Title
The Reality of Backward Induction

Author
Andrews, Ryan

Publication Date
2019

Undergraduate

The Reality of Backward Induction

Advisor: Ignacio Esponda

March 18, 2019

Abstract

There is a significant breadth of work focused on identifying the utilization of backward induction in game theoretical scenarios. Researchers have generally found that people do not use backward induction, rather opting for alternative heuristics. This paper aims at discerning the true causal factors behind these findings using the well-known Stick Game. Results suggest that believing that a rational solution exists has as much an effect on the success of a game as does the innate ability of an individual. The data is interpreted through the use of linear regression and difference-in-differences models.

Introduction

Backward induction (hereafter referred to as BI) is the process of reasoning backward in time. To solve a problem through BI is to start at the end of a problem and work your way back logically until you have strategy for every decision that you may come across. BI allows you to reason what the optimal decision is at every stage of a game and ultimately outsmart your opponent. Optimality refers to the payoffs that players receive and is determined by the criteria that is set. In the context of our experiment, an optimal decision refers to that which strictly benefits the player making the decision. If for example, you were deciding between two jobs, BI would have you look into the future and determine which job would make you happiest and then work your way back to the decision at hand. In game theory, BI can be used to find Subgame Perfect Nash Equilibrium[1]. In perfectly controlled scenarios, with a winner and a loser, the most effective and logical way to go about winning a game would be to employ BI. In lab settings however, this is not what researchers have observed.

Studying BI is important because of the insights that it can provide into different areas such as business management, global warfare, and advertising. There has been a significant breadth of work conducted on BI and whether participants exhibit the use of this heuristic[2] in game play. Across numerous studies, researchers have all found strikingly similar results. People do not backward induct. Whether prompted or otherwise, people generally strive to solve a problem step by step, rather than to stop and devise a strategy to tackle the problem as a whole. Steven D. Levitt et al. (2011) note that there are numerous possibilities why players do not exhibit the tendency to backward induct in game theoretical scenarios. Players may be risk averse to the loss of a potential payoff. They may have preferences for fairness, altruism, or cooperation. It is possible that there are so many players in the population who either make BI errors or choose not to backward induct, that not utilizing BI suddenly becomes rational.

The question that I attempt to answer is, why do people generally not exhibit the tendency to backward induct? There is a seeming lack of research delineating between the various factors that lead to this observation. Are the experiments fundamentally flawed in their design or delivery? I hypothesize that it is a culmination of factors, ranging from an inability to backward induct, to the lack of a belief in the existence of a rational solution.

My experiment is designed to test not only whether people backward induct, but also why they choose to employ BI or not. I make use of the Stick Game in my experiment. The Stick Game is an extensive-form game designed to have a winner-takes-all outcome. It's a turn-based 2 player game (player and computer). The game starts with a certain number of sticks and each player can take either 1 or 2 sticks during their turn. The goal of the game is to take the last stick. An example of how the game could play out is as follows. There are 5 sticks. Player 1 decides to take 1 stick. There are now 4 sticks left. Player 2 decides to take 1 stick. There are 3 sticks left. Player 1 decides to take 2 sticks. There is 1 stick left. Player two has now won the game regardless of whether he takes 1 or 2 sticks, because he will take the last stick inevitably.

This game can be extended outward into more complex versions by simply adding more sticks to the initial count. In this way, I can measure the percentage of people that backward induct

---

[1] A strategy profile is a subgame perfect equilibrium if it represents a Nash equilibrium of every subgame of the original game. Informally, this means that if the players played any smaller game that consisted of only one part of the larger game, their behavior would represent a Nash equilibrium of that smaller game.

[2] A heuristic enables a person to discover or learn something for themselves.

as the games become more complex. The 5 Stick and subsequent 10,13, and 16 Sticks Games that I have my participants play have a unique universal solution through the use of BI. If a participant beats multiple versions of the Stick Game, I can conclude with relative certainty that the individual has employed BI.

I provide half of the participants with an understanding of a simple version of the Stick Game, whereas I omit this information from the other half. I ensure that both groups read and understand their instructions through the use of comprehension questions following the treatment and control questions. The comprehension question is shown in detail in Appendix A. This way, I can delineate between the percentage of people who don't employ BI because they are incapable and those that don't because they don't think a logical solution exists. Once the treatment group is made aware of the solution and has the understanding that it can be extended to more complex versions of the Stick Game, their failure to beat the Stick Game can no longer be attributed to their lack of a belief in the existence of a logical solution. We have the participant play against a computer that is programmed to move optimally to win in every scenario. This leaves no room for cooperation on behalf of the participant.

I find that people are far more capable than researchers have traditionally come to believe. When presented with the certainty that a purely logical solution exists, the rate with which the participants beat the game rose dramatically. This effect wore off over time as the games grew more complex, indicating that in a certain percent of the cases, increased results were due to alternative heuristics rather than BI. Dufwenberg et al. (2010) refer to this use of mimicking/lack of understanding of the underlying mechanics of the game as a mini-epiphany, whereas the full understanding is referred to as complete epiphany.

The most significant limitation of my study was the number of data points that I could collect due to budget and time. I was able to obtain 64 participants' observations in the study and therefore had to limit the scope of my research to what I believed to be the two most significant causes behind BI failure. This is because if we utilize sample sizes smaller than 30, the chance with which we reject the null hypothesis grows increasingly slim.

My research is split up into 5 sections. Section 2 delves into the background of the research, section 3 deals with the design of the study, section 4 presents the empirical models and results, and section 5 concludes.

Literature Review

BI is a heavily researched branch of game theory and has existed since the inception of game theory itself. Game theoretic models are normally solved through the use of BI and there is no shortage of game theoretical literature. Determining what the causal factors are behind people's inability to backward induct involves understanding how people think. Grabiszewski and Horenstein (2018) identify that in addition to an ability to backward induct, an individual needs an ability to create an accurate model of the problem with which to reason through. What this means is that even if an individual has the capacity to solve a game, if there is any ambiguity in their ability to properly interpret what they are presented with, then it may be impossible to discern whether a failure resulted from a lack of understanding or lack of ability.

