ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

LEARNING IN A BLACK BOX: TRIAL-AND-ERROR IN VOUNTARY CONTRIBUTIONS GAMES

H.H. Nax, M.N. Burton-Chellew, S.A. West and H.P. Young

Number 653
April 2013
Revised February 2014
Revised January 2015

Manor Road Building, Manor Road, Oxford OX1 3UQ

Learning in a Black Box: Trial-and-Error in Voluntary Contributions Games*

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

First version: April 18, 2013
This version: January 27, 2015

*We thank Margherita Comola for help at an early stage of the project and Johannes Abeler, Guillaume Hollard, Richard Bradley, Fabrice Etilé, Philippe Jehiel, Thierry Verdier, Vince Crawford, Colin Camerer, Muriel Niederle, Edoardo Gallo, Amnon Rapoport, Guillaume Fréchette, Ozan Aksoy, Margaret Meyer, Roberto Serrano, Arno Riedl, Martin Strobel, Martin Cripps, anonymous referees, and participants of the Economics and Psychology seminar at MSE Paris, the Choice Group at LSE, the 24th Game Theory Workshop at Stony Brook, the CESS seminar at Nuffiel College, the 9th Tinbergen Institute Conference, and Economics seminar at the University of Maastricht, for helpful discussions, suggestions and comments. Correspondence to: H.H. Nax, ETH Zürich, Clausiusstrasse 37-C3, 8092 Zurich, Switzerland.

[†]Heinrich H. Nax, Department of Social Sciences, ETH Zurich. Nax's research was supported by the Office of Naval Research grant N00014-09-1-0751 and by the European Commission through the ERC Advanced Investigator Grant 'Momentum' 324247.

[‡]Maxwell N. Burton-Chellew, Nuffield College, Oxford & Department of Zoology, University of Oxford (The Tinbergen Building, S Parks Rd, Oxford OX1 3PS, United Kingdom). Burton-Chellew's research was supported by the European Research Council.

[§]Stuart A. West, Department of Zoology, University of Oxford (The Tinbergen Building, S Parks Rd, Oxford OX1 3PS, United Kingdom). West's research was supported by the European Research Council.

[¶]H. Peyton Young, Department of Economics, University of Oxford (Nuffield College, New Rd, Oxford OX1 1NF, United Kingdom). Young's research was supported by the Office of Naval Research grant N00014-09-1-0751 and the Air Force Office of Scientific Research grant FA9550-09-1-053.

Abstract

Many interactive environments can be represented as games, but they are so large and complex that individual players are mostly in the dark about others' actions and the payoff structure. This paper analyzes learning behavior in such a 'black box' environment, where players' only source of information is their own history of actions taken and payoffs received, the context being voluntary contributions games with high and low rates of return. We test four hypotheses regarding the players' adjustment dynamics: asymmetric inertia, search volatility, search breadth, and directional bias. Our findings provide behavioral support for key constituents of learning models that work with very limited information.

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

Keywords: black box, learning, information, public goods game

1 Introduction

Some games are so complex and involve so many players that in practice the participants have only a vague idea of what others are doing and what their payoffs are, even though their own payoffs are strongly influenced by the actions of the others. Nevertheless, mainstream research in experimental games has focused mainly on situations where the structure of the game is known and the actions of other players are observable. But what happens in situations where agents have very little information about the other players or the structure of the game? To address this question, we conducted a series of experiments in which subjects played a voluntary contributions game (Isaac et al., 1985; Isaac and Walker, 1988) but they had no information about the nature of the game, and they could not observe the actions or payoffs of the other players. In effect, subjects were inside a 'black box' and the only data available to them were the payoffs they received from taking different actions. We compared their learning behavior in the the black box environment with three other treatments where players were given more information about the nature of the game, and the payoffs and actions of others.

Standard learning models do not provide clear predictions about subjects' behavior in a black box environment. However, there is a growing body of theory that shows how near-equilibrium play can evolve from such dynamics.[1] A key feature of these dynamics

[1] Such behaviors are said to be *completely uncoupled* (Foster and Young, 2006; Young, 2009). (A

(variously known as 'exploration-exploitation', 'probe-and-adjust', 'trial-and-error', 'regret testing', and 'win-stay-lose-shift') is an asymmetry regarding responses to gains and losses. Players who experience high payoffs tend to stay with their current strategy, while those who experience low payoffs will search the strategy space for something better.[2] Related concepts are well-established in computer science (Eiben and Schippers, 1998; Bowling and Veloso, 2002), but the formal demonstration that such rules lead to Nash equilibrium in large classes of games is relatively recent (Foster and Young, 2006; Germano and Lugosi 2007; Young, 2009; Marden et al., 2009; Marden and Shamma, 2011; Pradelski and Young, 2012; Skyrms, 2012; Huttegger et al., 2014).

Learning experiments with very low information go back to Thorndike's (1898) studies of trial-and-error behavior in animals and the formulation of the 'law of effect' (Thorndike, 1911; Herrnstein, 1970). This law has inspired a variety of reinforcement learning models that have been tested experimentally (Bush and Mosteller, 1955; Suppes and Atkinson, 1959; Harley, 1981; Cross, 1983; Roth and Erev, 1995; Erev and Roth, 1998; Nevo and Erev, 2012; Erev and Haruvy, 2013).[3] Unlike standard belief-based models in game theory, these reinforcement models do not require subjects to have any knowledge of the game structure or to be able to observe others' actions and payoffs.[4] Such reasoning is impossible in black box environments.

In this paper, we test three key features of trial-and-error learning models that are related to their convergence properties:

(*i*) *Asymmetric inertia.* A player is more likely to stay with his previous contribution after a payoff increase than after a payoff decrease, that is, exploration is more likely when payoffs go down than when they go up.

(*ii*) *Search volatility.* Conditional on a search being triggered, the variance in the new contribution level is greater when the triggering event is a payoff decrease than

learning rule is *uncoupled* is it can depend on others' actions but not on their payoffs (Hart and Mas-Colell, 2003, 2006).)

[2]Exploration-exploitation is used in generalized reinforcement models (e.g. Nevo and Erev, 2012; Erev and Haruvy 2013). Probe-and-adjust is due to Skyrms (2010). Trial-and-error and regret testing, respectively, are introduced in Young (2009) and Foster and Young (2006).

[3]More sophisticated 'hybrid' learning models such as experience-weighted attraction (Camerer and Ho, 1999; Camerer et al., 2002) would also be reduced to reinforcement in a black box setting due to the inherent informational limitations.

[4]Such belief learning models are, for example, 'fictitious play' (Brown, 1951; Robinson, 1951; Cournot, 1960; Shapley, 1964; Fudenberg and Kreps, 1991, 1993; Krishna, 1991; Milgrom and Roberts, 1991; Monderer and Shapley, 1996; Fudenberg and Levine, 1996; Krishna and Sjostrom, 1995), 'imitation' (see Björnerstedt and Weibull, 1993; Stahl and Wilson, 1995; Nagel, 1995). The same applies to standard evolutionary dynamics based on myopic best reply (Kandori et al., 1993; Young, 1993; Blume, 1993). An exception is 'directional learning' (Selten and Stoecker, 1986; Selten and Buchta, 1998; Harstad and Selten, 2013), which we shall also interpret in a completely uncoupled sense.

when it is a payoff increase.

(*iii*) *Search breadth.* Conditional on a search being triggered, the expected amount by which a player changes his contribution is greater when the triggering event is a payoff decrease than when it is a payoff increase.

These features have not previously been tested in an experimental setting. In addition we examine a hypothesis that is specific to the one-dimensional strategy structure in voluntary contributions games:

(*iv*) *Directional bias.* If an increase in contribution results in a higher payoff this period, the player's expected contribution next period will be higher than if the increase results in a lower payoff. Similarly, if a decrease in contribution results in a higher payoff this period, the player's expected contribution next period will be lower than if the decrease leads to a lower payoff.

The presence of all four features is confirmed at high levels of statistical significance in the black box treatment. Moreover, even in the treatments where more information is available, none of the hypotheses can be rejected. We do not claim that these are the only relevant features of black box learning, nor do we attempt to fit a parametric model to the data. Rather our results identify a rich set of qualitative behaviors that can be examined in many other settings.

In the next section, we discuss the related literature. Section 3 describes the experimental set-up in detail, including treatments that provide subjects with increasing amounts of information. In Sections 4-5 we test the four conditions described above, both in the black box treatment and with more information. In section 6 we discuss the relationship between our findings and those of Nevo and Erev (2012), Erev and Haruvy (2013), and Bayer et al. (2013), which are the previous contributions most closely connected to ours. Section 7 summarizes our results and suggests topics for future research. The Appendix contains experimental instructions and supplementary regression tables.

