the linear probability model introduced in the previous section. In comparing the regression coefficients of on-mid to on-early teasers, it is clear that the midpoint teaser yields wins nearly twice as often as when the teaser is presented early in the game (34% and 18%, respectively). Examining the off-equilibrium equivalents reveals that a midpoint teaser payment actually reduces the probability of winning the game when it occurs off the equilibrium path of play by 8%, while the off-early teaser has almost no effect (0.2%) on the probability of winning the game.

As in the previous section, we now check average earnings for all subjects across experimental conditions to provide further evidence for our result. Subjects playing games with on-mid teasers earn 1,793.62 points ($11.96) on average, after accounting for the extra fifty points per game available in this setting. This is slightly better than the 1,732.32 points ($11.55) earned on average by those playing games with on-early teasers. In contrast, when the teaser is presented off the equilibrium path, subjects facing teasers early in the game fare slightly better than those who see teasers closer to the middle; in fact, they earn an average of 721.68 points ($4.81) and 523.62 points ($3.49) respectively. This comparison reinforces the result that mid-game teasers outperform early game teasers if the teaser also happens to occur on the equilibrium path of play.

Our final insight concerning the effect that variation in the structure of the Race to 21 game has on backward induction involves the strategy space a player faces:

RESULT 3: *Backward induction is made easier with fewer decision nodes, even if there are more choices to analyze at each node.*

To see this, first we pool the number of rounds won over all of the (1-3) and (1-4) strategy space treatments, respectively, from Table 3. This gives us the proportions shown in Table 7. On the whole, subjects win games where they choose from one to four at each of four decision nodes 38% of the time, versus 33% of the time for those games in which subjects choose from one to three at each of five nodes. The data here suggest that it is the number of decision nodes the player faces, rather than the strategy space at each node, which better predicts performance in backward induction tasks.

To lend further credence to this result, we break out the action space results by teaser treatment in a marginal effects probit model:

$$P(win_{it}) = \beta_0 + \beta_1(1to3) + \beta_2(on) + \beta_3(off) + \beta_4(on * early) + \beta_5(off * early)$$
$$+ \beta_6(off * early) + \beta_7(on * 1to3) + \beta_8(off * 1to3) + \beta_9(on * early * 1to3)$$
$$+ \beta_{10}(off * early * 1to3) + \varepsilon_{it}$$

where $P(win_{it})$ again refers to the binary outcome (1 if subject *i* wins the final stone in round *t*, 0 otherwise), and the regressors represent interactions among categorical dummies. As in the previous linear probability model, the coefficients must be appropriately combined to interpret the marginal effects of each teaser condition on the probability of winning the final stone. Stata gives us the raw output in Table 8. Recall that to interpret the marginal effects by treatment, we sum the coefficients as seen in Table 9 (using the alphabetical references to the regressors from Table 8). The cells in Table 9 correspond to the marginal effects of each teaser treatment relative to the no-teaser control condition within the same strategy space. For instance, having an on-early teaser in a (1-3) action space game increases the probability that a subject will win the game by 31% over the (1-3) game with no teaser. The third row in Table 9 is the most useful in providing further evidence in favor of Result 3. While the games played with on-early and off-mid teasers did not have statistically different outcomes based on action space, subjects won the on-mid teaser games 10% more often when asked to select from one to four stones, and they won the off-early teaser games 28% more often when tasked with strategy space (1-4).

In sum, we find that backward induction in the Race to 21 game is made easier by placement of on-equilibrium teasers near the middle of the game, and especially so when a subject has to contend with fewer decision nodes between the root and terminal node.

6. Conclusion

In this study, we introduce undergraduate student subjects to a controlled laboratory experiment in a pure test of backward induction. Making use of our Race to 21 game, we report several insights. We find that these subjects exhibit the ability to strategically backward induct substantially greater than random chance would explain. Further, we find that incentives placed within the structure of the game affect whether subjects are more or less likely to win the game depending on their location; specifically, subjects offered teaser payments on the equilibrium path near the midpoint of the game win more than twice as often as when they play the same game with no teaser available. The games in which subjects were offered teaser payments off the equilibrium path of play show variable results depending upon whether the teaser is offered earlier or later in the game. Further testing of off-equilibrium teasers should be done in order to draw more robust conclusions. Finally, we find that the length of the game matters for backward induction. Facing fewer nodes from beginning to end improves outcomes, regardless of the strategy space subjects must choose from.

Besides revealing these insights into incentivizing backward induction tasks, our work also offers a methodological contribution. It highlights the potential for computerized laboratory experiments to "train" subjects in backward induction via punishment for deviating from optimal play. It remains to be seen whether this type of training in backward induction can be generalized from one setting to another. We hope that future efforts will explore more fully other important dimensions of controlled laboratory experiments for training potential.

Chapter Two: Do incentives speed up or slow down learning depending on their positions relative to the equilibrium path and their locations within the game?

1. Introduction

In the previous chapter, we showed that human subjects are indeed capable of solving a decision task using backward induction techniques. Additionally, we reported that introducing an intermediate incentive payment into the structure of the decision task that our subjects faced had the propensity to enhance or distract them from using backward induction to strategically maximize their payoff in the game. Specifically, subjects playing the Race to 21 game won more than twice as often when offered teaser payments on the equilibrium path near the midpoint of the game as when they played the same game with no teaser payment available. We further concluded that the length of the game matters for backward induction. The fewer decision nodes a player has to contend with, the better the outcome.

The data we analyze in this essay is based on the same set of Race to 21 experiments that we used in the previous chapter. Given that we know our subjects are able to use backward induction to solve simple decision tasks such as those presented here, we test whether intermediate incentive payments affect the speed with which a subject learns the optimal path of play based on where the incentive is located within the structure of the game. We predict that the same conditions which were advantageous for backward induction in the first chapter will also prove to induce the fastest learning from one round to the next. Specifically, we hypothesize that on-equilibrium teasers will outperform off-equilibrium teasers, and that a teaser offered near the midpoint of the game will result in faster learning if it is simultaneously on the equilibrium path of play.

Before we can definitively say which conditions speed up or slow down learning, it is necessary to establish appropriate methods to measure the speed of learning. So then, what constitutes a good measure of speed of learning? In our Race to 21 game, we say a subject "learned" the game when he wins the final stone. Then the appropriate question is how long does it take a subject to figure out the path he must take to reach the final stone? To answer this, we will count the number of rounds a subject plays before he is successful in winning the game in the absence of a teaser payment. Finally, we want to know whether we can hasten or postpone subjects learning the optimal sequence of game play by offering an incentive (teaser) payment for taking an intermediate stone somewhere along the way. After analyzing the data through several different methods, we find that on-equilibrium teasers offered near the middle of the game result in faster learning than in games in which no teaser payment is offered, and further, that the on-mid combination is superior to all other teaser conditions tested.

2. Measuring the speed of learning

Since we are interested in measuring how long it takes an individual to learn the optimal (subgame perfect equilibrium) path between the root and terminal decision nodes, a logical first place to start is to count the number of rounds a subject played before he first reached the final stone in the Race to 21 game. In Table 10, we compare the average number of rounds subjects took to earn both the teaser stone and the final stone in each treatment. The non-italicized entries represent those averages conditional on subjects reaching the target stone at all, while the italicized entries refer to those averages calculated by what we call the unconditional method, defined immediately below. Additionally, Table 10 contains information on the number of rounds subjects take to get the teaser stone the first and second time, the number of rounds they take to reach the final stone the first and second time, the fraction of rounds in which these target stones were taken, and statistical tests for the differences among all of these numbers. Notice the no-teaser game results near the bottom of Table 10. These serve as our control treatments, against which we compare the various teaser treatments in order to draw conclusions about the effects of different incentives on the speed of learning.

From Table 10, we see that conditional on winning the game in any round, subjects who chose from one to three stones at each node first found the equilibrium path of play by (on average) round six of thirty. However, this result might be misleading. Only one-third of the subjects who played this particular treatment ever reached the final stone in the game even once, and they account for only 18% of all rounds played. So using this average to measure speed of learning is probably biased downward, making the game appear easier to solve than it really is. To mitigate this effect, we alternatively assume that for any subject who failed to win the game at least once in thirty rounds, he would have won the game in round 31. This unconditional assumption is second-best; we have no way of knowing how many rounds the subject really would have needed, much less if he would learn the game at all. Nevertheless, it better reflects the "tougher" nature of the decision task. After accounting for all subjects who played the no-teaser games and without conditioning on having won the game by round thirty, we find that on average subjects took the final stone in the (1-3) game by round 23, and the subjects who played the (1-4) version reached the optimal path by round 20.

Taking a closer look at the various teaser conditions, we likewise count the number of rounds it took our subjects to first reach the path to take the teaser stone. As we might expect, when the teaser was presented early in the game, conditional on subjects taking the teaser at all, they grabbed it in the very earliest rounds of the game—on average before round three[11]. In contrast, when the teaser occurred near the middle of the game, subjects took several additional rounds to take the teaser (on average by round seven). We begin to see more interesting patterns in the data once we compare the speed-to-teaser to the speed-to-final stone.

One caveat should accompany the discussion regarding our measure of learning:

RESULT 1: *The first time a subject takes a teaser or final stone does not necessarily constitute learning.*

[11] For the reasons discussed in footnote 25, it is impossible for subjects to grab the teaser in the off-early (1-4) treatment.

The prior probability that a subject would select the correct move at each decision node to win the Race to 21 game by random chance is 0.0041 in the (1-3) game and 0.0039 in the (1-4) version. One would think these probabilities small enough that actually observing a subject taking the final stone is more likely than not evidence that learning by backward induction has taken place. However, we have numerous examples in the data of a subject taking the final stone in one round, but then losing the next several rounds before consistently winning the final stone for the remainder of the game. For this reason, we also record the second round in which subjects won either the teaser or final stone. If the first time subjects take the teaser or final stone is indeed evidence of learning, then we would expect the second time to be the very next round. The columns in Table 1 labeled "delta$_x$," where *x* stands for either T (for teaser) or F (for final stone), represent the difference between the first and second times the subject took the teaser and final stones in each experimental condition. These calculations demonstrate that subjects average between one and six rounds between their first and second successful strategies[12]. Testing the null hypothesis that all delta=1, we are able to reject the null in all experimental treatments. Taking this into consideration, we will derive an alternate measure of learning later in this essay. But for now, we shall use the delta calculation to probe further the unique properties possessed by each teaser combination.

In order to draw conclusions about the effects that different incentive placements have on the speed of learning, we calculated the difference between the first round in which our subjects took the teaser and the first round in which they took the final stone. The most striking difference leads to the following insight:

RESULT 2: *Finding the path to the final stone once a subject has reached the teaser stone occurs instantaneously for the on-mid treatment and not for any other treatment.*

[12] Except for the on-mid treatment, which we will establish is different from the rest in Result 2.

From the rightmost column in Table 10, it becomes obvious that teaser payments have different effects on learning depending upon their position. We can see by looking at the on-early teaser treatment that on average subjects need five or six additional rounds after first taking the teaser stone before they find the optimal path to the final stone. In aggregate, this improves speed of learning relative to the no-teaser control condition by at least a couple of rounds. However, the most striking difference occurs when the teaser is offered on the equilibrium path near the midpoint of the game. In this case, our subjects need at most one round to find the path to the final stone after reaching the teaser. In fact, we are unable to reject the null hypothesis that the difference between the first round a subject found the teaser stone and the first time he reached the final stone are the same round in the on-mid teaser condition.

Further, this immediate teaser-to-win relationship is unique to the on-mid combination. If the important component of this condition is the fact that the teaser falls on the equilibrium path, we should expect to see similar results in our on-early teaser treatment. However, Table 10 shows that not only do subjects win fewer rounds in the on-early treatments, but it also takes them at least five extra rounds between getting the teaser and getting from the teaser to the final stone. Likewise, the fact that there appears to be a big difference between learning in the off-mid setting (taking at least eight rounds to reach the optimal path to the final stone after taking the teaser) and in the on-mid setting tells us that learning is not improved just by placing a teaser in the middle of the game.

