Observational Learning 
and Intelligence

Alexander Vostroknutov†‡ Luca Polonio Giorgio Coricelli§

July 2016

Abstract

We study experimentally how individuals learn from observing the choices of others in a non-stationary stochastic environment. The imitation choices of participants with low score in an intelligence test are driven solely by the value of imitation. High intelligence score participants, in addition, use choices of others to better understand the environment. They imitate more when other’s choices are stable, which makes them more optimal than low score participants. The knowledge that the observed other has high intelligence score affects behavior of only low score participants. Overall, intelligence predicts the usage of simple or sophisticated observational learning strategy.

JEL classifications: C91, C92
Keywords: observational learning, intelligence, bandit problems

†‡Corresponding author: a.vostroknutov@unitn.it
§Department of Economics, University of Southern California.
1 Introduction

Learning is an important and flexible process that allows humans to adapt to their environment. A first basic source of learning comes from personal experience. Humans interact directly with the environment and learn from the feedback they receive. A second source of learning comes from observing other people interacting with the same environment. In a world where we need to adapt quickly to the ever-changing circumstances (e.g., climate fluctuations, socio-political commotion), the ability to learn from others is fundamental because it reduces the effort of acquiring information. This is especially true when observational learning is based on simple reinforcement learning mechanisms, which is the case when an agent imitates others, evaluates the feedback she receives from the environment and chooses whether to keep imitating or not depending on the outcome. However, learning from others can be much more than a mere imitation of the observed behavior. A sophisticated way to learn from others includes the understanding of the rationale behind the observed choices. An agent integrates what she has observed with the feedback she has directly received from the environment and changes her behavior accordingly. This sophisticated observational learning process is particularly relevant when the environment changes and simple imitation becomes not reliable. This sort of learning can be very efficient but it is also more costly because it requires a higher level of attention and an ability to integrate information from diverse sources. In this respect, intelligence plays an important role as it makes it easier for the agent to gather and analyze information.

In this respect, intelligence may be associated with the ability of the agent to acquire and analyze information.

In this paper we study experimentally the mechanisms of how people learn from others. In particular, we are interested in what features of others’ behavior influence imitation and how this affects the performance in an uncertain environment. Moreover, we test hypotheses regarding the role of intelligence and look at the interactions between a measure of intelligence of our participants and the intelligence of the people they observe and learn from.

The knowledge of the interplay of observational learning and intelligence can be very important for policy and issues of economic efficiency. In any environment where people work and/or learn together in fixed groups (school classes, firms etc.) the problems or biases might arise in who learns what and from whom (given observed levels of intelligence). For example, Braaksma et al. (2002) study the increase in performance in a writing task among 8th graders as
dependent on the intelligence of the observed person. They find significant effects of the level of intelligence of the observed person on the success of the observer, which also depends on the observer’s intelligence. This raises interesting questions regarding, for example, the division of pupils into classes according to measures of intelligence.

To study observational learning in a changing environment we look at a two-armed bandit problem with independent non-stationary stochastic processes that determine the payoffs from the two arms. In our design, the performance depends on the ability of the participants to choose the action which gives a reward with higher probability. This is not trivial since the probability of getting a reward, associated with each action, changes over time. To understand how participants learn from others, we give them the possibility to observe the choices made by another person who previously interacted with the same environment, but not outcomes. We do not show the outcomes that the other obtained for two reasons: 1) observing outcomes would reduce the other to just additional information about the environment and 2) the optimality of the other would become evident. In addition, participants may or may not receive information about the intelligence of the observed person. In this way, we are able to evaluate the effect of information about intelligence on the level of imitation.

Our interest is the exact mechanism of the learning process and the way it is modulated by the knowledge about owns intelligence and the intelligence of the person observed. Non-stationary environment allows us to distinguish between agents who simply imitate observed behavior (simple imitation) and agents who use it to better understand the current state of the environment (sophisticated imitation). In non-stationary environment agents never stop learning, which makes it possible to study how they adapt throughout the experiment. Conversely, in stationary environment it is hard to disentangle imitation of observations of others and own learning curves of the observer and the observed.

When addressing the possible influences of intelligence on the mechanisms of how people learn from others, we decided to use a measure of fluid intelligence. In particular we were interested in non-verbal abilities and reasoning skills because our learning task is not in any particular verbal domain (spelling, reading, comprehension, etc.). Therefore, we used the Raven Advanced Progressive Matrices (RAPM), a measure of efficient problem solving and abstract reasoning which consists of a series of pattern matching tasks that do not require mathematical or verbal reasoning abilities (Raven et al., 1998). The performance in Raven test is linked to the ability to integrate information in sophisticated way. We also collect data about cognitive reflection, the ability to resist an immediate, intuitive and incorrect answer, executed with little deliberation, in favor of the search for the correct answer requiring a more complex reasoning (Cognitive Reflection Test, proposed by Frederick (2005)). The cognitive abilities measured by these tests are particularly relevant for the situations faced by participants in this experiment, as in our learning task participants can learn from the information about the observed action using simple and/or sophisticated strategies.
Our results can be summarized as follows. We do find both simple and sophisticated imitation effects in our participants. We find that participants with high Raven score use both simple and sophisticated learning strategies and participants with low Raven scores use only simple imitation. The imitation of the observed participant is increased when choosing the same action as the observed participant leads to an increase in earnings (simple imitation). We call this effect simple, because it naturally falls into the domain of a rather mechanical reinforcement learning paradigm where the values of actions are reinforced if they brought higher payoff in the past. In sophisticated imitation, participants track how stable the choices of the other are, make inferences about what payoffs the observed participant might have received, given her behavior, and imitate if they infer that the other is getting high payoffs. In addition, we find that high Raven participants choose how much attention to pay to the other depending on her own earnings in the past: the less they earn, the more imitation we observe.

We find an effect of information about the intelligence of the observed participant on the imitation strategy. When this information is not available, high Raven participants use both simple and sophisticated imitation strategies, whereas low Raven participants use only simple imitation. In case information is provided both high and low Raven participants use only simple imitation. This suggests that information about the intelligence of the observed participant is used as a signal of the ability to perform in the task.

We also study how imitation and intelligence of both the observer and the observed influence the optimality of the choices and the earnings. We find that participants with high Raven scores choose more optimally and earn more money than participants with low Raven scores. Only low Raven participants are affected by the information about Raven score of the observed. In particular, when knowing that the other has high Raven score, they increase imitation which leads to better performance. Interestingly, high Raven participants are not affected by this information. These observations suggest that participants with different intelligence levels use different observational learning strategies, which is particularly important considering the large number of social and economic contexts in which people can learn from others.

Finally, we find overall consistency of our data and the imitation choices. In particular, we find that higher Raven score implies less switching, which, in its turn, leads to higher optimality and higher earnings. On the behavioral side, we find that high Raven participants imitate more, the less switches they observe, thus learning to perform more optimally in the task. So, high Raven participants are able to correctly interpret what they observe, to learn from it and, thus, better performance, regardless whether the information about the Raven score of the other was provided or not. Conversely, low Raven participants need to know the information about the intelligence of the other in order to improve their performance.
2 Literature Overview

To our knowledge, we are the first in economic literature to study the mechanisms of observational learning in non-stationary environments. In previous studies this topic was addressed only in stationary situations and without taking into account the information about the intelligence of the observers and observed participants. The literature closest to our study is that of information cascades (Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sørensen, 2000). In these studies, agents try to learn the qualities of objects from observing the choices of other agents who are known to have imperfect signals about the qualities of these objects. Unlike this literature though, we consider the learning process from one other person, which is a distinguishing feature of our study. Observing and learning from the choices of one individual instead of sequence of single choices by many people can have important features overlooked by the information cascades literature. In particular, the characteristics of the choices of the observed other can have an effect on the learning process.