2 Related literature

The black box experimental design is novel in several aspects. First, by withholding information about the structure of the game, learning must be purely payoff-based (i.e. completely uncoupled), and one can therefore be sure not to confuse it with belief-based learning. Second, withholding information about others' actions and payoffs eliminates other concerns such as imitation, fairness and social preferences. Finally, the black box approach departs from the approach traditionally favored in experimental economics, which has been called the 'subjective expected utility correction project' (Gigerenzer and Selten 2001), by which behavior is explained via functional 'corrections' to underlying utility and uncertainty perceptions. Instead, the black box approach identifies behavioral regularities without assuming that they are driven by utility maximization. The black box framework is an approach to studying behavior in games that is inspired by experimental biology. Ido Erev and co-authors, notably Nevo and Erev (2012) and Erev and Haruvy (2013), have pursued the black box approach in their experiments on decisions under uncertainty (one-player games). In these experiments, subjects are faced with repeated black box gambling decisions to invest in/ to chose amongst gambles without any information beyond ex post payoff realizations of the lotteries as the gambling tasks are repeated. Nevo and Erev (2012) and Erev and Haruvy (2013) propose models that generalize simple reinforcement learning, and account for a large variety of behavioral regularities observed in black box (and non-black box) experiments. We shall discuss the connections between their findings and ours in section six. A key difference is our identification of asymmetric behavioral responses to gains and losses, which is a crucial feature of trial-and-error learning rules in the theoretical literature.[5]

In the context of multi-player games, the paper most closely related to ours is Bayer et al. (2013). They study a low-information treatment of voluntary contributions games that is directly comparable with one of our treatments. Their main contribution is the formulation of a directional learning model that predicts that contribution decisions follow the direction that led to success, and avoid the direction that led to failure. We shall discuss the connections between their results and ours in detail in a separate section later on, after presenting our results.

Apart from the above two experiments, our black box treatment provides subjects with even less information than older 'low-information' experiments. One is Rapoport et al. (2002) and Weber (2003), who provide information about the structure of the game but withhold information about the outcome of play and the distribution of players' types. By contrast, Friedman et al. (2012) withhold information about the payoff structure of the game but do provide information about other players' actions and payoffs. In Oechssler and Schipper (2003), players learn the payoff structure on the basis of their own payoffs and information about other's actions. Finally, there is Ben Zion et al. (2010), where players learn about consequences of actions not chosen in a portfolio allocation experiment. Even earlier experiments that explored minimal information learning are

[5] Related regret heuristics with inertia in low information settings are also proposed in Saran and Serrano (2014a,b).

the repeated two-by-two zero sum games used to test Markov learning models by Suppes and Atkinson (1959).

The overarching lesson from our experiments is that *asymmetric responses* with respect to gains and losses are a robust feature across all information treatments. Although this notion has antecedents in the classic experimental economics literature (Kahneman and Tversky, 1984), little is known about the corresponding behaviors in experimental game settings, owing to the fact that pure black box treatments have received almost no attention. Outside economics, these ideas are quite standard however. In particular, asymmetric inertia, search volatility and search breadth have been proposed in biology (Thuijsman et al., 1995; Motro and Shmida, 1995), organizational learning (March, 1991), computer science (Eiben and Schippers, 1998; Bowling and Veloso, 2002), and machine learning (Kraines and Kraines, 1989; Nowak and Sigmund, 1993). The basic idea is that an agent tends to keep playing a strategy that yields high payoffs, and switches when payoffs decrease.[6] One feature that has been considered in the prior literature is directional learning in games with a one-dimensional strategy set (Selten and Stoecker, 1986; Selten and Buchta, 1998; Harstad and Selten, 2013; Bayer et al. 2013; Laslier and Walliser, 2014).[7] Our directional bias hypothesis is similar in spirit.

The game context we choose for our black box investigation is the voluntary contributions game (Isaac et al., 1985; Isaac and Walker, 1988), which has been extensively studied in higher informational environments (Ledyard, 1995; Chaudhuri, 2011). In contrast to much of that literature, our aim is to study payoff-based learning in a black box setting, rather than to compare mechanisms or to disentangle learning from social preferences when more information is made available about the game's inherent individual-versus-collective incentive structure (e.g. Andreoni, 1993, 1995; Palfrey and Prisbey, 1996, 1997; Goeree et al., 2002; Ferraro and Vossler, 2010; Fischbacher and Gachter, 2010; Burton-Chellew and West, 2013).[8] Apart from the informational differences with previous experiments, an important feature of our set-up is that we examine learning behavior in contribution games with high and low rates of return. Thus we can compare subjects' behavior when the dominant strategy is to contribute fully, and when it is to contribute nothing.

[6] Traders in financial markets exhibit similar behavior (Coates, 2012).

[7] Harstad and Selten (2013) is a recent survey. For early experimental work on this issue, see also Tietz and Weber (1972), Roth and Erev (1995), Erev and Roth (1998), Erev and Rapoport (1998).

[8] Burton-Chellew and West (2012) use the same data set as in this paper, but their analysis focuses on the aggregate levels of contribution and rates of convergence rather than on period-by-period learning behavior.

3 Experimental set-up

A total of 236 subjects, in 16 separate sessions involving 12 or 16 subjects, participated in our public goods experiments yielding a total of 18,880 observations. Participants were recruited from a subject pool that had not previously been involved in public goods experiments.[9] The subjects were not limited to university students, but included different age groups with diverse educational backgrounds.[10] In the experiment, a group of subjects plays several voluntary contributions games, where each game is repeated for twenty rounds with randomly allocated subgroups in each round. Games differ with respect to two *rates of return* ('low' and 'high') such that either 'free-riding' or 'fully contributing' is the dominant strategy. These are played under three different information treatments ('black box', 'standard' and 'enhanced'). In this section, we shall describe the underlying stage game, the structure of each repeated game, and the different information treatments.

3.1 Linear public goods game

Consider the following linear public goods game, known as a "voluntary contributions game". Each player i in population $N = \{1, 2, ..., n\}$ makes a nonnegative, real-numbered contribution, c_i, from a finite budget $B > 0$. The vector of contributions is denoted by $\mathbf{c} = \{c_1, c_2, ..., c_n\}$. Given a *rate of return*, $e \geq 1$, the public good is provided in the amount $R(\mathbf{c}) = e \sum_{i \in N} c_i$, and split equally amongst the players.[11] Thus, given others' contributions $\mathbf{c}_{-i}$, player i's contribution c_i results in the payoff

$$\phi_i = \frac{e}{n} \sum_{i \in N} c_i + (B - c_i).$$

Write $\boldsymbol{\phi}$ for the payoff vector $\{\phi_1, ..., \phi_n\}$.

Nash equilibria. If the rate of return is *low* ($e < n$), an individual contribution of zero ('free-riding') is the strictly dominant strategy for all players. Similarly, if the rate of return is *high* ($e > n$), an individual contribution of B ('fully contributing') is the strictly dominant strategy for all players. The respective Nash equilibria consequently result in either nonprovision or full provision of the public good.

[9]The subjects were recruited through ORSEE (Online Recruiting System for Economic Experiments; Greiner, 2004) and subjects who listed prior participation in public goods games were excluded.

[10]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.

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

3.2 Repeated game

In each experimental session, the same population S (with $|S| = 12$ or 16) plays four 'phases' where each phase is a separate twenty-times repeated voluntary contributions game. In each period t within a given phase, players in S are randomly matched in groups of four to play the voluntary contributions game.[12] At the start of each period, each subject is given a new budget $B = 40$, of which he can invest any amount; however, subjects cannot invest money carried over from previous rounds. The rate of return is either *low* ($e = 1.6$) or *high* ($e = 6.4$) throughout a given phase.[13] Write N_4^t for any of the four-player groups matched at time t, and ρ_4^t for the partition of S into such groups. Given others' contributions $\mathbf{c}_{-i}^t$, each i receives a total of

$$\phi_i = \sum_{t=1,\ldots,20} \phi_i^t = \sum_{t=1,\ldots,20} \left(\frac{e}{4} \sum_{j \in N_4^t} c_j^t + (40 - c_i^t) \right). \tag{1}$$

For every i, ϕ_i represents a real monetary value that is paid after the game.[14] Note that, due to random rematching of players into groups throughout the experiment, for every player i, his relevant group N_4^t in the above expression 1 (i.e. such that $i \in N_4^t$) typically is a different one in each period t.

3.3 Information treatments

Each experimental session (involving four phases) is divided into two 'stages': phases one and two of the session constitute stage one, phases three and four constitute stage two. Each phase is a twenty-round voluntary contributions game with either the low ($e = 1.6$) or the high ($e = 6.4$) rate of return, and in each stage both rates of return are played. The two stages differ with respect to the information treatment.

Before the experiment begins, subjects are told that two separate experiments will be conducted, each stage consisting of two games. At no point before, during, or in between the two separate stages 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.[15]

[12] This follows random 'stranger' rematching of Andreoni (1988).

[13] The Nash equilibrium payoff of the stage game is 40 when $e = 1.6$, and 256 when $e = 6.4$.

[14] One hundred coins are worth £0.15. The maximal earnings over the whole experiment therefore amount to £19.20.

[15] See Appendix A for the full instructions and Appendix B for the output screens displayed during the experiment for each treatment.

All treatments. During each phase, the same underlying game is repeated for 20 periods. Each player receives 40 monetary units each period, of which he can invest any amount. After investments are made, each player earns 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 into real money.

Black box. No information about the structure of the games or about other players' actions or payoffs is revealed. Subjects play two voluntary contributions games (one with the high and one with the low rate of return). As play proceeds, subjects only know their own contributions and payoffs, but do not learn those of others.

