Learning in a Black Box*

H.H. Nax†, M.N. Burton-Chellew‡, S.A. West§, H.P. Young¶

April 18, 2013

Abstract

Many interactive environments can be represented as games, but they are so large and complex that individual players are in the dark about what others are doing and how their own payoffs are affected. This paper analyzes learning behavior in such 'black box' environments, where players' only source of information is their own history of actions taken and payoffs received. Specifically we study repeated public goods games, where players must decide how much to contribute at each stage, but they do not know how much others have contributed or how others' contributions affect their own payoffs. We identify two key features of the players' learning dynamics. First, if a player's realized payoff increases he is less inclined to change his strategy, whereas if his realized payoff decreases he is more inclined to change his strategy. Second, if increasing his own contribution results in higher payoffs he will tend to increase his contribution still further, whereas the reverse holds if an increase in contribution leads to lower payoffs. These two effects are clearly present when players have no information about the game; moreover they are still present even when players have full information. Convergence to Nash equilibrium occurs at about the same rate in both situations.

JEL classifications: C70, C73, C91, D83, H41

Keywords: learning, information, public goods games

*Correspondence to: H. H. Nax, 48 Boulevard Jourdan, Bureau A201, Paris School of Economics, 75014 Paris, France.

†Heinrich H. Nax, Paris School of Economics, École Normale Supérieure. Nax's research was in part supported by the Office of Naval Research grant N00014-09-1-0751 and under project NET of the Agence Nationale de Recherche.

‡Maxwell N. Burton-Chellew, Department of Zoology, University of Oxford.

§Stuart A. West, Department of Zoology, University of Oxford.

¶H. Peyton Young, Department of Economics, University of Oxford. Young's research was supported in part by the Office of Naval Research grant N00014-09-1-0751.

1 Introduction

Many interactive environments can be represented as games, but they are so complex and involve so many individuals that for all practical purposes the game itself is unknown to the players themselves. Examples include bidding in on-line markets, threading one's way through urban traffic, or participating in a group effort where the actions of the other members of the group are difficult to observe (guerrilla warfare, neighborhood watch programs, tax evasion). In each of these cases a given individual will have only the haziest idea of what the other players are doing yet but their payoffs are strongly influenced by the actions of the others. How do individuals learn in such environments and under what circumstances does their learning behavior lead to equilibrium?

In this paper we investigate these questions in a laboratory environment where the underlying game has the structure of a public goods game. Players take actions and receive payoffs that depend on others' actions, but (in the baseline case) they have no information about the others and they do not know what the overall structure of the game is. All they know is the amount that they themselves contributed and the payoff they received as a result. Learning in such environments is said to be completely uncoupled (Foster & Young, 2006; Young, 2009; see also Hart & Mas-Colell, 2003, 2006;). In this setting many of the standard learning rules in the empirical literature, such as experience-weighted attraction, k-level reasoning, and imitation, do not apply because these rules use the actions of the other players as inputs (Björnerstedt & Weibull, 1993; Stahl & Wilson, 1995; Nagel, 1995; Ho et al., 1998; Camerer & Ho, 1999; Costa-Gomes et al., 2001; Costa-Gomes & Crawford, 2006).[1]

Nevertheless our experiments show that learning 'inside the box' can and does take place.[2] The learning process exhibits certain distinctive patterns that have been identified in other settings. Two especially noteworthy features are the following:

Search volatility. A decrease in a player's realized payoff triggers greater variance in his choice of action next period, whereas an increase in his realized payoff results in lower variance next period.

Trend-following. If a higher contribution leads to a higher payoff this period, the player will tend to raise his contribution next period, whereas if a higher contribution leads to a lower payoff this period, he will tend to reduce his contribution next period. Similarly, if a lower contribution leads to a higher payoff the player will tend to lower his contribution, whereas if a lower contribution leads to a lower payoff he will tend to increase his contribution next period.

Learning rules with high and low search volatility have been proposed in a variety of set-

[1]We also note that social preferences cannot enter into subjects' behavior because they cannot compare their own payoffs with others' payoffs. Thus the black box environment eliminates some of the explanations that have been advanced for behavior in public goods games (Fehr & Schmidt, 1999; Fehr & Gachter, 2000, 2002; Fehr & Camerer, 2007).

[2]Oechssler & Schipper (2003) show that, in the context of 2x2 games, players can learn to identify the other player's payoff structure on the basis of their own payoffs and information about other's actions. In our set-up players may or may not learn the game, but even without information about the relationship between their own and others' payoffs they do learn to play Nash equilibrium.

tings, including biology (Thuijsman et al., 1995), organizational learning (March, 1991), computer science (Eiben & Schippers, 1998), and psychology (Coates, 2012). (In computer science and organization theory this type of learning is known as 'exploration versus exploitation', whereas in biology it is known as 'fast and slow' learning.) Search volatility is also a crucial feature of rules that have recently been proposed in the theoretical literature; in particular there is a family of such rules that lead to Nash equilibrium in generic n-person games (Foster & Young, 2006; Germano & Lugosi, 2007; Marden et al., 2009; Young, 2009; Pradelski & Young, 2012). Trend-following also has antecedents in the experimental literature, where it is sometimes called directional learning (Selten & Stoecker, 1986; Selten & Buchta, 1998). Directional learning describes behavioral dynamics based on local adjustments of strategy and aspiration (Sauermann & Selten, 1962; Cross, 1983). The underlying behavioral components are illustrated in a bargaining game: if a higher demand results in a higher realized payoff the agent increases his demand, whereas a decrease in realized payoff leads to a demand decrease (Tietz & Weber, 1972; Roth & Erev, 1995).

In this paper we show that these two patterns help explain the behavior of subjects who are learning in a 'black box' environment. Moreover, despite the lack of information and the simplicity of the adjustment rules, the behavior of the group approaches Nash equilibrium and does so at approximately the same rate as in the situation where subjects have full information (Burton-Chellew & West, 2012). As we shall see, this is true both when free riding is the dominant strategy, and also when full contribution is the dominant strategy. We conclude that even when subjects have virtually no information about the game, learning follows certain distinctive patterns that have been documented in a wide variety of learning environments, and in public goods games this behavior leads to Nash equilibrium at a rate that is very similar to learning under full information.

We structure the paper as follows. In section 2 we describe the experimental set-up in detail. In section 3 we present the empirical findings together with statistical tests of significance. We conclude in section 4.

2 Experimental set-up

Participants were recruited from a subject pool that had not previously been involved in public goods experiments. The subjects were not limited to university students, but included different age groups with diverse educational backgrounds.[3] Each subject played four separate contribution games, where each game was repeated twenty times. The twenty-fold repetition of a given game will be called a 'phase' of the experiment. This yields a total of 236 games and 18,880 observations. During sixteen separate sessions in groups of sixteen or twelve, subjects play the four phases in a row.

Games differ with respect to two *rates of return* ('low' and 'high') such that either 'free-

[3]The subjects were recruited through the Online Recruiting System for Economic Experiments (Greiner, 2004) with this request specifically made. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). All experiments were conducted at the Centre for Experimental Social Sciences (CESS) at Nuffield College, University of Oxford.

riding' or 'full-contributing' is the dominant strategy. Treatments differ with respect to the information that is revealed to the agents ('black box', 'standard' and 'enhanced'). Every player plays both black box games, and either the two standard or enhanced information games. Either two black box games are played first, or two games of one of the other two treatments.[4] The two first phases constitute stage one, the second two constitute stage two. Each possible information order is played in four separate sessions, each of which with a different order of the two rates of return.[5]

In this section, we shall describe the underlying stage game, the structure of each repeated game, and the different information treatments.