In addition to the information cascades literature there are other examples of observational learning studies that deserve mentioning. Armantier (2004) shows that observing the opponents’ private signals, bids and payoffs in a repeated common value auction homogenizes behavior and accelerates learning towards Nash equilibrium. Çelen and Kariv (2004) show that in a game with pure information externalities over time private information is ignored and decision makers become increasingly likely to imitate their predecessors. Merlo and Schotter (2003) show that in a complete information maximization problem participants can learn better by not doing but watching someone else’s behavior.

In neuroeconomic literature the most similar study is Burke et al. (2010). The authors use related learning task (though with stationary environment) and find the correlation of other’s action prediction error and brain activity in dorsolateral prefrontal cortex. The authors claim that observational learning is incorporated in standard payoff based learning and demonstrate this with a simple reinforcement learning model. The main difference between our approach and that of Burke et al. (2010) is that in their model observing other choosing an action automatically increases the probability of choosing that action. We show that this is not always the case and that it depends on the past payoffs and past behavior of the observed participant. Another related study is Nicolle et al. (2011). Here the authors show that people tend to be overly optimistic about rewards that they observe being received by someone else (rather than choosing for themselves). In this study, the payoffs of the observed participant where shown to participants, which makes it rather different from our study and the questions asked.

Some studies consider the connection between cognitive ability and performance in strategic environments. Gill and Prowse (2015) look at beauty contest games and find that high Raven score is associated with choosing actions closer to Nash equilibrium, with ability to converge faster to the equilibrium and to earn more than the low Raven score participants. Proto et al.
(2014) study the performance in the repeated Prisoner’s Dilemma. They find that, when participants are grouped into high Raven and low Raven score groups, the high Raven score groups are able to sustain cooperation, whereas low Raven groups cannot. High Raven groups earn more than low Raven groups. Even though our task does not involve games, the findings in both studies are in line with what we find: high Raven individuals perform better and earn more money.

3 Experimental Design

The study consisted of two experiments: a first experiment in which participants made choices in a 2-armed bandit problem without observing actions of others, and a second experiment in which participants made choices in the same environment with only difference that in a half of the trials they observed the choices made by one of the two participants selected from the first experiment. The purpose of the first experiment was to select two participants, one with high and one with low RAPM score, in order to use them as observed participants in Experiment 2.

The two observed participants were chosen using the following procedure. First, we divided the participants into deciles of Raven score. Then we calculated the median number of switches between actions for participants in the first and tenth decile. We chose two participants (one in the first and one in the tenth decile) who were closest to the median. The aim of this procedure was to select two participants who would have a prototypical behavior in terms of the number of switches in the two extremes of the RAPM dimension. We decided to use the number of switches parameter for two reasons: 1) it is an index related to the earnings of the participants and 2) we hypothesized that this is an important parameter that differentiates sophisticated and simple learners.

Participants in Experiment 2 were divided into four treatments with 2×2 design. The dimensions were: 1) the RAPM score of the observed participant (HighRaven or LowRaven) and 2) the information the participants received about the RAPM score of the observed other (Visible and Not-Visible). The Raven score of the observed participant could be high (28) or low (15). Only participants in the Vis-HighRaven and Vis-LowRaven treatments received this information. Participants in Novis-HighRaven and Novis-LowRaven treatments were matched with the corresponding observed participant without knowing his/her score on the RAPM test.

For the main experiment (Experiment 2), 4 Novis and 6 Vis sessions were conducted. In each session half of the participants were in the HighRaven condition and the other half in the

---

1 Among other studies on the connection between cognitive ability and strategic reasoning are Benito-Ostolaza et al. (2016); Fehr and Huck (2015); Hanaki et al. (2015); Kiss et al. (2016). An excellent review of broader literature is Rustichini (2015).

2 Switchiness reflects the need to explore alternative actions when previous outcomes were bad. Ihssen et al. (2016) show that the high number of switches is a sign of high sensitivity to bad outcomes.

3 The number of switches of the low and high Raven participants are equal to 49 and 20, respectively.
LowRaven condition. All participants were recruited from the subject pool of the Cognitive and Experimental Economics Laboratory at the University of Trento (CEEL). The dates of the sessions and the number of participants per session are reported in Table 10, Appendix F.

On average participants earned about €20.06, in addition to the €3 show-up fee. The presentation of the 2-armed bandit task was performed using a custom made program implemented in Matlab Psychophysical toolbox. The tests and questionnaires were administered with z-Tree software package (Fischbacher, 2007). A detailed timeline of the experiment and all instructions are reported in Appendix C.

3.1 Experiment 1

51 participants took part in Experiment 1. In the first part of the experiment participants made choices in a 2-armed bandit problem. In the second part they completed a RAPM test of 30 tables, the Holt & Laury Risk Aversion test, the Cognitive Reflection Test and the Empathy Quotient questionnaire (Baron-Cohen and Wheelwright, 2004).4

After entering the lab, participants were randomly assigned to a PC terminal and were given a copy of the instructions (see Appendix C). Instructions were read aloud by the experimenter, and then a set of control questions were provided to ensure the understanding of the 2-armed bandit problem.

The probabilities of getting a 10 cents reward from each of the two hands followed independent non-stationary stochastic processes. Figure 1 illustrates.5

![Figure 1: The actual probabilities of winning from two options in 200 trials.](image)

Participants were not aware of how the probabilities change but it was made clear that they would change slowly and independently of their choices, earnings and each other. The 2-armed bandit task included 200 trials divided into four blocks of approximately 50 trials each. At the end of the task participants were not informed about their earnings until after they completed

4The Empathy Quotient questionnaire was developed to assess emphatic abilities in adults with autism-spectrum disorders. This questionnaire was added to the study to assess whether emphatic abilities affect the way participants imitate others.

5The process is a decaying Gaussian random walk with parameters $\lambda = 0.8$, decay centre $\theta = 0.5$ and Gaussian noise with standard deviation 0.2 (see Wunderlich et al. (2009) page 17203 for the exact formula).
the second part of the experiment. In the second part of the experiment participants were given 20 minutes to solve 30 RAPM problems. They were told that they have 20 minutes to solve as many problems as they can and that they would earn 30 cents for each correct answer. If participants did not complete an item or their answer was incorrect they would earn 0 cents for that item. At the end of the RAPM test participants completed the Holt and Laury lottery task (with real incentives, see Appendix D), the CRT test and the EQ questionnaire (Appendices C and E). There was no time limit to complete these three tasks and no payment was provided for the CRT test and the EQ questionnaire. At the end of the Holt & Laury task a single lottery was selected at random and played by the computer.

At the end of the second phase, participants were paid according to their choices in the 2-armed bandit problem, their performance in the RAPM problems, the outcome of the selected lottery and a show-up fee of €3.

### 3.2 Experiment 2

In Experiment 2, 160 participants first completed the RAPM test, the Holt & Laury Risk Aversion test, the Cognitive Reflection Test and the Empathy Quotient questionnaire and then played in the 2-armed bandit task. The only difference with Experiment 1 (apart from the order of the tasks) was that participants in the 2-armed bandit problem, sometimes, and before making their choices, also observed the choices (but not the outcomes) made by one of the two selected participants from Experiment 1. The choices of the observed participant were provided in half of the trials (in 100 out of the 200 trials) between trial 10 and trial 200 in blocks of randomized length of 5 to 15 consecutive trials. It was made clear to the participants that the observed behavior was from a real person who took part in the experiment approximately one month before and that he/she chose in the same exact environment. Participants knew that the observed other has completed all the parts of the experiment, including the questionnaire on matrices problems that they completed at the beginning of the experiment. Participants also were informed that the observed other did not himself observe anyone while completing the 2-armed bandit problem task.

