We Are The Champions!
Overconfidence in Groups

María Paula Cacault* Manuel Grieder*

March 2, 2015

Abstract

In a controlled laboratory experiment we find that people are overconfident about the relative performance of their (randomly assigned) in-group on an intelligence test. Our design explicitly rules out any confounding effect of individual overconfidence in the self by excluding the self from the reference group. Although in-group overconfidence fades out once relevant information is provided, the speed of convergence is slow, because the belief-updating process is asymmetric (putting more weight on positive than on negative information about the in-group). Our results suggest that beliefs about group-level performance can be distorted in a similar ego-enhancing fashion as beliefs about individual performance. A bias in beliefs about group-level abilities can have important societal and economic consequences, e.g., for patterns of statistical discrimination in hiring contexts.

Keywords: Overconfidence, social identity, self-image, beliefs, updating, discrimination.
JEL codes: C92, D03, J71, C91, A03

1 Introduction

Is there a tendency to inflate beliefs in the skill and ability of one's in-group? Beliefs about group-level ability have important consequences when group membership is used to make economically relevant inferences about an individual's characteristics. The use of statistical discrimination, e.g., for hiring decisions, is a well-known example that has repercussions both on the societal level as well as directly for the concerned individuals (see, e.g., Coate & Loury, 1993). As demonstrated by theoretical models (Phelps, 1972; Arrow, 1973; Aigner & Cain, 1977), it is beneficial to use (accurate) group-level information when forming beliefs about an individual, as long as the information available on the individual level is incomplete. Psychological research shows that group-level information is indeed

*Faculty of Business and Economics (HEC Lausanne), University of Lausanne, Quartier UNIL-Dorigny, Internef, CH-1015 Lausanne, Switzerland, emails: mcacault@unil.ch, mgrieder@unil.ch.
We gratefully acknowledge financial support from the HEC Lausanne Research Fund. We thank Adrian Bruhin, Yan Chen, Lorenz Goette, Rafael Lalive, Dominic Rohner, Luís Santos-Pinto, Vera Shuman, Christian Zehnder, Ro'i Zultan, as well as seminar participants at the Micro Workshop and the OB Brownbag at the University of Lausanne, the 8th AEW workshop 2013 in Rome, SMYE 2014 in Vienna, THEEM 2014 in Stein am Rhein, and the Zurich Workshop on Experimental and Behavioral Economics Research 2014 for valuable comments.

often used to make judgments and inferences about individuals (e.g., Alport, 1954; Fiske, 1993).

Basing decisions on beliefs about group attributes is, of course, only advantageous to the extent that such beliefs are correct. This paper studies in a controlled laboratory experiment whether identification with a group leads to erroneous, overconfident beliefs in the skill and ability of this group. We further explore how objective information about the own and other groups is processed, and whether a potential group-bias persists after relevant information has been made available. Our results reveal that beliefs about group-level ability can be subject to systematic bias. Moreover, we also show that although this in-group bias can be reduced by providing relevant objective information about differences in skill and ability between groups, it disappears only slowly, because belief-updating is asymmetric: positive and negative information about the in-group are not treated the same way.

The remainder of the paper is organized as follows: section 2 discusses related literature and the main innovations of this paper, section 3 describes the experimental design in more detail. Section 4 presents the results. In section 5 we conclude and provide a brief discussion of the potential implications of our findings.

2 Related Literature

Existing empirical evidence shows that individuals are on average overconfident about their own skill or ability. When it comes to their relative standing within a reference group, people tend to believe that they are better than they really are (called "overplacement" by Moore & Healy, 2008 or the "better-than-average" effect by Alicke et al., 1995). The beliefs we form about ourselves seem to be generated by reconciling two (usually conflicting) objectives: making good decisions and protecting our self-image (see, e.g., Bénabou & Tirole, 2002). Koszegi (2006) formalizes this notion by proposing that people are motivated by ego utility. An ego-motivated agent derives utility from believing she has certain superior qualities, and the quest for ego utility produces overconfidence via information search, as agents strategically seek or avoid relevant information in order to maximize ego-utility. Recent empirical evidence is not in line with this exact proposition (Burks et al., 2013), but there is empirical support for the idea that ego-utility has a different way of influencing beliefs: ego-motivated agents overweight positive feedback, and discount negative feedback, in order to protect a positive self-image yielding ego-utility (Eil & Rao, 2011; Moebius et al., 2014).