Grabiszewski and Horenstein present their participants with two games that are identical except for that of their presentation. One game is laid out game theoretically, with the payoffs to the player appropriately labeled and every decision denoted by a line connecting the decision to its potential outcomes. The other version of the game mimics the first but requires the player to construct the game scenario mentally and does not provide a standard game tree model for the participants to interpret. They hypothesize that the probability of success is far greater when the problem is clearly laid out than when the problem requires both computation and mental reconstruction. The researchers found that a lot of the time, people do not reconstruct the problem properly, and therefore fail to understand the game at hand. This shows that for a proportion of the sample unable to solve the game that was not laid out game theoretically, their failure can be attributed to an inability to interpret the game rather than solve it.

Krockow (2016) draws parallels from a culmination of papers dealing with the centipede game from 1992-2016. The centipede game, similar to the Stick Game, is an extensive form game[3], but significantly more complicated. It demands interaction with another player and its game theoretical solution is not as clear as the Stick Game's solution. The use of forward induction[4] (rationalization of an opponent's non-game theoretical moves) in the centipede game can arguably supersede the use of BI.

In nearly every one of the centipede game experiments analyzed by Krockow, the results show that people do not behave game theoretically. Krockow attributes this to other-regarding preferences. Krockow's key insight is that people will cooperate when given the opportunity to. That is why I chose to run my experiment using a simple extensive-form game. Krockow shows that the centipede game gives too much of an incentive to cooperate as participants are able to look for patterns in the opponents' responses in order to communicate. In order to combat issues such as these, I have the participants play against a computer simulation. The computer is specifically designed to do what is optimal for it to win in every situation. This eliminates any incentive to cooperate and also reduces any uncertainty about the opponent's strategy.

Dufwenberg et al. (2010) points out that game theory research, and more specifically research into BI, can provide insights into many different areas. How we design and implement our educational systems, the long-term implications of targeted marketing, as well as the consequences of global warfare are all examples of situations that can benefit from constructing and solving game theoretical models using BI.

Design

My experiment uses four variants of the Stick Game. These four variants are the 5,10,13, and 16 Stick Games. The only difference between the four games is the number of initial sticks. The 5 Stick Game starts with 5 sticks and the players alternate picking sticks, the 10 Stick Game starts with 10 sticks, et cetera. I use these four variants so that I can model participants' behavior in the Stick Games as the games become increasingly more complex to solve.

The reason I chose to use the Stick Game is because it has certain characteristics that make identifying the use of BI much more apparent. The first advantage of using the Stick Game versus something more traditionally well-known such as the centipede game, is that there is no incentive to cooperate. The Stick Game has a winner-takes-all outcome. The player either won the game and

[3] An extensive-form game is a specification of a game in game theory, allowing (as the name suggests) for the explicit representation of a number of key aspects, like the sequencing of players' possible moves, their choices at every decision point, the (possibly imperfect) information each player has about the other player's moves when they make a decision, and their payoffs for all possible game outcomes.

[4] Forward induction reflects the idea that players rationalize their opponents' behavior whenever possible. In particular, players form an assessment about the future play of the game, given the information about the past play and the presumption that their opponents are strategic.

received the payoff, or lost and received nothing. In the centipede game and other popular BI models, both players' incentives increase as the game is played, and therefore, strict BI is not always the optimal solution.

The second advantage of using the Stick Game is that its complexity can easily be changed by increasing the number of initial sticks. By increasing the number of sticks, the game requires that the player go through an increasingly extensive thought process in order to arrive at the solution. With the ability to easily manipulate the difficulty of the game, it is possible to identify when players have developed heuristics that only solve a certain version of the game, versus a strategy that can effectively win at all the games. It is important to make this distinction because in the first case, the participant developed a strategy to win the game based on trial and error, whereas in the second scenario, the participant exhibited the use of BI. If you look at a participant's success in a series of Stick Games with a fixed number of initial sticks and you notice that at a certain point, the participant figures out the solution, while this may seem indicative of BI, it does not necessarily mean that the participant understood the game mechanics. In most cases, it just means that the participant stumbled upon a winning method and stuck to it.

It is vital to understand how BI works in the context of the Stick Game before diving into the intricacies of the experiment's design. The Stick Games, no matter the initial number of sticks, all function in the same manner. In order to solve the games, the participant has to start at the last stick and work their way back to the beginning stick. To ensure that the participant takes the last stick and that his opponent is unable to take the last stick, he must make sure that he takes the 4^{th} to last stick. If he takes the 4^{th} to last stick, the opposing player can at maximum take 2 sticks, in which case there is 1 stick left and the participant wins. The opposing player can at minimum take 1 stick, in which case there are 2 sticks left, and again, the participant wins. The participant can then treat the fourth stick as if it were the last and take the 7^{th} stick to ensure that he is able to take the 4^{th}. This logic can be extended backwards infinitely for any complexity of Stick Game.