Looking more closely at the relationship between the first and second rounds in which subjects take the teaser and final stones, we are able to glean the following observation:

RESULT 3: *Getting the teaser stone is somehow fundamentally different from winning the final stone in the on-mid teaser case.*

To see this, first we take the difference between the first and second times the subjects grabbed the teaser stone (column "delta_T"). Restating Result 1 above, if our subjects are learning to take the teaser rather than happening upon it randomly, then we should see very little (if any) time pass between the first and second rounds in which they take the teaser. We test the null hypothesis that the number of rounds between the first and second “takes” is less than the number of rounds between the start of the game and their first take. The results of this hypothesis test tell us that subjects are not getting the teaser by accident; however, subjects playing the on-mid teaser game take longer to get their second teaser than in any other experimental condition. Next, we compare the number of rounds between the first and second time subjects take the final stone to win the game (column “$delta_F$”). Here, the on-mid treatment when the action space is (1-3) shows a significant advantage over the no-teaser control, conditional on subjects learning the game at all (less than one round on average between the first and second wins, compared to over four rounds). Only the off-mid (1-4) treatment comes close, and that is complicated by the fact that the average first win comes over two rounds later in the game for the latter treatment. Similarly, although the differences between first and second wins for both the on-mid (1-4) and no-teaser (1-4) treatments are nearly identical, the subjects playing the on-mid (1-4) game first play the winning strategy over five rounds sooner. Finally, we test the null hypothesis that the $delta_T$-to-second-teaser ratio is the same as the $delta_F$-to-second-win ratio. In all cases we are able to reject the null hypothesis, although the difference between the ratios in the on-mid treatment is only weakly significant (p=0.096). So, we conclude from this analysis that the combination of an on-equilibrium teaser with the teaser located near the middle of the game offers significant advantages over every other condition we tested in the laboratory.

3. Ranking conditions for speed of learning

In Result 1 we asserted that just because a subject wins one round, does not necessarily mean he learned the optimal path. We recorded numerous observations in which a subject won his first round, and then failed to win for several rounds in a row before getting back to the equilibrium path of play. Then it cannot be the case that the first round the subject won provides evidence that learning has taken place. With that in mind, we need a more robust definition of learning in the context of the Race to 21 game. Our new and improved definition of learning has two parts: (1) once learned, there cannot be gaps of more than two consecutive rounds lost, and (2) learning starts with two sequential rounds won. According to this new definition of learning, we rank the various teaser conditions and report the results in Table 15.

Using Table 10 exclusively for all of our analysis leaves out a crucial part of the story. While it is advantageous for hypothesis testing, analyzing the data conditional on observing a subject taking a teaser and/or final stone fails to account for sample size in each treatment. To see why this is potentially problematic, notice that if a treatment is particularly difficult, then only the fastest (smartest?) fraction of subjects ever learn the path to either the teaser or the final stone. Taking the average of the first round in which that happens tells us nothing about how many were actually able to solve the game. Thus we cannot compare the average number of rounds to getting stone 21 across treatments. The average may be artificially low if the slower players are sorted out.

One way of correcting for this problem, mentioned earlier, requires calculating an unconditional average for each treatment. In doing so, we made the conservative assumption that any subject whom we did not observe ever playing the winning strategy would be arbitrarily assigned a win in round 31. While this assumption gave us higher average first-take rounds for some games than they were in the conditional analysis, we still have difficulty in assessing the magnitude of differences between treatments.

To get at the magnitude among rankings, we posit the following model. Consider a game without a teaser, and let $F(t)$ be the fraction of people who learn the game before or during round t. Then $1 - F(30)$ is the fraction of people who never win the game because they are cut off after 30 rounds. Next, suppose we add a teaser. We assume that $F^{T1}(t) = F(\beta^{T2}t)$, where the superscript Ts represent the particular teaser positions (*i.e*., on/off and early/mid) being compared. Now suppose that in the no-teaser treatment a fraction x of the n subjects learn the game by period t. These are the fast xn of the *n* subjects. With some teasers (*e.g.*, on-mid) it will take them less time to learn the game, so that same fraction *x* will learn the game by period β*t*, where, if the game is easier, β is less than one. If we took a harder treatment, like off-early, we might get β greater than one. Now let us assume that the β value is the same for *all* choices of *t*. So, if it takes twice as long for the fastest people, it also takes twice as long for the slow people. Likewise, if it takes half as long for the fast people, it takes half as long for the slow people.

If we have a harder treatment, everyone is slower, the slowest ones need more than thirty rounds, and we get fewer people solving the game. This might lead to a faster average learning time, though. To see why, suppose that in the "easy" treatment we get five people solving the game (one each in rounds one through four and one in round twenty). The average first-win occurs in round six. If the hard treatment makes everybody take twice as long, the fifth person never solves the game and we are left with four people who solve it in rounds two, four, six, and eight; thus, the average first-win round is five, which appears to be faster than in the easy treatment. However we know this cannot be the case, since it takes twice as long to solve the game in the harder treatment.

Now we have a way to estimate how difficult a game is. We do this by arranging the data in such a way to compare the earliest round in which the fastest player learned the game across treatments (according to our new definition of learning), then we find the first round in which the next-fastest player learned the game for each treatment, then do it again for the third-fastest player, and so on (Table 11).

According to our theory, in order to calculate the relative difficulty level between treatments (or β) we regress each experimental treatment (one column from Table 11) against its no-teaser strategy space analog, suppressing the constant term[13]. We test the null hypothesis that both treatments are of equal difficulty (β=1). If we reject in favor of β < 1, the experimental treatment is easier than the control

[13] Note the disadvantage to using this method: the OLS regression requires the same number of observations on the dependent and independent variables. Therefore not all of the data is used; i.e., when testing each treatment against its analogous no-teaser control, the observations are truncated at whichever has the lesser number.

because it takes less time for the same fraction of subjects to learn it. If we reject in favor of $\beta > 1$, the experimental treatment is harder than the control. The magnitude of the regression coefficient tells us how much easier or harder each teaser makes the decision problem, allowing us to rank the treatments.

The results of each regression are presented in Tables 12 and 13. In Table 12, we regress the teaser treatment on its no-teaser analog, under the assumption that all subjects who never solved the game in thirty rounds would have done so in the 31^{st} round. In both the (1-3) and (1-4) strategy space games, the on-mid teaser proves to speed up learning the most relative to the no-teaser option, followed by the on-early teaser condition. When subjects choose from one to three stones, the off-mid treatment is not significantly harder or easier to solve than its no-teaser analog. Likewise, the on-early, no-teaser, off-mid, and off-early teaser conditions are equally easy to solve in the (1-4) version. The relatively high R^2 values indicate that our theory works well with the data.

In Table 13, we treat the data slightly differently to test the robustness of our theory. Assigning a value of 31 to all subjects we never observed winning the final stone is useful to the extent that we do not artificially lower the average speed of learning; however, it turns out that in Table 12 we end up regressing a lot of 31s against 31s. This makes our coefficients closer to one, thus we have a lot of insignificant differences. Another alternative way to treat the data is to match the columns in Table 11 until both have coinciding 31s. This has the effect of reducing our sample size, but precludes the bias toward one in the coefficients. We report the results of this regression in Table 13. Notice that the on-mid teaser combination regression coefficients are identical in this treatment, and still win out over all other combinations tested.

These results, while helpful in that they offer evidence that some teaser treatments may make the game easier to learn than others, still suffer from the problem of including "false-positive" observations (i.e., assigning a win in round 31 to those whom we never observe winning). Then perhaps the appropriate comparison to make is among those subjects who would have learned the optimum strategy for any treatment. We shall call these the fastest twenty percent[14] of subjects in each treatment. Table 14 illustrates the top twenty percent of first-win observations taken from Table 11 above. Examining the fastest twenty percent of subjects, it seems that for the (1-3) version of the game, the on-early treatment proved easier than any other, followed closely by the on-mid and no-teaser treatments. Similarly, the fastest subjects found the off-early treatment easiest to solve, followed by the on-mid and on-early treatments. Table 15 provides a summary of our rankings under each method of analysis.

Under our new definition for learning, the (1-3) on-early and on-mid treatments swap ranks under the fastest twenty percent test. Additionally, the no-teaser, off-early, and off-mid treatments vary in their rank by the strategy space of the game and the method of data analysis. The most important result to take from this robustness check is that the on-mid treatment proves to speed up learning more than any other treatment under the majority of testing methods addressed, lending further support for the superiority of this teaser combination. The rankings illustrated in Table 15 allow us to summarize the data in the following way:

RESULT 4: *Offering a teaser payment in the on-mid position is the consensus fastest way to induce learning relative to the no-teaser control treatment in the Race to 21 game. However, rankings among all treatments are sensitive to the method of analysis.*

4. Conclusion

Looking at the data collected in our Race to 21 laboratory experiment, we report several insights into how the speed of learning may be affected based on the location of an intermediate incentive payment. We find that combining an on-equilibrium incentive payment with the mid-game location is unique for several reasons. First, once a subject learns how to find the optimal path of play from the first stone to the teaser stone via backward induction, learning the optimal path to solve the game occurs

[14] We chose to test the fastest 20% (rather than 10% or 30%) by counting the smallest number of observations we have for any treatment and dividing by the number of subjects who participated in said treatment. We only observed two out of eleven subjects ever winning the off-early (1-3) treatment, which represents 20% of those who played.

nearly instantaneously. The same is not true for any other experimental condition. Additionally, we find that for some reason, learning to get the teaser stone is somehow fundamentally different from learning to get the final stone in the on-mid teaser case. Subjects playing the on-mid game take longer to get their second teaser, while they need much fewer rounds to grab their second final stone, than any other experimental treatment tested. Finally, we show that although these results are sensitive to the method of analysis tested, the superiority of the on-mid teaser treatment holds in aggregate.

CHAPTER THREE: WHAT FACTORS PREDICT BACKWARD INDUCTION, AND TO WHAT DEGREE?

1. Introduction

In the previous chapter, we demonstrated that subjects learn how to solve the Race to 21 game over time. Further, we showed that the speed of learning is affected by the position of incentive payments within the structure of the game. Specifically, unlike any other teaser condition, placing an incentive on the equilibrium path near the midpoint of the game allows for backward induction to transfer almost instantaneously once a subject has learned the path to the teaser stone. Finally, we discussed how learning to get the teaser stone is somehow fundamentally different from learning to win the game for only the on-mid teaser experimental condition.

The data we use in this essay comes from the same set of Race to 21 experiments that we analyzed in previous chapters. As in previous chapters, we analyze the effect that various "teaser" incentive payments had on backward induction. However, in this study we add to the analysis demographic variables (i.e., age, gender, number of economics courses taken, and the number of economic experiments a subject previously participated in), as well as the results of a simple cognitive ability/patience battery to test whether, and if so to what degree, these factors predict backward induction in our decision task.

2. The Data

Subject Pool

Before reporting the results of demographic effects on backward induction, first it is necessary to have a good idea of the characteristics of our subject pool. As in the first two essays, we report on 179 undergraduate student subjects, of whom 74 are female and 105 are male. They range in age from eighteen to forty-three, with a mean age of 21.43 and median twenty. Figure 2 depicts age distribution via histogram.

In addition to age and gender, we asked subjects to report the number of economics courses they had taken at the university level, as well as whether they had participated in economics experiments previous to this one[15]. Nearly twenty-five percent of all subjects report they had never taken a single economics course, and the vast majority of our subjects (59.78%) have taken one economics course[16]. The remaining 15.64% of subjects report taking anywhere from two to fifteen[17] economics courses at the university level. Additionally, 141 of 179 subjects (78.77%) report having previously participated in an economics experiment.

Cognitive Reflection Test (CRT)

In addition to the demographic variables mentioned above, our subjects also faced a series of three questions designed to measure cognitive ability and patience[18], displayed in Figure 3. Frederick (2005) discusses the difference between the "intuitive" (and wrong) answers impulsive and/or unintelligent subjects give (ten cents, 100 minutes, and 24 days), and the correct responses given by subjects with a higher proclivity for computation (five cents, five minutes, and 47 days). His collective analysis of previous literature demonstrates a strong enough correlation between those who get the CRT questions correct and those with high scores on standardized tests of intelligence to conclude the CRT functions as a good test of cognitive ability. Additionally, his own results show that those who scored high on the CRT also made decisions which implied low discount rates of time preference, offering evidence of a positive relationship between cognitive ability and patience.

---

[15] We also have data on each subject's academic major as well as their responses to a 42-question personality-type survey. Future work along these dimensions should prove most enlightening.

[16] Apparently introductory economics is not a graduation requirement to earn an undergraduate degree from the University of Tennessee. Rather, there is a "social science" general education requirement that students may satisfy by taking two courses from a list of twenty-one offered in Africana, Anthropology, Child and Family Studies, Economics, Geography, Political Science, Psychology, Sociology, and Women's Studies. That more than half of our subjects have taken at least one economics course may be due to some economics instructors calling attention to the registration process for ORSEE (web-based lab management software) in class.