In the economics literature, the effects of ego-utility have so far only been considered at the individual level. A long line of research from psychology demonstrates, however, that the ego is not created in isolation. An individual's identity and self-image are decisively shaped by belonging to social groups (e.g., Tajfel & Turner, 1986).[1] Our study builds on these findings by hypothesizing that because of the importance of social identities, people may not only be motivated to hold positive beliefs about themselves individually,

[1] Akerlof & Kranton (2000) have integrated this insight into an economic model of identity, and there is a growing number of empirical studies within economics that demonstrate the general relevance of social identities for various economic domains such as, e.g., risk and time preferences (Benjamin et al., 2010), social preferences (Chen & Li, 2009), or cooperation and norm enforcement (Goette et al., 2006).

they may also derive a positive self-image (and thus ego-utility) from beliefs about the in-groups they belong to (see e.g., Crocker & Luhtanen, 1990; Aberson et al., 2000; Hewstone et al., 2002, for related theoretical and empirical work from psychology).

A number of studies in social psychology have investigated the effects of group membership and intergroup competition on individuals' judgments about the members of the in- and out-group. These studies show, e.g., that intergroup competition leads individuals to perceive (randomly assigned) in-group members as having more favorable personalities than members of the out-group (Wilson & Miller, 1961), or that the quality of the own group's work is perceived as systematically better than the work of other groups (Blake & Mouton, 1961). Directly speaking to our research question, it is found that group identity leads to more positive beliefs about the in-group's skill or ability (Bigler et al., 1997, in a study with elementary school children), and that a desire for positive group identity influences the way information about in- and out-groups is processed (Howard & Rothbart, 1980; Schaller, 1992). As it is common practice in psychology, these studies rely on non-incentivized self-report measures, and some of them also make use of deception.

Another line of psychological research that is relevant to our research question is the study of desirability bias (also called "wishful thinking," see Krizan & Windschitl, 2007a, for an overview). According to the desirability bias hypothesis, people's preferences influence their probability judgments. Specifically, they tend to overestimate the probability of an event if the occurrence of the event is in line with their preferences and provides them utility. A typical finding reports, e.g., that soccer fans overestimate the probability of winning of their favorite team (Babad & Katz, 1991). We are aware of two studies that examined wishful thinking linked to social identity: Price (2000) finds that people overestimate the probability of a fellow in-group member winning at darts against an out-group member, and Krizan & Windschitl (2007b) find that people are over-optimistic about an in-group member's relative performance on a trivia quiz for easy questions, but over-pessimistic for hard questions. Again, these studies do not use incentivized belief-elicitation methods, and they apply deception of participants.

Using the tools of experimental economics, the existence of overconfidence in relative group performance has been investigated in an unpublished study by Healy & Offenberg (2007). Their set-up differs importantly from ours as they study group overconfidence in real-existing groups in the lab (using campus fraternities and sororities), as well as at a scrabble tournament in the field (using groups of friends). Healy & Offenberg find that both students in the lab and scrabble players in the field display significant overconfidence in the relative performance of their group. While the use of real-existing groups has its advantages (see, e.g., Goette et al., 2012), it also comes with important methodological problems, especially for the study of overconfidence. In particular, the experimenter has no knowledge or control about the information structure that leads to the elicited belief. This is problematic because overconfidence can emerge as the result of rational information processing in certain environments (Benoît & Dubra, 2011). Moreover, Healy & Offenberg do not exclude the self when eliciting confidence about group-level performance. Their study does therefore not allow to disentangle the well-documented effect of individual overconfidence from overconfidence in group performance.

