

Self-efficacy beliefs and imitation: a two-armed bandit experiment

Stefania Innocenti, Robin Cowan

PII: S0014-2921(19)30005-4
DOI: <https://doi.org/10.1016/j.euroecorev.2018.12.009>
Reference: EER 3214

To appear in: *European Economic Review*

Received date: 17 July 2018
Accepted date: 28 December 2018

Please cite this article as: Stefania Innocenti, Robin Cowan, Self-efficacy beliefs and imitation: a two-armed bandit experiment, *European Economic Review* (2019), doi: <https://doi.org/10.1016/j.euroecorev.2018.12.009>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Self-efficacy beliefs and imitation: a two-armed bandit experiment ☆

Stefania Innocenti^{a,d,*}, Robin Cowan^{b,c,d}

^a *SSEE, University of Oxford, Oxford, UK*

^b *BETA, Université de Strasbourg, Strasbourg, France*

^c *IUF, France*

^d *UNU-MERIT, Maastricht University, Maastricht, The Netherlands*

Abstract

It is generally believed that individuals imitate others to gain status, minimise regret or simply ameliorate their performance. Psychology provides a complementary explanation: imitation becomes appealing when agents have little faith in their abilities. We investigate the extent to which self-efficacy beliefs affect agents' propensities to imitate others. We propose an experimental task, which is a modified version of the two-armed bandit. We measure participants' self-assessed self-efficacy, then study individual learning. Subsequently, we measure how individuals use the information they gather observing a randomly selected group leader. We find that, in stable environments, a 1% increase in individual self-efficacy reduces the propensity to imitate others by 3%.

Keywords: Learning, imitation, laboratory experiment, self-efficacy beliefs

☆We would like to thank Matthew Embrey, Eleonora Nillesen, Nico Rasters, Arno Riedl, Elisabet Rutstrom, Martin Strobel and Marie Claire Villeval and two anonymous referees for their suggestions. We also acknowledge the helpful comments of all the participants to the BEElab proposal meeting in Maastricht in February 2015, the ASFEE 2015 conference in Paris, the MLSE seminar in Maastricht in December 2015, the seminar at the University of Florence, IMEBESS 2016 conference, seminar at Nuffield college-Oxford in November 2016 and seminar at FLAME University (Pune -India) in April 2018. This research was supported by the Institut Universitaire de France. The experiment was conducted following the peer-approved procedures established by Maastricht University's Behavioral and Experimental Economics Laboratory (BEElab). Our study was approved at a public ethics review and project proposal meeting that is mandatory for all scholars wishing to use the BEElab facilities. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

*corresponding author

Email address: stefania.innocenti@smithschool.ox.ac.uk (Stefania Innocenti)

1. Introduction

When undertaking novel tasks, people can arrive at adaptive solutions by means of either individual learning or observational learning. The former involves consideration of one's own successes and failures; the latter considers those of others [54]. Amongst the most common forms of observational learning is imitation, which consists in copying the actions of others.¹ Imitation is widespread in the natural world: an animal can, for instance, avoid the costs related to sampling different feeding locations simply by imitating the actions of its conspecifics [60]. Humans also use imitation in a wide variety of circumstances [43].

Imitation can be motivated by preferences for conformism. When individual preferences bend towards or depend on those of others, motivational conformism might induce agents to behave like their peers [39]. Considerable evidence shows that peer effects can explain imitative behaviour regarding educational attainment [75], financial decisions [20, 23], criminal activities [16]. However, most theoretical and experimental work in economics tend to depict imitation as a simple "economizing behavior" [66, 3]. By copying the action of another decision-maker, the imitator minimises "the effort required to think through a choice problem" [66, p. 193] and maximises his or her chances of finding an optimal choice [65]. Thus, one of the standard arguments in economics is that imitative behaviour occurs due to payoff-enhancement motives: the decision maker wants to simultaneously enhance his or her performance and to reduce the decision costs associated with comparing alternative solutions.

The psychology literature provides a complementary explanation: imitation becomes appealing when individuals have little confidence in their own abilities to reach a designated goal, that is they have low self-efficacy beliefs [11]. The importance of strong self-confidence is widely acknowledged in the economics literature. Many studies have shown that very confident individuals are more likely to have a superior ability to collect information [41], to engage in entrepreneurial activities [50, 27] and undertake excessive asset trading [15]. Much less attention has been devoted to the case in which individuals have little faith in their own capacities, i.e. low self-efficacy beliefs. And, more specifically, no study has yet tested whether and to what extent, when making economic decisions, individ-

¹Concepts like imitation, emulation and mimicry are highly debated in several disciplines which include neuroscience, psychology, anthropology, sociology and animal behaviour. Hurley and Chater [46, pp. 1-52] provide a detailed review of the literature on the topic. Despite subtle differences, these concepts all represent forms of observational learning and mostly have to do with the act of copying another's action or performance. In this paper, we will focus on imitation as defined in the economics literature. We consider imitation as a choice method in which the agent makes a choice after having observed the actions and the outcomes of one other decision maker, thus having the chance to match or mimic such behaviour [65, 66].

ual self-efficacy beliefs complement payoff motives and affect the propensity to imitate others.

This paper addresses this gap and examines how individual variation in social learning dynamics is affected by introspective beliefs. We use a modified version of the common two-armed bandit game as the experimental setting. After gathering information on individuals' self-efficacy beliefs, agents make a series of choices from which they derive real payoffs. In a control treatment, we study individual learning patterns. In the following observational treatment, we analyse how individuals learn when observing the actions and rewards obtained by a randomly selected group "leader" who plays before everybody else. This specific frame transforms the leader into a pioneer. Our goal is to understand how people process the information acquired by observing the leader's experience and the extent to which introspective beliefs about one's self-efficacy affect the propensity to imitate others. Our hypothesis is that, holding payoff motivations constant, individuals with low self-efficacy beliefs are more inclined to imitate the leader's actions.

Our analysis runs along the following lines. We (i) posit that subjects employ one of a set of individual learning strategies, (ii) for each strategy we derive the conditional probability of behaviour, given the available information, (iii) use these probabilities to generate likelihoods of the laboratory data, and (iv) select the strategy that best fits the data. Having studied individual learning we (v) study, following the same approach, how individuals learn observing others. Lastly, (vi) we investigate whether individuals' propensity to imitate correlates with their self-efficacy beliefs, i.e. self-perceptions of incompetence. This is where our contribution lies.

We find significant evidence to support our hypothesis: particularly in stable environmental conditions, higher self-efficacy beliefs reduce the propensity to imitate others. Results are inconclusive for more volatile environments.

The remainder of the paper is organised as follows. Section 2 briefly reviews the literature on learning and imitation, making connections to the social psychology perspective. Section 3 presents the general design of our experiment. Section 4 presents our estimation strategy. We test our main hypothesis in section 5. Lastly, section 6 concludes and discusses the limitations of this study.

2. Related literature

Individual learning has been at the centre of numerous scholarly investigations. Many have examined reinforcement learning models [24, 25, 44, 49, 30, 7, 8] including modifications to account for recency and experimentation [67, 37, 38], and exploitation and inertia [35, 59, 36, 58].

These models, independently from their complexity and number of parameters, focus on the direct effects that personally obtained payoffs have on choice behaviours. These are among the most important drivers of human behavioural adjustments. Yet, as with most other animals, humans change their behavioural repertoire not only through personal experience but also by observing the actions of others [6]. Specifically, when it is possible to observe the actions and payoff obtained by someone else who is playing the same game as oneself, then following his or her behaviour becomes appealing.

The literature identifies several sources of imitative behaviour. Peer effects have been found to affect, for instance, people's decision to purchase financial assets, engage in risky [48, 42] or entrepreneurial activities [57]. This can happen simply because people want to gain social approval and/or status [1]. An individual might also decide to behave similarly to his or her peers due to social regret [48, 42], that is when disutility experienced from not taking an action is less intense if others have chosen the same. Alternatively, imitative behaviour might be undertaken to purposely extract information about the state of the world from the actions of others [39]. All these instances imply that individual preferences and utility depend on the actions of others. But there is another reason for imitation: agents can decide to imitate peers to simply enhance their performance and choose the strategy that yields the highest payoffs.

Most theoretical work focuses on this motivation [48] to explain imitative behaviour [73, 33, 68, 72]. Specifically, Vega-Redondo [73] and Selten and Ostmann [72] assume that people imitate the successful action carried out by one other player in the immediate previous period. More generally, Alós-Ferrer [2] assumes that people mimic the best performing action carried out in the past that the agent can recall. As for whom to imitate, people are often thought to follow the actions of the best performing competitor [73, 72], or of those that are just like them but playing against different opponents [68] or, if the players have some preference for conformity, they follow the average choice of all other players [17].

The experimental economic literature has provided evidence that when given the opportunity to do so, individuals tend to imitate others. Eckel and Wilson [31] show that social status affects social learning processes. They found that, in a coordination game, commonly observing an agent with high social status leads subjects to converge faster on the payoff-dominant equilibrium. Cooper and Rege [29] instead documented that imitative behaviour in lottery choice tasks might be related to some form of regret. As for the relative payoff assumption, Huck et al. [45] and Offerman et al. [61] provide support for the Vega-Redondo [73] model and both show that, in a Cournot game, when information on players' actions and payoffs is available, players tend to copy best performers and the environment becomes competitive. Offerman and Sonnemans [63] show that subjects match the forecasts of successful players. Apesteguia et al. [5] show that agents' propensity

to imitate the actions of more successful individuals is increasing in the difference between the highest payoff observed and own score. Merlo and Schotter [53] show that, in a Cournot game, observers learn better than subjects directly engaged in the task. Anderson and Holt [4], relying on the models of Banerjee [14] and Bikhchandani et al. [18], show that when people are able to retrieve information both personally and through the observation of others, information cascades can occur. Knowledge stops accumulating, individuals stop using private information and simply conform to the behaviour of their predecessors.

So summarising, “the assumption of relative payoff concerns is central in the literature” [48, p.82]. Observational learning is often associated with the discrepancies between one’s own payoffs and the payoffs of others, i.e. when someone is performing better than oneself, this can be a strong motivation for imitation. Players imitate others, or strategically use the information retrieved from their actions, to enhance their performance.