Participants were shown (and explained) a histogram of the number of RAPM problems solved by the 51 participants from Experiment 1 (see Figure 9 version a in Appendix C). In this way they had the possibility to compare their performance in the RAPM test, which they knew before starting 2-armed bandit task, with that of the group from which the person they are going to observe was chosen. The participants were only told that they will receive this information. No information about a possible connection between performances in Raven test and the learning task was provided.

---

6Participants were not told their total earnings at the end of the learning task, though, in principle, they could have calculated it by observing the outcomes after each trial.
The instructions were identical for all participants, except for the information that was given about the score obtained by the observed participant. In the Novis treatments (Novis-HighRaven and Novis-LowRaven) the score obtained by the observed participant remained unknown (only distribution of all Raven score was known). Conversely, in the Vis treatments (Vis-HighRaven and Vis-LowRaven) the score of the observed participant was marked in red on the histogram and also shown on the screen during the experiment (see Figures 9 versions b and c in Appendix C).

Participants were students from the University of Trento, Italy (mean age 23.17, SD 0.23). The study was approved by the local ethics committee and all participants gave informed consent.

4 Results

4.1 Summary Statistics

Before presenting our regression results we start with reporting the summary statistics for Experiment 1 and Experiment 2. Table 1 reports the scores obtained in the tests we conducted (average and standard deviation) and demographic information.

<table>
<thead>
<tr>
<th></th>
<th>Raven</th>
<th>CRT</th>
<th>HL</th>
<th>EQ</th>
<th>Gender</th>
<th>Age</th>
<th>Years study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>21.51</td>
<td>1.46</td>
<td>5.82</td>
<td>40.65</td>
<td>0.57</td>
<td>22.81</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.19)</td>
<td>(0.33)</td>
<td>(1.67)</td>
<td>(0.08)</td>
<td>(0.43)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>20.93</td>
<td>1.32</td>
<td>6.02</td>
<td>41.83</td>
<td>0.52</td>
<td>23.17</td>
<td>3.37</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.09)</td>
<td>(0.13)</td>
<td>(0.66)</td>
<td>(0.04)</td>
<td>(0.23)</td>
<td>(0.16)</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics for all data. Earnings reported are for the bandit task only.

A Kolmogorov-Smirnov test was used to test the hypothesis that the distributions of Raven, CRT, HL and EQ scores are different between the two experiments. None of the comparisons was found to be statistically significant, thus supporting the hypothesis that participants in the two experiments come from the same population. This is important because participants in Experiment 2 were presented the distribution of Raven scores of participants from Experiment 1 and the absence of significant differences means that they could compare their score with that of the observed participant in a reliable way.

4.2 Imitation

In this section we present results concerning factors that affect the degree of imitation of the observed participant. We define the variable \(i_m\), which equals to 1 if the participant chose the same action as the observed participant and 0 if she chose a different action.\(^7\) First, we want to

\(^7\)See Appendix A for a description of the variables used in the analyses.
look at the aggregate average levels of imitation in the four treatments (Novis-HighRaven, Vis-
HighRaven, Novis-LowRaven and Vis-LowRaven) for participants having high vs. low Raven
scores. However, there is a slight complication: there are two reasons why one would choose
the same action as the other. The first is pure imitation (the one we are interested in): the par-
ticipant observes what the observed participant has chosen and imitates her. The second is that
the participant and the other have the same preference-related behavior (the one we want to con-
trol): the participant chooses the same action as the observed participant because she thinks this
is the best thing to do regardless of what the observed participant does. To disentangle these
two effects we construct a new variable adjim which, for each participant, is equal to the av-
erage rate of imitation in periods when the other is observed minus the average rate of same
choice when the other is not observed. The reasoning behind this construction is the following:
when a participant chooses on his own without observing the other (in 100 over 200 trials), di-
rect imitation is not possible and the choice of the same action is the consequence of the same
preference-related behavior. For each participant, we estimate the average preference-related
behavior over 100 periods, when the other is not observed. Given large number of periods, this
measure should be a good estimate of the preference-related behavior during the periods when
other is observed. Thus subtracting this average imitation rate from the actual imitation during
the periods of observation should give us a good measure of pure imitation.8

Figure 2: The adjusted rate of imitation in Novis-HighRaven, Vis-HighRaven,

![Bar chart showing the adjusted rate of imitation in Novis and Vis treatments by terciles of Raven score.](chart.png)

Figure 2 shows the adjusted average rate of imitation in Novis-HighRaven, Vis-HighRaven,

---

8One might think that 0.5 is a good estimate of the average imitation rate in periods when the other is not
observed. This would be the case if both participant and the other chose randomly. However, since both are in the
same environment and both try to maximize their earnings the average imitation rate should be strictly above 0.5.
Indeed, the average imitation rate when the other is not observed is 0.63.
Novis-LowRaven and Vis-LowRaven treatments divided by the Raven score of the participants.\footnote{These results, by construction, represent the actual imitation rate in the four treatments. The interpretations of this figure do not change if we consider the not adjusted rate of imitation. See Figure 6 in Appendix B.}

To shed some light on how different information affects the imitation rate of high and low Raven participants, we run a regression reported in Table 7 in Appendix B. The dependent variable is the adjusted rate of imitation (\( \text{adjim} \)). Independent variables are indicators for treatments, terciles of Raven score and all interactions.\footnote{In order to better illustrate the link between reasoning ability and imitation we divide our pool of participants into terciles of Raven score. In the future analyses, for the sake of simplicity, we divide participants into below and above median Raven score.}

The results support the following observations: high Raven observed participant is imitated more in both Vis and Novis treatments. It means that participants in Novis treatments imitate high Raven other more than low Raven other even though they do not know the Raven score of the observed participant. Regression analysis shows that for the third tercile (high Raven score participants, Novis treatments) the difference in imitation rate is significantly higher than in first tercile (low Raven participants; sum of coefficients \( \text{obshigh} + \text{obshigh} \times \text{ravterc3}, 0.162, p = 0.004 \)).\footnote{The definitions of all variables used in the analyses can be found in Appendix A.}

This illustrates that high Raven participants in Novis treatments imitate high Raven other more than low Raven participants do.