Standard. The rules of the game are revealed, including production of the public good, high and low rate of return, and how groups form each period. As the game proceeds, players receive a summary of the relevant contributions in their group at the end of each period.[16]

Enhanced. In addition to the information available in the standard treatment, the payoffs of the other group members are explicitly calculated, and players receive a summary of their own and other players' payoffs in their group at the end of each period.

In each experimental session, every player plays two black box games (one with the high and one with the low rate of return), and either two standard or two enhanced information games (again one high and one low). Either the black box treatment occurs first, or a non-black box treatment occurs. Of the total of 18,880 observations, 9,440 are black box, 4,640 are standard, and 4,800 are enhanced. The order of high and low rates of return and of treatments is different in each session.[17] Sessions lasted between fifty and sixty minutes, and subjects earned between £6.20 and £15.50 (mean earnings were £12.40).

Black box versus grey box

Black box is our main treatment, and the analysis is based on sessions when black box is played first. Recall that we require the recruiting system (ORSEE) to select only 'first timers' of public goods experiments. Thus, when black box is played first, subjects are not likely to have prior knowledge of the structure of the game.

[16] This follows the standard information treatment design as in Fehr and Gächter (2000, 2002) and random 'stranger' rematching of subjects as introduced by Andreoni (1988).

[17] It turns out that the order of return rates does not lead to different conclusions regarding our learning model, more information can be found in Appendix C.

'Black box' versus 'grey box'. *We shall reserve the term 'black box' for the case when the no information treatment is played first. Sessions when it is played after the standard or enhanced treatments will be called 'grey box'.*

In grey box, subjects are told that a new and separate experiment will be conducted and that all information except for their own payoffs will be withheld. Since play in these sessions was preceded by two voluntary contributions games where subjects received full information about the structure of the game, they might (or might not) think that the grey box treatment has a similar structure. However, they will still be unable to infer others' contributions and the underlying rates of return from the information they receive. We make this distinction between black box and grey box because our black box learning model applies most evidently in the (pure) black box environment where more sophisticated learning models cannot apply because of the complete absence of information (about the structure and about others' actions).

4 Black box learning

In this section, we present and test our four hypotheses based on the analysis of the pure black box data, that is, those sessions when black box was played before the standard or enhanced treatments. There were eight sessions and 4,960 observations. Recall that subjects in the black box have no knowledge of the structure of the game and receive no information about others' actions or payoffs as the game goes on. Moreover, subjects are recruited to be 'first timers', that is, they have no prior experience playing public goods games.

4.1 The adjustment hypotheses

We begin by establishing some terminology.

Adjustment. The *adjustment* of player i in period $t+1$ is $c_i^{t+1} - c_i^t$. The adjustment is an *increase* if $c_i^{t+1} > c_i^t$, a *decrease* if $c_i^{t+1} < c_i^t$, and a *zero adjustment* if $c_i^{t+1} = c_i^t$.

Breadth. The *breadth* of an adjustment by player i in period $t+1$ is the absolute value of the adjustment, $|c_i^{t+1} - c_i^t|$.

Success versus failure. 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*.

Bias. The *bias* in adjustments of player i in period $t+1$ is the difference between the expected adjustment after success and the expected adjustment after failure, that is, $\mathbf{E}(c_i^{t+1} - c_i^t | \phi_i^t \geq \phi_i^{t-1}) - \mathbf{E}(c_i^{t+1} - c_i^t | \phi_i^t < \phi_i^{t-1})$.

Our hypotheses are summarized by a learning model, called **SEARCH**, which consists of the following four components:

(i) *Asymmetric inertia.* Zero adjustments are less likely after failure than after success.

(ii) *Search volatility.* Failure triggers a greater variance in non-zero adjustments than does success.

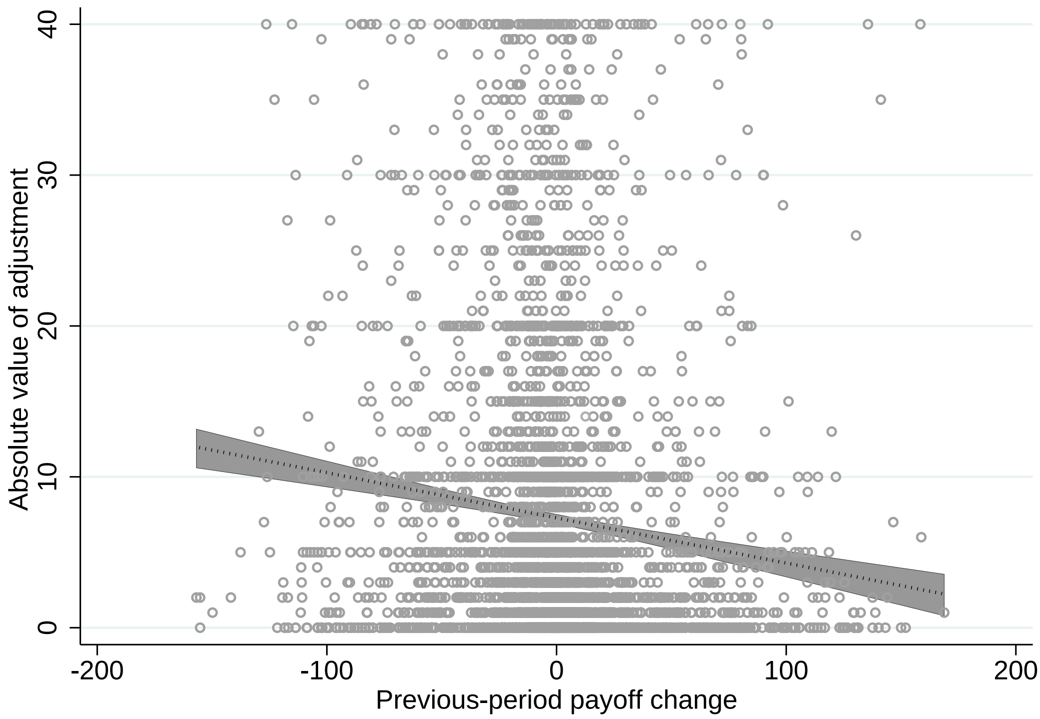
(iii) *Search breadth.* The absolute size of non-zero adjustments is larger after failure than after success.

(iv) *Directional bias.* If an increase in contribution resulted in success this period, the player's contribution next period will tend to be higher than if the prior contribution resulted in failure. Similarly, if a decrease in contribution resulted in success, the player's contribution next period will tend to be lower than if the prior contribution resulted in failure.

Consider the following example. Suppose an agent contributes 10 in one period and 20 in the next. Search volatility means that if 20 results in a higher realized payoff (success) to the player, then his next-period adjustment is drawn from some distribution with lower variance than if 20 resulted in a lower realized payoff (failure). Search breadth states that the absolute value of the adjustment, in case an adjustment is made, tends to be larger after failure than after success. Asymmetric inertia states that he is more likely to stick with 20 again in case of success than in case of failure. Finally, directional bias implies that he will tend to contribute more if 20 was a success than if 20 was a failure.

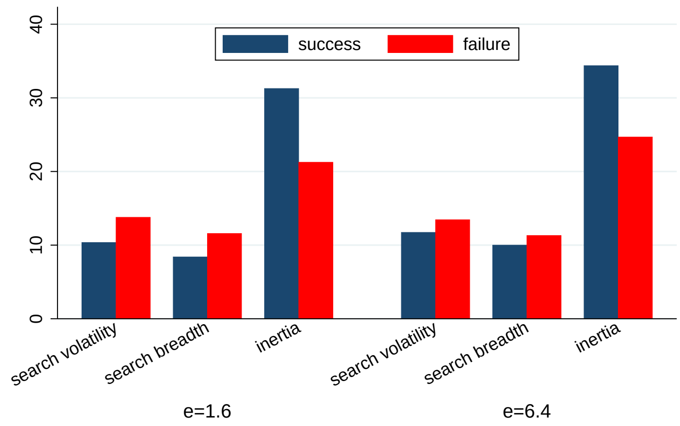
We find that all of these features are present in the data. Consider Figure 1, which plots the absolute value of adjustments against the performance of previous contributions. The scatterplot suggests that adjustments after failure are more volatile; larger adjustments occur more often after failure than after success; and lower or inertial adjustments are more likely after success than after failure. A summary of all four SEARCH components is given in Figures 2 and 3 . These patterns are confirmed by statistical tests of significance, which are summarized in Tables 1 and 2. Comparing columns (c) and (d) in Table 1, we make the following observations: (i) inertia is lower after failure than after success; (ii) search volatility is higher after failure than after success; and (iii) search breadth is higher after failure than after success. From columns (b) and (c) in Table 2, we see that (iv) directional bias is such that adjustments after increases (decreases) are higher

Figure 1: Black box adjustments.