Social psychology provides a complementary explanation. In general, it is thought that individuals do not necessarily engage in imitative behaviour as a result of a mere cost-benefit analysis. Humans might also decide to imitate other’s decisions and actions because of some specific cognitive needs. Festinger [40] posits that if individuals care about making the right choice but have no means to evaluate whether they are correct, they tend to anchor their decisions to those of others. Additionally, Bandura [11] argues that individuals’ decisions, including those to imitate someone else, depend on their introspective beliefs about their own capabilities. The latter are referred to as an agent’s beliefs in his or her own self-efficacy.² These beliefs are crucial for the decision to imitate others. According to Bandura [12, 9], apparently imitative behaviour is the result of

²Self-efficacy is concerned with a person’s perceived capabilities to achieve some goal. Self-efficacy differs from other concepts such as self-esteem, locus of control, or outcome expectancies. Self-esteem is a judgment of self-worth and according to Bandura [12, p. 11] “there is no fixed relationship between beliefs about one’s capabilities and whether one likes or dislikes oneself”. Even locus of control and self-efficacy are, according to Bandura [12, p. 20], “entirely different phenomena”. Locus of control is concerned with the beliefs that behavioural outcomes depend on one’s own actions or on forces beyond personal control and “cannot by any stretch of imagination be considered the same as beliefs about whether one can produce certain actions (self-efficacy)” [12, p.20]. Perceived self-efficacy is also different from outcome expectancies. Self-efficacy has to do with people’s confidence that they can perform a certain action if they wish or have to. Outcome expectancies instead identify one’s judgments about how performance affects outcomes. See Bandura [12] on this specific point. Outcome expectancies are closely linked to self-efficacy in tasks in which the effects are dependent on performance. Conversely, in situations where outcomes are only loosely tied to actions, self-efficacy and outcome expectancies are likely to be separate [10, 69]. This seems to suggest that in situations in which people think that their performance has no clear impact on the outcome obtained, self-efficacy matter less.

both individual and observational learning. While one's own evaluations remain part of any decision, copying the behaviour of others does become more likely when one perceives oneself as unable to figure out how to accomplish a specific goal. If the agent perceives him/herself as unable to take actions on his or her own, he or she will be more prone to rely on others' decisions or actions.³

Although Offerman and Schotter [62, p.461] argue that "those who feel the need to imitate must, by definition, either not be able to do all the necessary calculations" or lack the inclination to do so, the role of introspective beliefs concerning one's capabilities as possible motives to imitate has not been explored within the economics literature. This study takes a first step to fill this gap and test whether players' own self-evaluations of their capacities represent an additional motive for imitative behaviour. We hypothesise that if agents have low self-efficacy beliefs, i.e. they feel unable to understand and control the decisional environment, they might be more inclined to imitate others, regardless of their performance.

To test our hypothesis we use as an experimental task a modified version of the common two-armed bandit problem with finite time horizon, as explained in detail in the next section.

Only five studies in the literature have used this experimental task to investigate imitative behaviours specifically. Vostroknutov et al. [74] use a two-armed bandit setting to study how intelligence levels affect observational learning.⁴ They find that participants with high intelligence use choices of others strategically to better understand the decisional environment. Conversely, agents who score low on the intelligence test rely on simple mindless imitation. This study shares with ours the interest in uncovering the mechanism motivating imitation, but whereas Vostroknutov et al. [74] claim that this mechanism can be found in the knowledge of one's own intelligence and that of the person we are imitating, we posit that confidence in one's capacities is key. In the Burke et al. [21] study, students engaged in a two-armed bandit experiment while being administered an fMRI scan. Before participants made their own choice, they observed the behaviour of another player who faced the same options. They found that when participants are informed both about the option chosen and payoff obtained by a peer, they choose the higher paying option at a significantly higher rate. They also find that

³Indeed, the self-efficacy mechanism does not necessarily conflict with the payoff enhancement motive. Low self-efficacy could simply imply that a subject is less confident in his or her abilities to obtain a high payoff and thus more likely to imitate someone he or she believes to be more capable. Our goal is to test whether aside from payoff enhancement, status or regret, low self-evaluations of one's capacities can represent a possible reason to imitate others.

⁴Intelligence, defined as efficient problem solving and abstract reasoning, is measured using Raven Advanced Progressive Matrices.

agents tend to mimic a peer with a higher probability when only the choices made by the peer were observable relative to the case in which both actions and payoffs were observable. Nicolle et al. [60] use a two-armed bandit to study the relations between observational learning and optimism. As observers do not directly incur costs or benefits during the learning process, observational learning is associated with optimistic over-valuations of low-value options. Our experimental design is closer to the work of McElreath et al. [52, 51] who used two-armed bandit experiments to study social transmission of behaviour and culture. We modify their treatment set-up to assess whether self-efficacy beliefs are a good predictor of imitative behaviours. The general design of our experiment is explained in the next section.

3. Experimental design

Our experiment consists of three parts as summarised in Table 1. First, in the pre-game phase, subjects answered 10 questions aiming at measuring their problem-solving abilities and math skills. We selected these questions from the standard Graduate Record Examination (GRE) test.⁵ Subjects were allowed neither calculators nor pens and paper.

Subsequently, our subjects were presented with a questionnaire to assess their self-efficacy beliefs.⁶ This is a standard questionnaire used to rate participants' confidence to perform certain behaviours in a set of hypothetical situations [71, 70, 47]. It is not task-specific, it is rather meant to assess people's self-perceived potential and beliefs that actions are responsible for successful outcomes (e.g. "I can always manage to solve difficult problems if I try hard enough" or "If I am in trouble, I can usually think of a solution") . The questionnaire consists of 10 questions. Students were asked to rate the extent to which each question applied to them on a 4-points scale which ranged from "Not at all true" (1) to "Exactly true" (4). The final score for each participant was obtained summing up all questions' scores. We treat this information as static and exogenous. We

⁵In order to make sure that these questions were not too hard for undergraduate students, we ran a simple in-class pilot with 45 students. We administered the test to first year economics students enrolled in the International Economics course at Maastricht University. We checked the distribution of correct answers and fine-tuned the level of difficulty to make sure that on average people were able to answer at least 5 questions.

⁶One might be concerned that as a result of filling in a questionnaire about self-efficacy, subjects might change their own self-efficacy beliefs. Many studies have shown that this does not happen. Recording one's efficacy judgments has been shown not to affect subsequent behaviour. Simply having evaluated one's own self-efficacy did not change people's capacity to cope with threats, regulate motivations, tolerate pain, ameliorate cognitive performance, recover from coronary surgery, or commit to physical exercise [12, 13].

do not consider any feedback-loop between learning and self-efficacy beliefs. This seems reasonable given the short duration of the experiment.

Finally, subjects participated in the experimental game which followed a within-subject design.⁷ Our game is divided into 2 treatments whose order was not randomised and which differ from one another mainly in terms of the information participants received. In the baseline treatment, subjects sequentially chose between two alternative colours (1 and 2) for three sets of twenty rounds, thus making 60 binary choices. We will call a set of 20 rounds a sub-setting. The number of periods and sub-settings played were set to reduce the likelihood of participants growing bored and to have enough variation in the data to estimate learning patterns.

Table 1: Summary of the experimental design

Task	Pre-Game		Tr 1: Baseline	Tr 2: Observational
	Math test	Self-efficacy questionnaire	Binary choice (Colour 1 vs Colour 2)	Binary choice (Colour 1 vs Colour 2)
Information provided	None no calculators/pens	None	Individual choice made and reward obtained in previous round	1. Individual choice made and reward obtained in previous round 2. Leader's choice and reward he obtained in current round
Leader	-	-	No	Yes identity unknown, randomly selected and first mover. Leads a group of 4/5 people. Uniquely informed about personal choice made and reward obtained in previous round.
Variance payoff distributions	-	-	3 values in randomised order (High, Medium, Low)	3 values in randomised order (High, Medium, Low)
Number of repetitions	10 questions from GRE in 12 minutes	10 questions	3 sub-settings of 20 rounds (total=60 choices)	3 sub-settings of 20 rounds (total=60 choices)

To avoid people carrying their priors over sub-settings, which can happen in within-subjects designs, every 20 rounds we changed the two colours between

⁷The reasons for opting for a within-subject design rather than for a between-subject design are twofold. First, we want to study learning patterns and thus assess how people make repeated choices. Second, our goal is to investigate how individuals learn on their own as well as how they learn when they can observe the actions and payoffs obtained by a peer. We are not necessarily interested in the effects that the group exerts on learning processes. Therefore, this calls for a task in which subjects serve in more than one treatment.

which subjects had to choose. Each sub-setting had a “preferred” colour, in the sense that it yielded a higher expected payoff. The preferred colour was set randomly at the start of the sub-setting, was the same for all subjects, and remained unchanged during the sub-setting. Participants were informed that there was a difference in expected payoffs to the two colours, but not what the expected payoffs were, nor which was higher. They were informed that payoffs ranged between 1 and 18 units. Subjects played by selecting one colour in each round. After agents selected their colour they were informed about the score received, and reminded which colour they had chosen. Only the most recent choice and payoff were displayed. Before the first round, no information was displayed. The payoffs of both options were drawn from truncated normal distributions, with fixed mean and variance, bounded between 1 and 18.⁸ Thus a sub-setting can be characterised by a quadruple: $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$.

The mean of the more rewarding colour, for instance colour 1, was fixed and equal to $\mu_1 = 13$ units, while that of the less rewarding one was $\mu_2 = 10$ units. The variance of the payoff distributions changed across the three sub-settings but was set to be the same for both colours (i.e. $\sigma_1^2 = \sigma_2^2 \in \{0.25, 4, 16\}$). The higher the variance, the higher the payoff volatility. Each group of subjects played in all three sub-settings, but the sequence of variance values came in random order and changed across groups.⁹

The observational treatment of our experiment differs from the baseline in the information given to participants. As in the baseline treatment, participants had to select one of two colours for three consecutive sub-settings of twenty rounds each. Differently from baseline, subjects were randomly assigned to a group of 4 or 5 people whose identity remained unknown. Each group had a leader who was randomly selected and faced the same environmental conditions as everybody else. Leaders played as in the baseline treatment: they did not observe others’ behaviour — only their own choices and payoffs were shown to them, exactly as in the baseline treatment. Non-leaders had different information. For them this treatment resembles what Bikhchandani et al. [18] define as an “observable signals” scenario. Leaders played first. Non-leaders were immediately informed of their leader’s choice and payoff. This makes our treatment different from the one analysed by McElreath et al. [52, 51] wherein participants were given the possibility of clicking a button and observing the most recent decision, but not yield, of one single unknown member of their group. Starting from round two

⁸We checked that the mean and variance in the sample of payoffs were close to those of the underlying normal distribution.