When participants with low Raven score (first tercile) observe the action of a low Raven observed participant without knowing her Raven score (Novis-LowRaven) they tend to imitate her (first red bar from the left). By contrast, when they observe the action of a low Raven observed participant \textit{k}nowing her Raven score (Vis-LowRaven) they \textit{dis}imitate her. This difference is significant (\(-0.82, p = 0.032\)).\footnote{The difference between the two treatments is represented by the coefficient on \textit{vis} in OLS regression in Table 7.}

Similarly, when participants with low Raven score observe the action of a high Raven in Vis treatment, they imitate her significantly more (0.0975, \( p = 0.032 \)) comparing to Novis-HighRaven treatment.\footnote{The difference is represented by \( \text{vis} + \text{vis} \times \text{obshigh} \) in Table 7.}

Conversely, participants with high Raven score (third tercile) do not significantly change their rate of imitation upon observing high or low Raven observed participant in Vis and Novis treatments (0.0465, \( p = 0.406 \) for low other; \(-0.0314, p = 0.419 \) for high other). However, they still imitate the high Raven observed participant significantly more in both Novis (0.162, \( p = 0.004 \)) and Vis treatments (0.084, \( p = 0.033 \)).\footnote{The difference between high and low Raven other in Novis treatments is presented by \( \text{obshigh} + \text{ravterc3} \times \text{obshigh} \) in Table 7. In Vis treatment is presented by \( \text{obshigh} + \text{obshigh} \times \text{vis} + \text{ravterc3} \times \text{obshigh} + \text{ravterc3} \times \text{obshigh} \times \text{vis} \).}

To summarize, participants with low Raven score do strongly react to the information about the Raven score of the observed participant provided to them: the imitation rate drops significantly when they know that the observed participant has low Raven score and increases significantly when they know that the Raven score of the observed participant is high. On the contrary, participants with high Raven score do not seem to react much to the information about the Raven score of the other. This suggests that low and high Raven score participants use different mecha-
nisms of imitation. Finally, it should be noted that the same results can be obtained if we divide participants according to the CRT test instead of Raven. Figure 7 in Appendix B shows adjusted imitation of subjects by Vis and Novis treatments divided into two groups: below median and above median score on the CRT test. The conclusions are unchanged. This finding shows that our results are robust to the different measures of cognitive ability.

Now we look at the dynamics of the adjusted imitation rate. First, we consider 11 periods moving average of imitation across periods when the observed participant’s action is visible. Figure 3 illustrates the moving averages for low Raven and high Raven participants when they observe high Raven others in Vis and Novis treatments.

![Figure 3: The dynamics of the Adjusted imitation rate by low and high Raven participants (only for 100 periods when the action of the other is observed). Only graphs for high Raven others are shown. Ranges are ±1 SE.](image)

In Novis treatment with high Raven observed participant we see no particular difference in how high and low Raven participants imitate. However, when they have a “prior,” that observed participant has high Raven score, they exhibit different behavior. Low Raven participants react to the prior by significantly increasing their rate of imitation throughout the experiment (at least until 85 periods of observing the other). High Raven participants are affected by the prior only in the first 25 periods, and then they exhibit the same imitation rate as in nonvisible condition.\(^\text{15}\) We interpret this difference as follows. High Raven participants start with following the observed participant, as if they are testing whether and how much the behavior of the other was optimal. Then, they converge to imitating in the same way as in the Nonvis condition, as if they adapt the rate of imitation to the feedback they receive. On the other hand, low Raven participants keep blindly follow the observed participant almost until the end of the experiment. This is consistent with our findings shown in Figure 2.

For the low Raven observed participant we also confirm the results obtained in Figure 2. High Raven participants do not change their rate of imitation in the two conditions (Vis and Nonvis). Low Raven participants exhibit a drop in the level of imitation when they know that the other

\(^{15}\text{The effect of the prior was found also in different contexts, see, for example, Fouragnan et al. (2013).}\)
has low Raven score, though this effect is not that pronounced (see Figure 9 in Appendix B).

So far we found evidence showing that the imitation rate of high and low Raven participants in the four treatments differ depending on their intelligence and the information they received about the observed participant. Next we want to investigate in more detail what information about the choices of the observed participant is used by high and low Raven participants and what effect the revelation of the Raven score of the other has on the usage of different types of imitation. To do this we look at participants’ actual choices in each period and run a panel logit with two dimensions (participants and time) and four independent variables ($v_{im}$, $sw_{ioth}$, $vow$ and $absop$). The first three independent variables reflect possible effects that the choices of the observed participant and obtained payoffs in previous 10 periods might have on the imitation rate.\footnote{The choice of 10 periods is not random. We investigated another model with the same variables, which take into account 5 periods in the past. We compared the models using Akaike and Bayesian information criteria. Models with 5 and 10 periods have practically exactly the same IC’s. We decided to choose the model with 10 periods, since it accounts for more periods in the past.}

The first variable ($v_{im}$) tracks the average payoff obtained after choosing the same action as the observed participant last 10 times.\footnote{It should be mentioned that the last 10 times when observed participant was imitated might go beyond period $t − 10$. We decided to use last 10 times the observed participant was imitated instead of last 10 periods because in the latter case it often happens that no imitation takes place, in which case we have a missing observation.} This represents what we call simple effect on imitation: participants might imitate more if they got good payoffs from choosing the same action as the other in the past. We expect that both high and low Raven participants imitate more if they increase their earnings after imitation.

The second variable ($sw_{ioth}$) is the number of switches between actions that the observed participant made in the last 10 periods when the action of the observed participant was shown. This is constancy effect on imitation: participants might infer that switching between actions implies observed participant’s getting low payoffs and, thus, imitate less. We expect that high Raven participants do not imitate when the observed participant is switching and do imitate when the observed participant chooses the same action repeatedly.

The third variable ($vow$) is the participant’s average payoff in the last 10 periods. This is a second type of sophisticated imitation because it requires the participant to balance their level of imitation with their ability to be optimal in their interaction with the environment. This variable represents the attention effect: if participants get high payoff they might choose to not pay attention to the choice of the other just because it becomes irrelevant. We expect that only high Raven participants imitate when the previous payoffs are low but do not imitate when previous payoffs are high.

The last variable, $absop$, is the absolute value of the difference between the averages of the last 5 payoffs from both actions. This variable controls for the artificial effect of what might look like imitation, since both the participant and the other choose in the same environment (the probabilities of getting the reward are the same). The high value of $absop$ indicates an
environment in which one action has very high probability of reward and another very low probability. In these circumstances, it is conceivable that both the participant and the other will choose the same action just because it is clear that one is better than the other. Therefore, absop controls for the situations where there is no imitation, but just common cause for choosing the same action.

<table>
<thead>
<tr>
<th>Participant: Low Raven</th>
<th>High Raven</th>
<th>NOT VISIBLE</th>
<th>VISIBLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other: Low Raven</td>
<td>High Raven</td>
<td>b/se</td>
<td>b/se</td>
</tr>
<tr>
<td>vim</td>
<td>0.967**</td>
<td>1.502**</td>
<td>1.960**</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.262)</td>
<td>(0.372)</td>
</tr>
<tr>
<td>swioth</td>
<td>−0.526</td>
<td>−0.715*</td>
<td>−0.013</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.333)</td>
<td>(0.526)</td>
</tr>
<tr>
<td>vow</td>
<td>0.190</td>
<td>−1.033**</td>
<td>−0.185</td>
</tr>
<tr>
<td></td>
<td>(0.252)</td>
<td>(0.246)</td>
<td>(0.340)</td>
</tr>
<tr>
<td>absop</td>
<td>−0.031</td>
<td>0.611**</td>
<td>0.808*</td>
</tr>
<tr>
<td></td>
<td>(0.169)</td>
<td>(0.169)</td>
<td>(0.220)</td>
</tr>
<tr>
<td>cons</td>
<td>0.319</td>
<td>0.493*</td>
<td>−0.055</td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.215)</td>
<td>(0.321)</td>
</tr>
<tr>
<td>N</td>
<td>3367</td>
<td>3366</td>
<td>2451</td>
</tr>
<tr>
<td>Indep. N</td>
<td>37</td>
<td>37</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 2: Random effects logit regression of imitation choices. * − p < 0.05; ** − p < 0.01.

Resulting regressions are reported in Table 2.¹⁸ We start from noticing that the signs of the coefficients on all significant variables are consistent with our hypotheses presented above. Increase in switchiness of the observed participant decreases imitation. Increase in payoff received after imitating increases imitation and increase in own average payoff in the past decreases imitation.