[17] The economics major requires only ten economics courses to graduate. Methinks fifteen a dubious claim.

[18] These questions are borrowed from Frederick (2005). He finds that CRT scores are predictive of the types of choices that feature prominently in tests of decision-making theories, e.g. expected utility or prospect theory.

Our subjects were required to answer the CRT questions as part of a survey they completed after all rounds of the Race to 21 game were complete (the computer was programmed in such a way that their final score would not be displayed until all survey questions were accomplished). To make their responses salient, subjects were paid an additional fifty cents for each correct answer. Summary statistics on the CRT responses appear in Table 16. It is interesting to note that while nearly one-third of subjects failed to calculate a correct answer in any of the three questions, more than twenty percent in our sample came up with the correct answer to all three questions[19].

3. Estimation Strategy

We assume that if a subject learned how to solve the Race to 21 game via backward induction at all, then he should have won the final round of the game. Therefore, our first measure of learning backward induction is whether he won or lost the game in round thirty. Using this binary outcome as our limited dependent variable, we report the results of a series of four marginal effects probit regressions to measure the degree to which experimental condition, demographic information about subjects, and subject performance on a cognitive ability/patience battery predict backward induction in our decision task. As a test of robustness, we regress via ordinary least squares (OLS) all of the same right hand side variables against the total number of points scored in all thirty rounds of the game. Each model tested is presented below, where we substitute "Total Points" for $P(win_{i,30})$ on the left hand side in our second (OLS) set of tests:

- Teaser treatment effects:

$$P(win_{i,30}) = \beta_0 + \beta_1(1to3) + \beta_2(on) + \beta_3(off) + \beta_4(on*early) + \beta_5(off*early) + \beta_6(off*early) + \beta_7(on*1to3) + \beta_8(off*1to3) + \beta_9(on*early*1to3) + \beta_{10}(off*early*1to3) + \varepsilon_{i,30}$$

- Teaser treatment, demographic effects:

[19] The twenty percent who got all three questions here is slightly higher than the seventeen percent that Frederick found in his study, though his sample is far larger.

$$
\begin{aligned}
P(win_{i,30}) = \beta_0 &+ \beta_1(1to3) + \beta_2(on) + \beta_3(off) + \beta_4(on*early) + \beta_5(off*early) + \beta_6(off*early) \\
&+ \beta_7(on*1to3) + \beta_8(off*1to3) + \beta_9(on*early*1to3) + \beta_{10}(off*early*1to3) \\
&+ \beta_{11}(gender) + \beta_{12}(age) + \beta_{13}(prev.exp) + \beta_{14}(econ.courses) + \varepsilon_{i,30}
\end{aligned}
$$

- Teaser treatment, cognitive battery effects:

$$
\begin{aligned}
P(win_{i,30}) = \beta_0 &+ \beta_1(1to3) + \beta_2(on) + \beta_3(off) + \beta_4(on*early) + \beta_5(off*early) + \beta_6(off*early) \\
&+ \beta_7(on*1to3) + \beta_8(off*1to3) + \beta_9(on*early*1to3) + \beta_{10}(off*early*1to3) \\
&+ \beta_{11}(ball) + \beta_{12}(lake) + \beta_{13}(widget) + \varepsilon_{i,30}
\end{aligned}
$$

- Teaser treatment, demographic effects, cognitive battery effects:

$$
\begin{aligned}
P(win_{i,30}) = \beta_0 &+ \beta_1(1to3) + \beta_2(on) + \beta_3(off) + \beta_4(on*early) + \beta_5(off*early) + \beta_6(off*early) \\
&+ \beta_7(on*1to3) + \beta_8(off*1to3) + \beta_9(on*early*1to3) + \beta_{10}(off*early*1to3) \\
&+ \beta_{11}(gender) + \beta_{12}(age) + \beta_{13}(prev.exp) + \beta_{14}(econ.courses) + \beta_{15}(ball) + \beta_{16}(lake) \\
&+ \beta_{17}(widget) + \varepsilon_{i,30}
\end{aligned}
$$

The goal of this estimation method is twofold. First, we want to determine how much more variation each model is explaining over and above the baseline teaser treatment effects. Second, among the demographic and cognitive battery variables, we want to know precisely which regressor is picking up the biggest effect on backward induction. The results of our analysis will tell us the answers to both.

4. Results

We begin the analysis of our results by looking at how well our models explain the variation in the experimental data. Table 17 shows the respective R^2 values for each regression form. The baseline marginal effects probit model we posit explains 14% of the variation in the probability a subject will win round thirty, while the analogous baseline OLS regression accounts for 23% of variability in total points earned by subjects. To analyze how much more variation is explained away by adding demographic information, we take the difference between the second and first columns. This difference tells us that adding demographic information accounts for an additional 8.5% of variation in the probit, and an additional 10.21% in the OLS model. Likewise, we want to compare the baseline to the model in which we include the cognitive ability/patience battery. Here, the latter adds 17.97% explanatory power to the baseline probit, and adds 23.85% to the OLS baseline. These values suggest our first result:

RESULT 1: *The cognitive ability/patience battery is better able to predict backward induction than simple demographic information.*

Running a test on “all of the above” regressors indicates a marginal improvement in explanatory power over the cognitive ability probit model, and a larger but still modest increase over the cognitive battery in the OLS version[20].

Tables 18 and 19 provide the results of each probit and OLS regression run using the data specified in section 2 of this essay. The leftmost column identifies the pertinent regressors, and each subsequent column contains the coefficients on those independent variables included in each model. Recognizing that it is necessary to calculate linear combinations to ascertain the marginal effects of each teaser treatment on backward induction, we present those marginal effects that correspond to the “*f*(Treatment)” column in Tables 18a and 19a. These will serve as our baselines for comparison.

Recall that in the first essay we employed the same marginal effects probit model on all observations over thirty rounds of our race game. Comparing the results in Table 9 to Table 18a below (using data on round thirty only), we see the coefficients are very similar in direction and magnitude, allowing us to re-state the following result from our first essay:

RESULT 2: *On-equilibrium-path teasers make backward induction easier; more specifically, on-mid teasers are superior to any other combination; and fewer decision nodes are preferable to more, even if there are more alternatives to choose from at each node.*

The similarity between Tables 9 and 18a also implies that the relative degree to which different teaser treatments affect success in backward induction is unlikely to vary across the four models we present in this essay, regardless of which other independent variables we include. Therefore, for ease of exposition of the remaining probit results we will take the teaser treatment relationships as given and discuss only the

[20] It should be noted that adding more independent variables to a model necessarily increases R^2; however, the relatively large jump in R^2 we get from adding demographic and cognitive battery information to the baseline model suggests the chance we are overfitting the model is negligible.

demographic and cognitive battery effects.

Following our baseline regressions, we test whether four demographic variables (age, gender, previous experiments, and number of economics courses) predict backward induction. In both the marginal effects probit and the OLS regression, age proves a positive but insignificant predictor of backward induction. However, that is not to say that age cannot predict backward induction necessarily. Our results are complicated by the fact that our subject pool consists only of undergraduate students; recall that Figure 2 displays a tight age distribution of our subjects. In fact two-thirds of our sample is nineteen, twenty, or 21 years old. Likewise, having participated in previous economics experiments demonstrates a positive, but not significantly different from zero, prediction for backward induction in both the probit and OLS measures.

Analyzing Tables 18 and 19 for demographic variables that serve as significant predictors of backward induction leads to the following insight:

RESULT 3: *Females find solving the Race to 21 game via backward induction more difficult than do males.*

From Table 18, we see that when demographics are included on the right hand side with experimental conditions, female subjects are 37% less likely to successfully win the game than males. Similarly, when included in our "all of the above" model, females are 25% less likely to use backward induction than males. Our robustness test results in Table 19 show the same negative and highly significant relationship. Frederick (2005) found very similar gender effects in his CRT study, and suggests that the gender gap may exist because men might be better at quantitative endeavors[21].

Also significant is the coefficient on the number of economics courses a student has taken. Specifically, the results of our tests imply that taking economics courses negatively relates to the propensity

[21] Larry Summers' infamous and intensely controversial hypothesis regarding gender disparity in tenure-track science and engineering positions at top research institutions pointed to high variability in IQ scores between men and women, similarly suggesting that some men may have an intrinsic aptitude advantage.

to use backward induction. While these results are significantly different from zero in three out of four models tested, they are likely biased by the fact that 84% of our sample have zero or only one economics course under their belts. Since undergraduate students can hardly be taught how to think strategically in a single principles course (much less if they have never taken a class in economics at the university level), it is not likely that this variable tells us much that is meaningful.

Finally, analysis of the results on the inclusion of the cognitive ability/patience battery leads to the following result:

RESULT 4: *Correctly answering the "ball" and "lake" questions from Frederick's CRT survey predicts backward induction in the Race to 21 game.*

Specifically, the "lake" question is a far better predictor of backward induction than the "ball" question, although both coefficients are large and significant. Answering the lake question correctly increases the probability that a subject will win the final stone by over forty percent, compared to nearly 25% for answering the ball question correctly. Likewise, testing our OLS model for cognitive battery effects shows that calculating the correct answer to the lake problem leads to an increase of over 1,100 total points earned over the entire game, compared to an increase of over 700 points for providing the right answer to the ball question. Curiously, the prediction value of the "widget" problem in our Race to 21 game, while positive, proved insignificantly different from zero. One possible explanation for this is fatigue, as it is among the last questions our subjects are faced with in a fairly strenuous decision task lasting approximately 45 minutes. However, considering the lake question appeared just after the widget question, fatigue is unlikely to be the reason that particular question was missed so often. In the future, it would be interesting to switch the order in which subjects read the problems to see if this effect holds[22].

[22] Frederick (2005) does not report how subjects performed on each problem in his CRT studies—only how many problems subjects answer correctly out of three.

5. Conclusion

Analyzing our Race to 21 game data after accounting for demographic variables and the results of a cognitive ability and patience battery allows us to report several insights into what factors may predict backward induction. Our analysis reinforces the results of the first essay in this study, in that teaser placement within the structure of the race game predict backward induction based on their location relative to the equilibrium path of play. Further, we find that certain demographic variables predict backward induction, namely gender (females struggle with backward induction more than males) and to a lesser extent the number of economics courses one has taken at the university level (though for reasons we discussed, this effect may not be as significant as it appears). Finally, we find that two of the three problems on Frederick's (2005) cognitive reflection test are significant predictors of backward induction.

BIBLIOGRAPHY

Aumann, R., and Maschler, M. *Repeated Games with Incomplete Information*. Cambridge, MA: MIT Press, 1995.

Bell, David E.; Raiffa, Howard and Tversky, Amos. "Descriptive, Normative, and Prescriptive Interactions in Decision Making," *Decision Making: Descriptive, Normative, and Prescriptive Interactions.* Cambridge; New York and Melbourne: Cambridge University Press, 1988, 9-30.

Bolton, G. E. and Ockenfels, A. "Erc: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 2000, *90*(1), pp. 166-93.

Busemeyer, J. R.; Weg, E.; Barkan, R.; Li, X. Y. and Ma, Z. P. "Dynamic and Consequential Consistency of Choices between Paths of Decision Trees." *Journal of Experimental Psychology-General*, 2000, *129*(4), pp. 530-45.

Costa-Gomes, M. and Crawford, V. P. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," 2004, 1-53.

Costa-Gomes, M.; Crawford, V. P. and Broseta, B. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, 2001, *69*(5), pp. 1193-235.

Costa-Gomes, M. A. and Crawford, V. P. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 2006, *96*(5), pp. 1737-68.

Crawford, Vincent P.; Kreps, David M. and Wallis, Kenneth F. "Theory and Experiment in the Analysis of Strategic Interaction," *Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress. Volume 1.* Econometric Society Monographs, no. 26. Cambridge; New York and Melbourne: Cambridge University Press, 1997, 206-42.

Cubitt, R. P.; Starmer, C. and Sugden, R. "Dynamic Choice and the Common Ratio Effect: An Experimental Investigation." *Economic Journal*, 1998, *108*(450), pp. 1362-80.

Dufwenberg, M., Sundaram, R., Butler, D. "Epiphany in the Game of 21," University of Arizona, 2008, 22.
Fehr, E. and Schmidt, K. M. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 1999, *114*(3), pp. 817-68.

Fischhoff, B.; Slovic, P. and Lichtenstein, S. "Fault Trees - Sensitivity of Estimated Failure Probabilities to Problem Representation." *Journal of Experimental Psychology-Human Perception and Performance*, 1978, *4*(2), pp. 330-44.