⁹The variance order was randomised at the group level. Had we randomised at the individual level, patterns of imitation would have been difficult to detect and the learning parameters would not have been easily compared.

onwards, non-leaders also observed the outcome of their previous personal choice as in the baseline condition.

The leader is not necessarily the best performer, the most skilled or informed but simply the one who explores the decision environment first. Non-leaders were clearly aware of this (refer to 3 for details on instructions). We opted for this design because it is not our goal to make claims about leaders' characteristics but rather about those of the followers.

As in the baseline treatment, the payoffs were drawn from truncated normal distributions. The mean of the distribution from the most rewarding colour was set equal to 13, whilst the one for the worse option to 10. The pairs of colours between which the individuals had to make a choice changed every twenty rounds, at the end of each sub-setting. One of the two colours was always on average better than the other. The best option was randomly decided and changed every sub-setting, as were the colours. The variance of the payoff distributions, as in the previous treatment, was set to be either low, medium or high. The order was randomised across groups.

The experiment was programmed in PHP and administered via computer. It lasted about one hour. All instructions were displayed on the screen (refer to Appendix 3). One hundred and seventy five undergraduate students (74 females and 101 males) participated in all parts of this experiment. Monetary rewards, proportional to the total score each subject achieved, were given at the end of the session. Theoretical gains, which included a flat-rate show-up fee of 5 euros, ranged from a minimum of 15.1 to maximum 19.4 euros. The lowest realised payoff was 17.2 while the highest one was 19.1 euros. On average, students received 18.2 euros. The experiment was run at the BEELab at Maastricht University.

4. Theoretical models and estimation strategy

In order to test our hypothesis and study imitative behaviour, we have first to study individual learning patterns. We observed the choices made and payoffs obtained in each round by each agent in the baseline treatment. We select three plausible individual learning models. For each of the three prior-posterior updating rules, we estimate the individual learning parameter (β) and select the best fitting model. Subsequently, we move to the analysis of the observational treatment, the goal of which is to understand how people filter available information concerning the behaviour of others. We select two imitation models and estimate the related parameters (α or θ).

4.1. Choice model

In all the analysis that follows we use a standard logistic discrete choice model. The model assumes that each of the two possible strategies has a numerical evaluation, generally defined in the literature as “attraction score”, which is updated in response to experience [26]. We denote the attraction score of the option $i = \{1, 2\}$ at time t considering the choices made and payoffs obtained in each round as $A_{i,t}$. Thus the probability of choosing colour i in round t is written as:

$$Pr(i|\mathbf{A}_t, \Theta)_{t+1} = \frac{e^{\beta A_{i,t}}}{e^{\beta A_{1,t}} + e^{\beta A_{2,t}}} \quad (1)$$

where Θ is the set of free parameters, while $A_{1,t}$ and $A_{2,t}$ represent the attraction scores (see below for clarification) and more specifically the subjects’ current estimates of the mean payoff of colour 1 and 2 respectively.

The logit model initially proposed by Luce [49], is widely used in economics. Generally, it is meant to explain how best responding individuals maximise their expected payoffs based on the distribution of scores they obtained in previous periods. The parameter β is usually interpreted as a measure of rationality [19]. The larger is β , the smaller the probability that the individual will deviate from the best response. When $\beta = 0$ the agent chooses randomly between the two alternatives with probability 0.5. As β goes to infinity, the individual never deviates from the best response and the choice is, in that sense, optimal.

A different interpretation is possible. In our case, β can be interpreted as the strength of the belief about the estimated reward for each colour, or differently the strength of the belief of being correct. If $\beta = 0$, the agent has little faith in his or her own estimation, the difference in the mean of the payoff distributions is neglected and the choice is made randomly. If β goes to ∞ , the agent firmly holds onto his or her estimates and the colour which is thought to be, on average, the most rewarding is always chosen.

4.2. Baseline treatment: Individual learning

We use data from the baseline treatment to detect patterns of individual learning. This situation can be considered “a black box” [58]. Players take actions and receive payoffs.¹⁰ No information apart from the result of the individual performance is provided to the players. Consequently, learning is the result of an asocial process.

In Table 2 we report descriptive statistics concerning self-efficacy, math abilities, payoffs and number of switches between colours. In addition we count,

¹⁰Differently from Nax et al.’s [2016] baseline case, in our case the payoff structure does not depend on others’ choices.

backwards in time starting with the last round, the number of consecutive correct choices made in the baseline treatment. Number of switches can be considered a measure of exploration, number of consecutive correct choices can be considered a measure of exploitation.

Table 2: Basic summary statistics

	mean	sd	median	min	max	n
		Low	variance			
Payoff	249.43	9.68	252.00	201.00	261.00	175
#switches	4.47	2.79	4.00	0.00	15.00	175
# last consecutive correct choices	7.07	5.80	5.00	0.00	20.00	175
		Med	variance			
Payoff	236.59	9.40	236.00	203.00	256.00	175
#switches	6.02	3.72	6.00	0.00	16.00	175
# last consecutive correct choices	5.63	5.83	4.00	0.00	20.00	175
		High	variance			
Payoff	234.07	11.46	232.00	202.00	266.00	175
#switches	6.98	3.87	7.00	0.00	17.00	175
# last consecutive correct choices	4.16	4.62	3.00	0.00	20.00	175
Math Score	4.89	1.96	5.00	1.00	9.00	175
Self-eff. Score	31.38	3.53	32.00	19.00	38.00	175

These measures provide indications of the individual learning dynamics. For example, it appears that subjects explore more as the variance on payoffs increases, and, as it might be expected, their ability to exploit their learning at the end of the sequence declines. However, there are several methods individuals can use to learn in this condition. Following previous literature [52], we fit three alternative, widely used and minimally parametrised learning rules to the data. These models are presented in Table 3 as rules for updating attraction scores.

The first model, Running Average, is a standard reinforcement learning model [26, ch.6]. The second model, Memory Decay, is a weighted average of all past payoffs. If the ‘recency’ parameter r is equal to 0, information obtained from earlier rounds is completely ignored, if $r = 1$ the model reverts to a Running Average. Lastly, under the third model, individuals estimate the distribution means in a Bayesian fashion. Subjects sequentially update their attraction scores by combining their prior and the observed payoff using Bayes’ formula. It is worth noting that this inference process is influenced by the variance of the estimate of the mean of the payoff distributions ($\hat{\sigma}_{i,t}^2$) and the real long-run variance (σ^2)

Table 3: Theoretical models

Model	Updating rule	Free parameters
Running Average	$A_{i,t} = \frac{N_{i,t-1}A_{i,t-1} + y_{i,t-1}}{N_{i,t}}$	β
Memory Decay	$A_{i,t} = rA_{i,t-1} + (1-r)y_{i,t-1}$	β, r
Bayesian Updating	$A_{i,t} = \frac{\frac{A_{i,t-1}}{\hat{\sigma}_{i,t-1}^2} + \frac{y_{i,t-1}}{\sigma^2}}{\frac{1}{\hat{\sigma}_{i,t-1}^2} + \frac{1}{\sigma^2}}$ $\hat{\sigma}_{i,t}^2 = \left(\frac{1}{\hat{\sigma}_{i,t-1}^2} + \frac{1}{\sigma^2} \right)^{-1}$	β

which is assumed to be known. Low variance in the observed payoffs (i.e. a low $\hat{\sigma}^2$) will lead quickly to a relatively fixed belief of the estimated distribution mean.

For all models analysed, when colour i is not chosen at round t , its estimated payoff mean is assumed to stay equal to previously formulated attraction score ($A_{i,t} = A_{i,t-1}$).¹¹

All models share the unknown parameter β which can be estimated with maximum likelihood techniques. The Memory Decay model has an additional parameter, r , which needs to be estimated, again using maximum likelihood. Independently from the model analysed, the attraction scores are transformed into predicted choice probabilities using the standard logit model described above.¹²

¹¹We acknowledge that alternative assumptions could be formulated. However, this assumption seems the simplest and the most consistent with common assumptions about rationality. Additionally, it is the most common in the literature [26, 52].

¹²We also examined two other models: the ‘Win-Stay-lose-shift’ model; and a revised version of the I-SAW model. Neither model outperforms the Memory decay model. Additionally, for I-SAW, the number of free parameters increases and thus the risk of measurement problems is likely to bias our analysis. For this reason, we did not pursue these models any further.

4.2.1. Estimation of the individual learning parameter β

We can fit these models either on an individual basis — obtaining for each subject estimates for the parameters that maximise the likelihood of observing the vector of his or her choices — or across individuals pooling the data together, and obtaining one value of the parameter estimate for the entire population of subjects.

We begin by fitting the models on pooled data, assuming all participants use the same updating rule. This produces, for each model analysed, a negative log-likelihood of observing the true data under the assumption that the model is true: $-\log \mathcal{L}(\mathbf{D}|x, \Theta)$ for model x given the set of free parameters Θ and where \mathbf{D} denotes the data, a matrix containing the colours chosen by all participants over the 20 periods. The likelihood is defined as:

$$\mathcal{L}(\mathbf{D}|x, \Theta) = \prod_{t=1}^{20} Pr(D|\mathbf{A}_{t-1}, \Theta)_t \quad (2)$$

Taking the natural log of all conditional probabilities and summing them, we obtain:

$$-\log \mathcal{L}(\mathbf{D}|x, \Theta) = -\sum_{t=1}^{20} \log Pr(D|\mathbf{A}_{t-1}, \Theta)_t \quad (3)$$

We fit each model in Table 3 to the data to retrieve the values of the parameter that maximises the joint likelihood of observing the pooled data (Θ). Parameter estimation is done through a numerical grid search. We use flat homogeneous priors. We set the initial attraction scores equal to 9.5 for both colours. This is consistent with the experiment instructions that the possible payoffs ranged between 1 and 18 and that one of the two options was on average always more rewarding than the other. For the case of the Bayesian model, the assumption is that agents also know the long-run variance of the payoff distribution. Therefore, we set, for each sub-setting, σ^2 equal to the real variance of the payoff distribution.¹³

We report in Table 4 the fit of each model on the pooled data for the three variance values. The parameter estimates are shown together with the estimators of the standard error of our parameters.