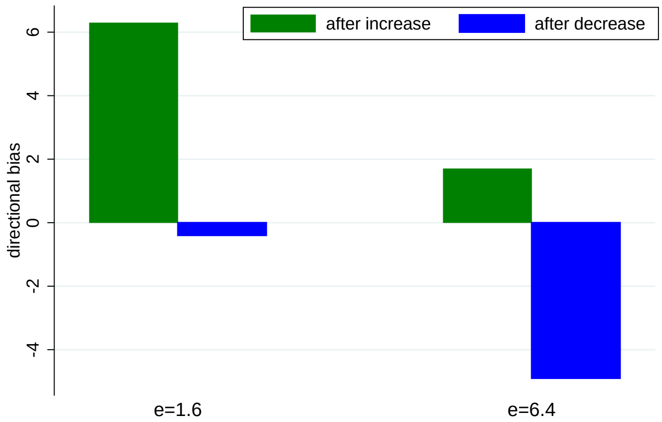
The figure shows absolute values of adjustments in the current period as a function of the last-period payoff change. The linear regression reveals that larger adjustments tend to lie on the side of previous-period failures, while smaller and inertial adjustments tend to lie on the success-side.

Figure 2: Asymmetric inertia, search volatility, and search breadth ($e = 1.6$ and $e = 6.4$).



The bar charts summarize asymmetric inertia, search volatility, and search breadth conditional on success versus failure under black box information for both rates of return. The y-axis are units of contribution for search volatility and search breadth, and percentage rates of zero adjustments for inertia.

Figure 3: Directional bias ($e = 1.6$ and $e = 6.4$).



The bar charts summarize the directional bias, that is, the difference between adjustments after success and adjustments after failure conditional on previous-period increase versus decrease under black box information.

(lower) in case of success than in case of failure.

The tests of significance were conducted as follows. For search volatility, we applied Levene's test, which is a nonparametric test for the equality of variances in different samples. The null hypothesis of equal variances following success and failure, which would imply no search volatility, is rejected with 99 percent confidence. This holds for both high and low rates of return and for different orders of the games (see Appendix C, Table 3, Tests 1-3).

Next, we test the hypothesis of equal search breadth conditional on success versus failure. We regress the absolute values of the non-zero adjustments conditional on success versus failure, controlling for phase, group, period and individual fixed effects with individual-level clustering. The average adjustment following success is smaller than after failure at a 1 percent level of significance (see Appendix C, Table 7). The breadth of search is significantly larger in periods 1-9 than in periods 10-20. Group, phase, and period fixed effects are not significant.

To test the hypothesis of equal inertia after success and failure, we use an ordered probit regression of the inertia rate controlling for phase, group, period and individual fixed effects with individual-level clustering. Success turns out to have a significantly positive effect on the inertia rate (see Appendix C, Table 5). Moreover, there is significantly less inertia in the first phase than in the second phase. Period fixed effects are negative until period six, suggesting that inertia tends to increase over time. Group fixed effects are not significant.

Table 1: Summary statistics for asymmetric inertia, search volatility, and search breadth (black box treatment)

(a) search component	(b)	(c) failure	(d) success	(e) difference (c)-(d)
search volatility				
standard deviation of adjustment	standard deviation	13.7	11.1	2.6**
	# observations	2266	2198	
search breadth				
absolute value of non-zero adjustments	mean	11.4	9.2	2.2**
	# observations	1746	1476	
inertia				
probability of zero adjustments	relative frequency	0.23	0.33	0.10**
	# observations	520	722	

*: $p < 0.05$; **: $p < 0.01$.

Table 2: Summary statistics for directional bias (black box treatment)

(a) mean non-zero adjustment	(b) success	(c) failure	(d) difference (b)-(c)
after *increase*	-1.6	-6.5	4.9**
after *decrease*	2.5	5.5	-3.0**

*: $p < 0.05$; **: $p < 0.01$.

Finally, we analyze the directional patterns of adjustments. Our hypothesis of directional bias states that (*i*) if an increase leads to success, the player's next-period contribution will tend to be higher than if it leads to failure; (*ii*) if a decrease leads to success, the player's contribution next period will tend to be lower than if it leads to failure. To test this hypothesis, we regress the difference between adjustments after success and adjustments after failure, controlling for the direction of the previous adjustment as well as for phase, group, period and individual fixed effects with individual-level clustering. The directional bias is confirmed, for both increases and decreases in prior-period contributions, at high levels of significance ($p < 0.01$). The details are given in Appendix C, Table 6.

4.2 Summary

Our black box findings may be summarized as follows. First, there is strong evidence for SEARCH: search volatility and search breadth are both larger after failure than after success; inertia is larger after success than after failure; and there is directional bias, meaning that increases (decreases) tend to be followed by larger (smaller) contributions after success than after failure. Despite subjects' lack of information, however, the overall pattern of the dynamics is in line with previous experiments. In free-riding games ($e = 1.6$), for example, contributions deteriorate at almost the same rate as in previous studies (Ledyard, 1995; Chaudhuri, 2011; Burton-Chellew and West, 2013).

5 Non-black box data

In this section, we assess our black box findings in light of the data from grey box and from the other two treatments, standard and enhanced. In particular, we shall investigate whether SEARCH describes only black box behavior or whether its components also persist in the other settings, that is, where players gain experience and/ or have explicit information about the structure of the game and others' actions (and payoffs).

Figure 4 illustrates play in the different treatments, averaged over individuals and sessions. Different information clearly makes a difference in the level of contributions and/ or convergence rates (see Burton-Cellew and West, 2013). However, the features of SEARCH persist in all sessions and in all treatments, as is summarized in Figure 5.[18]

[18] Appendix C, Figure 7 contains a full summary of all SEARCH components in all treatments for both rates of return separately.

Figure 4: Mean play all treatments (black box, standard, enhanced).

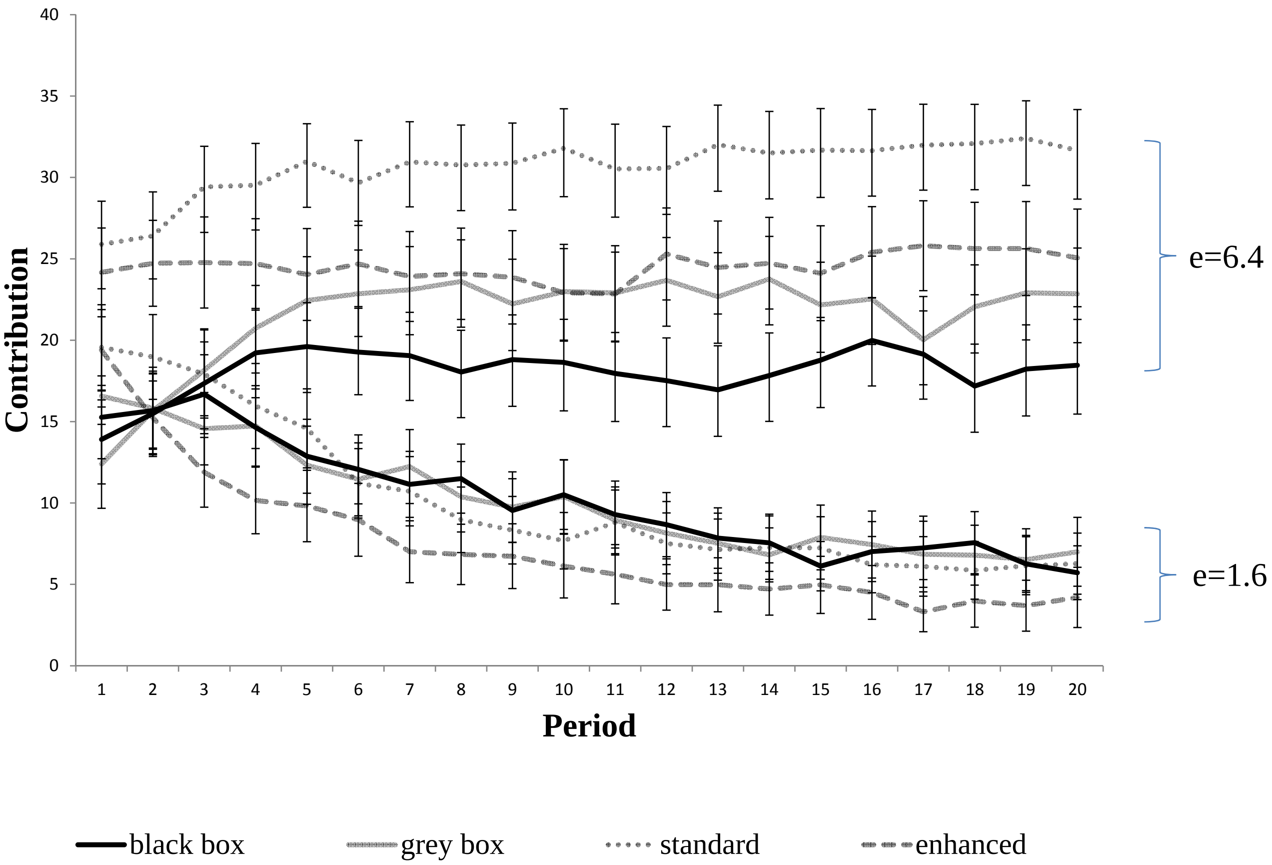
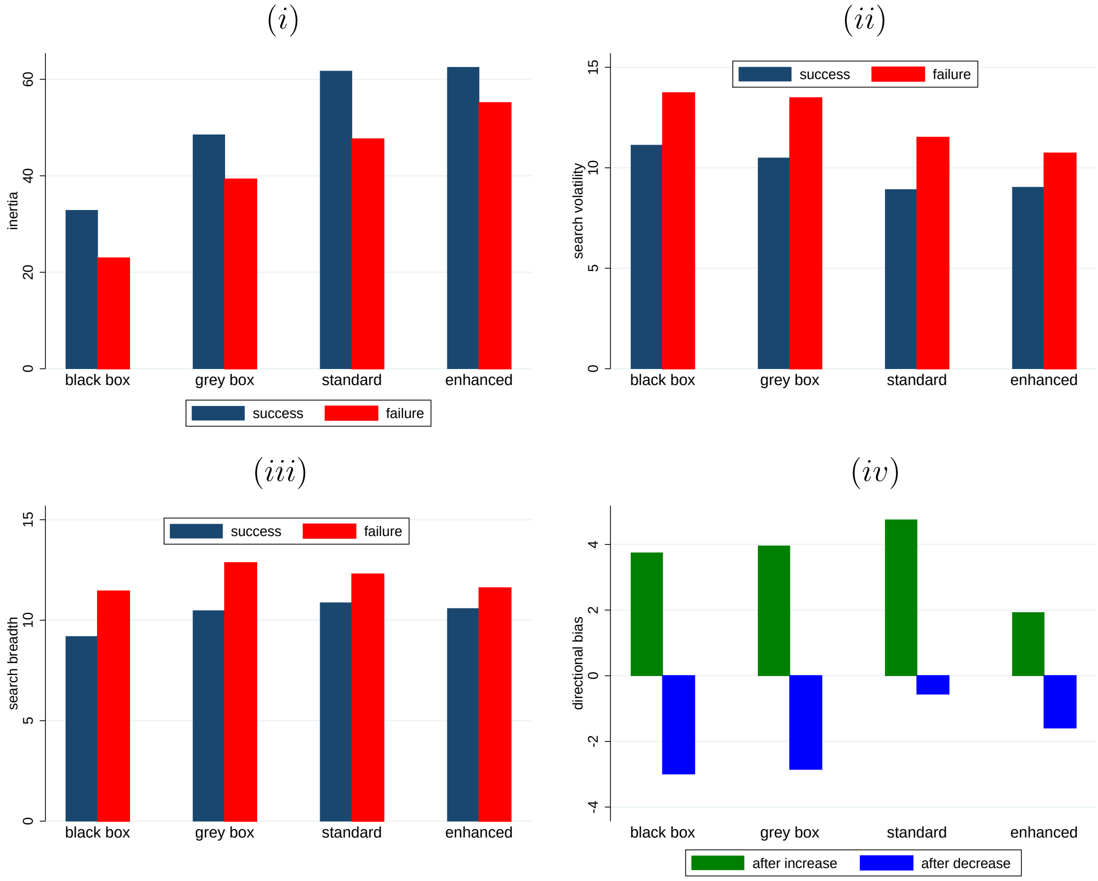