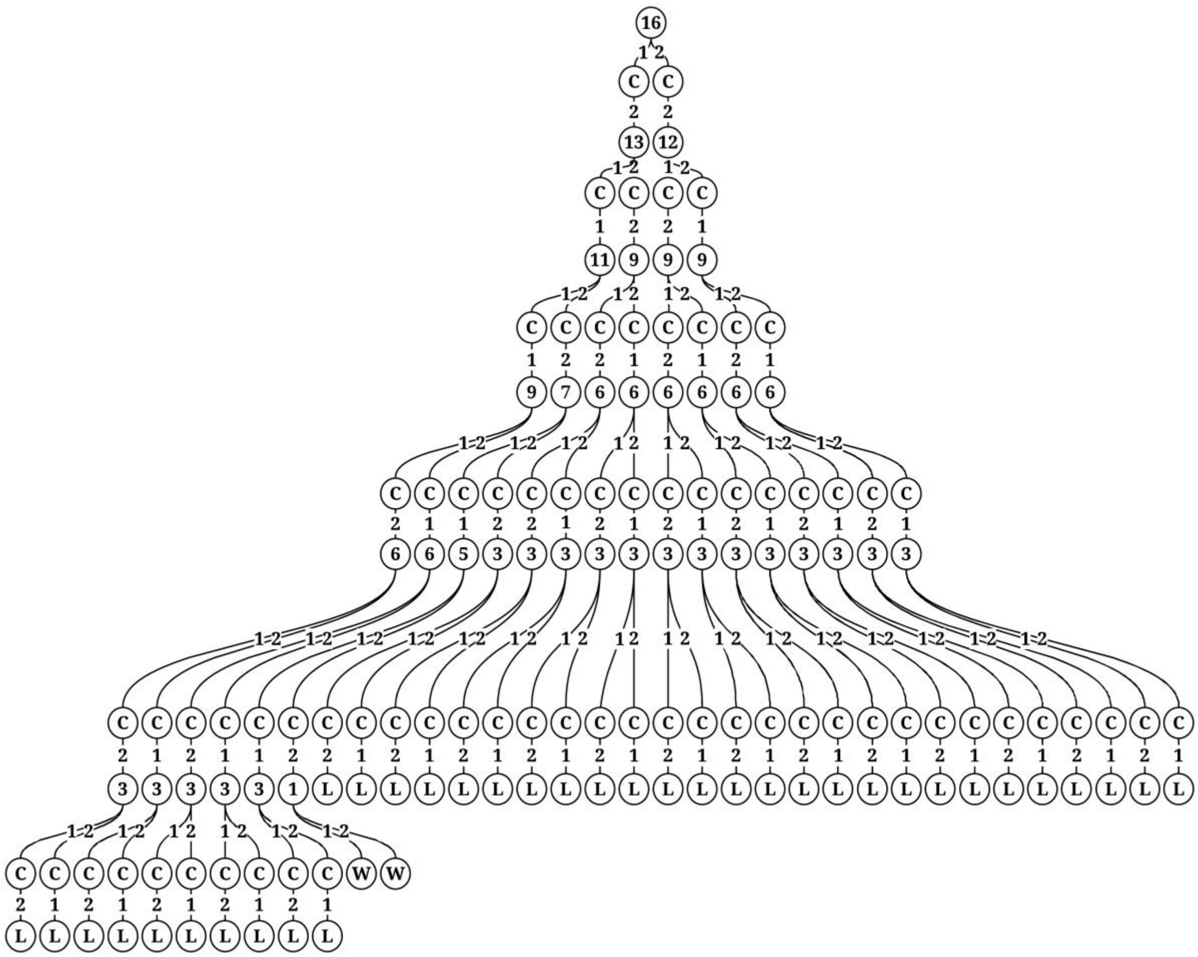
Because the Stick Games have only one unique solution, it is often lost on participants that their decisions at every stage of the game are equally important. Participants frequently rush through the game to get back to a point that they understand, but are often dumbfounded when they find that they can't get back to where they had originally desired to end up. The 16 stick game is modeled below in game tree form[5]. Every circle with a number inside of it represents a possible decision point for a participant and the associated number of sticks left at that point in the game. In practice, these are referred to as decision nodes[6]. Not all decision nodes will be reached in a single game, as the tree models all the possible outcomes of the game. The circles with a C in the middle denote a decision point for the computer. This is why, when the game reaches a player's decision node, there are two options, whereas when the game reaches a computer's decision node, there is only one predetermined response.

The computer is programmed to play the game perfectly, and therefore, already has a set input at every decision node. The computer plays perfectly by taking the 4^{th}, 7^{th}, 10^{th}, or 13^{th} stick if the participant does not make the optimal move at any point of the game. Something to note here is that taking the 16^{th} stick is also optimal and ensures victory, but as the participant goes first, this is impossible for the computer to achieve. For example, if the participant were to make an error initially and take 2 sticks instead of taking 1 stick (the 16^{th} stick), then the computer is programmed to take 2 sticks in response to this move. At this point the computer has won the game as it has

[5] In game theory, a game tree is a directed graph whose nodes are positions in a game and whose edges are moves.
[6] A node represents a "test" on an attribute and each branch represents the outcome of the test.

taken the 13th stick, and there is no combination of moves that the participant can make in order to correct his error. If the participant plays the game perfectly and the computer is faced with a situation where it is impossible to take the 4th, 7th, 10th, or 13th stick, it randomizes between choosing 1 stick or 2 sticks. This is to make it more difficult for participants to exploit a pattern in the computer's play. The game starts at the top of Image 1 with 16 sticks.

Image 1:



There is only one decision path that ultimately ends up in the player beating the game. If the player makes one wrong move at any stage of the game, then the computer is programmed to make losing the game an inevitability. This is the unique subgame perfect Nash equilibria[7].

In my experiment, I have the players play eight games in total. They first play the 5 Stick Game twice followed by the 10 Stick Game twice. After the participants have played the 10 Stick Game twice, half of them get a dummy question which does not inform them about anything of what is to follow. The other half of the participants get the treatment which is the solution to the

[7] A strategy profile is a subgame perfect equilibrium if it represents a Nash equilibrium of every subgame of the original game.

10 Stick Game. The treatment and control text is presented in Appendix A at the end of the paper. Using this approach, I differentiate between the two main causal factors behind failure to use BI. The two factors being an inability to realize that an analytical solution is possible, and an inability to perform the analytic to solve the problem. These factors can be simplified down to belief versus ability. After receiving the treatment or control question, the participants continue on to play the 13 Stick Game twice followed by the 16 Stick Game twice. I have the participants play each Stick Game twice to show their progression as they come to understand the mechanics of the game.