First, we would like to notice that the coefficients on the average payoff from imitation (vim) are very significant in all treatments except the case when above median Raven score participants know that they are shown the actions of the low Raven score observed participant (we discuss this case below). This is consistent with our hypothesis that for both high and low Raven participants the average payoff from imitation should have an effect on their choices (simple effect).

Next we look at the number of switches (swioth). Notice that in the Novis treatments the imitation choices of high Raven participants are significantly affected by the number of switches of the observed participant (constancy effect). We interpret this as an attempt of high Raven participants to find out whether and when it is advantageous to imitate the observed participant.

¹⁸In the table Low/High Raven participants are those below/above the median on Raven score.
Conversely, the imitation of the low Raven participants is not affected by the constancy effect. They do not seem to try to understand whether and when the observed participants is earning money, but just react to the average payoff they obtain from imitation.

Attention effect on imitation, or the average own payoff in the past ($v_{ow}$), is present in Novis treatments for high Raven participants only. This is in accordance with the hypothesis that only high Raven participants balance their level of imitation depending on their ability to choose optimally.

In Novis treatments the imitation of the high Raven participants is influenced by all three effects. For low Raven participants only simple effect ($v_{im}$) is present. This supports our hypothesis that high Raven participants choose to imitate (or not) in a more sophisticated way.

The next question is what happens when the information about the Raven score of the observed participant gets revealed. There are two cases possible: high Raven participants might still try to understand the rationale behind the choices of the other and imitate in the same way as in Novis treatments or they might take the information about the Raven score of the other as a signal of the ability to perform in the task and, thus, use only simple imitation. We see that the coefficients on $sw_{ioth}$ and $v_{ow}$ are insignificant (two rightmost columns in Table 2), which supports the second hypothesis. Moreover, the fact that the imitation of high Raven participants is no more influenced by constancy and attention effects supports the hypothesis that sophisticated imitation is used only when participants do not know whether to imitate the observed participant is optimal or not. Low Raven participants do not use sophisticated imitation in Vis treatments in the same way as they did not use it in Novis treatments (except significant coefficient on $v_{ow}$ when low Raven participants observe low Raven other, which can be interpreted as disimitation that was described above in Figure 2).

Overall, we observe the most difference in imitation between high and low Raven participants in Novis treatments. High Raven participants use much more information to choose whether and when to imitate than low Raven participants. In Vis treatments the situation is different. The data suggest that the information about the Raven score of the observed participant has an effect on the sophistication of imitation: with visible information imitation becomes simple. It is also worth noticing that high Raven participants, when knowing that the observed participant is low Raven, completely disregard his choices and do not even use average imitation payoff. This also supports an idea that Raven score is treated as informative of the ability. So participants do not need to understand whether and when it is optimal to follow the other when they know other’s Raven score.

### 4.3 Optimality and Earnings

Now we turn to the analysis of optimality of choices (choosing the action with the highest probability of reward) and the resulting earnings that participants make. We are mostly interested
in how the imitation strategies of different participants, as discussed in Section 4.2, are reflected in participants’ overall optimality and earnings. We also analyze the effects of visibility of the Raven score of the other and the effect of participants’ own Raven score on optimality and earnings.

For our analysis, we introduce new variables in the panel data set: $i_{sop}t$ is 1 if participant chose an optimal action and 0 otherwise; $i_{siv}l_{ow}$ is 1 for treatment Vis-LowRaven and 0 otherwise; $i_{siv}l_{igh}$ is 1 for treatment Vis-HighRaven; and $r_{av}l_{igh}$ is 1 if a participant is above median in a Raven task.

![Figure 4: Distribution of the average optimality rates for each participant (all data).](image)

Figure 4 shows the distribution of averages of $i_{sop}t$ for all participants in both treatments. The lowest optimality rate is 0.45 and the highest optimality rate is 0.87. It is important to notice that the optimality rates of the two observed participants are 0.83 for the high Raven score of 28 and 0.64 for the low Raven score of 15. These values lie in the tales of the distribution of Figure 4 which implies that imitating high Raven other should be mostly beneficial and imitating low Raven other should be harmful earnings wise.

Next we look at the logit regression of $i_{sop}t$ on the above/below median Raven score of the participant and its interactions with the treatments. We consider two regressions: one where only periods where observed participant was visible are taken into account and another where all periods are considered (Table 3).

The first observation is that participants with above median Raven score choose more optimally overall (Table 3, left column). The second observation is that below median Raven participants (baseline) significantly increase their optimality when observing high Raven other. As was mentioned above, this is the result of two effects: 1) low Raven participants increase imitation when observing the Raven score of the high Raven other and 2) the high Raven other has high

---

19Optimal action is the one which gives the reward with higher probability. Participants do not observe the probabilities of actions.
When we look at the high Raven participants there are no effects of either Vis-lowRaven or Vis-HighRaven treatments on optimality. This might be because high Raven participants already enjoy high optimality in Novis treatments, so observations of the Raven score of the high Raven observed participant does not significantly change their optimality. When high Raven participants observe low Raven other, we know from Section 4.2 that they pay little attention to the choices of this other, so their optimality does not decrease.

Finally, we look at the earnings of participants in the regression in Table 4. The variable feedback is equal to 1 for the periods in which participant earned a reward and 0 otherwise. We see that high Raven participants earn more than low Raven participants. This is not surprising since, as we just learned, high Raven participants are more optimal than low Raven participants. In addition, as expected, the increase in optimality of low Raven participants when they observe high Raven other leads to the significant increase of earnings. We see no significant effect of visibility of either low or high Raven score of the other on the high Raven participants. The reason for this is that high Raven participants already earn more than low Raven participants and the effect of imitation on their earnings is not significant.

Now we try to understand the mechanism behind low Raven participants being less optimal than high Raven participants. First, we check the connection between the number of switches and Raven score. One hypothesis is that low Raven score participants behave in “noisier” way

\[ isopt \]
\[ \text{Imitation Periods} \quad \text{All Periods} \]
\[ b/se \quad b/se \]
\[ \text{ravhigh} \quad 0.204^* \quad 0.161 \]
\[ (0.094) \quad (0.085) \]
\[ \text{vishigh} \quad 0.341^{**} \quad 0.214^* \]
\[ (0.103) \quad (0.093) \]
\[ \text{vislow} \quad 0.023 \quad 0.110 \]
\[ (0.115) \quad (0.105) \]
\[ \text{ravhigh} \times \text{vishigh} \quad -0.257 \quad -0.133 \]
\[ (0.158) \quad (0.143) \]
\[ \text{ravhigh} \times \text{vislow} \quad -0.124 \quad -0.133 \]
\[ (0.157) \quad (0.143) \]
\[ \text{cons} \quad 0.711^{**} \quad 0.711^{**} \]
\[ (0.066) \quad (0.060) \]

\[ N \quad 32000 \quad 32000 \]
\[ \text{Indep. } N \quad 160 \quad 160 \]

Table 3: Random effects logit regression of optimal choices. * – \( p < 0.05 \); ** – \( p < 0.01 \).

Optimality rate of 0.83.

Table 8 in Appendix B shows the same regressions only redone with the linear probability model (random effects OLS). Qualitatively, linear model shows significance for the exactly same variables. This demonstrates the robustness of our analysis.

Table 9 in Appendix B shows the same analysis with linear probability model. All coefficients are significant as in the logit model. This demonstrates that our analysis is robust.