Figure 4 illustrates play in all treatments.

5.1 Black box versus grey box

First, we shall investigate 'grey box' play. Recall that grey box means that black box is played after either enhanced or standard. Even though subjects are explicitly told that a separate experiment is started after the first stage of the experiment, players may or may not make inferences about the game structure. We shall investigate the consequences of this effect in comparison with (pure) black box behavior. Our analysis reveals that although there are differences in average levels of contribution in the two cases (Figure 4), all of the features of SEARCH are robust for both rates of return. In particular, the levels of search volatility and search breadth are not significantly different in the

Figure 5: All learning features (all treatments).



The bar charts summarize the four search components in all treatments. Except for inertia, the values are in units of contributions. The respective panels are: (i) asymmetric inertia (in percent), (ii) search volatility, (iii) search breadth, (iv) directional bias.

two cases (Appendix C, Tables 7 and 8). The differences in inertia rates conditional on success-versus-failure are similarly robust, however, the level of inertia is higher in grey box than in black box (Appendix C, Table 7). Directional bias is unchanged in terms of its size and significance levels compared to black box (Appendix C, Table 10). We conclude that grey box differs from black box with regard to the absolute level of inertia, but not qualitatively with respect to any of the SEARCH components.

5.2 Black box versus standard and enhanced

Finally, we consider whether the standard and enhanced treatments lead to different conclusions. The situation is summarized in Figure 5.[19] Search volatility, though smaller

[19]See Appendix C, Figure 7 for a summary for both rates of return separately.

in absolute size in the standard treatment and even smaller in the enhanced treatment, is higher after failure than after success, just as in the black box case (see panel (*i*) in Figure 5). Moreover, this difference is significant at the 1 percent level (see Appendix C, Table 7). Search breadth exhibits the same success-versus-failure difference as in black box (see panel (*ii*) in Figure 5). It continues to be statistically significant at the 1 percent level (see Appendix C, Table 8). Panel (*iii*) in Figure 5 illustrates that inertia rates decrease after failure in all treatments. This success-failure differential in inertia rates is statistically significant in all four treatments, and there is higher absolute inertia in non-black box treatment data (see Appendix C, Table 7). Finally, panel (*iv*) in Figure 5 illustrates the presence of directional biases across all four treatments. We qualitatively confirm the increase-decrease directional bias from the black box data when analyzing the non-black box data at levels that are statistically significant (see Appendix C, Table 10).

6 Discussion

Here we shall compare our findings with those of Nevo and Erev (2012) (see also Erev and Haruvy, 2013) and with those of Bayer et al. (2013). Nevo and Erev's (2012) learning model applies to one-player decisions under uncertainty when feedback is limited to payoff consequences of chosen actions, which is analogous to our 'black box' treatments. Their learning model features exploration/exploitation with inertia. During exploration, the agent chooses randomly amongst available strategies. The agent enters exploration with probability one in the first period, and with history-dependent (positive) probability thereafter. A high-payoff history leads to a switch to exploitation under which a strategy is chosen that maximizes the player's estimated payoff given random samples of relative payoffs from past actions. These samples are biased towards recent realizations. Moreover, exploration/exploitation is always subject to inertia, that is, there is a positive probability of no strategy adjustment in a given period. In Nevo and Erev's (2012) model, inertia rates go down when a player is "surprised" by a change in payoffs. This surprise is symmetric in success and failure in their model. Our analysis will reveal an important asymmetry concerning the direction of payoff changes.

Our learning model relates to Nevo and Erev's (2012) in the following ways. First, we confirm the presence of inertia, but inertia is asymmetric with respect to the direction of payoff changes. Indeed, we find an inertia base-rate of 0.07 when there is no payoff change in the black box treatments, compared with an average inertia rate of 0.23 after a strict payoff increase and 0.32 after a strict payoff decrease. Using Erev and Haruvy's (2013) terminology, an especially negative surprise "triggers change" in our game. This finding is robust across all our information treatments. Additionally, we find that search volatility and search breadth are larger after failure; these features are not part of Nevo and Erev's model.

Bayer et al. (2013) propose a directional learning model that applies to voluntary contributions games similar to ours.[20] At the heart of directional learning is an asymmetry in adjustment depending on whether previous adjustments lead to success or failure. In the model of Bayer et al., a player's contribution this period lies closer to whichever one of the two previous contributions led to a higher payoff. (This is a somewhat stronger form of directional bias than the one we consider.) For the treatments in our experiment that are directly comparable with those considered in Bayer et al. (2013) (i.e. for the pure black box data and for the low rates of return), this stronger directional adjustment model is validated.

There are a number of additional features that are worthy of discussion. First, we find that the strategy space - integers between zero and forty - is not explored randomly. Indeed, 3,081 out of a total of 4,960 black box decisions were chosen as multiples of five, and 2,498 were multiples of ten. We conjecture that the shape and support of the probability distribution vary among individuals and also with the underlying strategy space. Second, our model considers only adjacent periods, so presents a somewhat extreme view of the recency effect noted in Erev and Haruvy (2013). The restriction to adjacent periods is something we plan to relax in future studies. In order to interpret directional biases in this experiment, however, we restrict ourselves to neighbouring contribution adjustments. Finally, we can make a connection with the "hot stove effect" (Denrell and March, 2001; see also Erev and Haruvy, 2013), which states that players shy away from strategies that led to failure. This effect may explain the directional bias in the following sense: an increase which leads to failure turns high contributions into a 'hot stove' and higher contributions are avoided in the next period, similarly a failure-inducing decrease leads to avoidance of lower contributions.

7 Conclusion

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 the game proceeds. In this paper, by contrast, we have examined situations in which players have **no** information about the strategic

[20] There are some subtle differences related to confusion in their learning condition and in our black box. These may explain the slightly different aggregate contribution patterns.

environment. This takes us to the other end of the information spectrum. Players in such environments must feel their way to equilibrium based solely on the pattern of their own realized payoffs. To highlight the potential differences between adaptive learning and best-reply dynamics, we implemented a black box environment in a voluntary contributions game and compared the resulting behavior to play under intermediate and rich information treatments. We identified four key features of the learning dynamics – asymmetric inertia, search volatility, search breadth, and directional bias. Although these components have precursors in both psychology and biology, they have not been given the precise formulation that we propose here. It turns out all four features are validated at high levels of statistical significance in the black box treatment. Moreover, they are present even when players gain more experience and/ or have more information across information treatments for both rates of return in the voluntary contributions games. Whether this remains true for other classes of games is an open question for future research.