2.1 Linear public goods game

Consider the following linear public goods game. Every player i in population $N = \{1, 2, ..., n\}$ makes a nonnegative, real-numbered contribution, a_i, from a finite budget $B \in \mathbb{N}^+$. Given a resulting vector of all players' contributions, $\mathbf{a} = \{a_1, a_2, ..., a_n\}$, for some *rate of return*, $e \geq 1$, the public good is provided, $R(\mathbf{a}) = e \sum_{i \in N} a_i$, and split equally amongst the players.[6] Given others' contributions $\mathbf{a} \setminus a_i$, player i's contribution a_i therefore results in a payoff of

$$\phi_i = e/n \sum_{i \in N} a_i + (B - a_i),$$

where e/n is the marginal rate of return. Write $\boldsymbol{\phi}$ for the payoff vector $\{\phi_1, ..., \phi_n\}$.

Nash equilibria. Depending on whether the rate of return is low ($e < n$) or high ($e > n$), individual contribution of either zero (free-riding) or B (full-contributing) is the strictly dominant strategy for all players; the respective Nash equilibrium results in either nonprovision or full provision of the public good.[7]

2.2 Repeated game

In any session, the same population S (with $|S| = 12$ or 16) plays four phases. Each phase is a twenty-times repeated, symmetric public goods game. In each period t of each game, players in S are randomly matched to groups of four to play the stage game. The budget every period is a new $B = 40$ to every player of which he can invest any amount, but he cannot invest money carried over from previous rounds. The rate of return is either *low* ($e = 1.6$) or *high* ($e = 6.4$) throughout the game.[8] Write N_4^t for any of the four-player

[4] Of the 18,880 observations, 9,440 are black box, 4,640 are standard, and 4,800 are enhanced.

[5] This yields sixteen different scenarios each played in one of the sessions given by the possible orders and permutations; four (combinations) times four (orders) equals sixteen (sessions).

[6] If $e < 1$, $R(\mathbf{a})$ is a public bad.

[7] When $e = n$, any contribution is a best reply to any level of contribution by the others, hence any level of provision is supported by at least one pure-strategy Nash equilibrium.

[8] The Nash equilibrium payoffs of the stage games are 40 when $e = 1.6$, and 256 when $e = 6.4$.

groups matched at time t, and ρ_4^t for the partition of S into such groups. Given $\mathbf{a}^t$, ρ_4^t, N_4^t, in any period t such that $i \in N_4^t \in \rho_4^t$, each i receives

$$\phi_i = \sum_{t=1,\ldots,20} \phi_i^t = \sum_{t=1,\ldots,20} [e/4 \sum_{j \in N_4^t} a_j^t + (40 - a_i^t)].$$

For every i, ϕ_i represents a real monetary value that is paid after the game.[9]

2.3 Information treatments

Before the first game begins, each player is told that two separate experiments will be played. (From the analyst's point of view the experiment consists of both of these taken together.) At no point before, during, or in between the separate phases of the experiment are players allowed to communicate. Depending on which treatment is played, the following information is revealed at the start of each stage when a new treatment begins.

All treatments. Two separate but possibly different twenty-times repeated stage games are played. In each, the underlying stage game is unchanged for 20 periods. Each player receives 40 coins each period of which he can invest any amount. Moreover, after investments are made, each player gets a nonnegative return each period which, at the end of each game, he receives together with his uninvested money according to a known exchange rate.

Black box. No further information about the structure of the game, or about other players' actions or payoffs is revealed. Throughout the game, therefore, players only know their own investments and payoffs.

Standard. The rules of the game are revealed including production of the public good, rate of return, and how groups in the session form to play each stage game. In addition, at the end of each period, players receive a summary of the relevant contributions in their previous-period group; not knowing who these players are and what payoffs they receive.[10]

Enhanced. In addition to the information available in the standard treatment, players learn about the other group members' previous-period payoffs.

(See Appendix A for full black box instructions and Appendix B for the output screens displayed during the experiment for each treatment.)

3 Findings

First, we analyze the black box data when players have no prior experience of public goods games. Second, we assess the effects of experience. Third, we assess the effects of

[9] Hundred coins are worth £0.15 to pay £19.2 maximal individual earnings from the whole experiment.

[10] This follows the standard information treatment design and procedure of rematching subjects as introduced by Andreoni (1988).

information being revealed about the game structure and about other players' previous contributions (and payoffs). We carry out the analysis for both high ($e = 6.4$) and low ($e = 1.6$) rates of return. For each step of the analysis, we first use nonparametric tests to establish distributional differences in the different treatments and game, and then use regressions controlling for phase, period, individual, group and past payoff and contribution effects.

3.1 Black box

In this section, we analyze the black box treatments when played before the standard or enhanced treatments. There were eight such sessions and 4,960 observations. By choosing these sessions, we ensure that players not only have no explicit knowledge of the game structure and no information about others, but also have no "experience" of the game.[11] Unless stated otherwise, "significance" refers to a 95% confidence interval when rejecting an opposite null and to a 90% confidence interval when a hypothesis cannot be rejected.

3.1.1 Contributions

The mean initial contribution lies between one third and half the budget.[12] Using period-specific Mann-Whitney tests, we cannot reject the null of equal contributions for games with different rates of return and for phase one versus phase two, that is, similar observations are made in all games (with different rates of return) played in all initial black box treatment games. This is evidence of "restart effects" (observed throughout the literature).[13] The period-two adjustments to the initial contribution have means close to zero.[14] Notably, a large part of these contributions lie at 0, 10, 20, 30, 40, with several contributions made also in the range from 1 to 9; contributions at other numerical values are rare.[15] Using period-specific Mann-Whitney tests, we cannot reject the null of equal contributions for different rates of return for periods 1 through to period 3.

(Figure 1 illustrates.)

In the early periods of black box treatments, this is the behavior one would expect and that this is consistent with hundreds of previous experiments on public goods.[16] Since players have no knowledge of the differences in the game structures, they initially make random contributions and play different games in much the same way. Note that players initially contribute less than half of their budget, another observation consistent with previous experimental results. A possible explanations for this is ambiguity aversion

[11]Recall subjects were explicitly chosen to not have experience from previous experiments. In Section 4, we will test for experience and information effects by comparison with the rest of the data.

[12]The mean is 15.3 in the games with $e = 1.6$, and 13.9 in the games with $e = 6.4$.