---

\[ ^{20} \text{Table 8 in Appendix B shows the same regressions only redone with the linear probability model (random effects OLS). Qualitatively, linear model shows significance for the exactly same variables. This demonstrates the robustness of our analysis.} \]

\[ ^{21} \text{Table 9 in Appendix B shows the same analysis with linear probability model. All coefficients are significant as in the logit model. This demonstrates that our analysis is robust.} \]
Table 4: Random effects logit regression of earnings. * – $p < 0.05$; ** – $p < 0.01$.

than high Raven participants. To test this we construct a variable $swi$, which, for each participant, is equal to the number of switches. Table 5 (Column 1) shows the results.

Table 5: OLS regressions of switchiness, distance from optimal policy and amount earned by treatment. Errors are robust. * – $p < 0.05$; ** – $p < 0.01$.

The regression shows that low Raven score participants indeed switch more. The effect is rather large. According to the regression the participant with less correct Raven answers switches around 71 times, whereas the participant who scores perfectly on all 30 Raven questions switches 39 times (variable $raven$ is normalized to $[0, 1]$). Both numbers are higher than the optimal amount of switches (9 switches is optimal), however, high Raven participants are much closer to optimality.

In order to test whether high Raven participants actually choose in more optimal way than low Raven participants we construct the optimal policy using the actual probabilities of getting
rewards from two actions in each period and compare the actual choices that participants made to the optimal policy. We calculate the distance between optimal policy and the participants’ choices by penalizing by 1 point each choice which was not optimal. Table 5 (Column 2) shows the result. One can see that low switchiness brings participants closer to the optimal policy. Thus, higher Raven score should imply more optimal choice.\textsuperscript{22} To understand how much the value of distance changes across participants, Figure 8 in Appendix B provides a histogram. Next we notice the connection between switchiness and earnings, which should not be surprising at all since switchiness influences optimality. Column 3 of Table 5 shows that earnings of participants who switch a lot are lower (the coefficient on \( w_1 \) looks small, but \( w_1 \) ranges in \([1, 107]\), thus the maximal effect is around \( \$2.4 \)).

These three observations give us the following picture. Higher Raven score decreases the number of switches that participant makes, which in turn increases the optimality of choices, which leads to increases in earnings. To finish this analysis we look at Columns 4 and 5 of Table 5. From what we just concluded, we should have Raven score being positively correlated with earnings. This is indeed true for the whole sample.\textsuperscript{23} However, if we look separately at Novis and Vis treatments, we see that this significance comes from Novis treatments. How can we explain this? We think that this difference goes to the heart of all our analysis before. In our opinion, the difference comes from the fundamentally different ways that high Raven score participants and low Raven score participants learn in Vis and Novis treatments. High Raven score participants are evidently able to recognize the type of the observed participant (high performing or low performing) when they do not know the observed participant’s Raven score. Given this recognition, they manage to use this information in order to increase their earnings. Low Raven score participants fail to do that. Thus, this leads to a significant difference in earnings between high and low Raven score participants in Novis treatments. In Vis treatments this effect is not present: low Raven score participants now know who is who and consecutively change their imitation choices, which leads to the increase of their earnings, thus closing the gap with the high Raven score participants.

\textbf{4.4 Information Transfer}

Finally, we analyse what information gets transferred from the observed participants to the observers and whether our participants do learn correctly from observations. We look at the choices of the 51 participants from Experiment 1. As above, we consider 3 variables: Raven score, money earned and number of switches. First, number of switches is correlated with Raven score (Spearman’s \( \rho = -0.3350, p = 0.0163; \) OLS coefficient \( \beta = -48.44^*, \) intercept 82.81\textsuperscript{**}). Second, earnings are (weakly) correlated with the number of switches (Spearman’s \( \rho = -0.3294, p = 0.0182; \)

\textsuperscript{22}Indeed, the regression of distance on raven gives significant coefficient of \( \beta = -21.47^* \) with intercept 76.70\textsuperscript{**} (\( ^* - p < 0.05; \) \( ** - p < 0.01 \)).

\textsuperscript{23}For the whole sample the coefficient on raven is \( \beta = 1.137^* \) with intercept 11.30\textsuperscript{**} (\( ^* - p < 0.05; \) \( ** - p < 0.01 \)).
OLS coefficient is not significant). We would like to connect these findings with the imitation behavior observed in Experiment 2 and interpret our results in terms of the information about the Raven score of the observed others being inferred from their choices. Indeed, high Raven participants increase imitation when they observe less switches of the other. Given that lower number of switches is correlated with higher Raven score, we can conclude that high Raven participants imitate the other the more, the higher Raven score they have. In this sense, the information about the Raven score of the observed participant is transferred by observing the number of switches. This is also efficient, because lower number of switches is associated with higher earnings.

Lastly, we look at the effect of observation of other on optimality. Figure 5 shows the moving average of optimality across all 200 periods. We can see that the optimality of the participants in Vis-HighRaven treatment is almost always higher than that of the participants in Experiment 1 who do not observe anyone.

![Figure 5: The 11 points moving average of optimality of choices in Experiment 1 and Experiment 2, Vis-HighRaven treatment.](image)

In fact we can say more. Optimality of choices is the lowest in Experiment 1, followed by higher in Novis-LowRaven, yet higher in Novis-HighRaven, higher in Vis-LowRaven and still higher in Vis-HighRaven. Pairwise sign rank tests show significant $p < 0.0001$ for all comparisons, except Experiment 1 and Novis-LowRaven where $p = 0.0121$.

These findings suggest that just observing the other, regardless of the Raven score of the observer, improves the optimality of choice. The reason for this might be that, independently of the tendency to imitate, observation makes people understand the environment better and thus improve their performance.

## 5 Conclusion

We study observational learning in a 2-armed bandit task in which the probability of getting a reward from each hand changes over time. Participants in the experiment sometimes can observe the actions of one other person who chooses in the same environment. Out of two possible others, one has low and another high Raven score. In some treatments, participants know the Raven score of the other and in other treatments they do not. We also elicit Raven
scores of the participants in order to investigate the relationship between the intelligence of the observer and observational learning strategies.

We find that low Raven score participants use simple learning strategy: they imitate the other when the recent payoffs they received after imitation are high. High Raven score participants, on the other hand, use sophisticated learning strategy: in addition to simple learning their level of imitation is flexibly modulated by 1) the stability of the choices of the other (the number of switches between hands the other made in recent periods decreases imitation) and 2) their current performance (lower average payoff in recent periods increases imitation).

We also find an effect of the information about the Raven score of the other on the imitation strategy. When Raven score of the other is not available, high Raven participants use both simple and sophisticated learning strategies, whereas low Raven participants use only simple learning strategy. When information about the Raven score of the other is provided, both high and low Raven participants use only simple imitation. Thus, the information about the Raven score of the other is used as a signal of the ability of the other to perform in the task.

High Raven score participants are on average more optimal and earn more money than low Raven score participants. Optimality is affected by the information about the Raven score of the other. In particular, low Raven score participants choose more optimally when they know that the other has high Raven score. High Raven score participants are not affected by this information. This suggests that low and high Raven scores are associated with different observational learning strategies.

Finally, we find that participants, who use sophisticated learning strategies, do learn in optimal way. They correctly take into account the information about the choices of the other when the other does not switch too often, which is correlated with higher optimality and higher earnings.