Table 4 also displays some goodness of fit measurements (AIC, AICc, w , and Δ) which allow us to compare the models scrutinised [22].

¹³As it can be seen from the experimental instructions, participants were not aware of the real value of σ^2 , but could have made reasonable estimates if they were keeping track of their payoffs.

Table 4: Goodness of fit measures: pooled data

Variance level	Low	Medium	High
Run. Average			
-LogLik.	1516.89	1995.33	2140.91
$\hat{\beta}$	0.63 (0.02)	0.50 (0.02)	0.49 (0.02)
AIC	3035.77	3992.67	4283.82
AICc	3033.80	3990.69	4281.84
Δ	0.37	0.18	0.12
w	0.03	0.00	0.00
N	175	175	175
Mem. Decay			
-LogLik	1512.28	1852.81	1994.63
$\hat{\beta}$	0.60 (0.02)	0.42 (0.02)	0.28 (0.01)
AIC	3028.57	3709.61	3993.27
AICc	3025.64	3706.68	3990.34
Δ	0.38	0.24	0.18
\hat{r}	0.61 (0.05)	0.26 (0.03)	0.40 (0.02)
w	0.97	1.00	1.00
N	175	175	175
Bayes			
-LogLik	1520.69	2027.56	2237.38
$\hat{\beta}$	0.73 (0.02)	0.54 (0.02)	0.51 (0.03)
AIC	3043.39	4057.11	4476.76
AICc	3041.41	4055.14	4474.79
Δ	0.37	0.16	0.08
w	0.00	0.00	0.00
N	175	175	175

Note: AIC refers to the Akaike Information Criterion; AICc refers to the Akaike Information Criterion with a correction for small sample sizes, w are the Akaike weights; Δ is the ratio of the negative log-likelihood of the model analysed, and the log-likelihood of a model wherein individuals choose randomly. $\hat{\beta}$ and \hat{r} are our estimated parameters. Standard errors are in parenthesis

By all the goodness of fit measures, Memory Decay is the best model. The AIC and AICc values for this model are the lowest regardless the higher number of free parameters. Δ values are the highest for all variance values. Although this measure does not account for model complexity, it is able to provide a rough guide of the variance explained by this model. Moreover, the Akaike weights (w) show that the probability that people use the Memory Decay rule of updating, compared to the other available models, is approaching or equal to 1 for all variance values.

The estimates of β show that, for all models considered, choices become more random with the increase of the variance in the payoff distributions. This implies that the extent to which individuals believe in their estimates of the payoff means, given the scores obtained from the two colours, decreases — as demonstrated by the declining β — when the variance of the payoff distributions increases. This first result is in line with what was found by McElreath et al. [52].

As expected, and consistent with declining β values, subjects' abilities to determine which colour had the higher mean payoff falls with the variance of payoffs. This can be seen in Figure 1 which displays the proportion of correct decisions in each round.

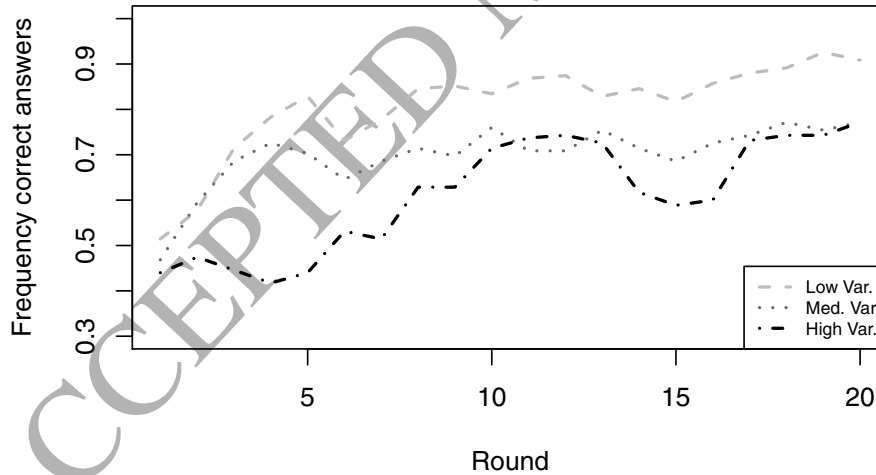


Figure 1: Share of correct answers per round in low, medium and high variance in the baseline treatment

We observe a consistent pattern when asking ex-post subjects to give their best guess concerning the average reward obtained by choosing each of the colours.

As can be seen from Figure 2, in case of low variance, the distribution of the answers is nicely peaked around the real means, indicated as a solid vertical line. Conversely, when the environmental volatility increases, some subjects have more difficulty in understanding which are the real, correct means of payoffs' distributions.

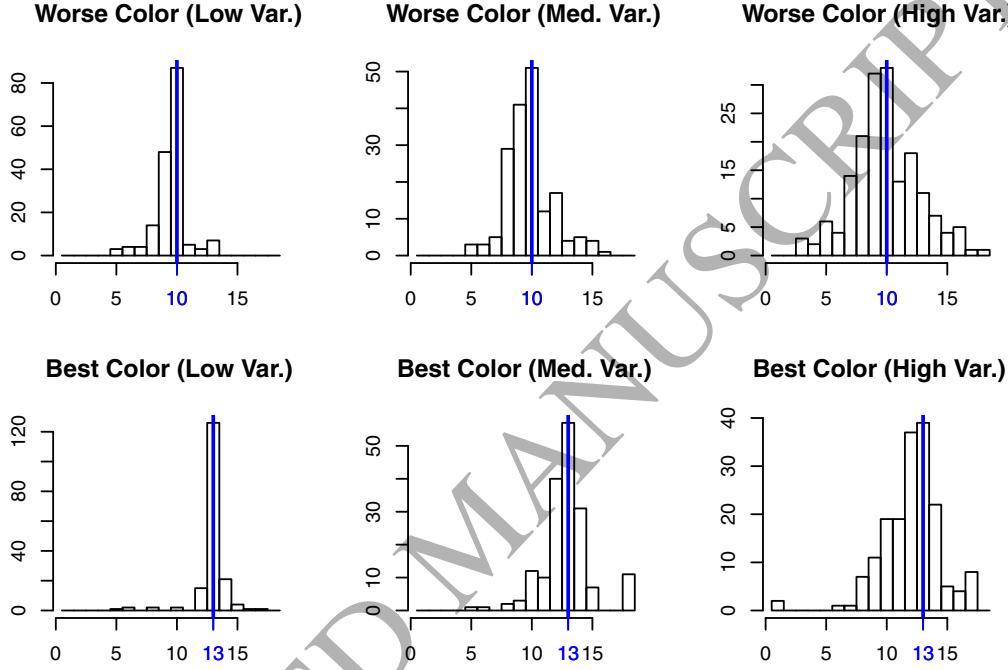


Figure 2: Distribution of posterior beliefs on the means of payoff distributions by variance value: Baseline

Table 4 also shows that the estimate of the r parameter — the additional unknown parameter for the Memory Decay model — declines, although not monotonically, with increasing variance. In the high variance case agents pay more attention to recent events than they do in the low variance case.

In summary, our pooled analysis confirms the results obtained by McElreath et al. [52].

However, differently, from them we also fit the models on an individual basis, treating the three models presented in Table 3 as possible alternatives to explain the choices made by each participant.

In this case, given the vector of choices made by each individual j over 20 rounds (\mathbf{D}_j) as well as the set of parameters to be estimated individually Θ_j , the log-likelihood function for model x can be written as follows:

$$\log \mathcal{L}(\mathbf{D}_j | x, \boldsymbol{\Theta}_j) = \sum_{t=1}^{20} \log Pr(\mathbf{D}_j | \mathbf{A}_{j,t-1}, \boldsymbol{\Theta}_j)_t \quad (4)$$

We estimate the individual learning parameter (β) for each model and each variance value, thus obtaining nine $\hat{\beta}$ per agent. In order to retrieve the values of the parameter that maximises the joint likelihood of observing each vector of individual data, we proceed numerically as we did in the pooled estimation case. We arbitrarily specify the upper bound of the grid within which search is conducted. It turns out that, in some cases, $\hat{\beta}$ takes values equal to the upper bound of this grid search, regardless of the value of the upper bound. These values might be divergent. This problem, which emerged particularly for the Memory Decay case, could be due to the fact that $\hat{\beta}$ is computed over only 20 rounds and thus unreliable estimates are produced. This implies that even if a true value of β exists, simply because of statistical variation, our estimation strategy would be unable to estimate it correctly. To test our estimation method, we run 20000 Montecarlo simulations. We set the true β and r equal to the estimates obtained fitting the Memory Decay model to the pooled data as reported in table 4. We use the payoff distributions used in our experiment, and create fictitious data. We then fit the same Memory Decay model to these data. In the case of low variance in the payoff distribution, in 26% of the cases $\hat{\beta}$ takes extreme values, diverging from the true β . In case of medium variance, the proportion of extreme values declines to 10%, and is 11% when the payoff volatility is high (see section 1 in the Online Appendix). These proportions are in line with what we observe in our experimental data. Our estimates are divergent in 37 cases (21%) in the low variance sub-setting, in 31 cases (18%) for medium variance and in 21 (12%) cases when the payoff variance is high. This is not a new problem in this literature. The individual learning literature is well aware that in many cases the estimation method is unable to recover the true parameter values even when this is among those considered [51]. Different studies have shown that the estimation of learning models can be disappointing if the number of possible strategies is small (e.g. 2 or 4) and the number of periods considered by the experiment is not long [26, ch.6]. Given these premises and the results of our simulations, we eliminate subjects whose estimated β took on extreme values, and perform our individual learning analysis with a reduced sample. Those excluded from the analysis do not differ much from those included, either in self-efficacy or in payoffs obtained (refer to section 2 in the Online Appendix). The main difference between the two groups concerns the number of times they switch from one colour to the other and thus the number of rounds taken to fix their choices. Those excluded make many fewer switches, making it difficult to arrive at a reasonable estimate of β . Additionally, we checked whether there exists a significant difference between self-efficacy for

excluded and non-excluded individuals for each variance value. On the basis of a Wilcoxon test, we conclude that the median self-efficacy of those excluded is not significantly different from the median self-efficacy of those included in our analysis.