In a recent experiment, Brookins et al. (2014) study the effects of group identity and group judgments on confidence in individual and group performance. Their design includes a condition with a group manipulation similar to the one used by Chen & Li (2009). Brookins and colleagues find that merely asking about the relative standing of one's group reduces overconfidence in individual performance, when participants are asked to rate their performance compared to in-group members. Most interestingly for our research question, they also find that their participants are on average overconfident about the relative standing of their group compared to a rational benchmark (even though experimentally induced group identity does not seem to have an effect). However, also in their design this result remains potentially confounded by overconfidence in individual performance, as they do not exclude the self from the group comparisons.

In order to address the issues present in earlier studies, we conduct an incentivized laboratory experiment in which we control the information structure and where we can closely monitor the process of belief formation. Moreover, we exclude the self from the reference groups about which beliefs are formed in order to avoid any confounding effect stemming from inflated self-confidence. Finally, we do not simply attempt to identify overconfidence by comparing elicited beliefs to a rational benchmark. Our design manipulates group identification exogenously and involves a control condition that allows to control for the possibility that belief formation (and elicitation) may be biased because of general cognitive limitations.

3 Experimental Procedure and Data Collection

Participants in the experiment were recruited from the participant pool for behavioral experiments at the University of Lausanne, Switzerland, using the online recruitment tool ORSEE (Greiner, 2003). The participant pool includes undergraduates from all disciplines at the University of Lausanne and the Federal Polytechnic School of Lausanne.

We ran 12 sessions with 16 participants each, yielding a total of 192 participants (46% or 89 participants were women). In four of the sessions we implemented the *group identification* condition, in four sessions we implemented the *group categorization* condition and in the remaining four sessions we implemented the *control* condition. Participants in each session were randomly assigned to groups of four members, and groups were matched bilaterally in each session (i.e., each "in-group" was matched to an "out-group"). Interactions in the laboratory were computerized using z-Tree (Fischbacher, 2007), and subject-subject anonymity was ensured for the entire duration of the experiment. Participants could therefore not know which other participants were part of their group. Sessions lasted for about an hour, and participants earned on average 20.10 Swiss Francs (around 19.10 US Dollars at the time), consisting of a show-up fee of 8 Francs plus the money earned during the experiment. Participants were paid in cash at the end of each session. Sessions were run in May and September 2013 (*group identification* and *control* conditions), as well as in November 2014 (*group categorization* condition).

3.1 Stages in the Experiment

Below we describe the different stages of the experiment (see Table 1 for an overview by condition). At the beginning of each stage, participants received detailed written instructions and, when appropriate, answered control questions to ensure that they had read and understood the instructions. A translation of the instructions can be found in the Supplementary Material of this paper.

A. Group Manipulation (*group identification* condition only)
After being randomly assigned to groups of four members, participants in the *group identification* condition performed a first task that consisted of designing the flag that would represent the group during the experiment. Participants found envelopes on their desks containing colored paper (each group had a different color) and scissors. They were instructed to cut out a shape of their free choice and to put it back in the envelope. Envelopes were collected after four minutes and an assistant outside the room pasted each of the four shapes into the corners of a white paper and photographed the resulting flags. In the meantime, participants read instructions for the next part. Pictures of the flags were uploaded into the system and in all remaining stages of the experiment, participants could see their in-group's and out-group's flags at the top of of their screens.

Next, participants played an interactive game that consisted in reducing the size of a colored-circle displayed on the screen by clicking on it. The color of the circle coincided with the color that represented the out-group. All members of the group could contribute to reducing the size of the circle by clicking as fast as they could. The task was presented as a competition with the other group: after 30 seconds, the group with the smallest circle would win a reward of 10 Monetary Units (MU).[2] Ties were broken randomly. However, participants never saw the other group's circle, nor were they informed of the outcome of the game until the very end of the experiment.

After the first clicking game, participants completed the IQ test (see stage B below).