The subject pool consisted of UCSB students who are registered online for the economics laboratory experiments. I recruited these participants January 30th, 2019 and the experiment was run on January 31st, 2019. Their exact instructions and the design of the experiment can be found at the end of the report in Appendix A. Additionally, verbal instructions mimicking those found in the Appendix were given to the participants.

Model/Results

To understand the effect that belief has on the outcome of the game, I first use a simple linear regression model:

Success = $\beta_0 + \beta_1$Treatment + β_2Male + β_3CRT + β_4Session + β_5Ethnicity + β_6Training (5 Stick) + β_7Training (10 Stick) + ε

Where:

- Success is whether or not the individual beat the computer
- Treatment is whether or not the solution to the 10 Stick Game was given
- Male is whether or not the individual is male or female
- CRT is a factor variable representing how many questions the participant got correct on the cognitive reflection test[8] out of a total possible 3
- Session is a factor variable for which of the four sessions the participant took place in
- Ethnicity is a factor variable for what ethnicity a participant is. It can take on the values of White, Black or African American, Asian, Native Hawaiian or Pacific Islander, Hispanic or Latino, or other
- Training (5 Stick) is a factor variable representing how many out of the two 5 Stick Games the participant beat the computer in
- Training (10 Stick) is a factor variable representing how many out of the two 10 Stick Games the participant beat the computer in

The summary statistics for this model are shown below as well as a table of the control variables showing the difference between the treatment and control groups, and their associated P-values. The participants play each of the 5, 10, 13, and 16 Stick Games twice. In table 2 I denote the participant's first attempt at each Stick Game by labeling it 1st and the second by labeling it 2nd.

[8] The cognitive reflection test (CRT) is a task designed to measure a person's tendency to override an incorrect "gut" response and engage in further reflection to find a correct answer.

Table 1 shows the summary statistics and Table 2 shows the controls variables of the experiment and their associated P-values for the difference between treatment and control groups.

Table 1:

Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
1st 5 stick game win percentage	64	0.594	0.495	0	1
2nd 5 stick game win percentage	64	0.812	0.393	0	1
1st 10 stick game win percentage	64	0.109	0.315	0	1
2nd 10 stick game win percentage	64	0.406	0.495	0	1
1st 13 stick game win percentage	64	0.359	0.484	0	1
2nd 13 stick game win percentage	64	0.422	0.498	0	1
1st 16 stick game win percentage	64	0.281	0.453	0	1
2nd 16 stick game win percentage	64	0.344	0.479	0	1
CRT score (out of 3)	64	1.734	1.073	0	3

Table 2:

	Treatment	*Control*	
	Fraction	Fraction	P-value of difference
1st 5 stick game	21/34	17/30	0.920
2nd 5 stick game	28/34	24/30	0.952
1st 10 stick game	2/34	5/30	0.734
2nd 10 stick game	10/34	16/30	0.631
0 correct in CRT	6/34	4/30	0.912
1 correct in CRT	10/34	7/30	0.897
2 correct in CRT	9/34	8/30	1.000
3 correct in CRT	9/34	11/30	0.834
Session 1	5/34	5/30	0.960
Session 2	10/34	7/30	0.889
Session 3	6/34	9/30	0.771
Session 4	13/34	9/30	0.865
Male	15/34	12/30	0.936
White	9/34	10/30	0.880
Black	2/34	1/30	0.904
Asian	11/34	13/30	0.826
Hawaiian	2/34	0/30	0.734
Hispanic	9/34	5/30	0.818
Other	1/34	1/30	0.984

Notes: Column 4 shows the P-value against the null hypothesis that the treatment and control are the same.

Tables 3 and 4 show perfect play in the Stick Games by round and by treatment and control. Round 1 is the first time a participant played a version of the Stick Game and round 2 is the second. While these tables aggregate the data, they are insufficient in determining the causal effect and therefor are more of a visual representation. It is worth noting that the chance of beating the 5 Stick Game by chance is 50%, the 10 Stick Game is 12.5%, the 13 Stick Game is 6.25%, and the 16 Stick Game is 3.1%.

Table 3:

Perfect play by round

Game	*Round 1*		*Round 2*		*Pooled*	
	Perfect games	Percentage	Perfect games	Percentage	Perfect games	Percentage
5 stick	38/64	59	52/64	81	90/128	70
10 stick	7/64	11	26/64	41	33/128	26
13 stick	23/64	36	27/64	42	50/128	39
16 stick	18/64	28	22/64	34	40/128	31
All rounds	86/256	34	127/256	50	213/512	42

Notes: Columns 2,4, and 6 show the number of games won for each game. Columns 3,5, and 7 show the associated percentages.

Table 4:

Perfect play by treatment and control

Game	*Control*		*Treatment*		*Pooled*	
	Perfect games	Percentage	Perfect games	Percentage	Perfect games	Percentage
1st 13 stick	6/30	20	17/34	50	23/64	36
2nd 13 stick	8/30	27	19/34	56	27/64	42
1st 16 stick	5/30	17	13/34	38	18/64	28
2nd 16 stick	8/30	27	14/34	41	22/64	34
All rounds	27/120	23	63/136	46	90/256	42

Notes: Columns 2,4, and 6 show the number of games won for the treatment and control groups. Columns 3,5, and 7 show the associated percentages.

Table 3 shows the perfect play in each of the Stick Games. Table 3 shows round 1, round 2, and the pooled average of games played perfectly. As was expected, when the participants advanced through to the second iteration of the 5,10,13, or 16 Stick Games, there is a noticeable increase in success rate. This is attributable to the participants getting better with practice. Over time, the margin between the participants' success rate between rounds shrank as the games became more complex.

Table 4 shows the perfect play in the Stick Game by treatment and control. Only the 13 and 16 Stick Games are shown as they were the only games played post treatment. The table shows the control group, treatment group, and pooled average of all the perfect games played. As the games become increasingly more complex, the success rate dropped significantly. The treatment group seemed to benefit greatly from the knowledge of a rational solution.