A simple count shows that the Memory Decay represents the most frequently used updating rule. As shown in Table 5 this applies to over 68% of individuals in case of low variance, 83% in medium variance and 66% in the high variance case. Bayesian updaters represent 24% of the sample population in low variance, 3% in medium variance and 20% in volatile environments. This result points to a wide use of the Memory Decay updating rule.

Table 5: Number of individuals using the different models

	Low Var.	Med. Var.	High Var.
Running avg.	9	20	20
Mem. Decay	95	120	102
Bayesian Updating	34	4	32
Total	138	144	154

As a robustness check, in Table 6 we show the correlations of individual estimates of β across models, excluding individuals whose β estimates took on extreme values.

The correlations are high for all variance values. More precisely, the individual estimates of β obtained fitting the Memory Decay model are highly correlated with the estimates obtained fitting the other two candidate models. In both the low and medium variance cases the correlation coefficients are significantly around 70%. In the high variance case, the correlation coefficients decline but remain above 50%.¹⁴

Given the simple count reported in Table 5 as well as the high correlation coefficients across estimated β , the Memory Decay model seems to have a clear advantage in predicting individual choices. Consequently we consider the individual $\hat{\beta}$ obtained by fitting this model to the baseline data to carry the rest of our analysis.¹⁵

We also analysed within model correlation coefficients in order to establish whether the learning patterns change depending on the level environmental

¹⁴As a robustness check we also repeated this same exercise while considering all individual β values, including the most extreme ones. The Pearson's correlation coefficients remain in line with those presented in Table 6.

¹⁵As a robustness check (not shown) we also carried our analysis with the individual $\hat{\beta}$ s obtained by fitting the Running Average model presented in Table 2.1 to the data. We observed that the general results do not change.

Table 6: Correlation Individual β across models by variance value - excluding extreme β values

	Low Var.		Med. Var.		High Var.	
	Run.Avg.	Mem. Dec.	Run.Avg.	Mem. Dec.	Run.Avg.	Mem. Dec.
Run.Avg.						
Mem. Dec.	0.75***		0.70***		0.69***	
Bayes	0.93***	0.69***	0.90***	0.81***	0.63***	0.51***

volatility. In Table 7 we report the results obtained after the extreme β values have been excluded. It can be seen that the individual learning parameter β varies with the variance of the payoff distributions. Thus, we continue our analysis for the three levels of environmental volatility.¹⁶

Table 7: Correlation Individual β within models by variance value - excluding extreme β values

	Run. Avg.		Mem. Dec.		Bayes	
	Low Var.	Med. Var.	Low Var.	Med. Var.	Low Var.	Med. Var.
Low Var.						
Med. Var.	0.10		-0.06		0.22***	
High Var.	0.07	0.24***	0.21**	0.28***	0.14*	0.17**

4.3. Observational Treatment: Imitation

The goal of the observational treatment is to understand whether and how people use newly available information concerning the choices and results obtained by a randomly-chosen group leader. As for the baseline treatment we investigate which of the candidate models of imitation best predicts the decision behaviours of our subjects.

This treatment can be seen as an “observable signals scenario” [18]. A leader plays before everybody else, and his or her actions and payoffs are observed by his or her group members from round one onwards. Agents take action after the leader, and from round two on they are also presented their own choice and outcome from the previous round.¹⁷

As shown in Table 8, we restrict our analysis to two models of social learning. Both models are very simple and share two characteristics. Each player acts after the leader has made his or her choice. This means that each group member, after his or her choices, has two observations from the same process with which to

¹⁶Indeed, as a robustness check, we calculated within model correlation coefficients also when the entire sample is considered. Even in this case we find that individual β are scarcely correlated within models.

¹⁷The leader sees only his or her own actions and payoffs.

revise the means of the payoff distributions. Moreover both models assume that individuals weight the signals gained observing the leader and possibly use them to validate or invalidate private information.¹⁸

Table 8: Theoretical models

Model	Updating rule	Parameters
Nested model	$Pr(i \mathbf{A}_{t-1}, \mathbf{\Theta})_t \propto (1 - \alpha)L_{i,t} + \alpha X_{i,t}$	α
Additive model	$A'_i = (1 - \theta)A_{i,t-1}^j + \theta y_{i,t}^z$ $Pr(i \mathbf{A}, \mathbf{\Theta})_t = \frac{e^{\beta A'_i}}{Z}$ $A_{i,t} = rA'_i + (1 - r)y_{i,t}^j$	θ

First, in line with McElreath et al. [52] this situation can be modelled using a nested probability model. Specifically, the probability of choosing colour 1 at time t is given by

$$Pr(1|\mathbf{A}_{t-1}, \mathbf{\Theta})_t \propto (1 - \alpha)L_{1,t} + \alpha X_{1,t} \quad (5)$$

where $X_{1,t}$ is an indicator variable taking the value 1 if the leader chose colour 1 at time t , and is 0 otherwise. $L_{1,t}$ is the probability that the subject would choose colour 1 at t were he playing alone, as defined in Equation 1. This probability is calculated in a conservative manner. We rely on our previous estimates from the baseline treatment, using the individual estimates of β and r obtained by fitting the Memory Decay model to the data from the baseline treatment.¹⁹ The relevant unknown parameter to be estimated in Equation 5 is α . This parameter measures

¹⁸The most common deterministic imitation rules in the literature assume that individuals, in a observational learning context, need understand neither why the observed person has made a certain choice nor why the choice made has generated the observed outcome. However, individuals, even when given the opportunity to observe another player, maintain their capacity to privately learn and thus revise the attractions scores associated with the two colours [9]. For this reason we depart from simpler imitation rules.

¹⁹As a robustness check (not shown), we also carried out our analysis using the best fitting individual learning model for each individual. The results are very similar. Therefore, we retain the ‘single-model’ analysis for simplicity.

the individual propensity to imitate. If α is zero, the model reduces to the simple individual learning process. The agent fully relies on individual learning and any information provided by the leader's action is dismissed. If α is 1, the individual mindlessly mimics the behaviour of the target subject. Intermediate values of α can be interpreted in two ways. α can represent the propensity of a subject to mimic the leader, which is equivalent to ignoring his or her own private information. Alternatively, equation 5 can represent a reduced form observational learning model, in which α captures the importance the subject puts on social observation. In either case, we hypothesise that the value of α will be related to a subject's self-efficacy beliefs.

Second, imitation is modelled as an additive process in which individuals add the information retrieved from the leader's action to their private information. The principle here is that the quality of the information gathered from the observation of the leader is exactly the same as that of any other agent. The leader's signal is not *a priori* any better than anyone else's. Thus, at the end of round one, group members hold two pieces of information from the same payoff distribution. The relevant unknown parameter to be estimated is, in this case, θ which measures again the individual propensity to imitate. More precisely, θ captures how much individuals value the experience of the leader. In this case, for estimating the parameters, the sequence we assume is the following. At the start of each round the subject has, for each colour, 1 and 2, a prior belief of the mean payoffs (A_1, A_2) . The leader (identified as subject z) plays and the non-leader (subject j) observes the leader's choice. Suppose that the leader has chosen colour 1, and has obtained a corresponding payoff $y_{1,t}^z$. According to the updating rule of Table 8, subject j updates his estimate of mean payoff for colour 1 to A_1' . Based on the pair (A_1', A_2) the subject makes his or her choice following the standard logit model presented in Equation 1 and reported in Table 8 where $Z = e^{\beta A_1'} + e^{\beta A_2}$. We use the previously estimated individual β to calculate this probability. Suppose j chooses colour 1. He observes the payoff he obtained ($y_{1,t}^j$) and updates his or her estimate of the mean payoffs to a new pair $(A_{1,t}, A_{2,t})$ following the Memory Decay rule (see Table 3). As discussed in the previous Section, statistically, the Memory Decay model works best for updating based on individual information. Thus, in this step we use that rule, applying for each subject the value of Memory Decay, r , fitted from the baseline, individual play, treatment. We assume that in period $t = 0$ our agents start off with flat priors (uniform on $[1, 18]$), on the unknown mean of the payoff distributions.

We fit the two models to the individual data from the observational treatment and estimate the two imitation parameters for each variance value.

First, we report simple information concerning the choices made by our subjects in the observational treatment.

When comparing the vector of “correct” answers that all individuals provided in the baseline treatment, (individual play), with those of the observational treatment, (play with a leader), it can be seen that on average performance increases (see Figure 3). The share of correct answers per round in the observational treatment is generally higher than that observed in the baseline treatment both in case of low and medium variance.²⁰

In the high variance case, the effect of playing with a leader is unclear. Signals are mixed, and while the payoff distributions had the same moments both with and without the leader, there is no temporal pattern. This is almost certainly driven by the fact that in the high variance case, subjects were simply unable to detect which colour was superior, due to the large noise in the signals they receive. Information could not be extracted either from their own signals or from the signals given by the leader’s play.

²⁰One might wonder whether the increase in performance is simply be due to learning the experimental setup, as the task in the observational treatment is very similar to that in the baseline treatment. Were this the case, though, we would observe a marked increase in performance also for the leader, for whom the task is identical in the treatment conditions. We do not observe this though, as seen in Figure 4 below. It could also be that leaders might behave differently simply because they knew they were being observed. However, in our context the leaders’ payoffs do not depend on the actions of the observers and no incentive is provided to them for performing differently. Moreover, we ran two Wilcoxon tests to compare the time leaders spent to complete the 2 parts of the experiment and the cumulative payoffs they obtained. In both instances, we conclude that the median time leaders spent in part one and the total payoff obtained is not significantly different from the median time spent and payoffs gained in part 2. Thus, we can reject these conjectures.