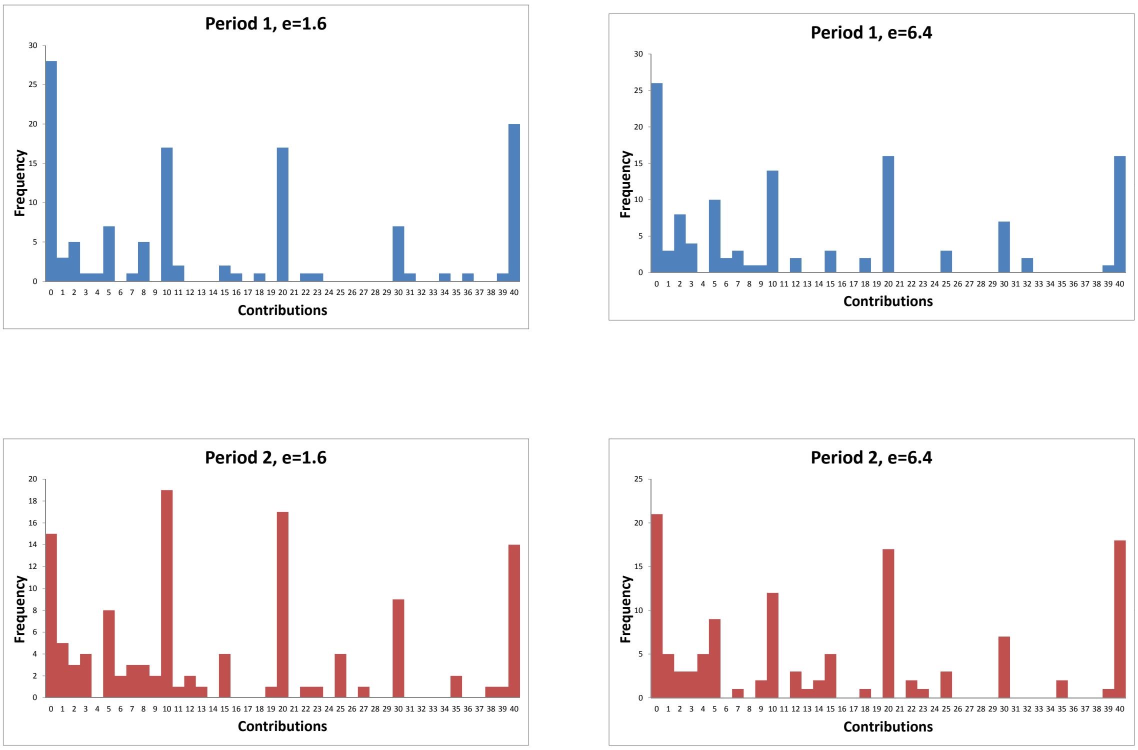
[13]See Andreoni, 1988; Croson, 1996; Cookson, 2000; Camerer, 2003.

[14]Means are 0.4 when $e = 1.6$ and 1.6 when $e = 6.4$.

[15]As is usually the case; see Camerer, 2003.

[16]As surveyed, for example, in Chaudhuri, 2011.

Figure 1: Initial contributions (phases 1,2).



Initial contributions often lie at 0, 10, 20, 30, 40, with several contributions made also in the range from 1 to 9; contributions at other numerical values are rare.

(Epstein, 1999).[17] At the beginning of the game, players cannot judge their initial payoffs to be particularly "good" or "bad", and therefore have no comparison concerning the performance ("success" vs. "failure") of different strategies. One would consequently not expect an obvious direction for the initial-period adjustments, and indeed they are similarly random until period 3 for the different rates of return.

(Figure 2 illustrates.)

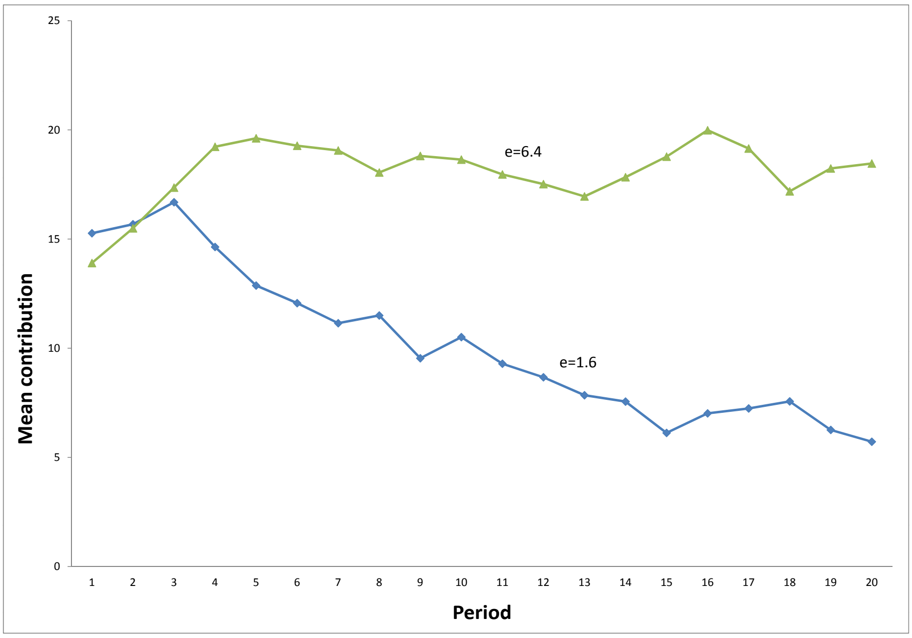
If players were not to learn at all, one would expect similar behavior throughout the games. Using period-specific Mann-Whitney tests, however, we reject the null of equal contributions for different rates of return in all periods after period 3. Moreover, play evolves following distinctly different patterns: the contributions towards the end of the game are closer to the respective stage game Nash equilibrium.[18] Simple linear regressions of contributions on a time trend (controlling for phase and group) reveal a trend of -0.6 when $e = 1.6$ (significant at 99%), and of +0.1 when $e = 6.4$ (significant at 90%). Note the trend is accentuated in the games with $e = 1.6$.[19] When $e = 6.4$, the mean contribution

[17]In Fehr & Schmidt, 1999, for example, mean contributions are ca. 40 percent of the budget; this would mean ca. 16 in our experiment.

[18]Means are 5.7 when $e = 1.6$ and 18.5 when $e = 6.4$.

[19]This pattern is consistent with previous experiments as, for example, noted in Ledyard, 1995. Note

Figure 2: Mean black box play (phases 1,2).



Contributions deteriorate in the free-riding game. This trend is accentuated compared to the full-contribution games which display a weak positive trend.

never exceeds twenty and thus remains more than "halfway" from the Nash contribution. Again, players' ambiguity aversion may explain this phenomenon.[20]

(Figure 3 illustrates.)

To summarize our observations regarding aggregate contributions in free-riding games, we reconfirm the standard findings from the literature regarding initial contributions, restart effects and contribution deterioration. Moreover, we note that in games where full-contributing is Nash, the observed trends to Nash equilibrium are substantially weaker. Higher contributions seem harder to learn.

3.1.2 Difference in variation

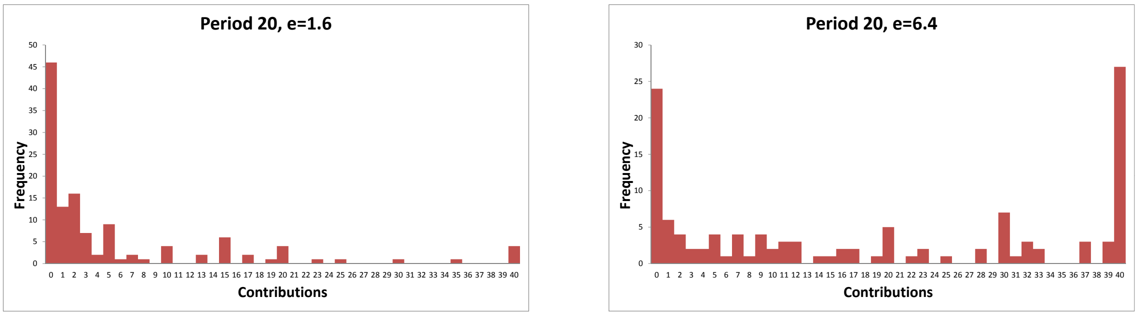
Adjustment. The *adjustment* of a player i in period t is $a_i^{t+1} - a_i^t$.

Success versus failure. A player i experiences *success* in period $t+1$ if his realized payoff does not go down ($\phi_i^{t+1} \geq \phi_i^t$); otherwise he experiences *failure*.

our contributions decline more sharply than in Bayer et al. (2013), possibly indicating a "within-game restart effect" in their experiments since players are lead to believe that underlying stage games are changing.

[20] A contribution decrease represents less ambiguity and is, ceteris paribus, preferred by ambiguity-averse agents, a contribution increase results in more ambiguity (Schmeidler, 1989).