Frederick, S. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 2005, *19*(4), pp. 25-42.

Gneezy, U., Rustichini, A., Vostroknutov, A. "I'll Cross That Bridge When I Come to It: Backward Induction as a Cognitive Process," San Diego: University of California at San Diego, 2007, 21.

Hansson, Sven O. "Decision Theory: A Brief Introduction," 2005.
Johnson, E. J.; Camerer, C.; Sen, S. and Rymon, T. "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining." *Journal of Economic Theory*, 2002, *104*(1), pp. 16-47.

Johnson, J. G. and Busemeyer, J. R. "Multiple-Stage Decision-Making: The Effect of Planning Horizon Length on Dynamic Consistency." *Theory and Decision*, 2001, *51*(2-4), pp. 217-46.

Kagel, J. H. and Levin, D. "Common Value Auctions with Insider Information." *Econometrica*, 1999, *67*(5), pp. 1219-38.

Kagel, John H. and Roth, Alvin E. eds. *The Handbook of Experimental Economics*. Princeton: Princeton University Press, 1995.

Kahneman, D. and Tversky, A. "Prospect Theory - Analysis of Decision under Risk." *Econometrica*, 1979, *47*(2), pp. 263-91.

Levitt, S., List, J., and Sadoff, S. "Checkmate: Exploring Backward Induction among Chess Players," *University of Chicago Working Paper.* University of Chicago, 2008, 18.

List, J. A. and Lucking-Reiley, D. "Bidding Behavior and Decision Costs in Field Experiments." *Economic Inquiry*, 2002, *40*(4), pp. 611-19.

McKelvey, R. D. and Palfrey, T. R. "An Experimental-Study of the Centipede Game." *Econometrica*, 1992, *60*(4), pp. 803-36.

Palacios-Huerta, I. and Volij, O. "Field Centipedes." *American Economic Review*, 2009, *99*(4), pp. 1619-35.

Rapoport, A. "Research Paradigms for Studying Dynamic Decision Behavior." *Utility, probability, and human decision making : selected proceedings of an interdisciplinary research conference*, 1975, *11*, pp. 347-69.

Tversky, A. and Kahneman, D. "Judgment under Uncertainty - Heuristics and Biases." *Science*, 1974, *185*(4157), pp. 1124-31.

APPENDIX A: FIGURES AND TABLES

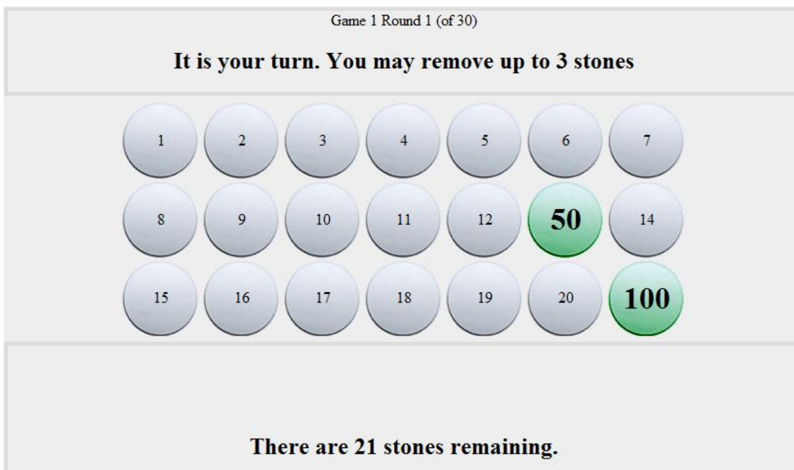


FIGURE 1: GUI SCREENSHOT

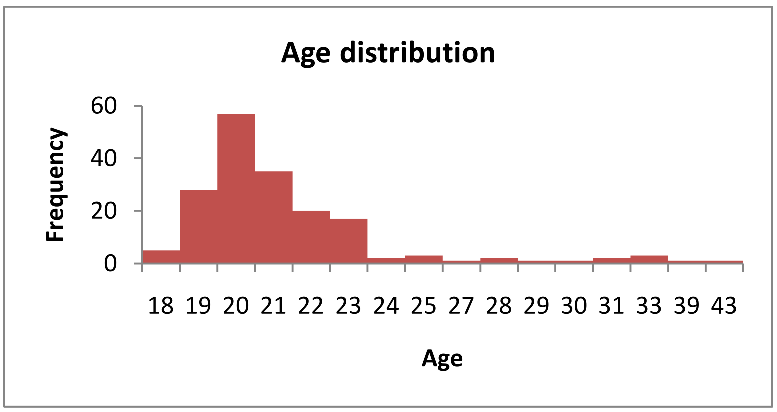


FIGURE 2: AGE DISTRIBUTION

1. A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?

2. If it takes five machines five minutes to make five widgets, how long would it take 100 machines to make 100 widgets?

3. 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 the lake?

FIGURE 3: CRT QUESTIONS

TABLE 1: TREATMENTS

Teaser	Game 1 Action Space (Teaser On/Off Equilibrium)
Control (No Teaser)	1-3
Control (No Teaser)	1-4
5	1-3 (On)
5	1-4 (Off)
6	1-3 (Off)
6	1-4 (On)
11	1-3 (Off)
11	1-4 (On)
13	1-3 (On)
13	1-4 (Off)

TABLE 2: CONDITIONAL PROBABILITIES

	Conditional Probabilities[23]				
Node (1-3) [1-4]	Final (17) [16]	Final-1 (13) [11]	Final-2 (9) [6]	Final-3 (5) [1]	Final-4 (1)
No T (1-3) (n=360)	.96	.88	.81	.58	.46
No T (1-4) [n=630]	.97	.88	.64	.54	

[23] The conditional probabilities at each node are not statistically different from one another depending on action space, except for those at node "Final-2", which are different at the 99% confidence level per two-sample test of proportions.

TABLE 3: PERCENTAGE OF ALL ROUNDS WON

<table>
<tr><th colspan="6">Percentage of all rounds won[24]</th></tr>
<tr><th colspan="3" rowspan="2"></th><th colspan="3">Teaser location within the game
(n)</th></tr>
<tr><th>Early
(1,950)</th><th>Mid
(2,430)</th><th>None
(990)</th></tr>
<tr><td rowspan="6">Teaser position
w.r.t. equilibrium
path
(n)</td><td rowspan="2">On
(2,520)</td><td>[1-3]</td><td>0.45</td><td>0.54</td><td rowspan="4"></td></tr>
<tr><td>[1-4]</td><td>0.41</td><td>0.65</td></tr>
<tr><td rowspan="2">Off
(1,860)</td><td>[1-3]</td><td>0.06</td><td>0.18</td></tr>
<tr><td>[1-4]</td><td>0.42[25]</td><td>0.16</td></tr>
<tr><td rowspan="2">None
(990)</td><td>[1-3]</td><td colspan="2" rowspan="2"></td><td>0.18</td></tr>
<tr><td>[1-4]</td><td>0.30</td></tr>
</table>

[24] The statistical significance of each comparison will be addressed in turn. All comparisons are made using a two-sample test of proportions, unless otherwise noted.

[25] In the (1-4) treatment, it turns out that putting the teaser payment on stone 5 actually serves no purpose but to frustrate the subject. Suppose the subject selects stone 1 to begin on the equilibrium path. Then the computer is programmed to take the teaser on 5, since it's available and the subject has not yet deviated from the equilibrium path. If, instead, the subject takes stones 2, 3, or 4, the computer will skip the teaser on 5 to move to the equilibrium path and take stone 6. Either way, it's impossible for the subject to get the teaser. Although a perfect backward inductor should see this game in the same way as a no-teaser setting, the fact remains that the early-off teaser yields better results than no teaser (42% > 30%). It could be that merely having a teaser present induces more/better backward induction vs. forward-looking behavior.

TABLE 4: ON-EQUILIBRIUM VS. NO TEASER

Teaser	% rounds won	n
on	0.52	2,520
no teas.	0.25	990
p-value	0.000	

TABLE 5: STATA OUTPUT 1.1

Dependent Variable: Win [0,1]	R^2=0.1108 *n*=5,370 P(*win*=1) = 0.359		
Independent Variables:	Marginal Effect	*t* stat	*p*>|t|
on	0.340	17.78	0.000
off	-0.083	-4.25	0.000
on*early	-0.162	-8.98	0.000
off*early	0.085	3.92	0.000
constant	0.255	17.69	0.000

TABLE 6: EARLY VS. MID VS. NO TEASER

Teaser	% rounds won	n	p-values relative to "none"[26]
mid	.39	2,460	0.000
early[27]	.35	1,560	0.000
none	.25	990	

[26] Mid and early are also significantly different from each other.
[27] Excluding (1-4) off-early, for the reasons discussed in an earlier footnote.

TABLE 7: ACTION SPACE

Action Space	% rounds won	n
(1-3)	0.33	2,400
(1-4)	0.38	2,970
p-value	0.000	

TABLE 8: STATA OUTPUT 1.2

Dependent Variable: Win	Pseudo R^2=0.1109 *n*=5,370 Observed P: 0.3594 Predicted P: 0.3385 (at x-bar)			
Independent Variables:	Marginal Effect[28]	*z* stat	p>|z|	x-bar
a) 1to3	-0.1317	-3.91	0.000	0.4469
b) on	0.3283	12.71	0.000	0.4693
c) off	-0.1634	-5.96	0.000	0.3464
d) on*early	-0.2023	-8.40	0.000	0.2291
e) off*early	0.3094	9.33	0.000	0.1341
f) on*1to3	0.0284	0.66	0.511	0.2346
g) off*1to3	0.1895	3.83	0.000	0.1453
h) on*early*1to3	0.1533	3.96	0.000	0.1229
i) off*early*1to3	-0.3367	-9.54	0.000	0.0615

[28] Marginal effect = dF/dx for discrete change of dummy variable from 0 to 1.

TABLE 9: MARGINAL EFFECTS

Action Space Comparison	Teaser Treatment			
	On-Early	On-Mid	Off-Early	Off-Mid
(1-3) vs. No Teaser	*(b+d+f+h)* 0.3077*	*(b+f)* 0.3567*	*(c+e+g+i)* -0.0012	*(c+g)* 0.0261
(1-4) vs. No Teaser	*(b+d)* 0.1260*	*(b)* 0.3283*	*(c+e)* 0.1460*	*(c)* -0.1634*
(1-3) vs. (1-4)	*(a+f+h)* 0.0500	*(a+f)* -0.1033*	*(a+g+i)* -0.2789*	*(a+g)* 0.0578

*Statistically different from 0 at 99% confidence level

TABLE 10: LEARNING-RAW COUNTS

		Teaser					Final								
		1st round taken	2nd round taken	delta_T[29]	$P_{\text{delta<1st}}$	% of rounds	1st round taken	2nd round taken	delta_F[30]	$P_{\text{delta<1st}}$	% of rounds	$P_{\text{1st=No Teaser}}$	delta/2nd T: delta/2nd F (p-value)	1st Final-1st Teaser	$P_{\text{diffT=diffF}}$[31]
On-Early	(1-3)	1.27	2.45	1.18	0.246	0.91	6.91	11.80	4.89	0.140	0.45	0.000	0.001	5.64	0.002
	(1-4)	2.32	3.53	1.21	0.058	0.84	8.58	12.26	3.68	0.019	0.41	0.000	0.048	6.26	0.001
On-Mid	(1-3)	6.90	8.35	1.45	0.182	0.54	7.05	7.94	0.89	0.202	0.54	0.720		0.15	
		6.90	*11.75*	*4.85*	*0.098*		*7.05*	*12.55*	*5.50*	*0.209*		0.000	*0.096*	*0.15*	*0.468*
	(1-4)	4.81	7.62	2.81	0.028	0.68	5.10	8.62	3.52	0.040	0.65	0.000		0.29	
		6.48	*9.52*	*3.04*	*0.026*		*7.61*	*10.30*	*2.69*	*0.014*		0.000	*0.050*	*1.132*	*0.040*
Off-Early	(1-3)	1.27	2.73	1.46	0.253	0.65	14.00	19.50	5.50	0.155	0.06	0.085		12.73	
							27.91	*29.73*	*1.82*	*0.000*		0.234	*0.001*	*27.91*	*0.000*
	(1-4)	---	---	---	---	---	5.90	10.78	4.88	0.378	0.42	0.021		---	
							11.69	*17.23*	*5.54*	*0.096*		0.053			
Off-Mid	(1-3)	6.82	9.10	2.28	0.017	0.25	17.20	18.75	1.55	0.001	0.19	0.009		10.38	
		13.27	*16.67*	*3.40*	*0.010*		*21.80*	*24.80*	*3.00*	*0.000*		0.833	*0.085*	*8.533*	*0.009*
	(1-4)	1.52	3.35	1.83	0.087	0.46	9.63	9.60	-0.03	0.381	0.16	0.419		8.11	
							23.30	*26.57*	*3.27*	*0.000*		0.272	*0.067*	*23.3*	*0.001*
No Teaser	(1-3)						6.00	10.33	4.33	0.356	0.18			6	
							22.67	*26.50*	*3.83*	*0.003*				*22.67*	*0.000*
	(1-4)						10.17	13.91	3.74	0.005	0.30			10.17	
							19.57	22.48	*2.91*	*0.000*				*19.57*	*0.000*

$P_{\text{delta<1st}}$: 1-tail, paired t-test
$P_{\text{1st=NoTeaser}}$: 2-tail, unpaired t-test
**Unconditional* = for any treatment in which a subject never got it, we conservatively assume they would have gotten it in round 31.

