Temptation, Learning, and Backward Induction in Sequential Decision Tasks

William Neilson
University of Tennessee
Knoxville, TN
wneilson@utk.edu

Michael Price
University of Tennessee
Knoxville, TN
mprice21@utk.edu

Mikhael Shor
Vanderbilt University
Nashville, TN
Mike.Shor@owen.vanderbilt.edu

Abstract:

Backward induction can be difficult, and research has shown that paying individuals for making steps along the optimal path can help. At the same time, businesses have exploited the difficulty by tempting individuals off the optimal path. Little is known, however, about what individuals actually learn from these interim payments.

This study uses a laboratory experiment to determine how interim payments, both on and off the optimal path, impact learning. Subjects play a game against a computer in which winning a prize requires six or more correct moves. Interim payments, smaller than the prize for winning, are inserted in some games, and treatment effects concern whether these interim payments lie on or off the optimal path, and whether they occur early or late in the game. After 30 rounds with the same game and the same interim payments, the game is changed so that a new path becomes optimal. Data from the first game allow us to determine whether the interim payments aid or impede learning, and data from the second game allow us to determine whether subjects learned the game-specific pattern or a backward induction process. Learning a game specific pattern implies that subjects will take time to learn a new pattern when the game changes, but learning to backward induct implies that learning in the second game should be quick, or even immediate.

We find that interim payments on the optimal path help subjects learn the game-specific pattern, but impede them from learning to backward induct. In contrast, interim payments off the optimal path hinder them from learning anything, but those who do learn are able to backward induct when the game changes. These results may have implications for social programs that try to reward good behavior instead of final results, because such programs may prove detrimental in a changing environment.

I. Introduction

In many important economic contexts, including consumption-saving decisions, capital investment programs, resource extraction settings, and multi-year contracting environments, optimization requires making the “right” choice at every decision node. While a significant body of research applies dynamic programming techniques or numerical methods to identify these optimal decision paths, little empirical research explores the ability of individuals to actually follow them. Casual observation of outcomes suggests that people have trouble, especially when one considers evidence on addiction or obesity rates or on insufficient retirement savings.

The marketplace provides further evidence suggesting that people have difficulty following optimal plans. Mortgage lenders, credit card companies, cable and satellite TV companies, health clubs, and mail-order book and record clubs all offer promotional "teaser" rates to lure customers to enter into long-term contracts.[1] All of these practices use monetary incentives to lure people off of the optimal path, thereby reducing their long-run welfare. Teasers need not always hurt welfare, though. Fryer (2010) presents evidence that paying students to read improves reading comprehension but paying them to do well on reading comprehension tests does not. Charness and Gneezy (2009) find that paying people to go to the gym can induce higher exercise levels even after the payments end. Individual development accounts (IDAs) are savings accounts for low-income households that provide matching funds when used for specific purposes like financial instruments, housing, or education. Mills et al. (2008) find that IDAs increase both homeownership rates and non-retirement financial assets, and conclude that IDAs are more effective than more traditional forms of assistance, such as income supports. Ashraf, Karlan, and Yin (2006) show that a commitment savings product offered by a Philippine bank greatly increased savings for those who used the product, and that the effect was long-lasting. Duflo, Kremer, and Robinson (2010) find that farmers in Kenya fail to invest in fertilizer in general, but do invest in response to small, time-limited discounts. All of these studies show that monetary incentives can be used to lure people onto a "more optimal" path, raising their long-term welfare.

These findings, coupled with the existence of market mechanisms to lure people away from the optimal path, raise some important questions. While interim payments entice people onto a better path, would they eventually have found their way to the better path without the inducement of interim payments? If the answer is yes the interim payments lead to short-term gains because they simply speed up the move to the better path, but if the answer is no then interim payments are the only way to get more people on the optimal path. The studies claim that beneficial interim payments lead to long-term beneficial changes in behavior. Do harmful interim payments lead to long-term detrimental changes in behavior, too? If the answer is no the interim payments simply delay people from finding the optimal path, but if the answer is yes then perhaps policy actions are warranted. More fundamentally, though, how exactly do interim payments impact people's ability to optimize in long sequential decision tasks? The purpose of this paper is to use laboratory experiments to address these questions.

Few recent experimental studies investigate the ability of subjects to identify and follow an optimal path in a complex, non-repeated sequential decision setting. Houser, Keane, and McCabe (2004) give subjects a 15-decision, analytically-unsolvable, human capital investment problem with the goal of creating a taxonomy of decision rules. Brown, Chua, and Camerer (2009) present subjects with a 30-round consumption-saving exercise to determine how and how well subjects

[1] DellaVigna and Malmendier (2004) identify ways in which firms can profit from time-inconsistent consumers and then provide evidence that firms use the identified pricing strategies including the ones listed here.

learn to save optimally. They also study the impact of temptation, primarily triggering time inconsistent, present-biased behavior.

In this paper we run an experiment to determine how temptation affects subjects' choices and learning in a relatively short, relatively simple sequential decision setting designed to avoid issues raised by heterogeneous preferences over the timing of events. The experiments use the race-to-21 game discussed by Dufwenberg, Sundaram, and Butler (2010) and Gneezy, Rustichini, and Vostroknutov (2010).[2] The game starts with 21 stones. The subject moves first, picking up between one and three stones, inclusive. The computer then picks up between one and three stones, the subject picks up between one and three stones, and so on until all the stones have been picked up. The subject wins a monetary prize by picking up the last stone. The game can be solved simply through backward induction, and to guarantee being able to pick up the 21st stone the subject must end his turns by picking up stones 1, 5, 9, 13, 17, and 21. To see why, note that stopping at 17 means that no matter what the computer chooses the subject can win on the next turn. To guarantee being able to stop on stone 17, the subject must first stop at stone 13 so that no matter what the computer chooses, the subject can stop at 17 on the next turn. Following this same logic reveals the optimal path as described above.