Intelligence plays an important role in learning from observation. We find that high intelligence participants improve performance in a non-stationary uncertain environment by observing the actions of others regardless of the information about others’ intelligence. At the same time, low intelligence participants are strongly influenced by the information about the intelligence of the observed individual. Thus, intelligence is associated with the ability to uncover the characteristics of the observed other, being objective about prior information and, therefore, learning optimally from observation.
References


Proto, E., Rustichini, A. and Sofianos, A. (2014). Higher intelligence groups have higher cooperation rates in the repeated prisoners dilemma, mimeo, University of Warwick and University of Minnesota.


Appendix

A Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>im</td>
<td>0/1</td>
<td>0 if participant chose different action than the other in a given period, 1 if he chose the same action</td>
</tr>
<tr>
<td>adjim</td>
<td>[0, 1]</td>
<td>for each participant, the average rate of imitation in periods when the other is observed minus the average rate of imitation when the other is not observed</td>
</tr>
<tr>
<td>swi</td>
<td>[0, 1]</td>
<td>for each period $t$, the average number of switches of the observed participant in the last 10 periods (when observed)</td>
</tr>
<tr>
<td>vim</td>
<td>[0, 1]</td>
<td>for each period $t$, the average payoff from choosing the same action as the observed participant for the last 10 times in the past</td>
</tr>
<tr>
<td>vow</td>
<td>[0, 1]</td>
<td>for each period $t$, the average payoff from the previous 10 periods</td>
</tr>
<tr>
<td>absop</td>
<td>[0, 1]</td>
<td>for each period, the absolute value of the difference of the average payoffs obtained from the last 5 times choosing each action in the past</td>
</tr>
<tr>
<td>isopt</td>
<td>0/1</td>
<td>for each period, is 1 if the choice was optimal (the action with highest actual probability of reward was chosen)</td>
</tr>
<tr>
<td>vislow</td>
<td>0/1</td>
<td>for each period, is 1 if the treatment is Vis-LowRaven</td>
</tr>
<tr>
<td>vishigh</td>
<td>0/1</td>
<td>for each period, is 1 if the treatment is Vis-HighRaven</td>
</tr>
<tr>
<td>rinvalid</td>
<td>0/1</td>
<td>for each period, is 1 if participant’s Raven score is above median</td>
</tr>
<tr>
<td>feedback</td>
<td>0/1</td>
<td>for each period, is 1 if reward was received (10 cents)</td>
</tr>
<tr>
<td>vis</td>
<td>0/1</td>
<td>for each participant, is 1 if participant is in VisHigh or VisLow conditions</td>
</tr>
<tr>
<td>obshigh</td>
<td>0/1</td>
<td>for each participant, is 1 if participant is observing high Raven other</td>
</tr>
<tr>
<td>ravterc2</td>
<td>0/1</td>
<td>for each participant, is 1 if she is in the second tercile according to Raven score</td>
</tr>
<tr>
<td>ravterc3</td>
<td>0/1</td>
<td>for each participant, is 1 if she is in the third tercile according to Raven score</td>
</tr>
<tr>
<td>swi</td>
<td>[0, 200]</td>
<td>number of switches from one action to another that a participant made</td>
</tr>
<tr>
<td>distance</td>
<td>[0, 200]</td>
<td>the number of periods in which a participant did not choose optimal action equals raven score divided by 30. The value of Raven score normalized to [0, 1]</td>
</tr>
<tr>
<td>earned</td>
<td>[0, 1]</td>
<td>the amount earned by participants in the bandit game. Normalized to [0, 1] from [9.1, 14.7]</td>
</tr>
</tbody>
</table>

Table 6: Variables used in analyses and regressions.
B Additional Analyses

Figure 6: The rate of imitation in Vis and Novis treatments by the terciles of the Raven score (q1 - the lowest Raven score; q3 - the highest). Blue bars represent the high Raven other, the red bars - the low Raven other. Spikes are ±1 SE.

Figure 7: The rate of imitation in Vis and Novis treatments by the median of the CRT score. Blue bars represent the high Raven other, the red bars - the low Raven other. Spikes are ±1 SE.
<table>
<thead>
<tr>
<th></th>
<th>b/se</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>vis</strong></td>
<td>-0.082* (0.038)</td>
</tr>
<tr>
<td><strong>obshigh</strong></td>
<td>0.075 (0.044)</td>
</tr>
<tr>
<td><strong>vis × obshigh</strong></td>
<td>0.179** (0.059)</td>
</tr>
<tr>
<td><strong>ravterc2</strong></td>
<td>-0.012 (0.033)</td>
</tr>
<tr>
<td><strong>ravterc3</strong></td>
<td>-0.031 (0.052)</td>
</tr>
<tr>
<td><strong>vis × ravterc2</strong></td>
<td>0.073 (0.054)</td>
</tr>
<tr>
<td><strong>vis × ravterc3</strong></td>
<td>0.128 (0.067)</td>
</tr>
<tr>
<td><strong>obshigh × ravterc2</strong></td>
<td>0.034 (0.057)</td>
</tr>
<tr>
<td><strong>obshigh × ravterc3</strong></td>
<td>0.087 (0.071)</td>
</tr>
<tr>
<td><strong>vis × obshigh × ravterc2</strong></td>
<td>-0.074 (0.083)</td>
</tr>
<tr>
<td><strong>vis × obshigh × ravterc3</strong></td>
<td>-0.257** (0.090)</td>
</tr>
<tr>
<td><strong>cons</strong></td>
<td>0.025 (0.023)</td>
</tr>
</tbody>
</table>

| N                         | 160     |

Table 7: OLS regression of adjusted imitation. Errors are robust. * – $p < 0.05$; ** – $p < 0.01$. 
<table>
<thead>
<tr>
<th>isopt</th>
<th>Imitation Periods</th>
<th>All Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b/se</td>
<td>b/se</td>
</tr>
<tr>
<td>ravhigh</td>
<td>0.043* (0.020)</td>
<td>0.035* (0.014)</td>
</tr>
<tr>
<td>vishigh</td>
<td>0.067** (0.021)</td>
<td>0.042** (0.013)</td>
</tr>
<tr>
<td>vislow</td>
<td>0.007 (0.024)</td>
<td>0.025 (0.013)</td>
</tr>
<tr>
<td>ravhigh × vishigh</td>
<td>−0.050 (0.033)</td>
<td>−0.027 (0.025)</td>
</tr>
<tr>
<td>ravhigh × vislow</td>
<td>−0.026 (0.033)</td>
<td>−0.030 (0.025)</td>
</tr>
<tr>
<td>cons</td>
<td>0.666** (0.014)</td>
<td>0.666** (0.008)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>32000</th>
<th>32000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indep. N</td>
<td>160</td>
<td>160</td>
</tr>
</tbody>
</table>

Table 8: Random effects OLS regression of optimality. Errors are clustered by session. * – $p < 0.05$; ** – $p < 0.01$.

<table>
<thead>
<tr>
<th>feedback</th>
<th>b/se</th>
</tr>
</thead>
<tbody>
<tr>
<td>ravhigh</td>
<td>0.025** (0.005)</td>
</tr>
<tr>
<td>vislow</td>
<td>0.020* (0.009)</td>
</tr>
<tr>
<td>vishigh</td>
<td>0.042** (0.009)</td>
</tr>
<tr>
<td>ravhigh × vislow</td>
<td>−0.017* (0.008)</td>
</tr>
<tr>
<td>ravhigh × vishigh</td>
<td>−0.038** (0.012)</td>
</tr>
<tr>
<td>cons</td>
<td>0.582** (0.007)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>32000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indep. N</td>
<td>160</td>
</tr>
</tbody>
</table>

Table 9: Random effects OLS regression of earnings. Errors are robust. * – $p < 0.05$; ** – $p < 0.01$. 
Figure 8: The distribution of the values of the variable distance, the distance of participants’ choices from optimal policy.