[29] In a simple paired t-test for means, we test the null hypothesis that delta=1. We reject the null in all treatments.
[30] In a simple paired t-test for means, we test the null hypothesis that delta=1. We reject the null in all treatments.
[31] In a simple paired t-test for means, we test the null hypothesis that (1st Final-1st Teaser)=0. We reject the null hypothesis in all treatments except on-mid (1-3).

TABLE 11: FINAL STONE-1ST "REDEFINED" WIN

Final stone-1st Redefined "win"									
On-Early		On-Mid		Off-Early		Off-Mid		No Teaser	
(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)
1	1	3	1	14	1	11	4	3	6
1	1	3	1	29	1	16	5	10	7
1	1	3	2	31	1	19	8	16	9
1	8	5	3	31	14	20	11	31	10
2	8	5	3	31	16	22	23	31	14
5	14	6	4	31	16	23	31	31	15
9	19	6	5	31	18	23	31	31	17
13	22	6	6	31	26	31	31	31	17
14	22	8	6	31	31	31	31	31	18
21	27	9	7	31	31	31	31	31	23
25	27	12	8	31	31	31	31	31	31
25	31	18	8	-	31	31	31	31	31
27	31	18	8	-	31	31	31	-	31
28	31	22	11	-	-	31	31	-	31
31	31	31	12	-	-	31	31	-	31
31	31	31	16	-	-	-	31	-	31
31	31	31	19	-	-	-	31	-	31
31	31	31	21	-	-	-	31	-	31
31	31	31	23	-	-	-	31	-	31
31	-	31	23	-	-	-	31	-	31
31	-	-	31	-	-	-	31	-	31
31	-	-	31	-	-	-	31	-	-
-	-	-	31	-	-	-	31	-	-

TABLE 12: REGRESSION RESULTS-UNCONDITIONAL

Regression results: Unconditional (31 for all unobserved subjects)[32] (*significantly different from 1)						
	y	x	β	SE	R^2	n
(1-3)	on-mid	noT	0.2676*	0.0392	0.8088	12
	on-early	noT	0.3987*	0.0862	0.6604	12
	off-mid	noT	0.8908	0.0578	0.9557	12
	off-early	noT	1.0575*	0.0919	0.9298	11
(1-4)	on-mid	noT	0.4885*	0.0522	0.8141	21
	on-early	noT	0.9903	0.0353	0.9776	19
	off-mid	noT	1.0973	0.0579	0.9472	21
	off-early	noT	1.1054	0.8122	0.9378	13

[32] Boxes surrounding treatments in this and all subsequent tables indicate indifference.

TABLE 13: REGRESSION RESULTS-CONDITIONAL

Regression results: Conditional (truncated at coinciding 31s) (*significantly different from 1)						
	y	x	β	SE	R^2	n
(1-3)	on-mid	noT	0.2676*	0.0392	0.8088	12
	on-early	noT	0.3987*	0.0862	0.6604	12
	off-mid	noT	0.8097*	0.0834	0.9309	8
	off-early	noT	1.3492	0.3703	0.8157	4
(1-4)	on-mid	noT	0.4885*	0.0522	0.8141	21
	on-early	noT	0.9743	0.0735	0.9410	12
	off-early	noT	1.1712	0.1097	0.9193	11
	off-mid	noT	1.4011*	0.1244	0.9269	11

TABLE 14: FINAL STONE-TOP 20%

	Final stone-1st round learned-Top 20%									
	On-Early		On-Mid		Off-Early		Off-Mid		No Teaser	
	(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)	(1-3)	(1-4)
	1	1	3	1	14	1	11	4	3	6
	1	1	3	1	29	1	16	5	10	7
	1	1	3	2	-	1	19	8	-	9
	1	8	5	3	-	-	-	11	-	10
	-	-	-	3	-	-	-	23	-	-
Avg	1	2.8	3.5	2	21.5	1	15.3	10.2	6.5	8

TABLE 15: SUMMARY RANKINGS

Summary Rankings					
Action	Rank	Conditional regression	Unconditional regression	Fastest 20%	Unconditional average 1st round (Table 1)
(1-3)	1	on-mid	on-mid	on early	on-mid
	2	on-early	on-early	on-mid	on-early
	3	off-mid	off-mid	no-teaser	off-mid
	4	no-teaser	no-teaser	off-mid	no-teaser
	5	off-early	off-early	off-early	off-early
(1-4)	1	on-mid	on-mid	off-early	on-mid
	2	on-early	on-early	on-mid	on-early
	3	no-teaser	no-teaser	on-early	off-early
	4	off-early	off-mid	no-teaser	no-teaser
	5	off-mid	off-early	off-mid	off-mid

TABLE 16: CRT-CORRECT ANSWERS

CRT: correct answers		
	Frequency	% of subjects
Lake	97	0.54
Ball	82	0.46
Widget	59	0.33
Ball+Lake	60	0.34
Lake+Widget	50	0.28
Ball+Widget	43	0.24
All 3 correct	38	0.21
Zero correct	56	0.31

TABLE 17: R^2-GOODNESS OF FIT

	R^2: Goodness-of-Fit			
	f(Treatment)	*f*(Treatment, demographics)	*f*(Treatment, cognitive battery)	*f*(Treatment, demographics, cognitive battery)
Probit (Win)	0.1407	0.2256	0.3204	0.3606
OLS (Total points)	0.2307	0.3328	0.4692	0.5122

TABLE 18: MARGINAL EFFECTS PROBIT

Marginal Effects Probit (Win=0,1)				
Independent Variable	*f*(Treatment)	*f*(Treatment, demographics)	*f*(Treatment, cognitive battery)	*f*(Treatment, demographics, cognitive battery)
(a) 1to3	-0.2414 (0.200)	-0.3280* (0.093)	-0.0338 (0.876)	-0.0983 (0.659)
(b) on	0.4444*** (0.006)	0.3752** (0.031)	0.5835*** (0.001)	0.5369*** (0.005)
(c) off	-0.2813* (0.072)	-0.4302*** (0.009)	-0.3629** (0.035)	-0.4609** (0.011)
(d) on*early	-0.3502** (0.035)	-0.3239* (0.071)	-0.3563* (0.068)	-0.3511* (0.089)
(e) off*early	0.3727** (0.019)	0.4350*** (0.008)	0.4964*** (0.002)	0.5093*** (0.002)
(f) on*1to3	0.0059 (0.982)	0.0531 (0.847)	-0.3186 (0.274)	-0.2846 (0.353)
(g) off*1to3	0.4336** (0.043)	0.5096** (0.014)	0.3832 (0.117)	0.4382* (0.066)
(h) on*early*1to3	0.2372 (0.289)	0.2320 (0.327)	0.1695 (0.505)	0.1905 (0.466)
(i) off*early*1to3	-0.5271*** (0.008)	-0.5663*** (0.002)	-0.5573** (0.010)	-0.5873*** (0.005)
gender	-	-0.3676*** (0.000)	-	-0.2486** (0.011)
age	-	0.0038 (0.748)	-	0.0026 (0.840)
prev. experiments	-	0.1355 (0.191)	-	0.0865 (0.450)
econ courses	-	-0.0389 (0.120)	-	-0.0550** (0.045)
ball	-	-	0.2497** (0.010)	0.2454** (0.017)
lake	-	-	0.4231*** (0.000)	0.4007*** (0.000)
widget	-	-	0.0822 (0.450)	0.0495 (0.666)

TABLE 18A

Action Space Comparison	Teaser Treatment			
	On-Early	On-Mid	Off-Early	Off-Mid
(1-3) vs. No Teaser	*(b+d+f+h)* 0.3373**	*(b+f)* 0.4503***	*(c+e+g+i)* -0.0021	*(c+g)* 0.1523
(1-4) vs. No Teaser	*(b+d)* 0.0942	*(b)* 0.4444***	*(c+e)* 0.0914	*(c)* -0.2813*

TABLE 19: OLS-TOTAL POINTS

OLS: Total points				
Independent Variable	*f*(Treatment)	*f*(Treatment, demographics)	*f*(Treatment, cognitive battery)	*f*(Treatment, demographics, cognitive battery)
(a) 1to3	-604.7619 (0.283)	-905.7305* (0.095)	161.1973 (0.738)	-130.2171 (0.786)
(b) on	2236.5420*** (0.000)	1875.2630*** (0.000)	2334.5440*** (0.000)	2125.8470*** (0.000)
(c) off	688.7164 (0.143)	313.8169 (0.490)	720.8602* (0.069)	534.8416 (0.177)
(d) on*early	-575.5149 (0.233)	-411.0410 (0.370)	-222.3133 (0.593)	-195.2261 (0.631)
(e) off*early	502.6756 (0.352)	593.9525 (0.248)	773.7315* (0.092)	741.7474* (0.099)
(f) on*1to3	100.9576 (0.891)	334.1922 (0.634)	-776.1969 (0.219)	-541.6589 (0.385)
(g) off*1to3	914.6170 (0.232)	1317.3930* (0.073)	300.1094 (0.642)	606.6432 (0.344)
(h) on*early*1to3	674.3785 (0.322)	625.0268 (0.333)	226.4560 (0.693)	303.7275 (0.587)
(i) off*early*1to3	-869.6453 (0.289)	-1149.7260 (0.143)	-549.4719 (0.434)	-721.5144 (0.297)
gender	-	-1053.2750*** (0.000)	-	-526.4939** (0.013)
age	-	42.4044 (0.212)	-	36.1500 (0.220)
prev. experiments	-	315.6775 (0.259)	-	55.6498 (0.819)
econ courses	-	-149.1424** (0.022)	-	-170.6020*** (0.003)
ball	-	-	711.0321*** (0.002)	707.9256*** (0.002)
lake	-	-	1215.9400*** (0.000)	1154.219*** (0.000)
widget	-	-	311.8999 (0.205)	188.2037 (0.435)
(j) constant	1604.762*** (0.000)	1360.9760* (0.090)	331.0680 (0.304)	208.6719 (0.767)

TABLE 19A

Action Space Comparison	Teaser Treatment			
	On-Early	On-Mid	Off-Early	Off-Mid
(1-3) vs. No Teaser	*(b+d+f+h+j)* 4041.126	*(b+f+j)* 3942.260	*(c+e+g+i+j)* 2841.126	*(c+g+j)* 3208.095
(1-4) vs. No Teaser	*(b+d+j)* 3265.789	*(b+j)* 3841.304	*(c+e+j)* 2796.154	*(c+j)* 2293.478

APPENDIX B: LABORATORY SCRIPT AND SCREENSHOTS

Laboratory script:

Welcome to the UT Experimental Economics Laboratory. My name is Kelly Padden, and joining me today is _________. We are researchers from the Department of Economics. We understand that many of you have busy schedules and really appreciate your willingness to participate.

Before we begin I need to go over a few lab rules. Once the experiment begins, please refrain from communicating with each other (talking, texting) and please do not open or play any games on the computer (no solitaire or minesweeper).

In this study, you will be asked to make a series of market-like decisions. Your earnings in this experiment are based on the decisions you make. The money you will be paid with comes from a research grant, and this money can only be used to pay experiment participants. You will be paid in cash after the experiment is completed.

The decision-making setting may be unfamiliar to you. This is common. Therefore, in writing the instructions for this experiment, we have done our very best to clearly describe to you all relevant information from which to base your decisions.