Figure 3: Final contributions.



Final contributions amass close to zero in the free-riding game, and split between contributions close to zero and 40 for the full-contribution game.

Table 1: Adjustments after success and failure (phases 1,2; pooled data).

statistic	success	failure
variance	123.9	188.5
mean	+0.8	-1.2

SEARCH: the variance after success is lower than after failure; the average adjustments are close to zero.

Based on a Levene's test (robust variance), the hypothesis of equal variances of adjustments following success versus failure is rejected with 99% confidence.[21] Furthermore, we regress absolute adjustments on success versus failure controlling for phase, period, individual, group, and two lags of contributions and payoffs. We obtain a significantly (99%) smaller coefficient of adjustment following success (by -1.2) than after failure. The phase dummies are not significant. None of the period dummies is significant after period 3. One of the group dummies is significant. The previous-period payoffs and contributions are significant, those lagged two periods are not significant. We conclude that players' adjustments of successful contributions have a smaller variation than adjustments of unsuccessful contributions. As discussed in sections 1 and 2 of this paper, the phenomenon of "different adjustments following success versus failure" is a feature of "search" in several recent learning models.[22] Subsequently, we shall refer to this feature as SEARCH.

(Table 1 summarizes.)

[21]The robust variance test is not rejected dependent on whether $e = 1.6$ (mean -0.5) or 6.4 (0.2).
[22]See Young, 2009; Marden et al., 2011; Pradelski & Young, 2012.

Table 2: TREND hypothesis.

Player i in period $t+1$	given success in period t ($\phi_i^t \geq \phi_i^{t-1}$)	given failure in period t ($\phi_i^t < \phi_i^{t-1}$)
given an increase in period t ($a_i^t > a_i^{t-1}$)	*increases contribution with respect to period* $t-1$	*decreases contribution with respect to period* t
given a decrease in period t ($a_i^t < a_i^{t-1}$)	*decreases contribution with respect to period* $t-1$	*increases contribution with respect to period* t

3.1.3 Directional learning

Our games have the following structure of which subjects in the black box are unaware. First, contributions range from 0 to 40, and the Nash equilibrium lies at either end. Second, an individual's payoff rises with the contributions of the other players but may rise or fall with his own adjustment depending on the game's underlying rate of return. Importantly, in varying his own contribution, the individual agent is unable to distinguish between a self-induced success versus failure, and one that is caused by others: in the interplay with others, incremental adjustments in either direction may lead to higher or lower payoffs.[23] The structure of the action space calls for a directional learning model (Selten & Stoecker, 1986; Selten & Buchta, 1998).

Given period t, a player i may increase or decrease his contribution relative to his contribution in the previous period. As a result, he experiences either success or failure. Table 2 shows the model of directional adjustments that we propose and shall subsequently refer to as TREND.

To test the TREND hypothesis, we regress adjustments relative to two lags of contributions period controlling for phase, period, individual, group, two payoff lags, and previous contribution for failure and success adjustments separately.

After failure, we obtain significant coefficients for decrease with respect to the previous-period contribution (-3.4; significant at 99%) following an increase, and for increase with respect to the previous-period contribution (+2.2; significant at 99%) following a decrease. None of the phase or group dummies is significant. None of the period dummies after period 4 is significant. Past payoffs have significant but marginal (less than 0.05) effects. The previous contribution has a negative effect (-0.6; significant at 99%). Trend reversal after failure therefore summarizes as follows:

- *If an increase is a failure in the current period, the average next-period adjustment relative to the current period –net of negative level effects– is a decrease instead. Similarly, if a decrease is a failure in the current period, the average next-period*

[23] The effects that a player's own actions and those of others have on one's own payoff are important features in Young, 2009; Marden et al., 2011; Pradelski & Young, 2012.

adjustment relative to the current period –net of negative level effects– is an increase instead.

After success, we obtain significant coefficients for increase with respect to contributions two periods ago (+7.1; significant at 99%) following an increase, and for decrease with respect to contributions two periods ago (-12.5; significant at 99%) following a decrease. The previous contribution has a negative effect (-0.2; significant at 99%). All other coefficients are insignificant as in the failure case. Trend following after success therefore summarizes as follows:

- *If an increase is a success in the current period, the average next-period adjustment relative to the previous period –net of negative level effects– is another increase. Similarly, if a decrease is a success in the current period, the average next-period adjustment relative to the previous period –net of negative level effects– is another decrease.*

Based on TREND, given two previous contributions and a success versus failure stimulus, the relative location of the expected next-period contribution is unknown. This is because TREND implies adjustment tendencies only with respect to the current period after failure, and with respects to the previous period after success. It is unclear where, given both current and previous contributions, the next-period contribution will lie after either success or failure is experienced.

(Figure 4 illustrates.)

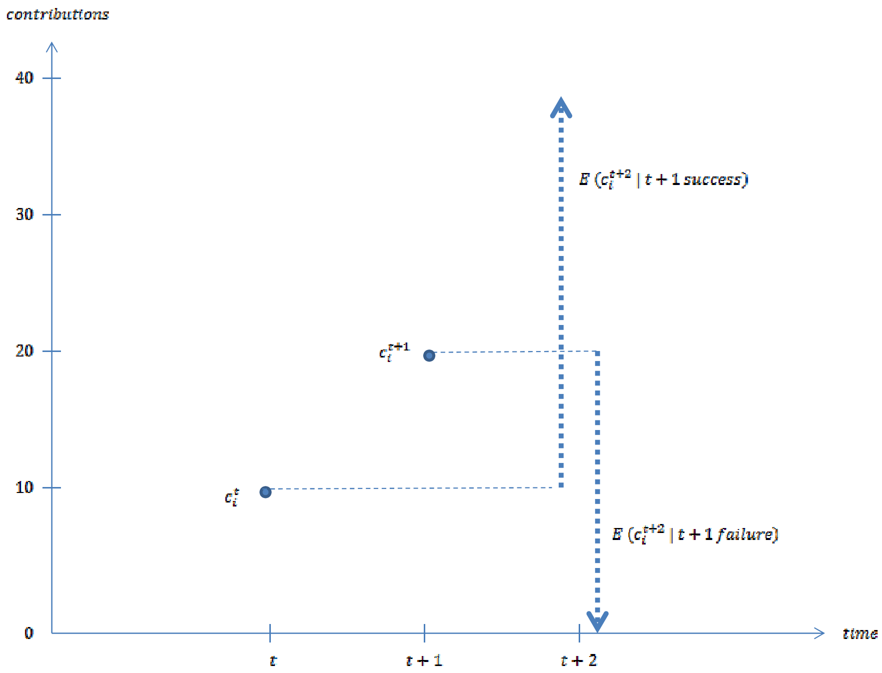
We shall illustrate this point with an example as illustrated in Figure 4, if an agent contributes 10 in one period and 20 in the next, on average he contributes at least 10 if 20 was a success, and less than 20 if 20 was a failure. The following (stronger) hypothesis has been considered (Bayer et al., 2013): given two previous contributions, the next contribution is, in expectation, closer to the contribution which resulted in a higher payoff. If the stimulus is trend-neutral, the expectation lies exactly at the average. As illustrated in Figure 5, if an agent contributes 10 in one period and 20 in the next, on average he contributes at least 15 if 20 was a success, and less than 15 if 20 was a failure.

(Figure 5 illustrates.)

We regress contribution adjustments on the four trends controlling for phase, period, individual, group, two payoff lags, previous contribution. This yields the following effects:

given a_i^t and a_i^{t+1} of some individual i, the net average adjustments, controlling for level effects, are

Figure 4: TREND without level effect.