The results from the linear regression both with and without the control variables are shown below:

Table 5:

Regression without controls

	Dependent Variable: Success in the 13 and 16 stick games			
	1st 13 stick game (1)	2nd 13 stick game (2)	1st 16 stick game (3)	2nd 16 stick game (4)
Treatment	0.300**	0.292**	0.216*	0.145
	(0.116)	(0.120)	(0.111)	(0.119)
Constant	0.200**	0.267***	0.167**	0.267***
	(0.085)	(0.088)	(0.081)	(0.087)
Observations	64	64	64	64
R-squared	0.097	0.087	0.057	0.023

Note: Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6:

	Dependent Variable: Success in the 13 and 16 stick games			
	1st 13 stick game	2nd 13 stick game	1st 16 stick game	2nd 16 stick game
	(1)	(2)	(3)	(4)
Treatment	0.362***	0.277**	0.214*	0.165
	(0.116)	(0.108)	(0.114)	(0.128)
Male	-0.031	-0.083	0.149	0.001
	(0.120)	(0.112)	(0.118)	(0.133)
1 correct in CRT	-0.159	-0.202	0.213	0.008
	(0.180)	(0.168)	(0.177)	(0.199)
2 correct in CRT	0.107	0.142	0.315*	0.093
	(0.183)	(0.171)	(0.180)	(0.203)
3 correct in CRT	0.192	0.422**	0.177	0.158
	(0.190)	(0.177)	(0.186)	(0.210)
Lab session 2	-0.014	0.201	0.296*	0.326*
	(0.172)	(0.161)	(0.169)	(0.191)
Lab session 3	-0.273	-0.390**	0.143	0.004
	(0.181)	(0.169)	(0.178)	(0.200)
Lab session 4	-0.184	-0.130	0.141	0.079
	(0.178)	(0.166)	(0.174)	(0.196)
Black	-0.453*	-0.518**	-0.372	-0.372
	(0.268)	(0.250)	(0.263)	(0.296)
Asian	-0.096	-0.269*	-0.027	0.058
	(0.149)	(0.139)	(0.146)	(0.165)
Hawaiian	-0.273	-0.109	0.044	-0.045
	(0.332)	(0.310)	(0.326)	(0.367)
Hispanic	-0.243	0.169	-0.033	-0.062
	(0.164)	(0.153)	(0.161)	(0.182)
Other	0.543	-0.177	0.006	-0.319
	(0.336)	(0.314)	(0.330)	(0.372)
1 win in 5 stick games	0.407	0.349	-0.077	0.275
	(0.338)	(0.315)	(0.331)	(0.373)
2 wins in 5 stick games	0.481	0.559*	0.280	0.466
	(0.338)	(0.315)	(0.332)	(0.374)
1 win in 10 stick games	-0.168	0.058	-0.039	-0.100
	(0.134)	(0.125)	(0.132)	(0.149)
2 wins in 10 stick games	0.321	0.107	0.316	0.316
	(0.256)	(0.239)	(0.251)	(0.283)
Constant	-0.045	-0.086	-0.324	-0.269
	(0.353)	(0.329)	(0.346)	(0.390)
Observations	64	64	64	64
R-squared	0.475	0.569	0.425	0.346

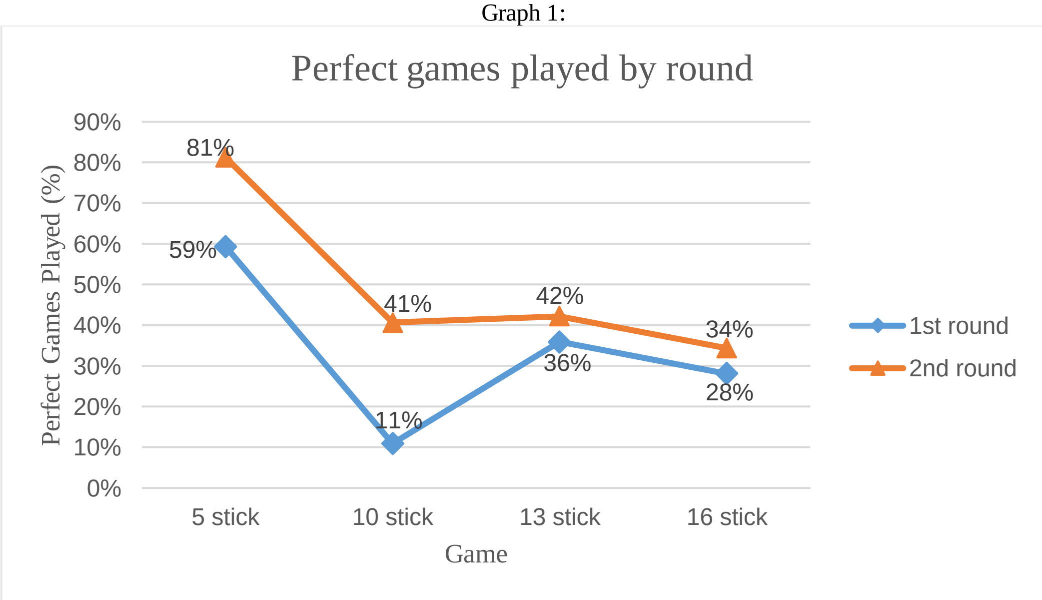
Note: Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

I measure the participant's cognitive response test score in order to control for any inherent intellectual differences that exist between the control and treatment groups. I control for wins in the first four games played in order to account of any inherent differences in the treatment and control's initial understanding of the Stick Games. Gender, ethnicity, and age were also controlled for. None of the controls had any consistent significant effect across all of the games.