Our study differs from previous studies of the race-to-21 game in two ways. First, previous studies were interested in the race-to-21 game as a game and so looked at how subjects played against each other, whereas we are interested in it as a decision problem and therefore look at how subjects play against a computer. Playing against a computer means that any deviations from the optimal path are disciplined immediately, since the computer is programmed to follow the winning path when it can. Furthermore, since the race-to-21 game has a solution, the problem is nonstochastic and so issues surrounding subject risk attitudes do not come into play. Second, their central question involved the learning process, whereas our central question involves temptation and learning transfer. Accordingly, their games did not involve teaser payments while ours do.

In the race-to-21 game described above, subjects earn $1 for taking the final stone, and to get there must end their turns on stones 1, 5, 9 13, and 17. Suppose that the subject can earn a 50¢ teaser payment by ending a turn on stone 5, which is on the optimal path. Such a teaser payment helps the subject figure out how to get to stone 5, and that learning process might help the subject figure out how to get to stone 21. On the other hand, a 50¢ teaser payment on stone 6 is off the optimal path. Furthermore, the computer will allow the subject to get stone 6 because the computer will always stop on stone 5 when it is available. The subject gets the 50¢ payment, but doing so makes it impossible to win the $1 for taking stone 21. We look at how teaser payments in various positions (stones 5, 6, 11, and 13), impact subjects' ability to find the optimal path and win the game.

We find that subjects can be categorized into three groups – transferable learners who figure out the general backward induction solution, nontransferable learners who figure out patterns on a case-by-case basis, and non-learners – and that these groups do not characterize the subjects per se, but rather characterize their play in a race-to-21 game. We find that on-path

[2] Levitt, List, and Sadoff (2010) use a race-to-100 game.

teasers help subjects recognize patterns, but not necessarily to backward-induct.[3] Off-path teasers inhibit pattern recognition and slow the process of learning the optimal path. Oddly, and contrary to the findings of the empirical studies, on-path teasers may have detrimental long-term effects when the environment changes. We find this by having subjects play a second game in which the teaser payment remains on the same stone but the number of stones chosen changes from 3 to 4. When players can pick up four stones at a time the optimal path becomes 1, 6, 11, 16, 21. A teaser on 13 is on-path for the 3-stone game but off-path for the 4-stone game, while a teaser on 11 is off-path for the 3-stone game but on-path for the 4-stone game. Our results show that teasers that enhance learning in the first game do so through non-transferable learning, and that games that are harder to win generate more transferable learning.

Overall, our findings confirm the field studies that interim payments impact behavior, and that on-path teaser payments help subjects find the optimal path. However, on-path teaser payments help subjects recognize patterns, not backward-induct, and therefore might actually be harmful in a changing environment when they apply the wrong pattern in the new game. Perhaps our most striking result is that an unattainable off-path teaser actually aids backward induction. It is impossible for a subject to end a turn on stone 5 when the subject can pick up four stones at a time. We find that this unattainable teaser leads more subjects to learn the game and to learn it more quickly. Based on these findings, the best policy might not be to pay people for the right steps on the path, but to induce people to think hard about what the early steps on the path should be.

II. The Game and Identification Procedure

Let D denote a sequential decision problem in the set of such problems Ω. Each decision problem in Ω is assumed to have a unique solution, and let $f(D)$ be the mapping which assigns to each D in Ω the sequence of actions $a_1,\ldots,a_T$ the individual should take to maximize her payoff. This notation abstracts away from the contingent nature of both the number of decision nodes and the optimal path, neither of which are germane for the task run in the experiment.

Because of its assumption of a perfect ability to maximize, standard economic theory predicts that individuals will always play the sequence $f(D)$. When actually confronted with a decision problem, though, a boundedly rational individual may or may not figure out the optimal sequence. Ultimately, the outcome depends on what the individual learns. Some individuals will not figure out the optimal sequence $f(D)$, and we label these *non-learners*. In contrast, *learners* figure out the optimal path. However, there are two ways that an individual can do this, either by learning the problem-specific pattern $f(D)$, which only applies to the precise task at hand, or by learning how the general class of problems Ω works, in which case she learns the entire mapping f. They can then use this mapping to solve any problem in Ω. We call a subject that learns the mapping f a *transferable learner* and one that only learns the specific pattern $f(D)$ a *nontransferable learner*. The purpose of the paper is to construct a methodology for separating the two types of

[3] There is also the possibility that on-path teasers simply shorten the game. We present evidence refuting this interpretation.

learners, and to explore how properties of the decision problem impact the type of learning subjects experience.

To do this we use a race game, which in its most basic form can be described using two parameters: the number of steps *s* that can be taken in a single turn, and the total number of steps *X* in the game. The decision-maker moves first choosing an integer number of steps between 1 and *s*, inclusive. The computer follows, and the players alternate, with the player ending on step *X* winning a prize. We identify a path according to where a turn ends, which allows us to identify a path that is not contingent on where the other player ended her turn. Using backward induction, the optimal path has a player ending turns (in reverse order) on steps

$$\begin{array}{c} X \\ X-(s+1) \\ X-2(s+1) \\ X-3(s+1) \\ \vdots \end{array} \qquad (1)$$

and so on to the beginning of the game. For example, when the race is to $X = 21$ and the maximum number of steps per turn is $s = 3$, the optimal path has the individual end turns on steps (1, 5, 9, 13, 17, 21), and when the race is to $X = 21$ and the maximum number of steps per turn is $s = 4$, the optimal path has her end turns on steps (1, 6, 11, 16, 21).