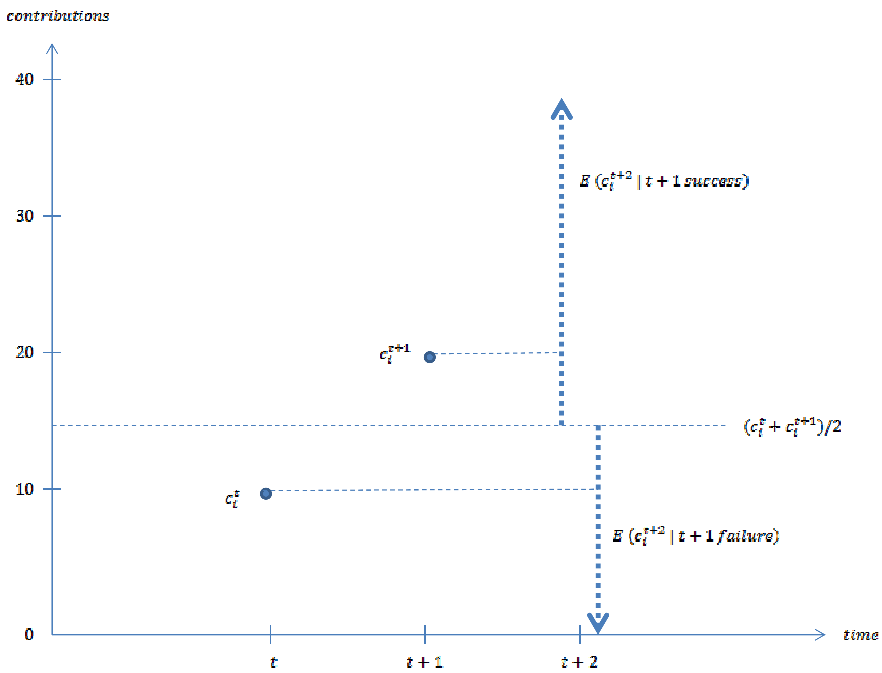
TREND without level effects: success versus failure leaves an intermediate range over which predictions are ambiguous.

$\mathbf{E}(a_i^{t+2}; a_i^{t+1}, a_i^t) - \frac{a_i^t + a_i^{t-1}}{2} =$	$\phi_i^t \geq \phi_i^{t-1}$	$\phi_i^t < \phi_i^{t-1}$
$a_i^t > a_i^{t-1}$	$+3.6^{**}$	$+0.3^{\circ\circ}$
$a_i^t < a_i^{t-1}$	-6.5^{**}	$-1.9^{*\circ}$

**: significant at 99%, consistent with the average hypothesis.
$^{\circ\circ}$: significant below 80%, inconsistent with the average hypothesis.
$^{*\circ}$: significant at 99%, inconsistent with the average hypothesis.

Note that the adjustments following success point in the "right" direction (following success) and are significant. The adjustments following failure point in the "wrong" direction (following failure) and are significant in case of decrease. As before, the previous

Figure 5: Average adjustment hypothesis (Bayer et al., 2013).



The average adjustment hypothesis ignores level effects and places further structure on TREND, making predictions over the intermediate range where TREND remained ambiguous.

contribution has a significant negative effect of -0.4 (with 99% confidence). Previous profits have a marginal but significant effect. If we drop the two payoff lags and the previous contribution from the regression (since they are not included in the average adjustment model), we improve in terms of the model's predictions but not regarding all aspects:

given a_i^t and a_i^{t+1} of some individual i, the gross average adjustments, not controlling for level effects, are

$\mathbf{E}(a_i^{t+2}; a_i^{t+1}, a_i^t) - \frac{a_i^t + a_i^{t-1}}{2} =$	$\phi_i^t \geq \phi_i^{t-1}$	$\phi_i^t < \phi_i^{t-1}$
$a_i^t > a_i^{t-1}$	$+1.4^{*}$	-3.3^{**}
$a_i^t < a_i^{t-1}$	-5.4^{**}	$-0.9^{\circ\circ}$

*: significant at 95%, consistent with the average hypothesis.
**: significant at 99%, consistent with the average hypothesis.
$^{\circ\circ}$: significant at 80%, inconsistent with the average hypothesis.

We conclude that there is strong evidence for TREND including negative previous-period contribution level effects. Ceteris paribus, a higher previous-period contribution leads to a negative adjustment. Depending on trend stimulus (success or failure), success trends are followed with respect to the contribution two periods ago, failure trends are reversed with respect to the previous contribution. The size of the adjustments relative to the previous two-period average is ambiguous and depends on the level of the previous-period contribution as well as the spread of the two previous contributions. A learning model assuming average adjustments without level effects gives consistent predictions in the success and increase directions; given a decrease-failure, however, it makes wrong predictions. It is worth noting that in the context where the average adjustment rule was originally formulated (Bayer et al., 2013), the problematic decrease-failure trends are relatively rare because only games where free-riding is Nash are played. The most common failure in those games comes from contributing more, not less.[24]

3.1.4 Experience

In this section, we analyze the whole black box data including sixteen sessions and 9,440 observations. In 4,480 observations, individuals have "experience" in the sense that they play the black box games in the second stage having previously played a non-black box treatment in the first stage where they received information about the game structure and about other players' past actions (and payoffs). Even though subjects are explicitly told that a separate experiment is started after the first stage of the experiment, there is evidence that experience matters.

First, we test for distributional differences in contributions using period-specific Mann-Whitney tests for the two rates of return separately. The following observations are made. In the initial period, restart effects are noted as previously. For both rates of return separately, the null hypothesis of equal contributions dependent on experience versus no experience cannot be rejected. Furthermore, there is no significant difference in between the two rates of return which is natural since players have no basis to distinguish between

[24]Conversely, in the games where full-contributing is Nash, decreases are the more common failures.

them.[25] If $e = 6.4$, the null hypothesis of equal contributions dependent on experience versus no experience can be rejected with 90% confidence after period 5 except for periods 16 and 17. If $e = 6.4$, it cannot be rejected in any period.

Second, we compare the linear time trends of contributions. Simple linear regressions controlling for phase and group reveal a similar trend as before if $e = 1.6$ (-0.5; significant at 99%) but a stronger drift $e = 6.4$ (+0.3; significant at 99%). Previously the respective trends were -0.6 (99%) and +0.1 (90%). Hence the contributions in games with high rates of return improve faster with experience, but there is no significant difference in play of games with low rates or return.

Finally, we perform the linear regressions with respect to SEARCH (absolute size of adjustments) and TREND (directional adjustments) from the previous section controlling for the same effects and adding experience dummies for both rates of return. We obtain the following results. The absolute size of adjustments (SEARCH) is decreasing in time (all period dummies after period 3 are negative and significant), and decreasing with experience. As before, the previous contribution is a positive and significant factor. Payoff level effects are significant but marginal. All other effects (including phase dummies) are not significant. With respect to directional adjustments (TREND), experience leads to larger contributions but the direction of adjustments is unchanged. As before, the level of previous contributions is a negative and significant factor. All period dummies after period 4 are negative. All other effects are as before.

(Figure 6 illustrates.)