References

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

J. Andreoni, “An experimental test of the public goods crowding-out hypothesis”, *The American Economic Review 83*, 1317-1327, 1993.

J. Andreoni, “Cooperation in public-goods experiments: kindness or confusion?” *The American Economic Review 85*, 891-904, 1995.

R.-C. Bayer, E. Renner and R. Sausgruber, “Confusion and learning in the voluntary contributions game”, *Experimental Economics 16,* 478-496, 2013.

U. Ben Zion, I. Erev, E. Haruvy, and T. Shavit, “Adaptive behavior leads to under-diversification”, *Journal of Economic Psychology 31*, 985-995, 2010.

J. Björnerstedt, and J. Weibull, “Nash Equilibrium and Evolution by Imitation”, in Arrow, K. and Colombatto, E. (Eds.), *Rationality in Economics*, New York: Macmillan, pp. 151-171, 1993.

M. N. Burton-Chellew and 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 and T.-H. Ho, “Experience-weighted Attraction Learning in Normal Form Games”, *Econometrica 67*, 827-874, 1999.

C. F. Camerer, T.-H. Ho and J.-K. Chong, “Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games”, *Journal of Economic Theory 104*, 137-188, 2002.

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.

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

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

V. P. Crawford, "Boundedly Rational versus Optimization-Based Models of Strategic Thinking and Learning in Games", *Journal of Economic Literature 51*, 512-527, 2013.

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 and C. Schippers, "On evolutionary exploration and exploitation", *Fundamenta Informaticae 35*, 35-50, 1998.

I. Erev and A. Rapoport, "Coordination, 'magic,' and reinforcement learning in a market entry game" *Games and Economic Behavior 23*, 146-175, 1998.

I. Erev and E. Haruvy, "Learning and the economics of small decisions", J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, forthcoming, Princeton University Press, 2013.

I. Erev and A. E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review 88,* 848-881, 1998.

E. Fehr and S. Gächter, "Cooperation and Punishment in public goods experiments", *American Economic Review 90*, 980-994, 2000.

E. Fehr and S. Gächter, "Altruistic punishment in humans", *Nature 415*, 137-140, 2002.

P. J. Ferraro and C. A. Vossler, "The source and significance of confusion in public goods experiments", *The B.E. Journal in Economic Analysis and Policy 10*, 53, 2010.

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

U. Fischbacher, and S. Gächter, "Social preferences, beliefs, and the dynamics of free riding in public good experiments", *The American Economic Review 100*, 541-556, 2010.

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

D. Friedman, S. Huck, R. Oprea and S. Weidenholzer, "From imitation to collusion: Long-run learning in a low-information environment", *Discussion Papers in Economics of Change SP II 2012-301*, WZB , 2012.

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

J. K. Goeree, C. A. Holt and S. K. Laury, "Private costs and public benefits: Unraveling the effects of altruism and noisy behavior" *Journal of Public Economics 83*, 255-276, 2002.

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

R. M. Harstad and R. Selten, "Bounded-Rationality Models: Tasks to Become Intellectually Competitive", *Journal of Economic Literature 51*, 496-511, 2013.

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

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

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

S. Huttegger, B. Skyrms and K J. S. Zollman, "Probe and Adjust in Information Transfer Games", *Erkenntnis 79*, 835-853, 2014.

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

M. Isaac, K. F. McCue and C. R. Plott, "Public goods provision in an experimental environment", *Journal of Public Economics 26*, 51-74, 1985.

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

D. Kahneman and A. Tversky, "Choices, Values, and Frames". *American Psychologist 39*, 341-350, 1984.

D. P. Kraines and V. Y. Kraines, "Pavlov and the Prisoner's Dilemma", *Theory and Decision 26*, 47-79, 1989.

J.-F. Laslier and B. Walliser, "Stubborn Learning", *Ecole Polytechnique WP-2011-12*, 2011.

J. O. Ledyard, "Public Goods: A Survey of Experimental Research", in J. H. Kagel and 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.

U. Motro and A. Shmida, “Near-far search: an evolutionary stable foraging strategy”, *Journal of Theoretical Biology 173*, 15-22, 1995.

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

I. Nevo and I. Erev, “On surprise, change, and the effect of recent outcomes”, *Frontiers in Cognitive Science*, 2012.

M. Nowak and K. Sigmund, “A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner’s Dilemma game”, *Nature 364*, 56-58, 1993.

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

T. R. Palfrey and J. E. Prisbrey, “Altruism, reputation and noise in linear public goods experiments”, *Journal of Public Economics 61,* 409-427, 1996.

T. R. Palfrey and J. E. Prisbrey, “Anomalous behavior in public goods experiments: how much and why?”, *The American Economic Review 87,* 829-846, 1997.

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

A. Rapoport and A. M. Chammah, *Prisoner’s Dilemma: A Study in Conflict and Cooperation*, University of Michigan Press, 1965.

A. Rapoport, D. A. Seale, and J. E. Parco, “Coordination in the aggregate without common knowledge or outcome information”, in: R. Zwick and A. Rapoport (Eds.): *Experimental Business Research*, pp. 69-99, 2002.

S. Rick and R. A. Weber, “Meaningful learning and transfer of learning in games played repeatedly without feeback”, *Games and Economic Behavior 68*, 716-730, 2010.

A. E. Roth and 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.

R. Saran and R. Serrano, "Ex-post regret heuristics under private values (I): fixed and random matching", *Journal of Mathematical Economics*, forthcoming, 2014.

R. Saran and R. Serrano, "Ex-post regret heuristics under private values (II): 2-by-2 games", *Journal of Mathematical Economics*, forthcoming, 2014.

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

R. Selten and J. Buchta, "Experimental Sealed Bid First Price Auction with Directly Observed Bid Functions", in: Budescu, D., Erev, I., Zwick, R. (Eds.), *Games and Human Behavior,* Essays in Honor of Amnon Rapoport, pp. 79-104, 1998.

R. Selten and 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.

B. Skyrms, *Signals: Evolution, Learning and Information*, Oxford University Press, 2010.

B. Skyrms, "Learning to signal with 'probe and adjust' ", *Episteme 9*, 139-150, 2012.

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

P. Suppes and A. R. Atkinson, *Markov Learning Models for Multiperson Situations*, Stanford University Press, 1959.

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

R. Tietz and 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.

R. A. Weber, "Learning' with no feedback in a competitive guessing game", *Games and Economic Behavior 44*, 134-144, 2003.

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

Appendices

Appendix A: Instructions

Participants received the following on-screen instructions (in z-Tree) at the start of the game. The set of instructions in standard and enhanced were the same, different instructions were given in black box. In black box, participants had to click an on-screen button saying, “I confirm I understand the instructions” before the game would begin. The same black box instructions were used for both rates of return. The standard/ enhanced instructions differ with respect to the relevant numbers for the two rates of return, and the example is adequately modified.

Black box

The following instructions were used in black box.

Beginning of instruction.

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 (click to confirm)

End of instructions.

Standard and enhanced

Here, we present the instructions for the rate of return $e = 1.6$ (one example of an instruction slide is given in Figure 6). The same instructions apply to standard and enhanced treatments. Equivalent instructions apply for the rate of return $e = 6.4$.

Beginning of instructions.

Instructions

Welcome! You are about to participate in an experimental study of human decision making. Thank you for your participation in our study. Please pay careful attention to the instructions on the following screens. If you wish to return to a previous screen, press the left arrow key. You are now taking part in an economic experiment. If you read the following instructions carefully, you can, depending on your decisions, earn a considerable amount of money. It is therefore very important that you read these instructions with care.

Everybody has received the same instructions. It is prohibited to communicate with the other participants during the experiment. Should you have any questions please ask us by raising your hand. If you violate this rule, we shall have to exclude you from the experiment and from all payments.

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

We are researching the decisions people make.

This part of the experiment is divided into separate rounds. In all, this part of the experiment consists of 20 repeated rounds. In each round the participants are assembled into groups of four. You will therefore be in a group with 3 other participants. The composition of the groups will change at random after each round. **In each round your group will therefore probably consist of different participants.**

In each round the experiment consists of two stages. At the first stage everyone has to individually decide how many credits they would like to contribute to a group project. These decisions have consequences for people's earnings. At the second stage you are informed of the contributions of the three other group members to the project and how many credits you have received from the group project. New groups are then randomly formed and the process repeats itself, with everyone again deciding how many credits they would like to contribute. This process will repeat 20 times.

At the beginning of each round, each participant receives 40 credits. In the following we call this your endowment. Your task is to decide how to use your endowment. You have to decide how many of the 40 credits you want to contribute to a group project and how many of them to keep for yourself. The consequences of your decision are explained in detail in the following slides.

Please note:

- The set-up is anonymous. You will not know with whom you are interacting.
- You will interact with a random set of 3 players in each round.
- Your decisions, and your earnings will remain anonymous to other players, even after the session has ended.

We now provide an animated illustration of a hypothetical scenario to demonstrate how the experiment works. In the following demonstration, we will use these green 'disks' to represent 'credits'. 1 disk equals 1 credit.