A more complicated race game places an interim teaser payment on an intermediate step $T < X$. We call the reward for ending a turn on *X* (and winning the game) the *final payoff* and the reward for ending a turn on *T* the *teaser payoff*. An individual can only collect the teaser payoff by ending a turn on step *T*, and passing through step *T* means that no one collects the teaser payoff. Some games do not have teaser payments, and to allow for this possibility we use the convention $T = 0$ for such a case. To make it truly a teaser, the teaser payoff must be strictly less than the final payoff, and we make this assumption throughout the paper. Note that under the maintained assumption that the final payoff exceeds the teaser payoff, the optimal path for the game is still described by (1) no matter where the teaser payoff is.

Holding the two payoffs constant, a decision problem *D* can be described by the triple (X,s,T). So, for example, (21,3,13) corresponds to the race-to-21 game with a maximum step length of 3 and a teaser on step 13, and (21,3,0) corresponds to the same game but without a teaser payoff. The mapping $f(X,s,T)$ is the rule described by system (1) above, which is independent of *T*. When faced with a specific race game, such as (21,3,13), a transferable learner will figure out rule (1), but a nontransferable learner will only figure out the particular path (1, 5, 9, 13, 17, 21). A non-learner will choose something other than the optimal path, in which case the computer will choose the optimal path and the non-learner will be unable to end a turn on step 21.

If an individual only faces a single race game, it is possible to distinguish learners from non-learners simply by observing whether they successfully end the game on step *X* and win the final payoff. A single race game does not allow one to distinguish between transferable and nontransferable learners, though. To do so requires a second race game with either a different endpoint *X* or a different step length *s*, because changing either of those alters the path identified in (1). A transferable learner will simply plug the new values into the function *f*(*X,s,T*) and solve the game, while a nontransferable learner will have to identify the new path from scratch.

Our identification procedure relies on the following process. Subjects first play the race game (21,*s*,*T*) for 30 rounds and then play the game (21,*s'*,*T*) for 15 rounds, where *s* and *s'* are either 3 or 4. With these pairings the game length is the same in both treatments and the teaser placement is the same in both, but the step length differs between treatments. Changing the step length is sufficient to change the optimal path as identified by (1). A learner is a subject who is able to obtain the final payoff in the last several rounds of the first race game, and a non-learner is a subject who cannot obtain the final payoff. A transferable learner is able to obtain the final payoff in the last several rounds of the first race game *and* in the first several rounds of the second race game. A nontransferable learner obtains the final payoff in the last several rounds of the first game but not in the first several rounds of the second game.

We use this identification procedure to study the impact of teaser placements, in particular their position either on or off the optimal path, and their position either early or in the middle of the game. It is possible to think of rationales for opposing predictions. One rationale works from the argument that when something is easy more people get it right but learn it less well, while when something is hard fewer people get it right but those who do really figure it out. On-path teasers make the game easier, so should increase learning. By this rationale, a larger fraction of the learning should be nontransferable. By the same logic, off-path teasers make the game more difficult, so fewer subjects should learn it. Those who do, though, really learn the game and are more likely to be transferable learners.

An opposing rationale comes from thinking about on-path teasers in a different way. With an on-path teaser, subjects must first learn to get the teaser, and then learn how to get from there to the final step. Thus, the subject must learn two games instead of just one. Because the component games are shorter, they are easier to learn. Also, because there are two of them, learning is more transferable. Consequently, this rationale predicts more transferable learning, not less, when the teaser is on the optimal path.

When the teaser is off the optimal path, the two-shorter-game logic fails. Getting the teaser makes it impossible to win the game. Furthermore, off-path teasers are easy for the subject to get, because once the subject is off the optimal path the computer always follows the optimal path and leaves the teasers available for the subject.[4] "Learning" to get an off-path teaser is not learning at all, and so has no impact on transferability. Consequently, the fraction of learners who are transferable learners should be the same in a game with an off-path teaser and a game with no teaser at all.

[4] There is one exception. In the game (21,4,5), which we run, the subject, who moves first, cannot reach the teaser on the first move. The optimal path is (1,6,11,16,21). If the subject chooses 1 on the first turn, the computer will take the teaser by ending on 5. If the subject chooses anything else, the computer ends the turn on 6 to get on the winning path, bypassing the teaser. Either way, the subject cannot get the teaser.

Both rationales lead to identical predictions for learning: on-path teasers promote learning and off-path teasers deter it. This leads to the following hypothesis.

> *Hypothesis 1.* The fraction of subjects learning the game is greatest when there is a teaser on the optimal path and smallest when there is a teaser off the optimal path. Games without teasers will lead to learning rates between these two.

The rationales lead to opposing predictions for transferability. If teasers make the game easier or harder and difficulty promotes transferability, then on-path teasers should lead to larger shares of nontransferable learners and off-path teasers should lead to larger shares of transferable ones. On the other hand, if on-path teasers allow subjects to solve two games but off-path teasers only allow them to solve one, transferable learning should be more prevalent with on-path teasers. These considerations lead to two alternative hypotheses regarding transferability, with the first stemming from the easy/hard task rationale and the second from the two-game/one-game rationale.

> *Hypothesis 2a.* Among those who learn, the largest fraction of transferable learners arises with off-path teasers, then no teasers, then on-path teasers.

> *Hypothesis 2b*. Among those who learn, the largest fraction of transferable learners arises with on-path teasers, and off-path teasers and no teasers lead to identical shares of transferable learners.