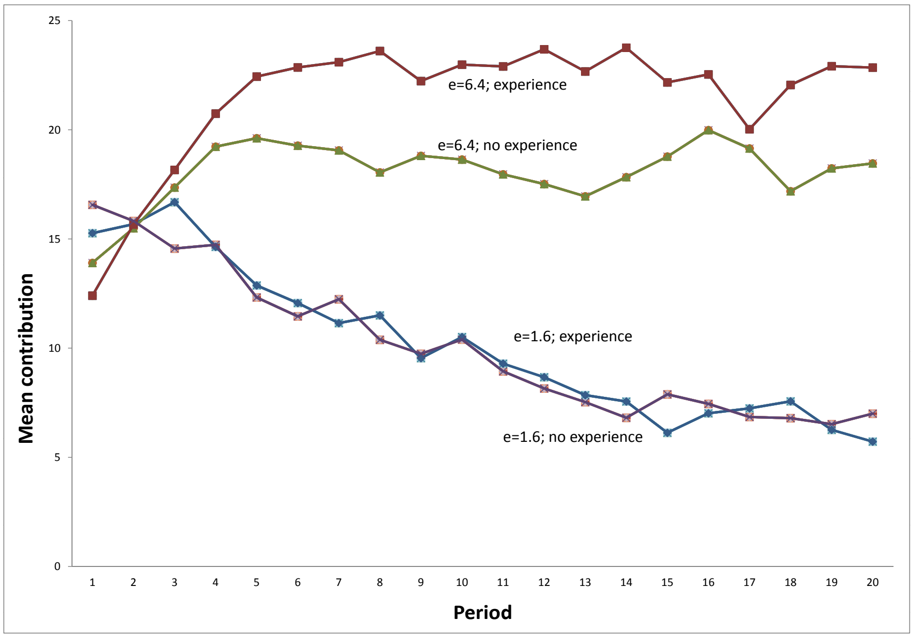
We conclude that experience has two main effects. On the one hand, experience decreases the SEARCH effect but preserves its sign. On the other hand, experience leads to larger contributions over time when $e = 6.4$, but has no significant effect when $e = 1.6$. In much a different setting, our findings complement the findings concerning experience in games with low rates of return from Marwell & Ames (1981), Isaac et al. (1984), Isaac et al. (1988) by analysis of games with high rates of return and by identifying a common SEARCH feature when standard information is withheld. In fact, it turns out that experience has a more significant effect in games where free-riding is not Nash (when $e = 6.4$) in which case experienced players contribute more and learn Nash equilibrium at a faster rate. A possible explanation for this is that experience reduces the agents' perceived ambiguity of the game.

3.2 Information

In this section, we shall investigate whether SEARCH and TREND are only features of black box behavior or whether they persist if players have more information. We shall find that SEARCH is very robust and persists in all information settings, while TREND is fairly robust but more sensitive to the information setting and rate of return.

[25] The initial contributions for both rates of return lie between 14 and 15 in all four phases of the experiment.

Figure 6: Mean black box play (phases 1,2 versus 3,4).



Experience plays no visible role in the free-riding games. Experience has a positive effect in the full-contribution games.

3.2.1 Distributional differences

Using period-specific Mann-Whitney tests we test for differences in contributions in the three information treatments for both rates of return.[26] It turns out that initial black box contributions are significantly lower than non-black box contributions for both rates of return. There are no differences in initial contributions between standard and enhanced treatments but significant differences with the black box treatment.[27] After period two, all treatments are significantly different from each other in all periods except standard and black box play when $e = 1.6$ which are not significantly different in any period. When $e = 6.4$, standard treatment contributions significantly exceed enhanced treatment contributions which exceed black box contributions. When $e = 1.6$, standard and black box contributions are not significantly different from each other but exceed enhanced treatment contributions. In the final period, of all three treatments, enhanced information is closest to Nash equilibrium when $e = 1.6$, while standard and black box are closer to Nash when $e = 6.4$.

(Table 3 and Figure 7 illustrate.)

[26]This yields three tests for each period for both high (when $e = 6.4$) and low (when $e = 1.6$) rates of return as summarized in Table 3.

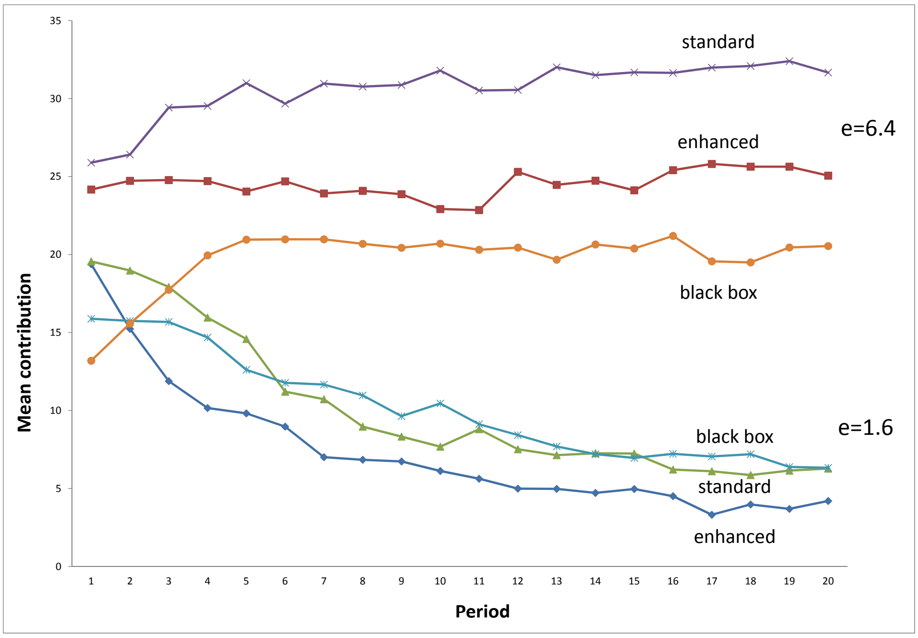
[27]Recall that initial black box contributions lie below 20 and show no differences for the two rates of return (see section 3.1.1). Initial contributions are higher for both standard and enhanced treatments;

Table 3: Mann-Whitney tests of equal contributions.

comparison		period						
$e =$	treatments	1	2	3	4	5	...	20
1.6	enhanced vs. standard	=	**	**	**	**	...	**
	black box vs. enhanced	**	=	***	***	***	...	***
	black box vs. standard	**	*	=	=	=	...	=
6.4	enhanced vs. standard	=	=	**	**	***	...	***
	black box vs. enhanced	***	***	***	***	***	...	***
	black box vs. standard	***	***	***	***	***	...	***

Significance: *: 90%; **: 95%; ***: 99%; =: not significant.

Figure 7: Mean play all treatments (black box, standard, enhanced).



Black box: initial contributions are low; trend when $e = 1.6$*, weak trend when* $e = 6.4$*. Standard and enhanced: initial contributions are higher; enhanced trend steeper than standard trends when* $e = 1.6$*, weak trends if* $e = 6.4$*.*

3.2.2 SEARCH

Testing for SEARCH, we perform our linear regressions (see section 3.1.2) for each rate of return separately, further controlling for black box, standard and enhanced information effects. We make the following observations. First, (more) information leads to less variation in adjustments in between time periods: for both rates of return, variation in adjustments between time periods in enhanced treatments is lower than in standard treatments which is lower than in black box treatments. The standard dummy is negative and significant at 95% but less negative than the enhanced dummy which is also significant at 95%. In all treatments and for both rates of return, success in the previous period leads to less variation in adjustments between time periods than failure in the previous period, and previous-period contributions have a positive level effect. None of the period dummies is significant, but all of the phase dummies are negative. In later phases, therefore, experience and information jointly further reduce variation in adjustments between time periods.