Figure 9: Graphs of imitation rate by low and high Raven participants. Only graphs for low Raven observed participants are shown. Ranges are ±1 SE.
C Experimental Instructions

The following is a translation of the original instructions in Italian. The experimenter read the instructions aloud to the participants while they followed along their own copy. Original instructions are available upon request. We decided to leave the instructions as is, which means that the numbering of figures in the instructions goes separately from the main text. In what follows, the numbers of figures refer only to figures in the instructions.

INSTRUCTIONS

Dear student you are about to participate in an experiment on decision making. Your privacy is guaranteed: results will be used and published anonymously. Your earnings will depend on your performance in the experiment, according to the rules which we will explain to you shortly. You will be paid privately at the end of the experimental session. Other participants will not be informed about your earnings. The maximum amount you can earn in the experiment is €35.85 and the minimum is €3.10.

The experiment is divided in 3 parts. Each part of the experiment will be described in detail below.

Part One

COMPLETION OF MATRICES PROBLEMS

In this task you will be asked to solve 30 different test items. In each test item we ask you to identify the missing element that completes a pattern of shapes. The patterns are presented in the form of a 3x3 matrix. An example of a test item is shown here.

The shapes included in the matrix follow some regular patterns, for example, in the first column there is a circle, a rhombus and a square. In the second column there is a rhombus, a square and a circle. In the third column there is a square and a circle. Therefore, the symbol that completes the sequence in the third column is a rhombus.
You can identify regular patterns also by observing the shapes by row. For example in the first row each shape includes a dashed line. In the second row each shape includes two dashed lines, and in the third row each shape includes three dashed lines. Therefore, the symbol that completes the sequence in the third row includes three dashed lines.

Regular patterns can be identified also by observing the shapes by diagonal. For example, in the first diagonal the dashed lines move from top-left to bottom-right, in the second diagonal the dashed lines are vertical, in the third diagonal the dashed lines move from bottom-left to top-right, in the fourth diagonal the dashed lines move again from top-left to bottom right, and in the fifth diagonal the dashed lines are again vertical. Therefore, the symbol that completes the sequence in the third diagonal includes lines that move from bottom-left to top-right.

According to the patterns described above the shape that completes the matrix is number 5. A rhombus with three dashed lines that move from bottom-left to top-right.

You will have 20 minutes to complete the largest possible number of items and you will earn 30 cents for each correct answer. If you do not complete an item or your answer is incorrect you will earn 0 cents for that item.

In this part of the experiment, you can earn between €0 and €9.

**Part Two**

**QUESTIONNAIRES**

See Appendix E.

**Part Three**

**CHOICE TASK**

In this part of the experiment you will face 200 trials divided in four blocks of 50 trials each. On each trial you will be asked to choose between two symbols. The two symbols are like two slot machines that give you a reward of 10 cents with a certain—unknown—probability (0 otherwise). You do not have to pay before to choose your symbol and your goal will be to try to choose the symbol that gives you the 10 cents with the highest probability. However, you do not know the probability associated with the two symbols of getting the 10 cents, and you will need to figure out what is the most convenient symbol to play following the feedback you will receive from time to time. In fact, after choosing one of the two symbols you will be notified about the outcome of your bet (if you have won 10 cents or 0 cents).

The experiment proceeds as follows:

When a red silhouette appears on the screen (Figure 1), it means that you have to choose one of the two symbols. After you press the answer button your choice will appear under the selected symbol (Figure 2) and you will receive a feedback about the outcome of your choice (Figure 3 appears if you have won 0 cents and figure 4 if you have won 10 cents).
Probability of winning associated with the two symbols
At the beginning of the experiment a certain probability of winning will be associated with each of the two symbols. The probability of winning will change slowly in the course of the experiment (Figure 5). For example, if at the beginning of the experiment the probability of receiving the 10 cents is equal to 0.8 (you receive the reward 8 times out of 10), this probability may increase or decrease slightly in the subsequent trial and so on until the end of the experiment. It is important to keep in mind that the probabilities of winning associated with the two symbols are independent, therefore if the probability of winning associated with one symbol increases from one trial to the next, the probability of winning associated with the other symbol does not necessarily decrease.

Observational phase
In addition to the feedback about the outcome of your choice, you will receive additional information during the experiment. Before making your choice, you often have the chance to observe the choice made (in the same trial) by a participant carrying out this same experimental task. The only difference between you and this participant is that he/she could not observe the actions of another participant. This participant will be represented with a green silhouette (Figure 6) and from time to time, before making your choice, you will see the choice made by this participant in the trial you are about to play (Figure 7). After you have observed the choice of this participant,
Timeline of the experimental phases
At the beginning of the experiment you will not see the choices made by the other participant. Then, in different time phases, and always before making your choice, you will see the choice made by the other participant. The other participant had the same goal as you, which is to choose the option that would guarantee the highest probability of winning.

In Figure 8 you can see a diagram that summarizes the experimental structure.

Random selection of the observed participant (Non-visible condition)
The participant you will observe has completed (like you) all the parts of the experiment, including the questionnaire on matrices problems that you completed at the beginning of the experimental session. The chart below (Figure 9) shows the scores obtained by all participants who participated in this study in the previous sessions (December 2015). The participant that you will observe was selected among these 51 participants.
Random selection of the observed participant (Visible low condition)
The participant you will observe has completed (like you) all the parts of the experiment, including the questionnaire on matrices problems that you completed at the beginning of the experimental session. The chart below (Figure 9) shows the performance achieved by all participants who participated in this study in the previous sessions (December 2015). The score achieved by the participant that you will observe is colored in red.

Random selection of the observed participant (Visible high condition)
The participant you will observe has completed (like you) all the parts of the experiment, including the questionnaire on matrices problems that you completed at the beginning of the experimental session. The chart below (Figure 9) shows the performance achieved by all participants who participated in this study in the previous sessions (December 2015). The score achieved by the participant that you will observe is colored in red.
D Holt and Laury Lotteries

The following figure shows how the Holt and Laury task was presented to the participants. The payoffs were identical to the payoffs in Holt and Laury (2002). The participants did not know how much they earned until the end of the experiment.
E Empathy Quotient

The Empathy Quotient consists of 60 questions aimed at assessing autistic traits. It is designed for healthy adults (Baron-Cohen and Wheelwright, 2004). The questions can be found here: http://docs.autismresearchcentre.com/tests/EQ.pdf.

F Details of the Experiment

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 11, 2015</td>
<td>21</td>
</tr>
<tr>
<td>Dec 14, 2015</td>
<td>16</td>
</tr>
<tr>
<td>Dec 18, 2015</td>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novis</td>
<td></td>
</tr>
<tr>
<td>Jan 22, 2016</td>
<td>22</td>
</tr>
<tr>
<td>Jan 22, 2016</td>
<td>20</td>
</tr>
<tr>
<td>Feb 16, 2016</td>
<td>20</td>
</tr>
<tr>
<td>Feb 16, 2016</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis</td>
<td></td>
</tr>
<tr>
<td>Jan 29, 2016</td>
<td>20</td>
</tr>
<tr>
<td>Jan 29, 2016</td>
<td>12</td>
</tr>
<tr>
<td>Feb 04, 2016</td>
<td>18</td>
</tr>
<tr>
<td>Feb 04, 2016</td>
<td>20</td>
</tr>
<tr>
<td>Feb 17, 2016</td>
<td>4</td>
</tr>
<tr>
<td>Feb 17, 2016</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 10: Summary of experimental sessions.