These two hypotheses are obviously contradictory. Importantly, though, it is possible to reject them both, for example if the no-teaser games lead to the most transferable learning.

III. EXPERIMENTAL DESIGN

(Text to be added)

Treatment	teaser stone location	action space (max stones per turn)	N
On path teaser	5	3	22
	6	4	19
	11	4	23
	13	3	20
			84
Off path teaser	5	4	13
	6	3	11
	11	3	16
	13	4	24
			64

No teaser	3	12
	4	21
		33

Treatments (by first game subjects play).

IV. Learning and Temptation

The most fundamental issue is whether subjects learn the first race-to-21 game they face (rounds 1 – 30) and achieve the winning stone. Table 2 shows the winning rates for subjects in the first game. When there is no teaser, subjects follow the optimal path in 25% of the rounds. A teaser on the optimal path increases performance greatly, with subjects following the optimal path in 52% of the rounds, and the differences is statistically significant at the $p = 0.001$ level using either a t-test or a Mann-Whitney test.[5] On the other hand, off-path teasers do not have much effect, reducing the fraction of rounds won by only a statistically insignificant 4 percentage points. This suggests that failures to achieve the winning stone cannot be explained solely by subjects chasing the off-path teaser, because chasing the teaser would make subjects win less often than in the no-teaser treatment. Furthermore, subjects only achieved the teaser in 48% of the off-path rounds even though the computer lets them have the teaser once they leave the optimal path.

Table 2. Summary of first-game activity

	Achieve teaser	Achieve winning stone
On-path	74%	52%
Off-path	48%	21%
No teaser		25%

Understanding how on-path teasers lead to higher success rates requires a closer look at the data. When there is no teaser, 45% of the subjects make the right first move (by choosing 1) in the first fifteen rounds, as shown in Figure 1. Only 20% make the correct second move (by choosing 5 when they can pick up 3 stones or 6 when they can pick up 4), and only 15% make the third move correctly. Despite the fact that there is a higher-than-random frequency of the correct first move, play deteriorates as the round progresses. However, subjects do learn, and the fraction of correct first moves increases to 61% in the last 15 rounds, but with the same pattern of deterioration.

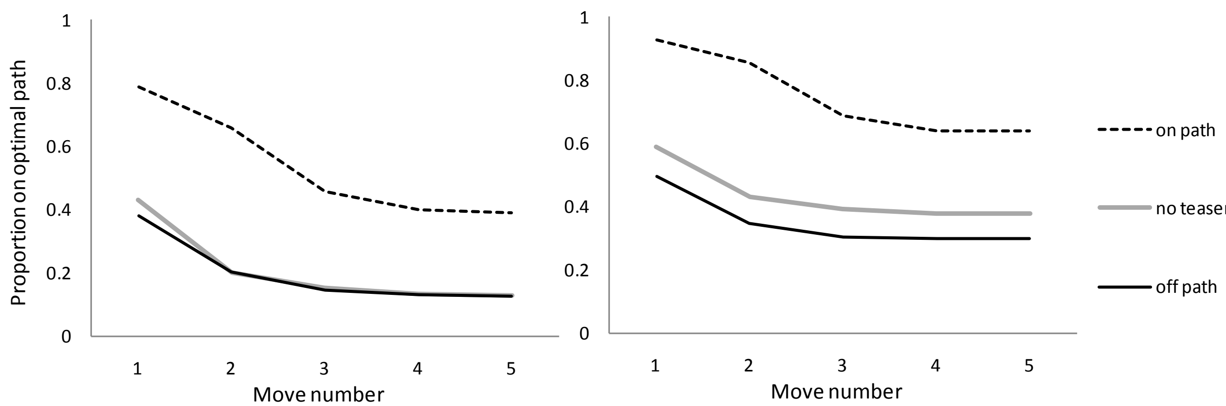
An on-path teaser shifts the curve upward in both the first fifteen rounds and in the last fifteen. Furthermore, in the first fifteen rounds subjects make the correct first choice 80% of the time. The same pattern of deterioration shows as in the no-teaser case, so the primary benefit of

[5] For these significance tests, the unit of observation is the fraction of wins for a single subject.

the on-path teaser comes in helping subjects make the right first move. Surprisingly, the same vertical shift persists into the last fifteen rounds, so the impact of the on-path teaser is long-lasting and is not made up for through learning.

As Figure 1 shows, an off-path teaser has no impact on the ability to follow the optimal path in the first fifteen rounds. It does have an effect in the last fifteen, though, and this effect comes through a decrease in learning, as the off-path curve represents a smaller upward shift than the no-teaser curve. The deterioration pattern remains unchanged, however, so the primary impact of off-path teasers is to impede the ability of subjects to learn the correct first move.