3.2.3 TREND

Testing for TREND, the relevant linear regressions (see section 3.1.3) become somewhat difficult to interpret. Net of level effects, the basic constituents of TREND continue to be significant. The main differences concern the size of the level effects and effects from earlier periods. In the black box treatments, only contributions in the previous period but not two periods ago are significant. In the non-black box treatments, however, contributions two periods ago are also significant in some but not all of the games. With 95% confidence, the level of contribution two periods ago has a negative effect on contributions in games with $e = 1.6$ under the enhanced treatment but not under the standard treatment. If $e = 6.4$, the level of contribution two periods ago has a positive effect (with 95% confidence) on contributions under the standard treatment but not under the enhanced treatment. This suggests longer "memory" effects in some but not in all of the non-black box treatments, and not necessarily in those with more information. Further difference with respect to the previous-period contribution effects are observed. These are significant at the 90% confidence interval. Enhanced information has a positive fixed effect on games with $e = 6.4$ but not on games with $e = 1.6$ (players contribute more), while standard information has a negative effect on games with $e = 1.6$ (players contribute less). Overall, therefore, enhanced information leads to play over time that gets closer to the Nash equilibrium than black box and standard treatment play in the games with $e = 1.6$, while standard information play gets closer to the Nash equilibrium than standard and black box treatment play in the games with $e = 1.6$.

3.2.4 Summary

Comparing black box with standard and enhanced treatments, we observe the following distributional differences. Initial contributions are higher in both standard and enhanced

close to 20 when $e = 1.6$, and above 20 when $e = 6.4$.

treatments (and not different from each other), and this finding includes the games with $e = 1.6$ where such contributions are worse replies in terms of the game's Nash equilibrium. After period two, there are significant differences between all three treatments when $e = 6.4$. When $e = 1.6$, there are no differences between black box and standard treatments, and enhanced treatment contributions are significantly lower. In fact, enhancing the information relative to standard treatments leads to lower contributions in all games; these are higher than black box contributions when $e = 6.4$ and lower than black box contributions when $e = 1.6$.

Let us briefly summarize our results regarding SEARCH and TREND. SEARCH is robust in all treatments. More experience and more information reduce the size of this effect but preserve its sign. TREND including level effects is robust in black box treatments with or without experience. Inside the box, adjustments are made dependent on whether the previous adjustment was successful or not. There is weaker evidence for TREND adjustments in the non-black box data and further contribution lags are significant. TREND behavior is likely to interact and depend on other factors such as social preferences.

4 Conclusion

In many real-world situations, individuals have no information about the other players' actions or payoffs. In such environments, behavior must be completely uncoupled from such information, and instead be based entirely on agents' own past actions and payoffs. Outside game theory, these rules have a long tradition dating back to Thorndike (1898). In our black box treatment, we enforce completely uncoupled behavior in the context of public goods games. Our two main findings identify search features and behavioral elements of directional learning. Experience and information reduce the magnitude of the search feature but it is still present at a statistically significant level. Moreover, players follow directed adjustment dynamics based on their previous contribution levels and previous success-versus-failure stimuli. This suggests that, even with experience, players' learning inside the box displays no significant evidence of acquiring a global notion of one action dominating another. Some but not all of the directional adjustment tendencies change when more information is revealed.

Recent game-theoretic models provide explanations for how different concepts of equilibrium can learned based on completely uncoupled behavior. In particular, search volatility is a key feature of the trial and error learning model proposed by Young (2009), which implements Nash equilibrium. Variants of these rules implement welfare-maximizing states (Marden et al., 2011) and welfare-maximizing Nash equilibria (Pradelski & Young, 2012). Completely uncoupled learning dynamics can also be applied to cooperative games, where they implement the core (Nax et al., 2012) and the Nash bargaining solution (Nax, 2013).

Much of the prior empirical work on learning in games has focussed on situations where players have a substantial amount of information about the structure of the game, and they can observe the behavior of others as learning proceeds. In this paper by contrast, we have examined situations in which players have no information about the strategic environment, and they must feel their way based solely on the pattern of realized payoffs. We identified two key features of such completely uncoupled learning dynamics – search volatility and directional adjustments, both of which have antecedents in the psychology literature. Search volatility in particular has not been examined previously in the context of experimental game theory and turns out to be very robust even when players gain experience and information. Whether this remains true for other classes of games is an open question for future research.

Appendices

Appendix A

Black box instructions

Participants received the following on-screen instructions (in z-Tree) at the start of the Black Box game and had to click an on-screen button saying, "I confirm I understand the instructions" before the game would begin:

Instructions

Welcome to the experiment. You have been given 40 virtual coins. Each 'coin' is worth real money. You are going to make a decision regarding the investment of these 'coins'. This decision may increase or decrease the number of 'coins' you have. The more 'coins' you have at the end of the experiment, the more money you will receive at the end.

During the experiment we shall not speak of £Pounds or Pence but rather of "Coins". During the experiment your entire earnings will be calculated in Coins. At the end of the experiment the total amount of Coins you have earned will be converted to Pence at the following rate: 100 Coins = 15 Pence. In total, each person today will be given 3,200 coins (£4.80) with which to make decisions over 2 economic experiments and their final totals, which may go up or down, will depend on these decisions.

The Decision

You can choose to keep your coins (in which case they will be 'banked' into your private account, which you will receive at the end of the experiment), or you can choose to put some or all of them into a '*black box*'.

This '*black box*' performs a mathematical function that converts the number of coins inputted into a number of coins to be outputted. The function contains a random component, so if two people were to put the same amount of coins into the '*black box*', they would not necessarily get the same output. The number outputted may be more or less than the number you put in, but it will never be a negative number, so the lowest outcome possible is to get 0 (zero) back. If you chose to input 0 (zero) coins, you may still get some back from the box.

Any coins outputted will also be 'banked' and go into your private account. So, your final income will be the initial 40 coins, minus any you put into the '*black box*', plus all the coins you get back from the '*black box*'.

You will play this game 20 times. Each time you will be given a new set of 40 coins to use. Each game is separate but the '*black box*' remains the same. This means you cannot play with money gained from previous turns, and the maximum you can ever put into the '*black box*' will be 40 coins. And you will never run out of money to play with as you get a new set of coins for each go. The mathematical function will not change over time, so it is the same for all 20 turns. However as the function contains a random component, the output is not guaranteed to stay the same if you put the same amount in each time.

After you have finished your 20 turns, you will play one further series of 20 turns but with a new, and potentially different '*black box*'. The two boxes may or may not have the same mathematical function as each other, but the functions will always contain a random component, and the functions will always remain the same for the 20 turns. You will be told when the 20 turns are finished and it is time to play with a new black box.

If you are unsure of the rules please hold up your hand and a demonstrator will help you.

I confirm I understand the instructions

Appendix B

On-screen output in each treatment

Supplementary Figures: the post-decision feedback information that participants received: (a) in the black-box treatment; (b) the first feedback screen in both the standard and the enhanced-information treatments; and (c) the second feedback screen, which differed between the two treatments (not shown in black-box treatment). Dashed lines border the information that was only shown in the enhanced-information treatment.

(a)

GAME SUMMARY	Number of Coins
Initial Coins	40
Minus (-) your input	15
Plus (+) the output returned	24
Your final number of coins	49

This screen lists your decisions and the results, along with your income (for this turn).

(b)

Player Name	Contribution (0-40)
You	15
player A	10
player B	30
player C	5
Total contributions	60
Total after growth	96

This screen lists the decisions of you and the other players (in random order) in your group (for this round). Remember the terms Player A, B, & C **are meaningless** as you play with randomly selected people in each round. The group's total contributions and the new total after the 'growth' stage are also shown.

(c)

Player Name	Contribution (0-40)	Player Income =	Credits retained	+ Credits returned	= Total credits
You	15	For you =	25	24	49
player A	10	player A	30	24	54
player B	30	player B	10	24	34
player C	5	player C	35	24	59

Your income from your group's contributions and subsequent 'growth' is shown.