There are two important protocols in experimental economics that we would like you to be aware of. First, the instructions contain only true information. There are no hidden tasks, and the experiment works exactly as stated in the instructions. Second, your decisions are confidential. What this means is that you have been randomly assigned an ID number. All decisions you make will be associated with this ID number and not your name. Therefore, when we analyze the data and present results, your name will in no way be affiliated with this study.

We have provided everyone with a pencil, calculator, and paper. Use these items, if you wish, as you make your decisions. But please do not write on the instructions.

Has everyone had a chance to read the informed consent sheet? Is everyone comfortable with the risks involved with participation in this experiment? If you would, please raise your hand to indicate you have read the Informed Consent Sheet and you agree to participate in the experiment.

Today, you will play two experimental games and then answer a short questionnaire. We will proceed by reading the instructions for the first game. I will read the instructions aloud and ask that you follow along on your copy.

Let's begin…<read instructions for Race to 21 game>

After you are finished with the final round of the first game, please DO NOT proceed to the second game. We will go through the instructions to the second game as a group before proceeding.

If any question should arise during the experiment, please raise your hand and one of us will address your question privately. Good luck, and we hope you earn lots of money!

<u>Race to 21 game instructions</u>:

Welcome

Thank you for participating in this experiment. You will be playing games against a computer for money. You will be paid for your participation in cash, immediately at the end of the experiment. How much you earn depends on your decisions. A research foundation has contributed the money for this study.

It is very important that you read all instructions carefully and that you strictly follow the rules of this experiment. If you disobey the rules, you will be asked to leave the experiment. You should never use the browser's forward or back button. All navigation through this experiment should be done by hitting the "proceed" button on each screen.

Your id number for this experiment is ________.

Do not press the PROCEED button until instructed to do so.

Rules of the Game

You will be playing a simple game against a computer opponent. You will see 21 stones, numbered 1 through 21. You will move first. On each of your turns, you may remove between 1 and <3, 4> consecutive stones beginning with the lowest-numbered remaining stone. You will remove stones by mousing over the stones you want to remove and clicking on the last stone you want to remove. After your turn, the computer will remove between 1 and <3, 4> consecutive stones. The stones removed by the computer will flash briefly, and then it will be your turn again. You and the computer alternate moves until all of the stones have been removed.

Profit

Most of the stones in the game are colored gray but one or more may be colored green. You make money when you, and not the computer, click on a green stone. Green stones have boldface numbers equal to the number of points that you earn for removing that stone. At the end of the experiment, points will be converted into dollars at the rate of $1.00 for every 150 points. If you remove a green stone in passing but without clicking on it, you do NOT earn points on that turn.

Practice

On your screen is a panel to help you practice removing stones. In the practice panel, there are only 8 stones. The last stone is a green stone and has a value of 100 points. You may remove between 1 and <3, 4> on each turn. Take this opportunity to try removing different numbers of stones. Notice that moving your mouse over a stone highlights it and the remaining stones with lower values, but that if you try to take more than <3, 4> stones, nothing happens. If you would like more tries, hit the RESET button and continue the practice round.

To sum up, you will play a game against a computer opponent in which you and the computer will alternate removing between 1 and 4 stones until all 8 stones have been removed. If, on your turn, you click on a green stone, you will earn the amount written on that stone.

You will play 30 rounds of this game. Then, you will play 15 rounds of a different game.

Do not press the "Proceed" button until instructed to do so.

Game 2

You have finished the first game. You will now play 15 rounds of game 2.

In game 2, there will be a total of 21 stones. You and the computer will alternate removing between 1 and <3, 4> stones on each turn. You will go first.

Practice

If you wish to practice removing stones, the practice panel on your screen contains 8 stones. You may remove between 1 and <3, 4> on each turn. If you would like more tries, hit the RESET button and continue the practice round.

Do not press the "Proceed" button until instructed to do so.

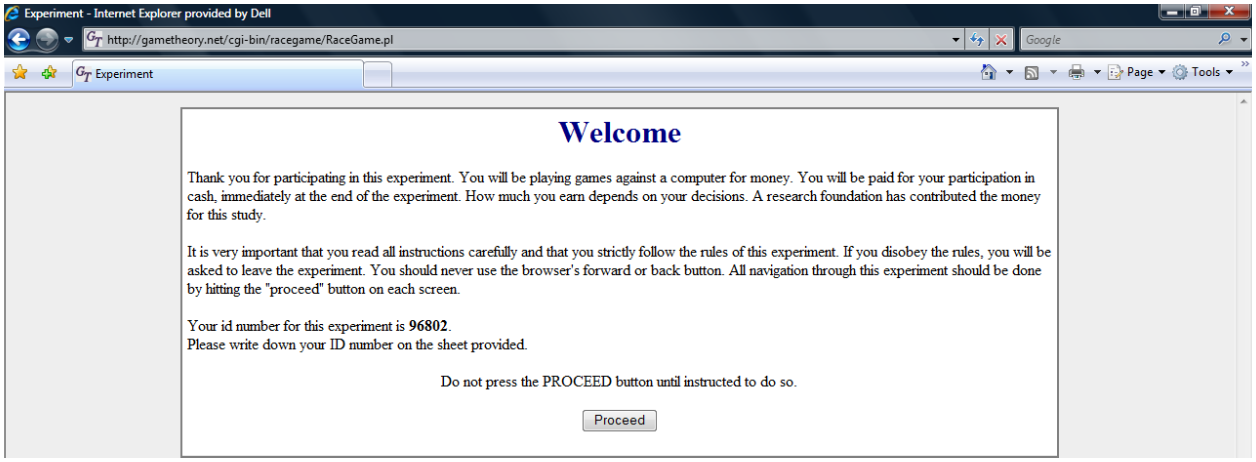
Survey

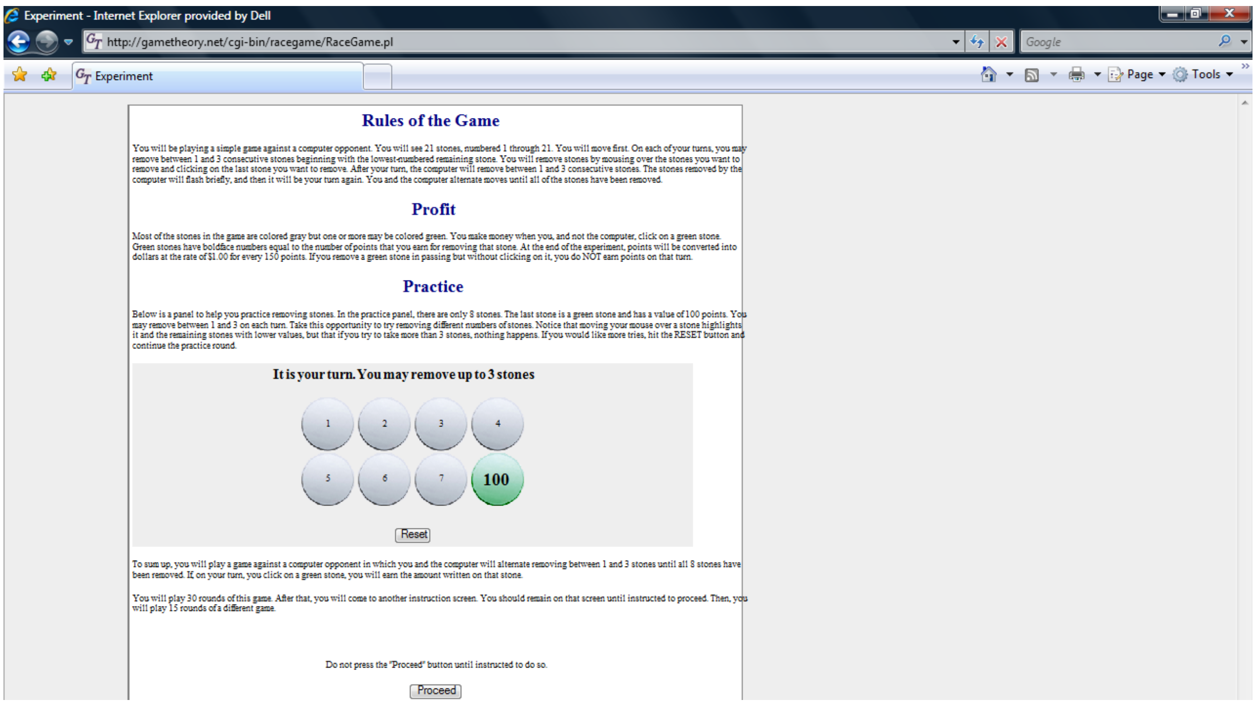
Once you have finished the second game, you will complete a brief survey.

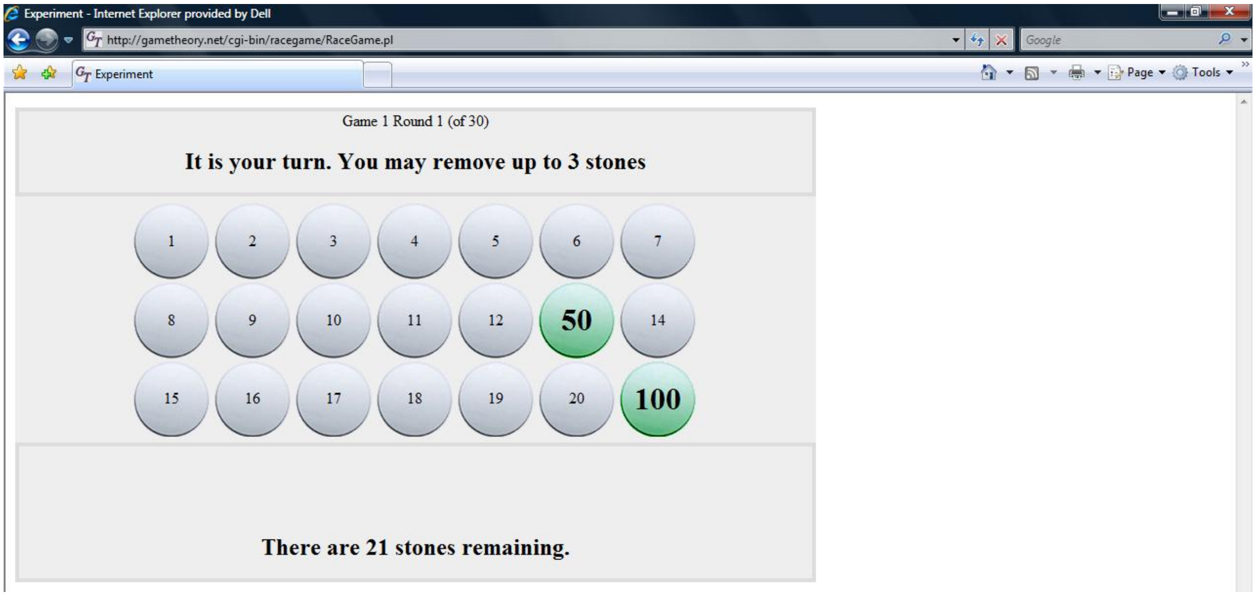
These questions will be used for statistical purposes only.

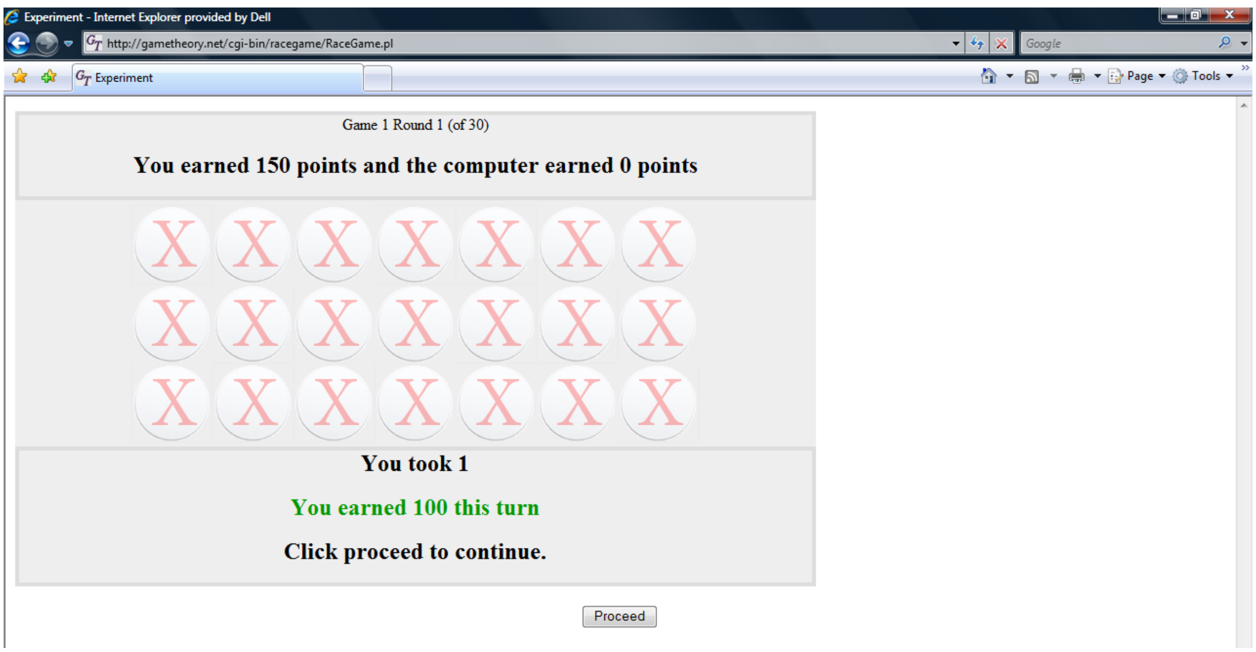
THIS INFORMATION WILL BE KEPT STRICTLY CONFIDENTIAL and WILL BE DESTROYED UPON COMPLETION OF THE STUDY.