The final task intended to induce group identification in this condition was another competitive game that consisted of enlarging the size of a circle by clicking on it. The color of the circle coincided this time with the color that represented the in-group. In order to make the task slightly more interesting, this time the circle was moving around on the screen. All members of the group could enlarge the size of the circle by clicking on it. After 30 seconds, the group with the biggest circle would win a reward of 10 MU. Again, no feedback about the outcome of this game was given and participants could not see the out-group's circle.[3]

B. IQ-Test
All participants completed an IQ test divided into three sections of eight questions each. The questions were taken from Catell's (1940) culture-free intelligence test, and consisted mostly of finding the image that would fit a certain pattern. One point was added per correct answer to the individual score. Participants had 90 seconds per section to answer as many questions as possible.

In the *group identification* and *group categorization* conditions, payoffs of this stage

[2]MU were exchanged into Swiss Francs (CHF) at a rate of 5 MU to 1 CHF (around 0.95 USD at the time of the study).

[3]The group manipulation tasks in the *group identification* condition were purposefully identified in a way that avoided any real interaction between participants. This was done in order to avoid that participants could infer their fellow group members' intelligence from interacting with them in this stage (as it would have been likely, e.g., had we used a manipulation involving a group chat like Chen & Li, 2009). Section SM1 in the supplementary material provides supporting evidence that participants were not able to infer group intelligence from stage A of the experiment.

were determined at the group-level; participants received a reward of 10 MUs if their in-group's score (the sum of the four members' individual scores) was higher than the score of the out-group against which they were matched. In the *control* condition, participants received the reward if their individual score situated them within the top 50% of all participants in the session. No feedback was given about individual nor about group scores until the very end of the experiment. The scores from the IQ quiz are our measures of individual intelligence and the sum of individual scores are our measures of group-intelligence. In the following stages of the experiment, beliefs about group-intelligence are elicited.

C. Elicitation of Prior Beliefs

Since our interest is to measure overconfidence about group performance without any confounding influence of self-confidence, we excluded the individual from the reference group for which beliefs about performance were elicited. Participants in the *group identification* and *group categorization* conditions were asked to state the probability that the *other* three members of their in-group scored better than three (randomly selected) members of the out-group in the IQ test of stage B.[4]

Participants in the *control* condition compared the score of three randomly selected participants in the room with the score of three other randomly selected participants in the room. Notice that this comparison is equivalent to the one made by participants in the *group identification* and *group categorization* conditions, considering that groups were formed randomly and that no interaction that could reveal anything about the fellow group members' intelligence took place before the elicitation of beliefs.

The only difference between the *control* and the *group categorization* condition is the use of the word "group" to refer to the participants whose scores are being compared. Hence, the *group categorization* condition involves a perfectly minimal group manipulation.[5] It allows us to measure the effects of group assignment and categorization. The additional feature of the *group identification* condition consists of the group-manipulation tasks in stage A designed to strengthen group-identification. The *control* condition, finally, allows us to identify any biases in belief formation that are not caused by group framing or identification, but may, e.g., be linked to the belief-elicitation mechanism, or to general cognitive limitations.

To elicit beliefs we implemented the mechanism suggested by Karni (2009) and called the "crossover" mechanism by Moebius et al. (2014). Participants stated a value μ above which they preferred to earn a reward of 10 MU with probability $\mu\%$ rather than earning the same reward if their in-group scored better than the out-group. A random number $y \in [0, 100]$ was drawn and participants were paid the reward with probability $y\%$ if $\mu \leq y$, else they were paid the reward if their in-group scored better than the out-group. They did not receive information about realized earnings until the end of the experiment.

The value μ stated by the participants corresponds to the subjective probability that

[4]The comparison with three members of the out-group serves to keep the size of the reference groups equal.

[5]According to the criteria laid out by Tajfel & Turner (1986), also our *group identification* condition can be considered a minimal group manipulation, as groups are formed randomly and there is no social interaction taking place between group members (see Chen & Li, 2009, for a a discussion of these criteria). However, the *group categorization* condition is arguably even more minimalistic.