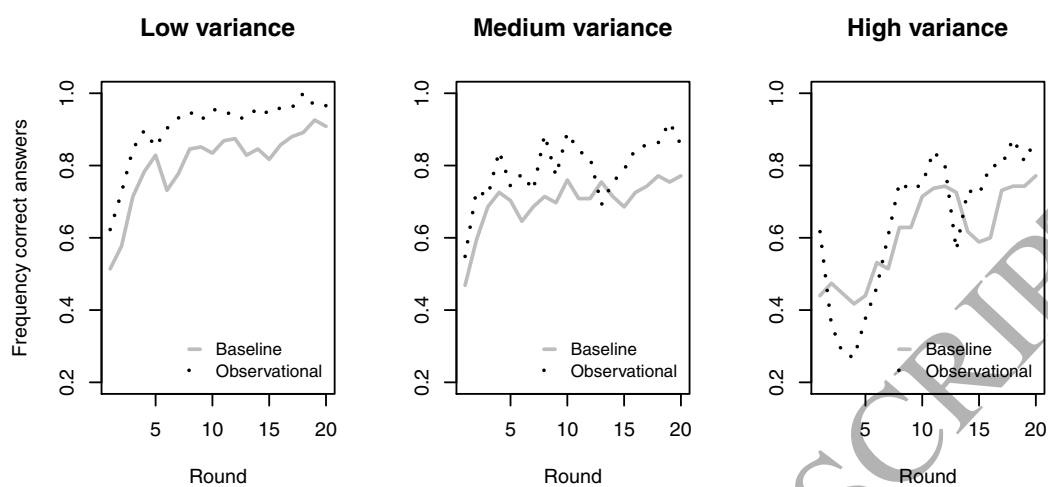


Figure 3: Comparison correct answers per round by variance: baseline and observational treatment

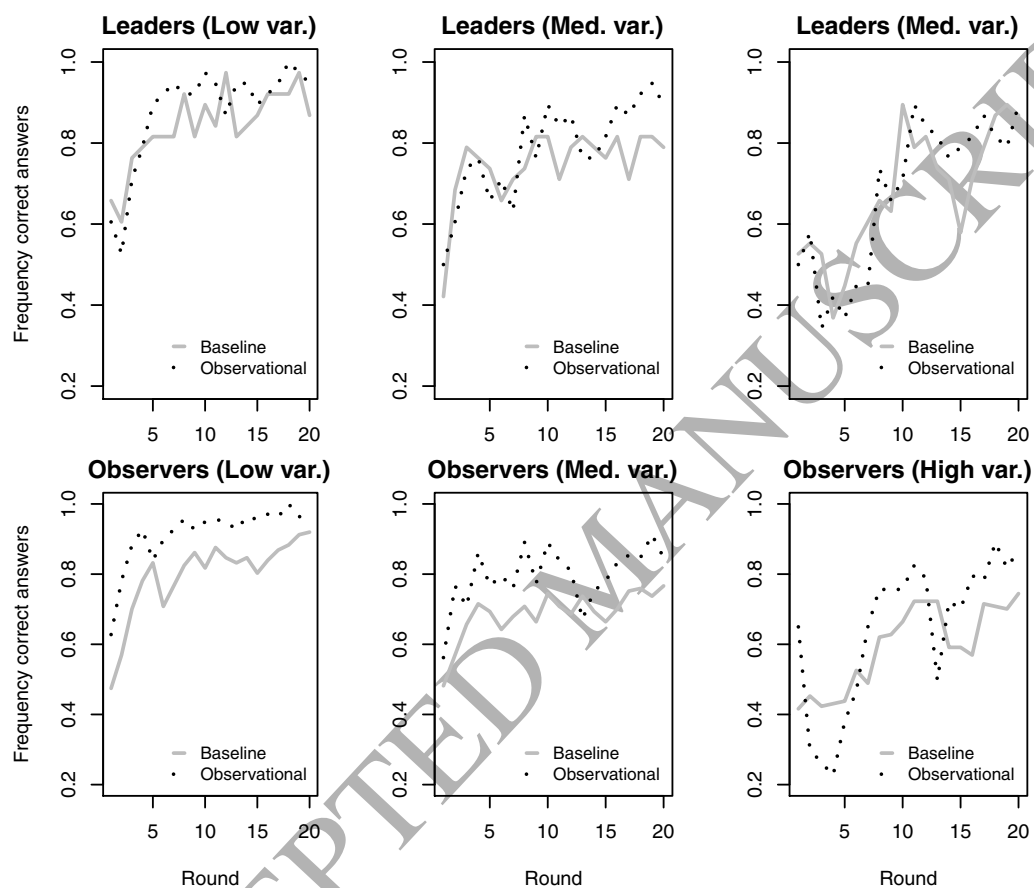


Figure 4: Comparison correct answers per round for leaders and observers in both baseline and the observational treatment

To look further at whether the performance increase observed in Figure 3 is really due to social learning, in Figure 4 we separate leaders from non-leaders (i.e. observers) and we compare the share of correct answers for the two groups in both treatments. The figure clearly shows that whereas the performance of observers in the observational treatment increases visibly — especially in the low and medium variance cases — the leaders’ performance in this second treatment of the experiment follows the same trend observed in baseline.²¹ Thus, we believe that this shows that observers have clearly improved their performance due to observational learning.²²

A nice question concerns the relationship between self-efficacy and payoffs.²³ If it is true that self-efficacious people “march to their own drummer” then it is possible that they will ignore the information available from observing another person, in this case the leader. If so, this could harm their performance in the observational treatment. When excluding the leaders and considering the low

²¹Additionally, as a robustness check, we conducted the following placebo exercise. We created groups of four or five observers, and randomly selected a leader per each group. Of course, the observers have no means to learn from the leaders in this case. As a consequence when comparing the observers’ performance in the second treatment to the performance of the leaders in the same treatment, we should observe no effect. This is exactly what we find.

²²Following the interesting suggestion of an anonymous reviewer, we also compared the baseline treatment and the observational treatment controlling for the total number of signals available. More specifically, we compared the share of correct answers in the two treatments as a function of number of information signals participants hold. The share of correct answers follows the same pattern across the two treatments, however in the second treatment subjects seem to make more correct choices especially in the low and medium variance case. This may be because in the observational treatment the 20th observation is closer in time (it occurs after 10 rounds) than it is in the individual treatment (where it takes 20 rounds and so more time). This might imply that it is easier for individuals to recall past choices, and so make better inferences about which colour is superior, in the observational treatment. In the high variance context a less clear pattern emerges. In this case, holding both individual and socially retrieved signals does not guarantee improved performance. This is, as mentioned, most probably due to the noisiness of the signals participants receive but also possibly related to their tolerance for ambiguity. While specifically investigating the relation between self-efficacy, performance and task complexity, a previous study has found that the accuracy of self-efficacy beliefs in predicting performance is significantly lower in complex environments when tolerance for ambiguity is low [34]. This seems to suggest that, in line with rational inattention argument, the complexity of the task trumps the effect that self-efficacy exerts on the social learning process.

²³We gratefully acknowledge an anonymous referee who posed this question to us. The Pearson’s correlation coefficient between self-efficacy and the weight given to the leader’s action stands at -0.158 (p-value=0.064) or -0.124 (p-value=0.147) depending on whether respectively θ or α are considered. The correlation coefficient between θ or α and the total score obtained in the observational treatment is 0.21 (p-value=0.009) or 0.11 (p-value=0.173) respectively. Conversely, when correlating self-efficacy and payoffs we obtain a coefficient equal to 0.13 (p-value=0.112).

variance setting, we do find a (weak) negative correlation between self-efficacy and the weight given to observation of others; and a (weak) positive correlation between the weight given to others and the payoffs obtained in the second treatment. This suggests that more self-efficacious people, by being confident in their own knowledge and abilities, under-value the experience of others as sources of information, and so tend to receive lower payoffs. However, while the two steps of the casual chain independently give corroborating correlations, the overall suggested relation, between higher self-efficacy and lower payoffs, was not strong enough to overcome other factors that must be at play (giving a non-significant correlation of positive sign).

As in baseline, after 20 rounds, participants were asked to give their best guess of the average rewards given by the 2 colours they had to choose from. The same pattern emerges here: as in the baseline condition, in more volatile environments fewer individuals correctly guess the real means of the payoff distributions.

5. Self-efficacy and imitation

The last step of our analysis uncovers the relation between the propensity to imitate and self-efficacy beliefs. As mentioned above, drawing on the economics and psychology literature, once the payoff-enhancement motive is accounted for, we expect a negative relation between propensity to imitate and self-efficacy beliefs.

In Table 9, we report the results of our preferred OLS regressions. Our dependent variable, i.e. the logarithm of θ or α , captures individuals' propensity to imitate, or the weight subjects give to leaders' actions in making their own choices. The logarithm of the results of the self-efficacy questionnaires represent our main regressor. We include two controls. The first is math ability. If a problem-solving environment is difficult for a subject, we might expect a higher reliance on a leader for information and learning. This would be in line with the work of Vostroknutov et al. [74] according to which imitation depends on individuals' intelligence. Therefore, to control for subjects' ability to solve the problems implicit in the experimental environment, we used the scores they achieved on the math test we administered. Second, in line with standard payoff-enhancement argument within the economic literature, we control for the score obtained by the group leader. Last, we include a dummy variable for gender.²⁴

In this regression analysis we excluded the 38 leaders (since they are playing individually and have no leader to follow or not) and those whose β estimates

²⁴We also controlled for risk-attitudes using the answers collected in our final questionnaire. Nonetheless, we found that risk does not significantly affect people's propensities to imitate others notwithstanding the volatility of the environment.