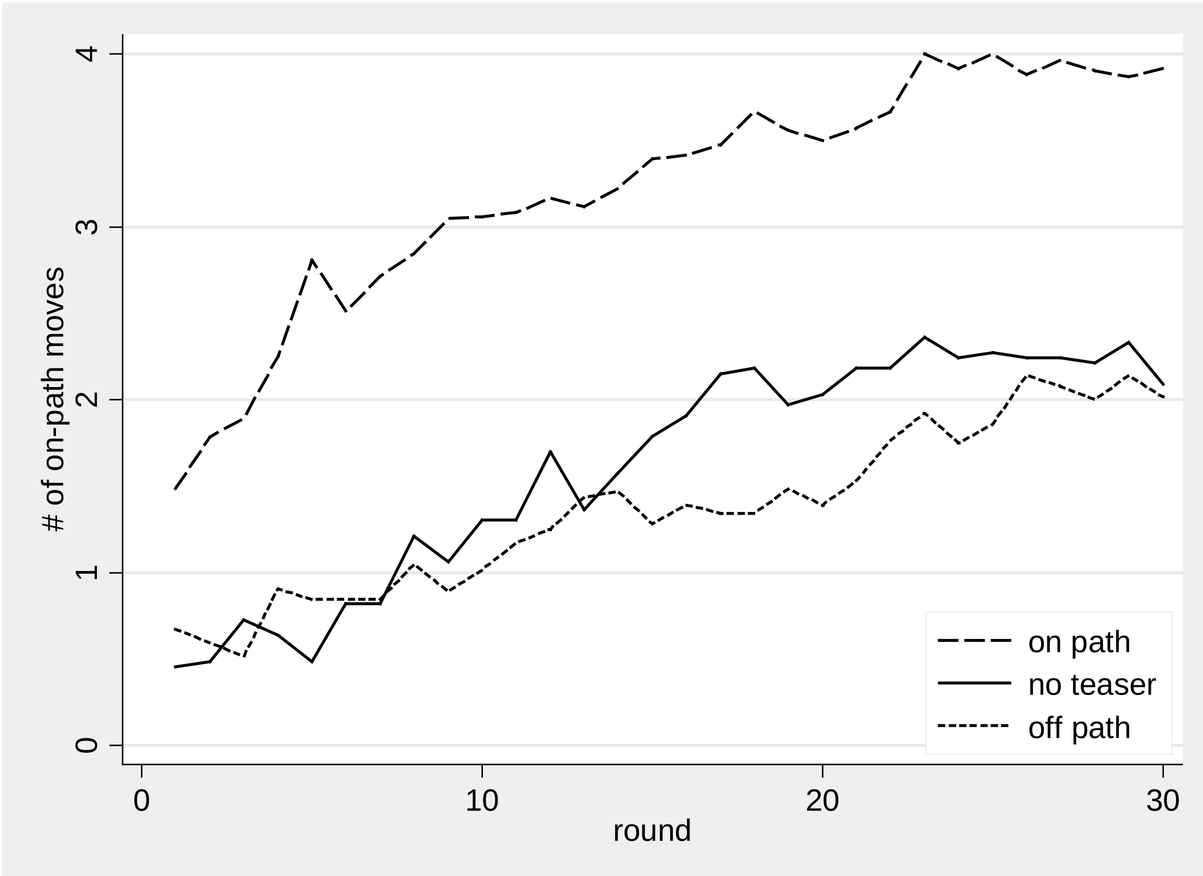
FIGURE 1. THE EFFECTS OF TEASER PLACEMENT



Rounds 1-15 Rounds 16-30

Figure 2 graphs the data differently to show both the existence of learning and the supremacy of on-path teasers. Each curve shows the average number of correct moves in each of the 30 rounds. On-path teasers have an immediate impact, doubling the average number of correct moves in the very first round. All three curves have an upward slope, reflecting the learning that takes place as the game is played repeatedly. Off-path teasers have no immediate impact, as the off-path and no-teaser curves show very similar patterns at the beginning. Midway through the repetitions, though, off-path teasers seem to provide a distraction, but only a temporary one and both the off-path and no-teaser curves end in the same place in round 30.

FIGURE 2: LEARNING BETWEEN ROUNDS



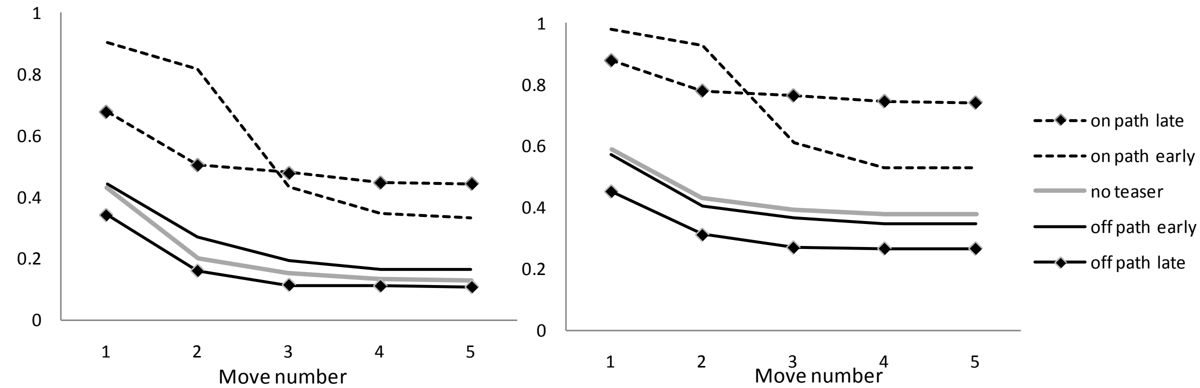
Figures 1 and 2 do not tell the entire story for off-path teasers, though. In the first round, 82% of subjects take the off-path teaser, and *every* subject takes the teaser at least once by round 5.[6] The teaser-taking trend is down with time, but in each and every round over 50% of subjects take the teaser, and nearly two-thirds of subjects take the teaser in a majority of rounds. Thus, teasers not only slow learning (which would imply that subjects eventually give up going for it), but also seem to preclude learning for some subjects (or, alternately, subjects who will not learn just take the money).

One can gain further insight by comparing behavior when the teaser payments come early (stones 5 or 6) or late (stones 11 or 13). Achieving an early on-path teaser requires two correct moves, while achieving a late one requires three or four correct moves, depending on the game configuration. Figure 3 reprises the information in Figure 1, but with the teaser treatments broken down by the timing of the teaser. As Figure 3 shows, early on-path teasers increase the probability that subjects make the right first two moves, but lead to a large drop-off for the third move. Late on-path teasers, on the other hand, show very little deterioration, especially in the last fifteen

[6] These numbers ignore data from the (21,4,5) game, because when subjects move first the teaser on stone 5 is unattainable.

rounds. Late teasers are harder to get, but when subjects get them they learn the game and have a greater chance of achieving the winning stone.