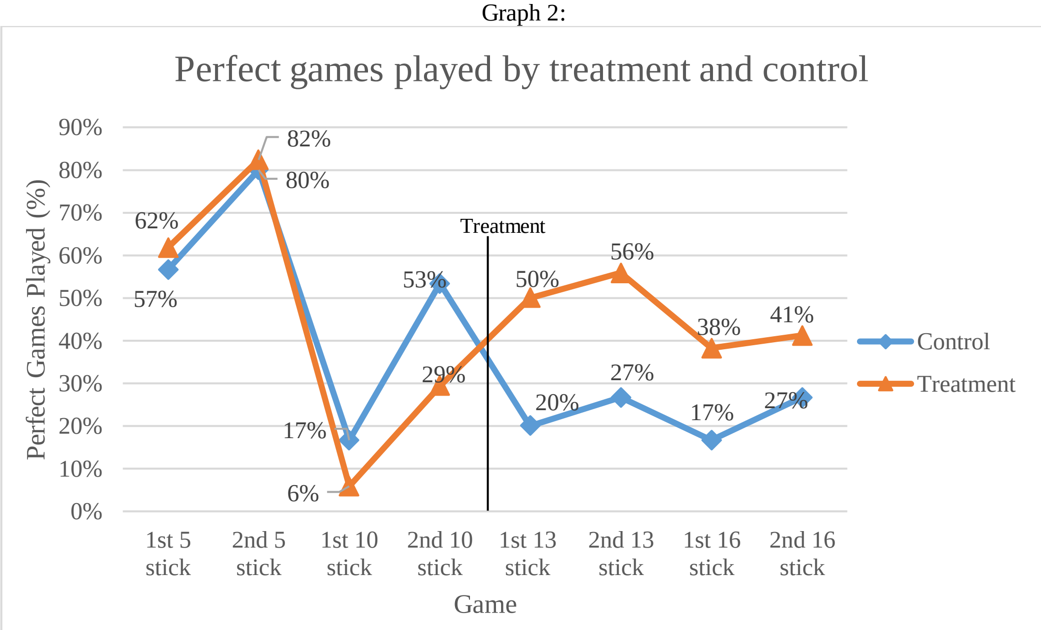
In the regression that includes the controls, the treatment effect did have a significant impact on the outcome of the 13 and 16 Stick Games. The effect was significant at the 99%

confidence level for the 1st 13 Stick Game, at the 95% confidence level for the 2nd 13 Stick Game and at the 90% confidence level for the 1st 16 Stick Game. The treatment did not have a significant effect on the 2nd 16 Stick Game. This is due to the limited sample size and that the participants in the control group differed in performance by less of a margin in the more complex games than in the less complex games. Due to the limited sample size, it is difficult to discern any real pattern in the control variables. There was no control variable that retained its significance throughout both the 13 and 16 Stick Games. Graphic illustrations of these results are shown below:

Graph 1:



Graph 2:



In Graph 1, the treatment effect can be observed despite the fact that it was only given to half of the sample. The success by round has an overall negative slope due to the games becoming more complex; the only notable exception being the 13 Stick Game as it came directly after the treatment. Additionally, the participants performed better on their second attempt at each game in every round.

Graph 2 depicts the inherent differences between the control and treatment groups. The line in the middle of the graph denotes the point at which the treatment was administered. Prior to the treatment, the control group was nearly twice as successful at the 2nd 10 Stick Game. Therefore, had we not controlled for the 5 and 10 Stick Games, we would have underestimated the effect of our treatment, as the control group had an initial understanding of the game greater than that of the treatment group.

The treatment is implemented between the 2nd 10 Stick Game and the 1st 13 Stick Game. Directly following the 2nd 10 Stick Game, the success rate of the treatment group rose immediately above that of the control group. This remained the case for the remainder of the games. The control group did however close the success rate gap due to an increased learning curve. I hypothesize that this is because the control group slowly learned of both their own capability and the existence of an analytical solution. The treatment group on the other hand, was fully aware of an analytical solution from the treatment given to them, and therefor benefited less over time.

The participants were administered the cognitive response test after completing the stick games. This was to ensure that the CRT did not influence their performance in the Stick Games in any way. The CRT measures a person's ability to override their gut response and cognitively reflect to find a solution. I hypothesized that the treatment effect would have less of an effect on the participants who scored higher on the CRT. I interacted the CRT with the treatment effect from the previous regression. The regression model I use is presented below. The results are shown in in Table 7.

Success = $\beta_0 + \beta_1$Treatment + β_2Male + β_3CRT + β_4Session + β_5Ethnicity + β_6Training (5 Stick) + β_7Training (10 Stick) + β_8CRT*Treatment + ε

Where:

- Success is whether or not the individual beat the computer
- Treatment is whether or not the solution to the 10 Stick Game was given
- CRT is a factor variable representing how many questions the participant got correct on the cognitive reflection test out of a total possible 3
- CRT*Treatment is the interaction term between the CRT and the treatment
- Male, session, and ethnicity are all the same control variables used in the previous regression
- Training (5 Stick) is a factor variable representing how many out of the two 5 Stick Games the participant beat the computer in
- Training (10 Stick) is a factor variable representing how many out of the two 10 Stick Games the participant beat the computer in

Table 7:

	Dependent Variable: Success in the 13 and 16 stick games			
	1st 13 stick game	2nd 13 stick game	1st 16 stick game	2nd 16 stick game
	(1)	(2)	(3)	(4)
Treatment	0.558**	0.457**	0.280	0.224
	(0.231)	(0.215)	(0.229)	(0.258)
Male	-0.029	-0.081	0.150	0.002
	(0.121)	(0.113)	(0.119)	(0.135)
1 correct in CRT	-0.087	-0.136	0.236	0.030
	(0.195)	(0.182)	(0.193)	(0.217)
2 correct in CRT	0.238	0.262	0.359	0.132
	(0.227)	(0.212)	(0.225)	(0.253)
3 correct in CRT	0.377	0.591**	0.239	0.213
	(0.267)	(0.250)	(0.265)	(0.298)
Lab session 2	0.019	0.231	0.307*	0.336*
	(0.176)	(0.164)	(0.174)	(0.196)
Lab session 3	-0.257	-0.375**	0.148	0.009
	(0.182)	(0.170)	(0.180)	(0.203)
Lab session 4	-0.174	-0.120	0.145	0.082
	(0.178)	(0.166)	(0.176)	(0.199)
Black	-0.472*	-0.535**	-0.378	-0.378
	(0.269)	(0.251)	(0.266)	(0.300)
Asian	-0.084	-0.257*	-0.022	0.062
	(0.149)	(0.140)	(0.148)	(0.167)
Hawaiian	-0.264	-0.100	0.047	-0.042
	(0.332)	(0.310)	(0.329)	(0.371)
Hispanic	-0.278	0.137	-0.044	-0.072
	(0.168)	(0.157)	(0.167)	(0.188)
Other	0.461	-0.253	-0.022	-0.344
	(0.347)	(0.324)	(0.343)	(0.387)
1 win in 5 stick games	0.322	0.270	-0.106	0.249
	(0.349)	(0.326)	(0.346)	(0.390)
2 wins in 5 stick games	0.391	0.477	0.251	0.439
	(0.350)	(0.327)	(0.347)	(0.391)
1 win in 10 stick games	-0.179	0.048	-0.043	-0.103
	(0.135)	(0.126)	(0.134)	(0.151)
2 wins in 10 stick games	0.286	0.075	0.305	0.306
	(0.259)	(0.241)	(0.256)	(0.289)
CRT*Treatment	-0.113	-0.103	-0.038	-0.034
	(0.115)	(0.107)	(0.114)	(0.128)
Constant	-0.083	-0.120	-0.336	-0.281
	(0.355)	(0.332)	(0.352)	(0.397)
Observations	64	64	64	64
R-squared	0.486	0.577	0.426	0.347

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The results from Table 7 neither confirm nor refute the original hypothesis that the effect of the treatment was less for those who scored higher on the CRT. The CRT*Treatment coefficient shows the interaction between the CRT and the treatment effect. While the coefficient is negative for all of the 13 and 16 Stick Games as hypothesized, it is not significant for any of them. We would expect the coefficient to be negative because an individual that scores higher cognitively would likely not benefit from the treatment to as much of a degree as would a participant who scored lower. The lack of significance is attributed to a small sample size as there were no more than ten people in any of the groups of participants scoring either a zero, one, two, or three on the CRT. Given that the coefficients were negative across all four games played post treatment, it is likely that with a larger sample size, the effect of the treatment would have had a significantly diminished effect on those who scored highly in the CRT as opposed to those who did not.

The final model that I utilized to fit the data was a difference-in-differences model. It is represented by:

Success = $\beta_0 + \beta_1$Treatment + β_2Post-Treatment + β_3Treatment$*$Post-Treatment + ε

Where:

- Treatment = is whether or not the treatment was given
- Post-Treatment = is whether the observation came after the treatment
- Treatment$*$Post-Treatment = the interaction term that is equal to the treatment effect

The difference-in-differences model allows us to analyze the data and treatment effect over all of the outcomes of the study and gain an aggregate treatment effect. The results of this regression are shown below:

Table 8:

Dependent Variable: Aggregate success in the 5,10,13, and 16 stick games	
Difference-in-Difference Estimation	
Treatment	-0.068
	(0.060)
Post-Treatment	-0.292***
	(0.062)
Treatment*Post Treatment	0.306***
	(0.085)
Constant	0.517***
	(0.044)
Observations	512
R-squared	0.049

Note: Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In the difference-in-differences model, the variable treatment shows the difference between the treatment and control group's performances before the treatment was administered. It is negative, which means that the treatment group performed worse before the treatment than did the control group. The post-treatment variable shows how the control group performed in the games following the treatment as compared to the games prior to the treatment. Because it is negative, it shows us that the success rate of participants in the control group fell as they played more complex games. The interaction term of treatment*post-treatment shows the treatment effect.

In the difference-in-differences regression, the treatment effect is significant at the 99% confidence level. This means that on average, the belief in a rational solution to the Stick Game resulted in roughly a thirty percent increase in the success rate of the games post treatment. The increase in success rate refers to all of the games after the treatment not just the 1st 13 Stick Game, 2nd 13 Stick Game, and 1st 16 Stick Game. This lends credibility to the aforementioned hypothesis that the lack of significance of the treatment effect on the 2nd 16 Stick Game was due to sample size and the decreasing margins of success rates between the treatment and control groups from repeated trials.

Conclusions

The majority of BI research focusses on whether or not participants exhibit the use of BI when it is optimal to do so. Research has shown that people tend not to use BI. A limited amount of research however, focusses on the reasons behind this apparent lack of game theoretical behavior. Researchers generally treat BI as a singular ability that an individual possess rather than a complex set of mental processes necessary in order to solve a problem.

I focus my research on the various components that make up BI. I hypothesize that there are numerous factors underpinning the apparent inability of people to use BI in games. It is necessary to account for other-regarding preferences to ensure that participants don't cooperate. The presence of an ability to forward induct must be accounted for as it can make determining what is truly game theoretical significantly more difficult. I use the Stick Game because it is simple, effective, and above all else, provides no room for misinterpretation or cooperation.

I first have the participants play a series of training rounds followed by half of the sample either receiving the treatment or receiving a control question. I then analyze the effect that the belief in the existence of an analytical solution has on the participants' ability to BI properly. We find that the effect of this belief on a participant's outcome is significant at the 99th, 95th, and 90th percentile for the first three out of the total four games played post treatment. The lack of significance for the fourth regression is attributed to a small sample size and a decreasing margin between the control and treatment groups as the games became more complex. The control group was successful in beating the computer in the Stick Games following the treatment approximately 22.75% of the time. The treatment effect increased a participant's chances of winning the game by 25.45%. Therefore, the treatment increased a participant's probability of success by approximately 112% of what it originally was. The treatment had a very large impact on a participant's likelihood of beating the Stick Game.