Table 9: The relation between imitation and self-efficacy beliefs

Low Variance						
dep. var.	$\log(\theta)$	$\log(\theta)$	$\log(\theta)$	$\log(\alpha)$	$\log(\alpha)$	$\log(\alpha)$
Const	8.86 (4.33)*	9.22 (4.48)*	-230.71 (64.29)***	6.19 (4.25)	6.66 (4.29)	-203.4 (56.48)***
log(self-eff.)	-3.08 (1.27)**	-3.24 (1.36)**	-3.51 (1.36)**	-2.57 (1.23)*	-2.78 (1.27)*	-3.12 (1.24)**
log(math)		0.12 (0.38)	0.14 (0.37)		0.16 (0.34)	0.14 (0.33)
log(scoreleader)			43.54 (11.70)***			38.23 (10.19)***
gender (male=1)			-0.37 (0.36)			-0.66 (0.32)**
R^2	0.04	0.04	0.15	0.03	0.03	0.16
obs.	109	109	109	109	109	109
Medium Variance						
Const	2.91 (5.64)	4.01 (5.83)	-2.69 (32.36)	-1.89 (4.84)	-0.77 (4.9)	-97.45 (27.06)***
log(self-eff.)	-1.43 (1.64)	-1.96 (1.72)	-2.15 (1.71)	-0.27 (1.40)	-0.80 (1.45)	-1.20 (1.36)
log(math)		0.47 (0.38)	0.41 (0.37)		0.48 (0.33)	0.40 (0.32)
log(scoreleader)			1.43 (5.84)			17.95 (4.90)***
gender (male=1)			-0.55 (0.34)			-0.35 (0.33)
R^2	0.00	0.03	0.05	0.00	0.02	0.04
obs.	113	113	113	113	113	113
High Variance						
Const	-3.65 (4.11)	-2.84 (4.12)	-57.27 (18.32)***	-4.67 (3.94)	-4.41 (3.88)	-42.11 (18.35)**
log(self-eff.)	0.58 (1.2)	0.15 (1.22)	0.18 (1.20)	0.62 (1.14)	0.48 (1.14)	0.54 (1.15)
log(math)		0.45 (0.32)	0.24 (0.30)		0.14 (0.30)	0.02 (0.32)
log(scoreleader)			10.00 (3.27)***			6.68 (3.32)**
gender (male=1)			-0.32 (0.30)			-0.06 (0.31)
R^2	0	0.02	0.09	0.00	0.00	0.04
obs.	122	122	122	122	122	122

Robust standard errors in parenthesis *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

were at the extreme values, as discussed in Section 4.2.1.²⁵ Thus, the sample size reduces to 109, 113 or 122 people depending on the variance value.

The results of the log-log regressions can be easily interpreted as elasticities.²⁶ As can be seen in the upper part of Table 9, a 1% increase in self-efficacy beliefs corresponds to about a 3% decrease in the weight people give to the leader's action and thus in their propensity to imitate.²⁷ This result is stable regardless of the social learning model used.²⁸ Its significance and magnitude is not affected

²⁵In some cases, the individuals whose β values took on extreme values in the baseline treatment are also leaders in the observational treatment.

²⁶The results of linear, log-linear and linear-log specifications are available upon request.

²⁷This result could be biased if self-efficacy were correlated with the Memory Decay parameters ($\hat{\beta}$ and \hat{r}). However, essentially no correlation exists between these parameters and self-efficacy scores.

²⁸As a slightly more direct estimation procedure, we tested whether a joint estimation of both individual learning and imitation parameters would change our main results. Using the data from the low variance observational learning treatment we optimised the three parameters (β , r and α or θ) jointly, and then ran the final regression. Our results remain quite stable — their magnitude and sign do not change. Controlling for gender, math abilities and score of the leader, a 1% increase in self-efficacy reduces the propensity to imitate by 2.5 or 3.4 percent,

when controlling for math abilities or the leader's score. We observe that math skills seem to have no effect, suggesting that subjects were able to cope with the logic implicit in the task. Subjects were also observant enough to identify when leaders were doing well, and gave more weight to observations of leaders whose payoffs were high. Specifically, a 1% percent increase in the leader's score induces roughly a 40% increase in the weight individuals give to the information gained by observing the leader. While this elasticity looks very large, when translated into behaviour change it is much smaller. When using the nested model presented in Table 8, the probability that the agent makes the same choice as the leader, following a 1% increase in his leader's score, increases by 6.3%.²⁹ This confirms the general claim that agents tend to imitate well-performing individuals and, comparing the elasticities estimated in Table 9, that payoff concerns exert a larger effect than self-efficacy on the weight people give to the actions taken by others. However, introspective beliefs of inability represent a reason for imitation which clearly complements the standard economic argument. In the case of low volatility, holding the leader's score constant, weakly self-efficacious individuals are more apt to follow the leader's action.

In the medium and high variance cases instead, we cannot reject the null hypothesis. Several factors can contribute to this result. First, higher variance implies that there can be a strong conflict in the signals about the value of an action. Even if the variance is small, results for one action will differ from time to time and player to player. But, observationally, the strength of the divergence of the signals will be small. When variances are high, the payoffs to a single action vary widely, and so subjects will perceive that the experience of the leader can be quite different from their own. Thus the perceived usefulness of the signal of the leader's experience is reduced. This factor is likely to be more relevant when estimating α , where the subjects are assumed to use their observations on the leaders experience as information equivalent to those of their own experiences. Additionally, given the noise in the signals received, subjects might conclude that the leader is as ill-equipped to understand the decisional environment as they are and thus imitation might not be perceived as valuable. This is confirmed by the

depending on whether imitation is measured by α or θ respectively.

²⁹This rate of change of the the probability that an agent chooses colour 1 if the leader does has been calculated using the nested model as follows. Suppose that a subject's propensity to imitate (α) is equal to 0.2, and that the probability that he would have chosen colour 1 if alone (L), is equal to 0.5. In this case, the probability that the agent chooses colour 1 if his leader did is: $(0.8*0.5+0.2*1)/(0.8*0.5+0.2*1+0.8*0.5+0.2*0) = 0.6$. Suppose that the leader's score, is 1% higher. This change in the leader's score will increase α by roughly 38%, (as estimated in the regression) i.e. by 0.076. As a consequence in this case the probability that agent chooses colour 1, following the leader is: $(0.724*0.5+0.276*1)/(0.724*0.5+0.276*1+0.724*0.5+0.276*0) = 0.63$.

fact that, at times, even the leader's score loses its significance in our regression analysis.

Moreover, as mentioned above, in tasks wherein outcomes are only loosely tied to actions, self-efficacy and outcome expectancies become separable [10] and possibly offset each other. Self-efficacy governs the effect that outcome expectancies, i.e. the judgements that actions will produce the expected outcome, have on the performance of a behaviour [12]. Therefore, even if self-efficacy is high, the performance of a behaviour might be unlikely if outcome expectancies are very low. In a nutshell, the effect of self-efficacy can be neutralised by low outcome expectancies, and any relationship between self-efficacy and behaviour weakens possibly fading away [28]. Thus, self-efficacy only has power when agents have strong outcome expectancies. When outcome expectancies are very weak, efficacious or not, subjects do not feel they can control the outcome of their actions, and so whether or not they feel efficacious will explain much less of any observed patterns.

6. Conclusions

Individuals can imitate others to gain status or to minimise regret. Additionally, imitation can be driven by simple instrumental reasons. Agents might feel compelled to imitate co-players if they think this will help them obtain higher payoffs. The goal of this study was to go beyond these explanations and test whether agents might engage in imitative behaviours because of their low self-efficacy beliefs.

In the laboratory our subjects played a modified version of the two-armed bandit problem. Our results point to a negative relation between self-efficacy beliefs and imitative behaviours. Regardless of the imitation model used, more self-efficacious agents are less likely to follow their group leader. This establishes a nexus between individuals' introspective beliefs and the way in which agents learn when given the possibility to observe a peer's actions and outcomes. These effects are strongest in low volatility environments, effectively disappearing when volatility of the payoff distribution gets too high. Nonetheless, taking into account human cognitive capacities seems to reveal another additional micro-mechanism able to explain why people imitate others. Imitation occurs not only because of the payoff enhancement that players envision, their quest for status or regret, but also because of low self-evaluations of their capacities.

We should mention two *caveats* to these results. The first is that we have taken self-efficacy beliefs as fixed over the course of the experiment. Beliefs were assessed at the start of the experiment by a questionnaire, and assumed not to change. But in fact we know that self-efficacy beliefs can change in light of positive or negative experiences. Bandura himself claims that whilst self-efficacy affects

behaviour and learning, learning processes feedback into self-efficacy beliefs. We have assumed that the experiment was short enough that this feedback had little effect. Second we restricted our analysis to a case in which people observe a peer and possibly imitate his or her action. We did not consider the effect of social interactions in the form of communication for example. We also did not allow for subjects to select the co-player to be imitated. Doing so would demand a different, and probably more complex, experimental design, but taking this into account would represent an interesting next step in formulating a model of how people decide to imitate others.

Imitation is sometimes referred to as “the poor man’s rationality” [62, p.461]. Offerman and Schotter [62] argue that in case a decision maker were fully rational and capable of effortlessly carrying out all necessary calculations, he would not feel any need to imitate anyone. Our finding strengthens this statement as we find that, after controlling for their abilities to process simple information and the leader’s payoffs, self-efficacious people are less apt to follow the leader’s action.

This result clearly uncovers the importance of individual introspective beliefs in the context of imitation and thus contributes to the experimental and theoretical economics literature. Additionally, we contribute to the literature on experts’ behaviour whose main concern has been to explain how confidence triggers rent-seeking and risk-taking behaviour while disregarding that low confidence and self-efficacy beliefs might trigger imitation. This paper also contributes to the cultural evolution literature whose goal, lately, has been to detect individual variation in social learning dynamics and the related consequences in behavioural transmission [55, 56, 32]. We believe that our finding suggests that in communities characterised by low volatility and where the general population has low self-efficacy beliefs, people are more inclined to follow the behaviour of prominent individuals regardless of the goodness of their actions. As a consequence low self-efficacious people will tend to follow the behavioural path set by somebody else, thus conforming to the past, promoting vertical institutional reproduction [64] and perpetuating possibly inefficient behaviours.