FIGURE 3: THE EFFECTS OF TEASER TIMING



Rounds 1-15 Rounds 16-30

Off-path teasers have only small effects, but in the opposite direction. Early teasers encourage learning, with more subjects achieving the final stone with an early off-path teaser than with a late one. Also, it becomes apparent that most of the learning difference between off-path and no-teaser treatments found in Figure 1 can be explained by the poor performance with the late off-path teasers. In fact, the probabilities of making correct moves in the last fifteen rounds with a late off-path teaser are indistinguishable from the probabilities in the *first* fifteen rounds with *no* teaser. Late off-path teasers provide the biggest distraction from learning.

TABLE 3. Probit analysis of conditional probability of a correct move

	Move 1 on optimal path	Move 2 on path (given move 1 is)	Move 3 on path (given move 2 is)	Move 4 on path (given move 3 is)
Teaser is:				
on path early	1.612***	1.303***	-0.800***	-0.522*
on path late	0.790***	0.693***	0.814***	0.199
off path early	0.000	0.150	-0.084	-0.226
off path late	-0.293	-0.051	-0.23	0.506
action space =4	0.077	0.232	-0.085	0.209
round #	0.032***	0.036***	0.030***	0.032***
constant	-0.523***	-0.444*	0.607**	0.923***
N	5430	3513	2744	2146
X^2 p-value	<0.001	<0.001	<0.001	<0.001

Probit analysis, standard errors clustered by subject. Dependent variables are whether move *m* is on the optimal path, given move *m*-1 was on the optimal path.

Probit analysis verifies the above patterns, as shown in Table 3. The dependent variable is the conditional probability of making the correct move given that one is on the optimal path, while explanatory variables are the round number, the action space with a choice of one to three stones as the omitted category, and the teaser placement with no teaser as the omitted category. As expected, conditional probabilities increase with the round number, showing learning. The size of the action space has no significant impact. Early on-path teasers help enormously for the first two moves, but those who make the second move correctly with either no teaser or an off-path teaser are significantly more likely to make a correct third or fourth move than someone who faced an early on-path teaser. Lat on-path teasers, in contrast, aid subjects in making the correct third move.

Ultimately, our hypotheses regard learning, which we measure by whether a subject wins the game. Table 4 shows the results of a probit analysis on the probability of following the optimal path the whole way through and winning the round. Model (1) includes as explanatory variables only whether the game had an on-path or off-path teaser, with the omitted category no teaser. It shows that on-path teasers help subjects learn the game. Model (2) breaks the teasers down by timing as well as location, showing that late on-path teasers lead to more learning than early on-path teasers. Model (4) adds the action space and the round number, and shows evidence of learning as the game progresses through the 30 rounds. Model (4) separates out the treatment (21, 4,5) which has an unattainable teaser because no matter what the subject's first move is, the computer either takes the teaser (if the subject chooses 1) or bypasses the teaser to get to the optimal path (if the subject chooses 2, 3, or 4). This unattainable teaser leads to the highest learning rate, and an attainable early off-path teaser leads to the lowest learning rate. Unachievable teasers may be a very good learning tool, but early achievable off-path teasers are have a negative effect comparable to the best positive impact of an on-path teaser.

TABLE 4. Probit analysis of probability of winning the game

	(1)	(2)	(3)	(4)
Teaser is:				
on path	0.698***			
on path early		0.488*	0.576*	0.549*
on path late		0.900***	1.005***	0.988***
off path	-0.138			
off path early		0.003	0.028	-0.927*
off path late		-0.230	-0.231	-0.239
unachievable				1.412**
action space =4			0.212	0.062
round #			0.047***	0.048***
constant	-0.660***	-0.660***	-1.580***	-1.500***
N		5430	5430	5430
X^2 p-value		<0.001	<0.001	<0.001

Probit analysis, dependent variable="win", standard errors clustered by subject.

Hypothesis 1 regarding learning in the first race-to-21 game stated that the highest learning rate occurs when there is a teaser on the optimal path and the lowest occurs when the teaser is off the optimal path. The data support the hypothesis that on-path teasers increase learning compared to games without teasers. The hypothesis for off-path teasers was not supported. Early attainable off-path teasers hurt learning, but early unattainable off-path teasers aid learning. Teasers in the middle of the game have no statistically significant effect.