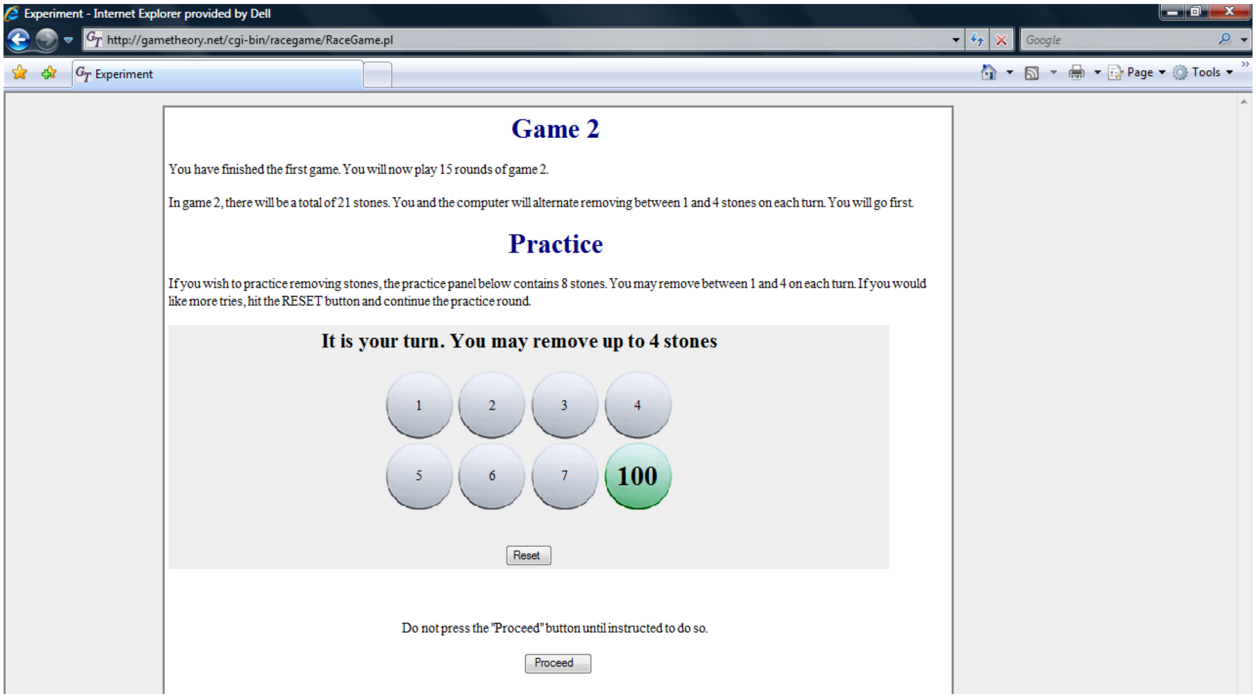
<u>Screen shots of the Race to 21 game interface:</u>

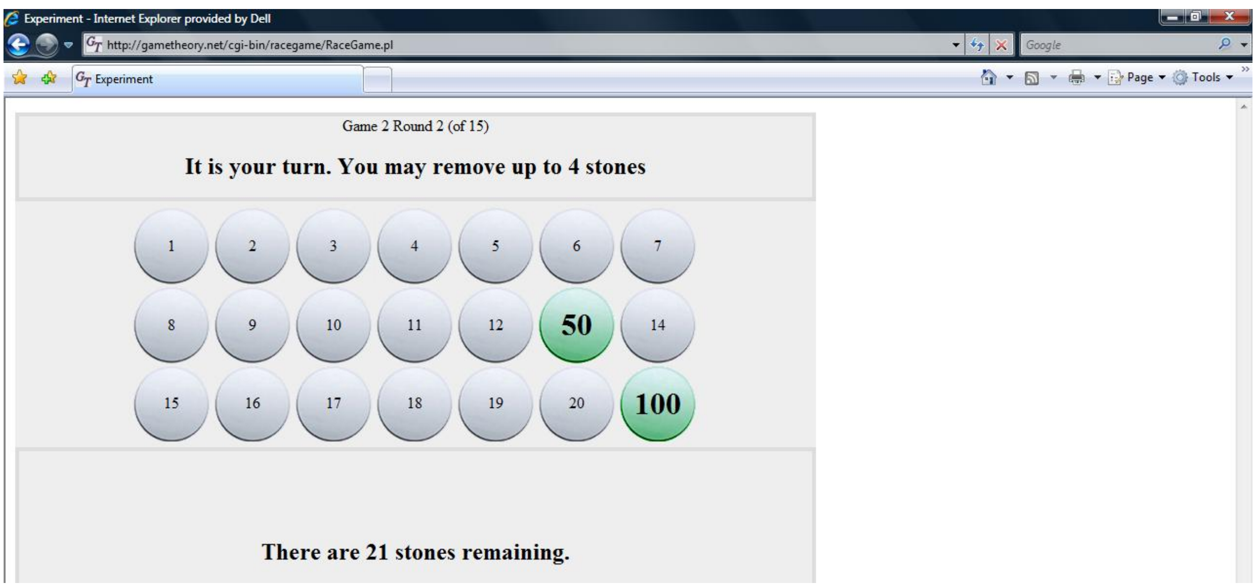


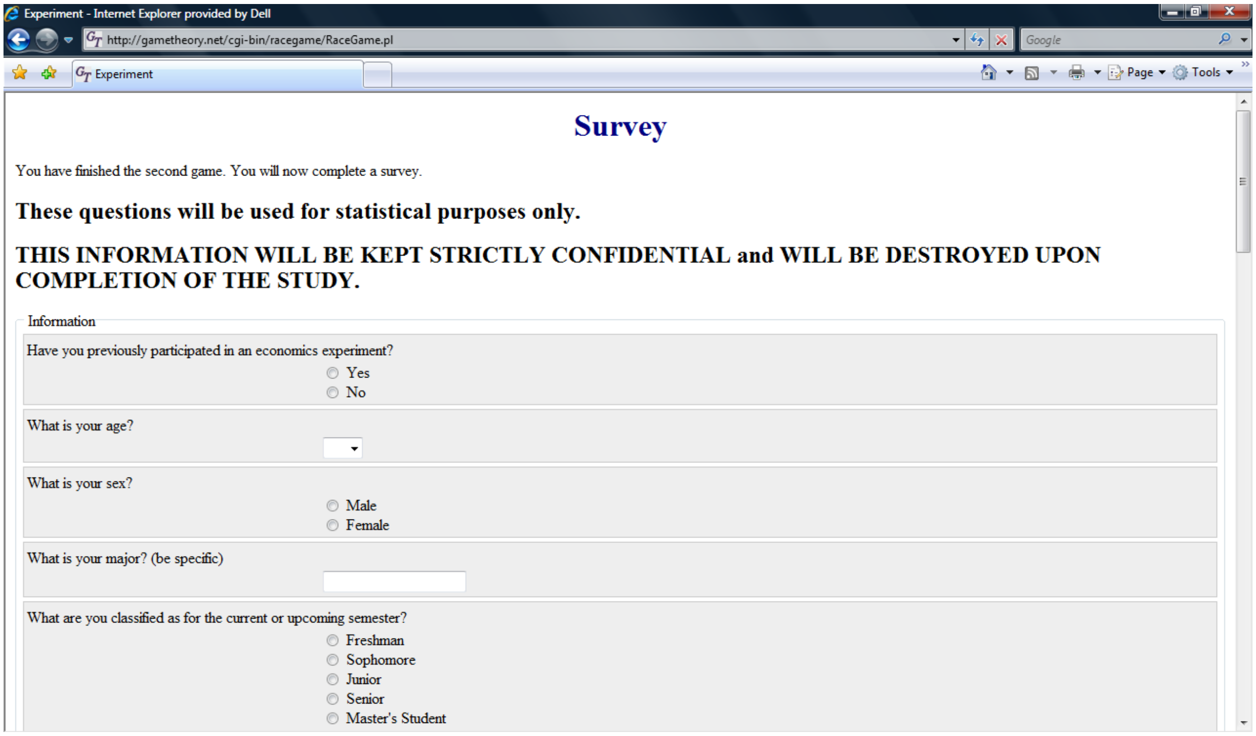



Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Experiment
Game 1 Round 1 (of 30)
It is your turn. You may remove up to 3 stones
1 2 3 4 5 6 7
8 9 10 11 12 50 14
15 16 17 18 19 20 100
There are 21 stones remaining.


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Experiment
Game 1 Round 1 (of 30)
You earned 150 points and the computer earned 0 points
You took 1
You earned 100 this turn
Click proceed to continue.
Proceed


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Google
Experiment
Page
Tools
Game 2
You have finished the first game. You will now play 15 rounds of game 2.
In game 2, there will be a total of 21 stones. You and the computer will alternate removing between 1 and 4 stones on each turn. You will go first.
Practice
If you wish to practice removing stones, the practice panel below contains 8 stones. You may remove between 1 and 4 on each turn. If you would like more tries, hit the RESET button and continue the practice round.
It is your turn. You may remove up to 4 stones
1
2
3
4
5
6
7
100
Reset
Do not press the "Proceed" button until instructed to do so.
Proceed


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Google
Experiment
Page
Tools
Game 2 Round 2 (of 15)
It is your turn. You may remove up to 4 stones
1
2
3
4
5
6
7
8
9
10
11
12
50
14
15
16
17
18
19
20
100
There are 21 stones remaining.


Experiment - Internet Explorer provided by Dell

http://gametheory.net/cgi-bin/racegame/RaceGame.pl

Survey

You have finished the second game. You will now complete a survey.

These questions will be used for statistical purposes only.

THIS INFORMATION WILL BE KEPT STRICTLY CONFIDENTIAL and WILL BE DESTROYED UPON COMPLETION OF THE STUDY.