Example

Remember! Credits = real money, and the more credits a player has at the end of the experiment, the more money that player will receive.

There are 4 players in each group, For the sake of convenience, we refer to these players as Player A, Player B, Player C, and Player D.

These 4 players then each receive an endowment of 40 credits.

Each of the players can then choose to make a contribution to the group project. They can contribute anything from zero to 40 credits. Non contributed credits are kept in the player's private account. They do this at the same time and anonymously.

Each player is then informed of the decisions of all their group members, although no one will know who the players are and they will randomly change in each round.

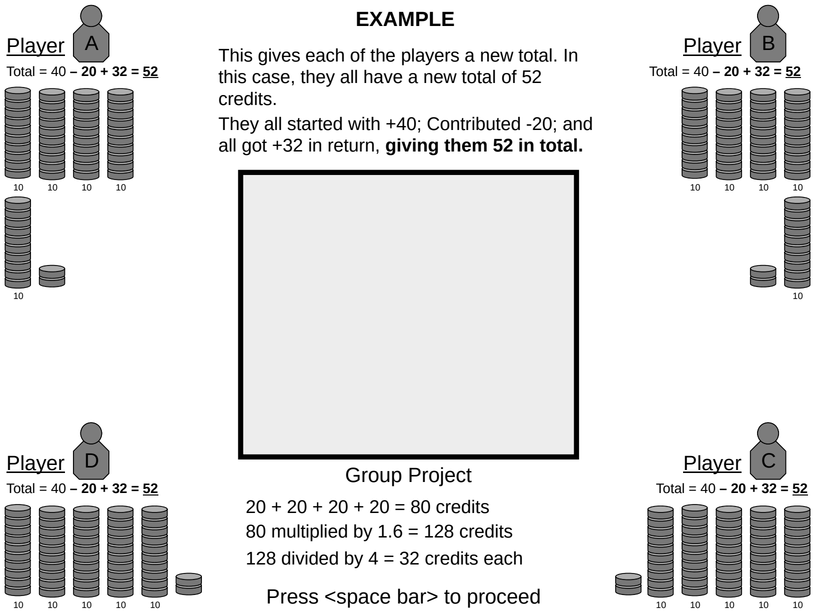
After all 4 players have made their decision to contribute or not, and by how much, the resulting total of contributed credits is automatically MULTIPLIED.

In your experiment, in every round, and in this example here, the total will be multiplied by 1.6. So for an imaginary total of 10 credits, this would result in 16 credits.

This new total is then always shared out equally between all 4 players. So, after the multiplication occurs, each player receives one quarter of the credits that are in the group project.

In this example, each player chooses to contribute 20 of their 40 credits. This means the group project has 80 credits; 20 + 20 + 20 + 20 = 80 credits. Which when multiplied by 1.6 results in 128 credits; 80 multiplied by 1.6 = 128 credits. These 128 credits are then shared out equally, giving 32 credits back to each player; 128 divided by 4 = 32 credits each. This gives each of the players a new total. In this case, they all have a new total

Figure 6: **Standard and enhanced instructions; example slide for** $e = 1.6$**.**



of 52 credits.

They all started with +40; Contributed -20; and all got +32 in return, **giving them 52 in total.**

That is the end of the demonstration.

Remember, this was just one of many possible scenarios. In the rounds you will now play, all players are free to choose how much they wish to contribute to the pot.

End of instructions.

Appendix B: On-screen output in each treatment

The following three tables summarize the post-decision feedback information that participants received in z-Tree under the three treatments: (a) feedback in black-box; (b) feedback screen #1 in standard and enhanced (identical in both treatments); (c) feedback screen #2 in standard and enhanced (dashed lines border the information that was only shown in enhanced).

(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.

Appendix C: Regression outputs

Table 3: **Search volatility (black box).**

(a) Test 1: *search volatility* null rejected ($W = 83.6$, $p < 0.01$)	(b) success	(c) failure	(d) total
standard deviation	13.7	11.1	12.6
mean	+0.8	-1.2	-0.2
frequency	2,266	2,198	4,464
Test 2: *rate of return* null not rejected ($W = 0.1$, $p > 0.87$)	$e = 1.6$	$e = 6.4$	total
standard deviation	12.5	12.6	12.6
mean	-0.6	+0.2	-0.2
frequency	2,232	2,232	4,464
Test 3: *phase* null not rejected ($W = 0.3$, $p > 0.59$)	phase 1	phase 2	total
standard deviation	12.7	12.4	12.6
mean	-0.2	-0.1	-0.2
frequency	2232	2232	4464
Test 4: *diminishing volatility* null rejected ($W = 47.0$, $p < 0.01$)	periods 1-9	periods 10-20	total
standard deviation	13.7	11.6	12.6
mean	-0.1	-0.2	-0.2
frequency	1,984	2,480	4,464

We use black box data from phases 1 and 2. We perform Levene's robust variance tests for search volatility (Test 1), differences in distribution for the two rates of return (Test 2), for phases one and two (Test 3), and diminishing volatility (Test 4). W is the Levene's test statistic.

Table 4: **Search breadth (black box)**.

	Coefficient (test statistic)
1 if "failure"	2.73 (9.77)**
1 if "period"<10	2.01 (2.43)*
1 if "phase"=1	2.32 (10.81)**
Group dummies	not significant
Individual fixed effects	not listed
Periods	not significant
Constant	2.42 (3.52)**
Observations	4,464
Adjusted R^2	0.25

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform an OLS regression of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering.

Table 5: **Inertia (black box)**.

	Coefficient (test statistic)
1 if "failure"	-0.52 (-10.39)**
1 if "period"<10	0.22 (1.37)
1 if "phase"=1	-1.21 (-50.38)**
Group dummies	not significant
Individual fixed effects	not listed
Periods	negative until period 6, not significant thereafter
Cut	0.04 (0.47)**
Observations	4,464
Adjusted pseudo-R^2	0.40

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform an ordered probit regression of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering. The pseudo-R^2 indicates the level of improvement over the intercept model (not interpretable as percentage of variance explained).

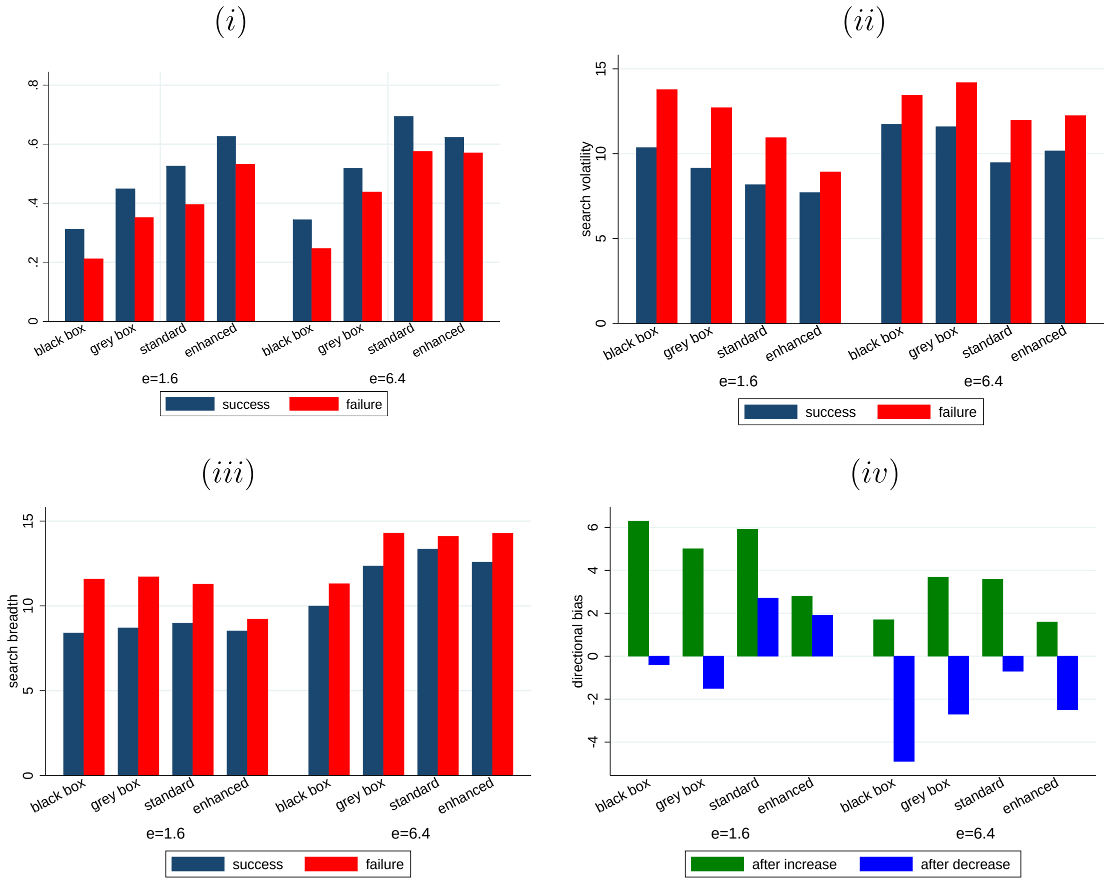
Table 6: **Directional bias (black box)**

	Coefficient (test statistic)
1 if "up"	-10.48 (8.93)**
1 if "up" and "success"	5.06 (5.20)**
1 if "down"	3.62 (2.41)*
1 if "down" and "success"	-3.33 (3.37)**
Individual fixed effects	not listed
Periods	negative, significant
Group dummies	not significant
1 if "phase"=1	-24.90 (13.16)**
Observations	3,222
Adjusted R^2	0.14

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform OLS regressions (without constant) to test for the directional bias of adjustments for each directional success/ failure impulse controlling for phase, group, period and individual fixed effects with individual-level clustering.