V. TRANSFERABILITY OF KNOWLEDGE

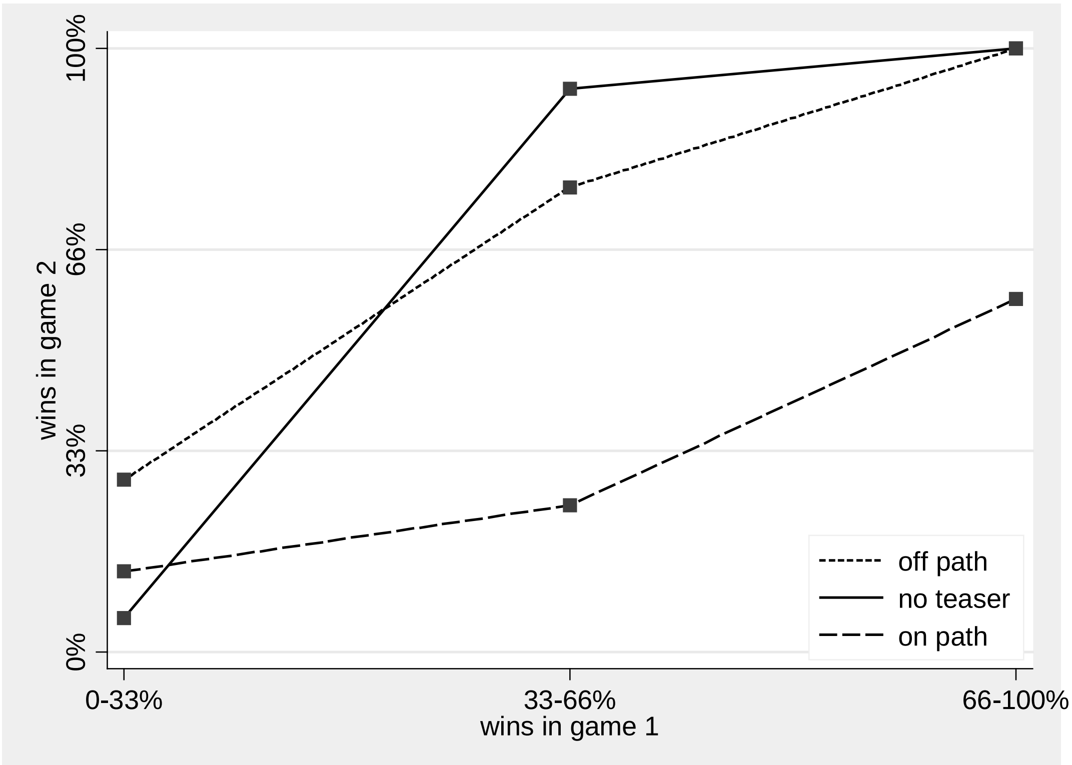
The preceding section showed that on-path teasers aid learning, as suspected. But what, exactly, do subject learn? Do they learn that the game can be solved through backward induction so that when the game changes they can still win, or do they simply learn the optimal path for this particular game so that they must start from scratch when the game changes? The former learning is transferable, while the latter is non-transferable. This section explores the transferability of the knowledge subjects gain when they learn the first race game during the first 30 rounds.

To address these questions we compare performance in the first 30 rounds to performance in the last 15, with the game changing between rounds 30 and 31 when the action space switches from drawing 3 stones to drawing 4, or vice versa. Figure 4 provides a first look at the data. The horizontal axis measures the fraction of rounds won in the first game (rounds 1 – 30) and the vertical axis measures the fraction rounds won in the second game (rounds 31-45). Outcomes for the first game are sorted into three discreet groups: those who won fewer than 10 of the 30 games, those who won 10 to 19, and those who won at least 20. The graph plots three curves corresponding to the placement of the teaser in the first game. Importantly, changing the action space moves on-path teasers off the optimal path, and off-path teasers onto the optimal path. No-teaser games still have no teaser when the action space switches.

To understand the graph, look first at the solid curve for the no-teaser treatment. The left-most point shows that those who did poorly at the first game, winning fewer than 10 rounds, also did poorly in the second game, winning only 8% of the rounds. In contrast, those who won at least 10 rounds in the first game were very successful in the second, and those who won at least 20 rounds in the first game won virtually every round in the second game. This suggests that learning is highly transferable for no-teaser games.

The off-path curve shows a different pattern, and this curve pertains to subjects who faced an off-path teaser in the first game and an on-path teaser in the second. The left-most point is higher for the off-path curve than the no-teaser game, reflecting the fact that the on-path teaser game (which is the second game) is easier than the no-teaser game. Some of the subjects who performed poorly with the original off-path teaser game were, nevertheless, able to figure out the subsequent on-path teaser game. The curve then slopes upward, and subjects who won between 10 and 19 rounds in the off-path game subsequently won 72% of the rounds in the on-path game. Those who won at least 20 rounds in the off-path game went on to win virtually all rounds of the second game. Combining this finding with those of the preceding section suggest that although games with off-path teasers are more difficult to learn, those who learn them gain highly transferable knowledge.

FIGURE 4. Winning rates across the two games



Both of these findings, for the no-teaser and off-path teaser games, contrast markedly with the results from the on-path teaser games. As shown in the preceding section, the on-path teaser games are much easier to learn, with subjects winning the final stone about twice as often than in the other treatments. However, those who win the on-teaser game between 10 and 19 times only win the off-teaser game about 28% of the time, and those who won the on-teaser game at least 20 times only won the off-teaser game about 60% of the time. The on-teaser game may be much easier to learn, but these data suggest that the knowledge gained is not very transferable.

Analysis based on Figure 4 masks the impact of the difficulty of the second game. To differentiate the difficulty of game 2 from the transferability of game-1 learning, we examine the *immediate* transferability of backward induction. Defining "learned game 1" as winning two of the last three rounds of game 1 and "knowing game 2" as winning two of *first* three rounds of game 2, we can operationalize the definitions of the three types of learners described in Section II. Non-learners do not "learn game 1." Transferable learners both "learn game 1" and "know game 2," while non-transferable learners "learn game 1" but do not "know game 2." Table 5 classifies subjects according to these definitions.