Using a difference-in-differences regression, I find that across all versions of the game implemented post treatment, the effect of the treatment is significant at the 99% level. Participants greatly benefited not only from the understanding that the game was solvable, but also from the understanding that the solution can be extended outward infinitely. I also found that in individuals that scored higher on the CRT, the effect of the treatment was lessened. While this effect was not significant, this is likely due to the limited sample size.

Life presents situations so convoluted that it is often subgame perfect for an individual to do what is best in the moment rather than attempt to reason backward through an infinitely complex series of events. In the absence of an ability to accurately construct or solve situations as they present themselves, people reconstruct their situations based on the information available to them and solve the simplified versions of their reality game perfectly.

When the absence of BI is observed, the appropriate response is not to conclude that people are incapable of BI, which is in fact inaccurate, but is to develop a model of the human psyche that can explain this phenomenon. I focus my research on what I believe to be two of the most important factors in this equation, belief and ability. I claim that people are far more capable than game theorists normally give them credit. Even the most advanced computing software fails to mimic the intricacies inherent to the human brain and necessary in order for us to accurately interpret our surroundings. Why then should it be assumed that people treat reality in a binary fashion. Rather, we should allow for the inherent complexity of life in our models and do our best to deduce the true causal factors behind the reality of backward induction.

References

Steven Levitt, et al. Checkmate: Exploring Backward Induction Among Chess Players. *American Economic Review 101 (April 2011): 975–990*, 2009.

Martin Dufwenberg, Ramya Sundaram, David J. Butler. Epiphany in the Game of 21. *Journal of Economic Behavior & Organization 75 (2010) 132–143*, 2010.

Grabiszewski Horenstein, et al. Tree Construction and Backward Induction: A Mobile Experiment." By Jikun Huang, Ruifa Hu, Scott Rozelle, Fangbin Qiao, Carl E. Pray :: SSRN, 31 May 2018, papers.ssrn.com/sol3/papers.cfm?abstract_id=3181740.

Krockow, Eva. "Cooperation in Repeated Interactions." Researchgate, Jan. 2017, www.researchgate.net/profile/Eva_Krockow/publication/323028589_Cooperation_in_repeated_interactions/links/5a7d5f3ea6fdccc013f5a42e/Cooperation-in-repeated-interactions.pdf.

Appendix A

{Subjects were given the following instructions. The participants played each game twice. The following was given to all participants}

Instructions:
In order to familiarize yourself with how the primary game in this study works, you will first play a simplified version. You will play this training game twice. This game is called the 5 Stick Game. It's a turn-based 2 player game (player and computer). The game starts with 5 sticks and each player can take either 1 or 2 sticks during their turn. The goal of the game is to take the last stick.

There are 5 sticks. How many sticks would you like to take?

- 1
- 2

There are now X sticks. How many sticks would you like to take?

- 1
- 2

et cetera:

The computer took X sticks. Because the computer took the last stick, you lost.

Or:

Congratulations you won!

In the next section of this experiment, you will play six games. At the end of the experiment, one of these games will be randomly selected by a dice. If you beat the computer in this randomly selected game, you will earn 6 dollars.

You will now play the 10 Stick Game. It's a turn-based 2 player game (player and computer). The game starts with 10 sticks and each player can take either 1 or 2 sticks during their turn. The goal of the game is to take the last stick.

There are 10 sticks. How many sticks would you like to take?

- 1
- 2

There are now X sticks. How many sticks would you like to take?

- 1
- 2

et cetera:

The computer took X sticks. Because the computer took the last stick, you lost.

Or:

Congratulations you won!

{The following was given to the treatment group}

Now that you have played the 10 Stick Game twice, you will be given the solution on how to beat the game.

The solution to the 10 Stick Game is to erase the tenth, seventh, and fourth to last sticks. Erasing the fourth to last stick ensures that the computer can't erase the last stick and that you can. Erasing the seventh to last stick ensures that the computer can't erase the fourth to last stick and that you can. Finally, erasing the tenth to last stick ensures that the computer can't erase the seventh to last stick and that you can.

This same logic that is used to beat the 10 Stick Game can be used to beat longer and more complex Stick Games.

To demonstrate that you have read and understood the paragraph above, what is the optimal strategy to beat the computer in the 10 Stick Game? If you get this question right you will earn 1 dollar.

- Erase the tenth, seventh, and fourth to last sticks
- Erase only the odd sticks
- Erase only the third, sixth, and ninth to last sticks
- Erase one stick every time

{The following was given to the control group}

You have finished playing the 10 Stick Games.

You will now play the 13 and 16 Stick versions of the game. To demonstrate that you have read and understood this, choose the versions of the sticks games that you will play next. If you answer correctly, you will earn 1 dollar.

- 12 and 15
- 17 and 18
- 15 and 14
- 13 and 16

{the following was given to all participants}

You will now play the 13 Stick Game. It's a turn-based 2 player game (player and computer). The game starts with 13 sticks and each player can take either 1 or 2 sticks during their turn. The goal of the game is to take the last stick.

There are 13 sticks. How many sticks would you like to take?

- 1
- 2

There are now X sticks. How many sticks would you like to take?

- 1
- 2

et cetera:

The computer took X sticks. Because the computer took the last stick, you lost.

Or:

Congratulations you won!

You will now play the 16 Stick Game. It's a turn-based 2 player game (player and computer). The game starts with 16 sticks and each player can take either 1 or 2 sticks during their turn. The goal of the game is to take the last stick.

There are 16 sticks. How many sticks would you like to take?

- 1
- 2

There are now X sticks. How many sticks would you like to take?

- 1
- 2

et cetera:

The computer took X sticks. Because the computer took the last stick, you lost.

Or:

Congratulations you won!

{This is the CRT administered to all participants}

You will now answer a series of questions. For every question that you get correct, you will earn 1 dollar.

A bat and a ball cost 22 dollars in total. The bat costs 20 dollars more than the ball. How many dollars does the ball cost?

If it takes 5 machines 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets?

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 many days would it take for the patch to cover half of the lake?