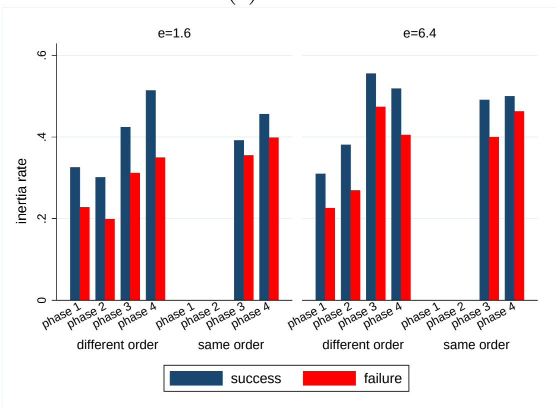
Figure 7: SEARCH (all treatments, both rates of return).



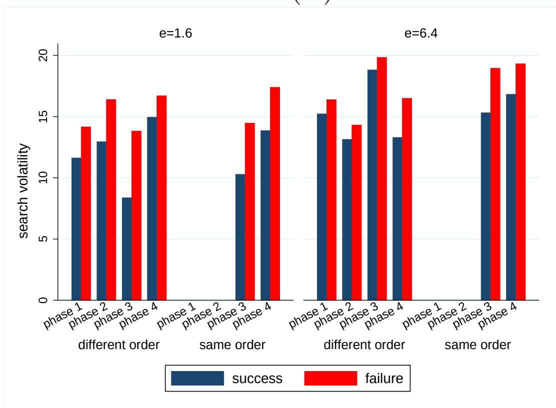
The bar charts summarize the four search components in all treatments for both rates of return separately. The respective panels are: (i) inertia, (ii) search volatility, (iii) search breadth, (iv) directional bias. For (ii) – (iv), the y-axis are units of contributions; for (i), the y-axis are probabilities.

Figure 8: SEARCH (all treatments, order effects).

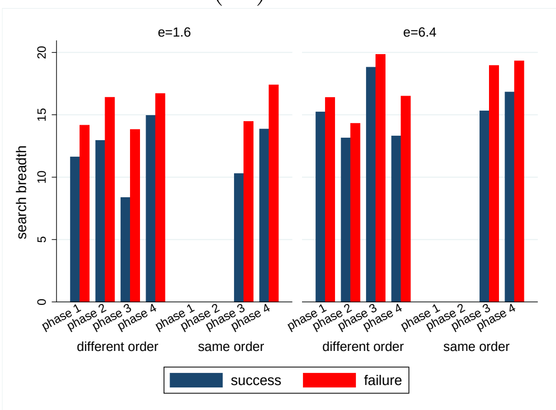
(*i*)



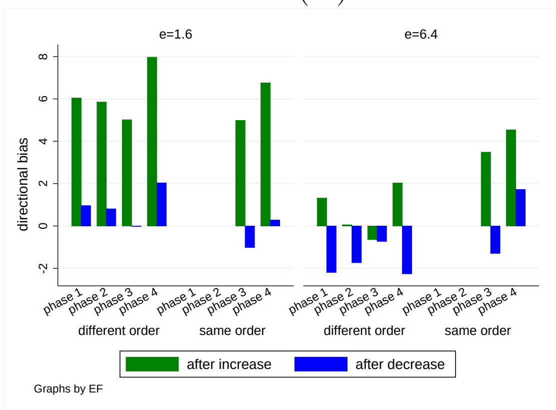
(*ii*)



(*iii*)



(*iv*)



The bar charts summarize the four search components in all treatments for both rates of return and for different orders separately. The case when the "same order" from the first two phases is repeated in the last two phases is distinguished from when a "different order" is played. The respective panels are: (i) inertia, (ii) search volatility, (iii) search breadth, (iv) directional bias. For (ii) – (iv), the y-axis are units of contributions; for (i), the y-axis are probabilities. Note that all four elements of SEARCH are present in all cases.

Table 7: **Search volatility (non-black box).**

Test	success	failure	total
Test 1: *standard* null rejected ($W = 83.1$, $p < 0.01$)			
standard deviation	8.9	11.5	10.3
mean	+1.5	-2.4	-0.2
frequency	2,314	1,862	4,176
Test 2: *enhanced* null rejected ($W = 30.6$, $p < 0.01$)			
standard deviation	9.0	10.7	9.9
mean	+0.6	-1.2	-0.3
frequency	2,189	2,131	4,320
Test 3: *grey box* null rejected ($W = 61.3$, $p < 0.01$)			
standard deviation	10.5	13.5	12.1
mean	+1.0	-1.2	-0.0
frequency	2,045	1,987	4,032

We use non-black box data; phases 3 and 4 for grey box, phases 1-4 for standard and enhanced. We perform Levene's robust variance tests for search volatility in standard (Test 1), enhanced (Test 2), and grey box (Test 3. Recall W is the Levene's test statistic.

Table 8: **Search breadth (non-black box)**.

	(i) **Grey box: Coefficient (test statistic)**	*(ii)* **Standard: Coefficient (test statistic)**	*(iii)* **Enhanced: Coefficient (test statistic)**
1 if "failure"	2.10 (5.10)**	1.64 (4.34)**	1.10 (2.73)**
1 if "period"<10	not significant	not significant	not significant
Phase dummies	"phase"=3, positive	not significant	if "phase"=2, negative
Group dummies	not significant	not significant	not significant
Individual fixed effects	mostly negative	mostly negative	mostly negative
Periods	mostly not significant	mostly not significant	mostly not significant
Constant	16.28(11.44)**	36.83 (24.38)**	12.27 (8.63)**
Observations	2,259	1,861	1,777
Adjusted R^2	0.42	0.41	0.54

*: $p < 0.05$; **: $p < 0.01$.

We use non-black box data; phases 3 and 4 for grey box (panel *(i)*), phases 1-4 for standard (panel *(ii)*) and enhanced (panel *(iii)*). We perform OLS regressions of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering.

Table 9: **Inertia (non-black box)**.

	Standard & Enhanced coefficient (test statistic)	**Grey box coefficient (test statistic)**
1 if "failure"	-0.23 (2.28)**	-0.29 (2.72)**
1 if "enhanced"	0.02 (0.08)	n/a
Phase dummies	5.11 (26.49)** for "phase"=1 others not significant	5.17 (26.59)** for "phase"=3
Group dummies	not significant	not significant
Individual fixed effects	not listed	not listed
Periods	not significant	not significant
Cut	6.13 (20.12)**	6.48 (20.78)**
Observations	8,464	4,032
Adjusted pseudo-R^2	0.40	0.34

*: $p < 0.05$; **: $p < 0.01$.

We use non-black box data; phases 1-4 for standard and enhanced (panel *(i)*), and phases 3 and 4 for grey box (panel *(ii)*). We perform an ordered probit regression of absolute adjustments on success-versus-failure adjustments controlling for treatment effects phase, group, period and individual fixed effects with individual-level clustering. The pseudo-R^2 indicates the level of improvement over the intercept model (not interpretable as percentage of variance explained).

Table 10: **Directional bias (non-black box)**

	Treatment	**Adjustment: coefficient (test statistic)**
1 if "up"	black box	-10.44 (8.83)**
1 if "up" and "success"	black box	5.07 (5.18)**
1 if "down"	black box	3.59 (2.37)*
1 if "down" and "success"	black box	-3.31 (3.34)**
1 if "up"	grey box	-10.61 (7.37)**
1 if "up" and "success"	grey box	6.55 (5.73)**
1 if "down"	grey box	4.69 (3.16)**
1 if "down" and "success"	grey box	-2.73 (2.70)**
1 if "up"	standard	-4.81 (3.86)**
1 if "up" and "success"	standard	9.85 (8.56)**
1 if "down"	standard	10.82 (5.33)**
1 if "down" and "success"	standard	-0.06 (0.04)
1 if "up"	enhanced	-5.96 (3.82)**
1 if "up" and "success"	enhanced	3.25 (2.49)*
1 if "down"	enhanced	8.04 (4.10)**
1 if "down" and "success"	enhanced	-1.10 (0.76)
Group dummies		not significant
Phase dummies		positive, significant
Period dummies		not significant
Treatment dummies		black box negative, others positive
Return rate dummies		negative, significant
Individual fixed effects		not listed
Observations		9,119
Adjusted R^2		0.15

*: $p < 0.05$; **: $p < 0.01$.

We use all the data. We perform OLS regressions (without constant) to test for the directional bias of adjustments for each directional success/ failure impulse in each treatment controlling for phase, group, period and individual fixed effects with individual-level clustering.