References

- [1] Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics* 115(3), 715–753.
- [2] Alós-Ferrer, C. (2004). Cournot versus walras in dynamic oligopolies with memory. *International Journal of Industrial Organization* 22(2), 193–217.
- [3] Alós-Ferrer, C. and K. H. Schlag (2009). Imitation and learning. In P. Anand, P. K. Pattanaik, and C. Puppe (Eds.), *The handbook of rational and social choice : an overview of new foundations and applications*, pp. 13–43. Oxford: Oxford University Press.
- [4] Anderson, L. R. and C. A. Holt (1997). Information cascades in the laboratory. *The American Economic Review* 87(5), 847–862.
- [5] Apesteguia, J., S. Huck, and J. Oechssler (2007). Imitation–theory and experimental evidence. *Journal of Economic Theory* 136(1), 217 – 235.
- [6] Aragonés, E., I. Gilboa, A. Postlewaite, and D. Schmeidler (2005). Fact-free learning. *The American Economic Review* 95(5), 1355–1368.
- [7] Arthur, W. B. (1991). Designing economic agents that act like human agents: A behavioral approach to bounded rationality. *The American Economic Review* 81(2), 353–359.
- [8] Arthur, W. B. (1993). On designing economic agents that behave like human agents. *Journal of Evolutionary Economics* 3(1), 1–22.
- [9] Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review* 84(2), 191.
- [10] Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist* 37(2), 122.
- [11] Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist* 44(9), 1175–1184.
- [12] Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York : W.H. Freeman.
- [13] Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan and F. Pajares (Eds.), *Self-efficacy beliefs of adolescents*, pp. 307–337. IAP.
- [14] Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics* 107(3), 797–817.

- [15] Barber, B. and T. Odean (1999). Do investors trade too much? *The American Economic Review* 89(5), 1279–1298.
- [16] Bayer, P., R. Hjalmarsson, and D. Pozen (2009). Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics* 124(1), 105–147.
- [17] Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy* 102(5), 841–877.
- [18] Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100(5), 992–1026.
- [19] Blume, L. E. (1995). The statistical mechanics of best-response strategy revision. *Games and Economic Behavior* 11(2), 111–145.
- [20] Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner (2008). Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance* 63(3), 1509–1531.
- [21] Burke, C. J., P. N. Tobler, M. Baddeley, and W. Schultz (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences* 107(32), 14431–14436.
- [22] Burnham, K. P. and D. R. Anderson (1998). *Model selection and multi-model inference: a practical information-theoretic approach*. Springer Science & Business Media.
- [23] Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica* 82(4), 1273–1301.
- [24] Bush, R. R. and F. Mosteller (1951). A mathematical model for simple learning. *Psychological Review* 58(5), 313–323.
- [25] Bush, R. R. and F. Mosteller (1953). A stochastic model with applications to learning. *The Annals of Mathematical Statistics* 24(4), 559–585.
- [26] Camerer, C. (2003). *Behavioral game theory : experiments in strategic interaction*. Princeton, N.J.: Princeton University Press.
- [27] Camerer, C. and D. Lovallo (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review* 89(1), 306–318.

- [28] Conner, M. and P. Norman (1998). Health behavior. In A. S. Bellack and M. Hersen (Eds.), *Comprehensive Clinical Psychology*, Volume 8, pp. 1 – 37. Oxford: Pergamon.
- [29] Cooper, D. J. and M. Rege (2011). Misery loves company: Social regret and social interaction effects in choices under risk and uncertainty. *Games and Economic Behavior* 73(1), 91 – 110.
- [30] Cross, J. G. (1973). A stochastic learning model of economic behavior. *The Quarterly Journal of Economics* 87(2), 239–266.
- [31] Eckel, C. C. and R. K. Wilson (2007, Sep). Social learning in coordination games: does status matter? *Experimental Economics* 10(3), 317–329.
- [32] Efferson, C. and S. Vogt (2018). Behavioural homogenization with spillovers in a normative domain. *Proceeding of the Royal Society B* 285(1879), 20180492.
- [33] Ellison, G. and D. Fudenberg (1993). Rules of thumb for social learning. *Journal of Political Economy* 101(4), 612–643.
- [34] Endres, M. L., S. Chowdhury, and M. Milner (2009). Ambiguity tolerance and accurate assessment of self-efficacy in a complex decision task. *Journal of Management and Organization* 15(1), 31–46.
- [35] Erev, I., E. Ert, A. E. Roth, E. Haruvy, S. M. Herzog, R. Hau, R. Hertwig, T. Stewart, R. West, and C. Lebiere (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making* 23(1), 15–47.
- [36] Erev, I. and E. Haruvy (2013). Learning and the economics of small decisions. In J. H. Kagel and A. Roth (Eds.), *The Handbook of Experimental Economics*, Volume 2, pp. 638–700. Princeton University Press.
- [37] Erev, I. and A. E. Roth (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review* 88(4), 848–881.
- [38] Erev, I. and A. E. Roth (2007). Multi-agent learning and the descriptive value of simple models. *Artificial Intelligence* 171(7), 423 – 428.
- [39] Fatas, E., S. P. H. Heap, and D. R. Arjona (2018). Preference conformism: An experiment. *European Economic Review* 105, 71–82.
- [40] Festinger, L. (1954). A theory of social comparison processes. *Human Relations* 7(2), 117–140.

- [41] Gervais, S. and T. Odean (2001). Learning to be overconfident. *Review of Financial Studies* 14(1), 1–27.
- [42] Gioia, F. (2017, Mar). Peer effects on risk behaviour: the importance of group identity. *Experimental Economics* 20(1), 100–129.
- [43] Herrmann, E., J. Call, M. V. Hernández-Lloreda, B. Hare, and M. Tomasello (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* 317(5843), 1360–1366.
- [44] Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior* 4, 267–272.
- [45] Huck, S., H.-T. Normann, and J. Oechssler (1999). Learning in cournot oligopoly—an experiment. *The Economic Journal* 109(454), 80–95.
- [46] Hurley, S. L. and N. Chater (2005). *Perspectives on Imitation: from neuroscience to social science*. Cambridge, Mass. ; London : MIT.
- [47] Judge, T. A. and J. E. Bono (2001). Relationship of core self-evaluations traits - self-esteem, generalized self-efficacy, locus of control, and emotional stability - with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology* 86(1), 80–92.
- [48] Lahno, A. M. and M. Serra-Garcia (2015). Peer effects in risk taking: Envy or conformity? *Journal of Risk and Uncertainty* 50(1), 73–95.
- [49] Luce, R. (1959). *Individual Choice Behavior a Theoretical Analysis*. New York : Wiley.
- [50] Malmendier, U. and G. Tate (2005). Ceo overconfidence and corporate investment. *The Journal of Finance* 60(6), 2661–2700.
- [51] McElreath, R., A. V. Bell, C. Efferson, M. Lubell, P. J. Richerson, and T. Waring (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 363(1509), 3515–3528.
- [52] McElreath, R., M. Lubell, P. J. Richerson, T. M. Waring, W. Baum, E. Edsten, C. Efferson, and B. Paciotti (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior* 26(6), 483–508.

- [53] Merlo, A. and A. Schotter (2003). Learning by not doing: an experimental investigation of observational learning. *Games and Economic Behavior* 42(1), 116–136.
- [54] Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior* 32(5), 334–342.
- [55] Mesoudi, A., L. Chang, S. R. Dall, and A. Thornton (2016). The evolution of individual and cultural variation in social learning. *Trends in Ecology & Evolution* 31(3), 215–225.
- [56] Muthukrishna, M., T. J. Morgan, and J. Henrich (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior* 37(1), 10–20.
- [57] Nanda, R. and J. B. Sørensen (2010). Workplace peers and entrepreneurship. *Management Science* 56(7), 1116–1126.
- [58] Nax, H. H., M. N. Burton-Chellew, S. A. West, and H. P. Young (2016). Learning in a black box. *Journal of Economic Behavior & Organization* 127, 1–15.
- [59] Nevo, I. and I. Erev (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology* 3(24), 1–9.
- [60] Nicolle, A., M. Symmonds, and R. Dolan (2011). Optimistic biases in observational learning of value. *Cognition* 119(3), 394 – 402.
- [61] Offerman, T., J. Potters, and J. Sonnemans (2002). Imitation and belief learning in an oligopoly experiment. *The Review of Economic Studies* 69(4), 973–997.
- [62] Offerman, T. and A. Schotter (2009). Imitation and luck: An experimental study on social sampling. *Games and Economic Behavior* 65(2), 461 – 502.
- [63] Offerman, T. and J. Sonnemans (1998). Learning by experience and learning by imitating successful others. *Journal of Economic Behavior & Organization* 34(4), 559–575.
- [64] Patterson, O. (2010). The mechanism of cultural reproduction: explaining the puzzle of persistence. In J. R. Hall, L. Grindstaff, and M.-C. Lo (Eds.), *Handbook of Cultural Sociology*, pp. 139–151. Oxford: Routledge.

- [65] Pingle, M. (1995). Imitation versus rationality: An experimental perspective on decision making. *The Journal of Socio-Economics* 24(2), 281–315.
- [66] Pingle, M. and R. H. Day (1996). Modes of economizing behavior: Experimental evidence. *Journal of Economic Behavior & Organization* 29(2), 191–209.
- [67] Roth, A. E. and I. Erev (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior* 8(1), 164–212.
- [68] Schlag, K. H. (1998). Why imitate, and if so, how? *Journal of Economic Theory* 78(1), 130–156.
- [69] Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist* 19(1), 48–58.
- [70] Schwarzer, R., J. Bäßler, P. Kwiatek, K. Schröder, and J. X. Zhang (1997). The assessment of optimistic self-beliefs: comparison of the german, spanish, and chinese versions of the general self-efficacy scale. *Applied Psychology* 46(1), 69–88.
- [71] Schwarzer, R. and M. Jerusalem (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, and M. Johnston (Eds.), *Measures in health psychology: A users portfolio. Causal and control beliefs*, pp. 35–37. Windsor, UK: NFER-NELSON.
- [72] Selten, R. and A. Ostmann (2000). Imitation equilibrium. Technical report, Bonn Econ Discussion Papers.
- [73] Vega-Redondo, F. (1997). The evolution of walrasian behavior. *Econometrica* 65(2), 375–384.
- [74] Vostroknutov, A., L. Polonio, and G. Coricelli (2018). The role of intelligence in social learning. *Nature-Scientific reports* 8(1), 6896.
- [75] Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics* 85(1), 9–23.