This screen lists the decisions and earnings of you and the other players (in random order) in your group (for this round).

Remember the terms Player A, B, & C **are meaningless** as you play with randomly selected people in each round.

References

J. Andreoni, "Why Free Ride? Strategies and Learning in Public Goods Experiments", *Journal of Public Economics 37*, 291-304, 1988.

R.-C. Bayer, E. Renner & R. Sausgruber, "Confusion and learning in the voluntary contributions game", *Experimental Economics,* forthcoming, 2013.

J. Björnerstedt, & J. Weibull, "Nash Equilibrium and Evolution by Imitation", in Arrow, K. and Colombatto, E. (Eds.), *Rationality in Economics*, New York: Macmillan, 1993.

M. N. Burton-Chellew & S. A. West, "Pro-social preferences do not explain human cooperation in public-goods games", *Proceedings of the National Academy of Science 110*, 216-221, 2013.

C. F. Camerer, *Behavioral Game Theory: Experiments on Strategic Interaction,* Princeton University Press, 2003.

C. F. Camerer & T.-H. Ho, "Experience-weighted Attraction Learning in Normal Form Games", *Econometrica 67*, 827-874, 1999.

A. Chaudhuri, "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature", *Experimental Economics 14*, 47-83, 2011.

J. Coates, "The Hour Between Dog and Wolf. Risk Taking, Gut Feelings, and the Biology of Boom and Bust", *Penguin Press USA*, 2012.

R. Cookson, "Framing effects in public goods experiments", *Experimental Economics 3*, 55-79, 2000.

M. Costa-Gomes, V. Crawford & B. Broseta, “Cognition and behavior in normal-form games: An experimental study”, *Econometrica 69*, 1193-1235, 2001.

M. Costa-Gomes & V. Craword, “Cognition and Behavior in Two-Person Guessing Games: An Experimental Study”, *American Economic Review 96*, 1737-1768, 2006.

R. Croson, “Partners and Strangers Revisited”, *Economics Letters 53*, 25-32, 1996.

J. G. Cross, *A Theory of Adaptive Economic Behavior,* Cambridge University Press, 1983.

A. E. Eiben & C. Schippers, “On evolutionary exploration and exploitation”, *Fundamenta Informaticae 35*, 35-50, 1998.

L. G. Epstein, “A Definition of Uncertainty Aversion”, *The Review of Economic Studies 66*, 1999.

E. Fehr & S. Gachter, “Cooperation and Punishment in public goods experiments”, *American Economic Review 90*, 980-994, 2000.

E. Fehr & S. Gachter, “Altruistic punishment in humans”, *Nature 415*, 137-140, 2002.

E. Fehr & K. M. Schmidt, “A Theory of Fairness, Competition and Cooperation”, *Quarterly Journal of Economics 114*, 817-868, 1999.

E. Fehr & Camerer, “Social neuroeconomics: the neural circuitry of social preferences”, *Trends in Cognitive Sciences 11*, 419-427, 2007.

U. Fischbacher, “z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, *Experimental Economics 10*, 171-178, 2007.

D. Foster & H. P. Young, “Regret testing: Learning to play Nash equilibrium without knowing you have an opponent”, *Theoretical Economics 1*, 341-367, 2006.

F. Germano & G. Lugosi, “Global Nash convergence of Foster and Young’s regret testing”, *Games and Economic Behavior 60*, 135-154, 2007.

B. Greiner, “An Online Recruitment System for Economic Experiments”, in: K. Kremer & V. Macho (Eds.): *Forschung und wissenschaftliches Rechnen 2003.* GWDG Bericht 63, 79-93, 2004.

S. Hart & A. Mas-Colell, “Uncoupled Dynamics Do Not Lead to Nash Equilibrium”, *American Economic Review 93*, 1830-1836, 2003.

S. Hart & A. Mas-Colell, “Stochastic Uncoupled Dynamics and Nash Equilibrium,” *Games and Economic Behavior 57*, 286-303, 2006.

T.-H. Ho, C. Camerer & K. Weigelt, “Iterated Dominance and Iterated Best-response in p-Beauty Contests”, *American Economic Review 88*, 947-969, 1998.

M. Isaac, J. Walker & S. Thomas, “Divergent Expectations on Free Riding: An Experimental Examination of Possible Explanations”, *Public Choice 43*, 113-149, 1984.

M. Isaac & J. Walker, “Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism”, *Quarterly Journal of Economics 103*, 179-199, 1988.

J. O. Ledyard, "Public Goods: A Survey of Experimental Research", in J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics,* Princeton University Press, 111-194, 1995.

J. G. March, "Exploration and Exploitation in Organizational Learning", *Organization Science 2*, 71-87, 1991.

J. R. Marden, H. P. Young, G. Arslan, J. Shamma, "Payoff-based dynamics for multi-player weakly acyclic games", *SIAM Journal on Control and Optimization 48,* special issue on "Control and Optimization in Cooperative Networks", 373-396, 2009.

J. R. Marden, H. P. Young, L. Pao, "Achieving Pareto Optimality Through Distributed Learning", *discussion paper,* University of Colorado and University of Oxford, 2011.

G. Marwell & R. Ames, "Economists Free Ride, Does Anyone Else?", *Journal of Public Economics 15*, 295-310, 1981.

R. Nagel, "Unraveling in Guessing Games: An Experimental Study", *American Economic Review 85*, 1313-1326, 1995.

H. H. Nax, "Completely uncoupled bargaining", *mimeo.*, 2013.

H. H. Nax, B. S. R. Pradelski & H. P. Young, "The Evolution of Core Stability in Decentralized Matching Markets", *Department of Economics Discussion Paper 607,* University of Oxford, 2011.

J. Oechssler & B. Schipper, "Can you guess the game you are playing?", *Games and Economic Behavior 43*, 137-152, 2003.

B. S. R. Pradelski & H. P. Young, "Learning Efficient Nash Equilibria in Distributed Systems", *Games and Economic Behavior 75*, 882-897, 2012.

A. E. Roth & I. Erev, "Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term", *Games and Economics Behavior 8,* 164-212, 1995.

H. Sauermann & R. Selten, "Anspruchsanpassungstheorie der Unternehmung", *Zeitschrift für die Gesamte Staatswissenschaft 118*, 577-597, 1962.

D. Schmeidler, "Subjective Probability and Expected Utility without Additivity", *Econometrica 57*, 571-587, 1989.

R. Selten & J. Buchta, "Experimental Sealed Bid First Price Auction with Directly Observed Bid Functions", in: Budescu, D., Zwick, I.E. & R. (Eds.), *Games and Human Behavior,* Essays in Honor of Amnon Rapoport, 1998.

R. Selten & R. Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach", *Journal of Economic Behavior and Organization 7*, 47-70, 1986.

D. Stahl & P. Wilson, "On Players' Models of Other Players: Theory and Experimental Evidence", *Games and Economic Behavior 10*, 218-254, 1995.

F. Thuijsman, B. Peleg, M. Amitai & A. Shmida, “Automata, matching and foraging behavior of bees”, *Journal of Theoretical Biology 175*, 305-316, 1995.

E. Thorndike, “Animal Intelligence: An Experimental Study of the Associative Processes in Animals”, *Psychological Review 8*, 1898.

R. Tietz & H. Weber, “On the nature of the bargaining process in the Kresko-game”, in: Sauermann, H. (Ed.), *Contributions to experimental economics*, Vol. 3, pp. 305-334, 1972.

H. P. Young, “Learning by trial and error”, *Games and Economic Behavior 65*, 626-643, 2009.