TABLE 5. Classification of subjects by learning type

Treatment	Fraction Non-learners	Fraction Learners: Non-transferable	Fraction Learners: Transferable	Fraction of learning that is transferable
No teaser	60.61	3.03	36.36	92.3%
Off path	64.06	10.94	25.00	69.6%
Early	58.33	12.50	29.17	70.0%
Late	67.50	10.00	22.50	69.2%
Achievable	70.59	9.80	19.61	66.7%
Not achievable	38.46	15.38	46.15	75.0%
On path	29.76	57.14	13.10	18.7%
Early	39.02	41.46	19.51	32.0%
Late	20.93	72.09	6.98	8.8%

The first column of Table 5 provides an alternative to the analysis of Section IV, but with the same results. In that section "learning" was defined on a round-by-round basis, whereas here it is defined as winning two of the last three rounds of the first game. The table shows that games with on-path teasers are easier to learn than games with off-path teasers, and later on-path teasers generate more learning than early ones do. Games with off-path teasers or no teasers are more difficult to learn, with the exception of the game with an unachievable off-path teaser, which 62% of subjects learned.

In the theory section we proposed two competing, but non-exhaustive, hypotheses. Hypothesis 2a is based on the notion that games that are harder to learn generate more transferable learning, while Hypothesis 2b is based on the idea that on-path teasers allow subjects to solve two games instead of one, and figuring out two games makes learning more transferable. Accordingly, Hypothesis 2a predicts that games with lower learning rates yield higher fractions of learners for whom the knowledge is transferable. In contrast, Hypothesis 2b predicts that, compared to games without on-path teasers, games with on-path teasers generate higher fractions of learners for whom the learning is transferable.

Treatments without teasers generate the highest transferability rate, that is, the highest fraction of learners for whom the knowledge is transferable, at 92%. Off-path teaser treatments produce the second-highest transferability rate at 70%, and on-path teasers yield by far the lowest transferability rate at 19%. These numbers contradict Hypothesis 2b and the notion that on-path teasers generate more transferable learning, but they are consistent with Hypothesis 2a and the notion that more difficult tasks generate a better quality of learning.

Table 5 shows a strong positive correlation between the first and last columns, that is, the fraction of non-learners and the fraction of learners who are transferable learners. Higher numbers in the non-learner column mean that the first game is more difficult, so the positive correlation suggests that institutions that encourage winning discourage true "learning" in that they encourage subjects to identify the optimal path but not to understand the game. Thus, teasers that help find the optimal path encourage the quantity but not the quality of learning. Furthermore, early on-path teasers generate a 50% increase in the fraction of subjects who learn for the first game compared to no-teaser treatments, but a 50% *decrease* in the fraction of the overall subject pool whose learning is transferable, and a 13-fold increase in the fraction whose learning is non-transferable. These effects are even more pronounced for late on-path teasers, which double the fraction of learners but cut the fraction of transferable learners by 80% while increasing the fraction of non-transferable learners by a factor of 24.

The qualitative results in Table 5 are robust to many alternative ways of measuring learning and transferability. Also, as a robustness check, we can also ask what each type of learner did in the last three periods of game 2. 100% of transferable learners won in at least two of the last three rounds, 66% of non-transferable learners did so, and 21% of non-learners. This suggests that behavior in the first three periods of game 2 was not by luck or accident, but really was driven by what subjects learned in game 1.

The highest transferable-learning rate, at 46% of all subjects, arose in a game in which the teaser was unattainable. When subjects can pick up 4 stones at a time the optimal path is (1,6,11,16,21). If the teaser is on the fifth stone subjects cannot possible attain it. but the teaser is on the fifth stone, the computer will not allow the subject to reach stone 5. If the subject follows the optimal path by choosing stone 1, the computer will take the teaser. if the subject chooses stone 2, 3, or 4 instead, the computer bypasses the teaser to get on the winning path with stone 6. This game yields both a relatively high learning rate at 62% and a high transferability rate at 75%. Perhaps the thought process subjects follow in trying to figure out how to get the (unattainable) teaser leads them to learn how the game works. In essence, these subjects face a sequence of three different games, not two. First, they play the game in which they attempt to reach the teaser. They learn something from that game, and what they learn is highly transferable. They then apply it to the whole race-to-21 game, and are able to win. The learning remains transferable, and they are able to win the new game that has a different action space.

This game-within-a-game logic might also explain behavior in the on-path teaser games. Subjects play a sequence of three games. They first learn to get the teaser, which is highly non-transferable. They then learn to get the final stone, which is again non-transferable. They have still learned little of use by the time they face the third game in round 31, and transferability rates are low.

VI. References

Ashraf, Nava, Dean Karlan, and Wesley Yin (2006), "Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines," *Quarterly Journal of Economics* 121, 635-672.

Brown, Alexander L., Zhikang Eric Chua, and Colin F. Camerer (2009), "Learning and visceral temptation in dynamic saving experiments," *Quarterly Journal of Economics* 124, 197-231.

Charness, Gary and Uri Gneezy (2009), "Incentives to exercise," *Econometrica* 77, 909-931.

DellaVigna, Stefano and Ulrike Malmendier (2004), "Contract design and self-control: theory and evidence," *Quarterly Journal of Economics* 119, 353-402.

Duflo, Esther, Michael Kremer, and Jonathan Robinson (2010), "Nudging farmers to use fertilizer: theory and experimental evidence from Kenya," mimeo, MIT.

Fryer, Roland G., Jr. (2010), "Financial incentives and student achievement: evidence from randomized trials," NBER Working Paper No. 15898.

Houser, Daniel, Michael Keane, and Kevin McCabe (2004), "Behavior in a dynamic decision problem: an analysis of experimental Bayesian type classification algorithm," *Econometrica* 72, 781-822.

Mills, Gregory, William G. Gale, Rhiannon Patterson, Gary V. Engelhardt, Michael D. Eriksen, and Emil Apostolov (2008), "Effects of individual development accounts on asset purchases and saving behavior: Evidence from a controlled experiment," *Journal of Public Economics* 92, 1509-1530.