Information

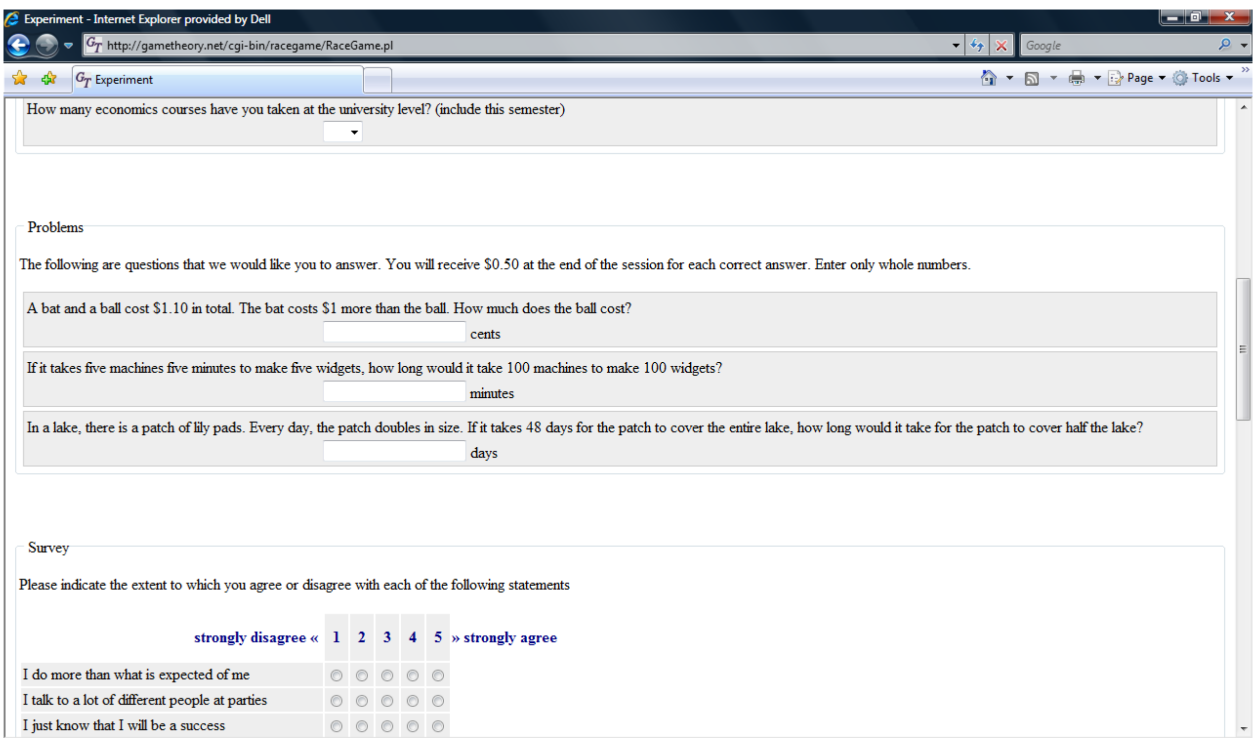
Have you previously participated in an economics experiment?
- Yes
- No

What is your age?

What is your sex?
- Male
- Female

What is your major? (be specific)

What are you classified as for the current or upcoming semester?
- Freshman
- Sophomore
- Junior
- Senior
- Master's Student


Experiment - Internet Explorer provided by Dell

http://gametheory.net/cgi-bin/racegame/RaceGame.pl

How many economics courses have you taken at the university level? (include this semester)

Problems

The following are questions that we would like you to answer. You will receive $0.50 at the end of the session for each correct answer. Enter only whole numbers.

A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?
cents

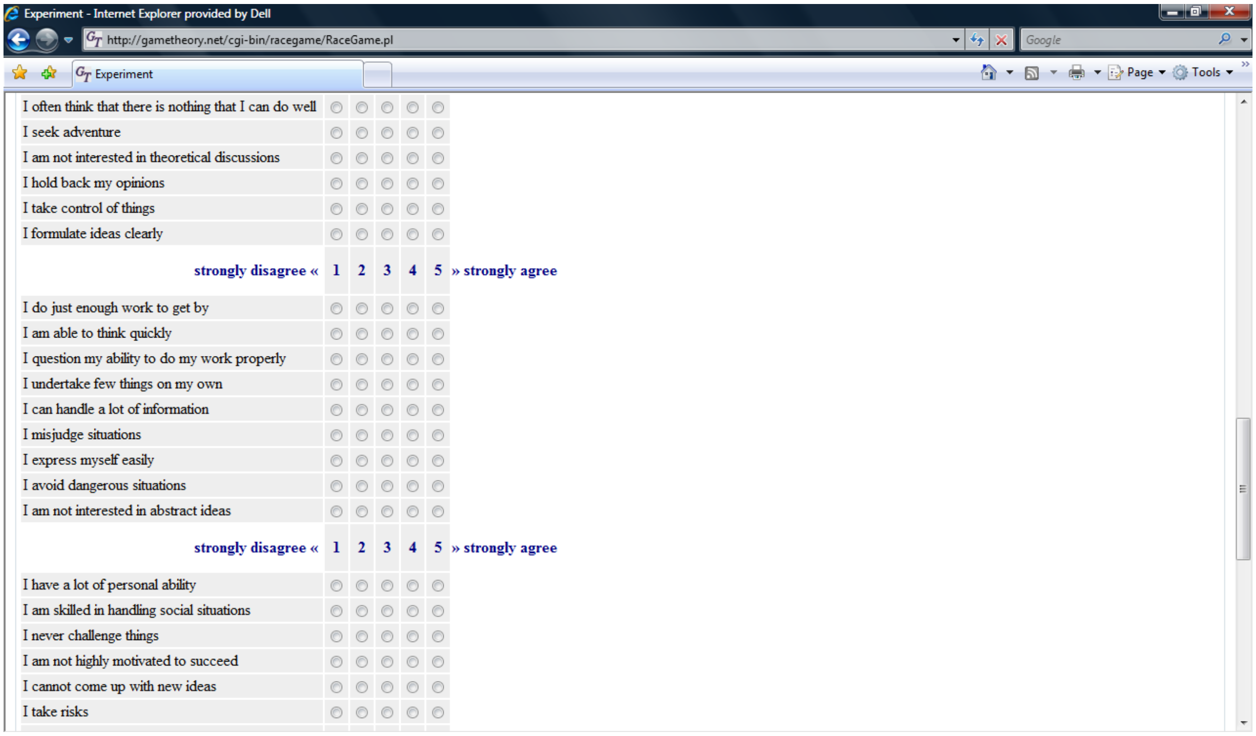
If it takes five machines five minutes to make five widgets, how long would it take 100 machines to make 100 widgets?
minutes

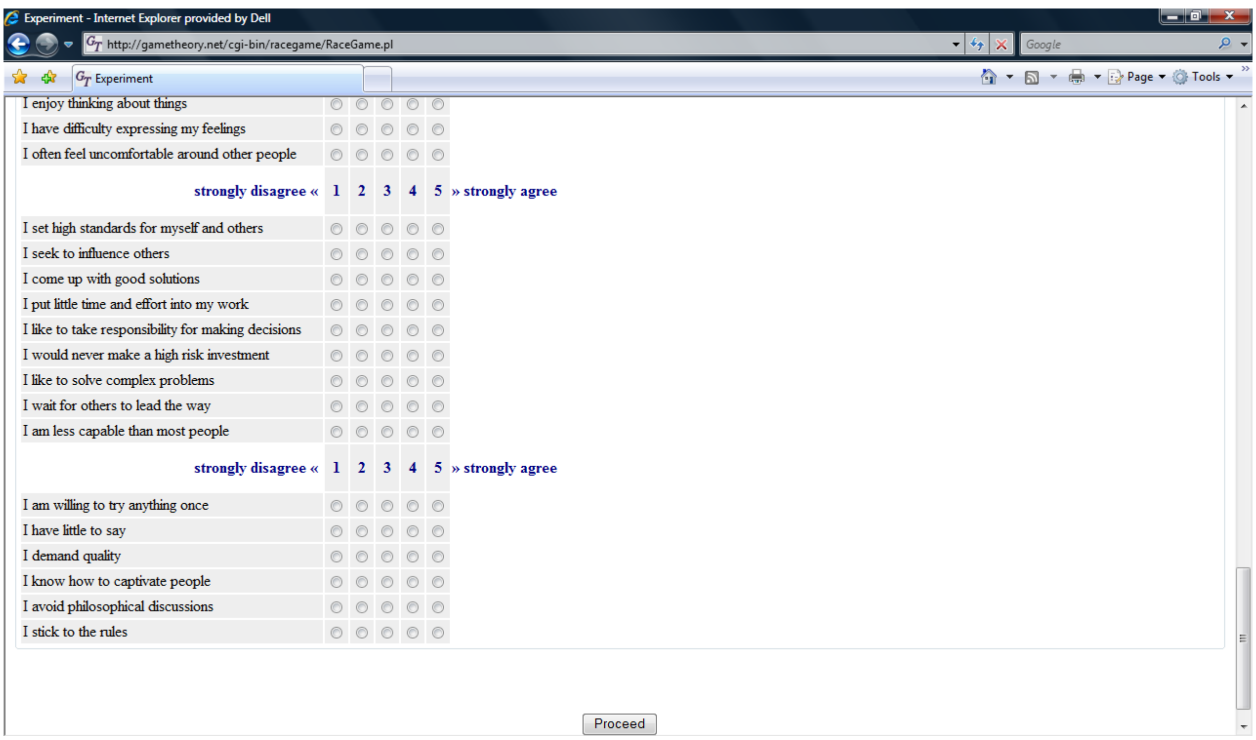
In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake?
days

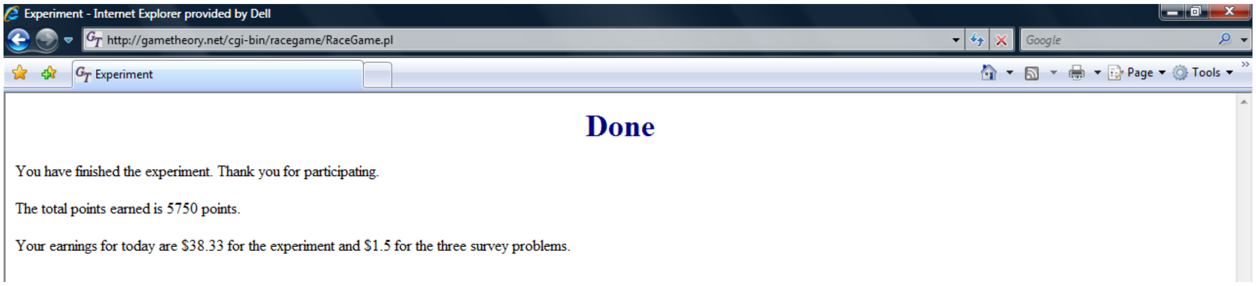
Survey

Please indicate the extent to which you agree or disagree with each of the following statements

strongly disagree «	1	2	3	4	5	» strongly agree
I do more than what is expected of me	○	○	○	○	○	
I talk to a lot of different people at parties	○	○	○	○	○	
I just know that I will be a success	○	○	○	○	○	


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Google
Experiment
Page
Tools
I often think that there is nothing that I can do well
I seek adventure
I am not interested in theoretical discussions
I hold back my opinions
I take control of things
I formulate ideas clearly
strongly disagree « 1 2 3 4 5 » strongly agree
I do just enough work to get by
I am able to think quickly
I question my ability to do my work properly
I undertake few things on my own
I can handle a lot of information
I misjudge situations
I express myself easily
I avoid dangerous situations
I am not interested in abstract ideas
strongly disagree « 1 2 3 4 5 » strongly agree
I have a lot of personal ability
I am skilled in handling social situations
I never challenge things
I am not highly motivated to succeed
I cannot come up with new ideas
I take risks


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Google
Experiment
Page
Tools
I enjoy thinking about things
I have difficulty expressing my feelings
I often feel uncomfortable around other people
strongly disagree « 1 2 3 4 5 » strongly agree
I set high standards for myself and others
I seek to influence others
I come up with good solutions
I put little time and effort into my work
I like to take responsibility for making decisions
I would never make a high risk investment
I like to solve complex problems
I wait for others to lead the way
I am less capable than most people
strongly disagree « 1 2 3 4 5 » strongly agree
I am willing to try anything once
I have little to say
I demand quality
I know how to captivate people
I avoid philosophical discussions
I stick to the rules
Proceed


Experiment - Internet Explorer provided by Dell
http://gametheory.net/cgi-bin/racegame/RaceGame.pl
Google
Experiment
Page
Tools
Done
You have finished the experiment. Thank you for participating.
The total points earned is 5750 points.
Your earnings for today are $38.33 for the experiment and $1.5 for the three survey problems.

VITA

Kelly Padden Hall graduated from Stephen F. Austin High School, Austin, Texas, in 1996. She earned her Bachelor of Science degree from the University of Texas at Austin in 2000 and Master of Business Administration from Texas State University-San Marcos in 2004. Upon earning her MBA, she received her commission as an officer in the United States Air Force. In her first assignment, she served as the Deputy Financial Services Officer for the 11th Wing at Bolling Air Force Base, District of Columbia. Next, she served a ten-month tour as a budget analyst for the Assistant Secretary of the Air Force, Financial Management and Comptroller, at the Pentagon in Arlington, Virginia. Her program portfolio included all Headquarters Air Force organizations and Global War on Terrorism supplemental appropriations.

Kelly entered the University of Tennessee, Knoxville, under the aegis of the Air Force Institute of Technology civilian institutions program. Upon graduation, she will join the Air University faculty at Maxwell Air Force Base in Montgomery, Alabama. Her duties as an instructor at the Defense Financial Management and Comptroller School will include teaching graduate-level economics to senior military and civilian comptrollers and resource managers from across all services within the Department of Defense.