Table 1: Characteristics of implemented experimental conditions

	No group framing	Group framing
No group manipulation stage	*Control*	*Group categorization*
Group manipulation stage	-	*Group identification*

the in-group scored better than the out-group. This is our main variable of interest. An unbiased prior should be $\mu_0 = 50\%$ because of random assignment to groups and exclusion of the self from the reference group.

D. Signals and Posterior Beliefs

After eliciting prior beliefs, three binary signals about actual group performance on the IQ test were drawn and communicated to participants. Beliefs were elicited again after each one of them, using the same mechanism as described above. Signals were constructed by randomly selecting three questions from the IQ quiz (without replacement) and comparing groups' scores on these three questions. A positive (negative) signal meant that the in-group scored better (worse) than the out-group on the selected questions. Ties were broken randomly, and participants were informed in the instructions about the tie-breaking procedure.

The number of positive and negative signals was completely balanced by design, because a positive signal for one participant meant a negative signal for another participant. On average, unbiased posteriors should therefore also always be at 50%.

E. Beliefs about individual performance

In order to get a measure of individual overconfidence, in this last stage, participants stated their beliefs about the probability that they obtained an individual IQ score that placed them among the top-half of all participants in the session. Beliefs were again elicited using the mechanism by Karni (2009).[6]

4 Results

4.1 Data Description

The average number of correct answers in the IQ test was 12.4 out of 24, with a minimum of 6, a maximum of 20, and a standard deviation of 2.5. At the level of the groups of

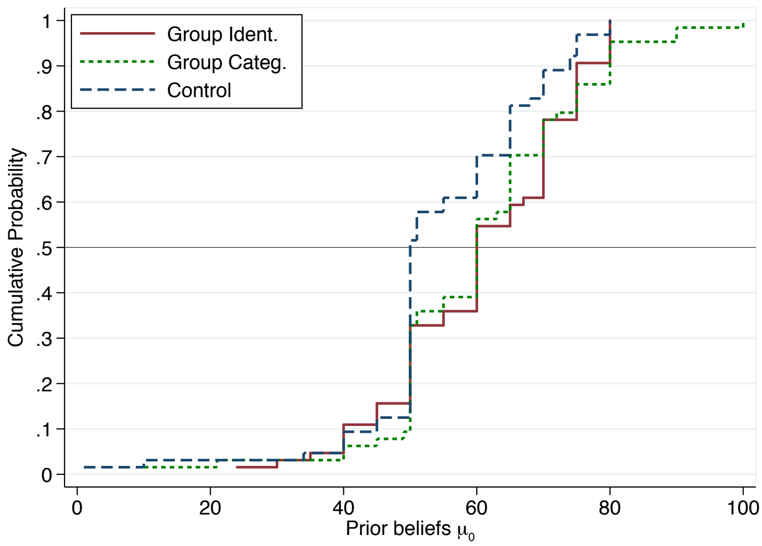
[6]Participants in the *group identification* condition completed three more stages after the elicitation of beliefs: a hiring stage in which they decided to "hire" a member of either the in- or the out-group and were paid according to the performance of the selected member; a stage where beliefs about the relative performance of the remaining groups in the room were elicited (intended originally to be a within-subject control condition); and a willigness-to-pay stage where participants chose whether or not to pay to discover the relative performance of their in-group and/or their individual relative performance. We abandoned the idea of a within-subject control condition because of apparent carry-over effects in the data. After redesigning the experiment into a between-subjects design, we dropped the other stages mentioned before in order to save time and to meet budget constraints.

three participants whose scores were compared in the belief-elicitation stages, the average group-score was 37.2 with a minimum of 26, a maximum of 51, and a standard deviation of 4.7. There was thus enough variance between individuals and groups on the intelligence measure to obtain meaningful comparisons.

A manipulation check at the end of the experiment confirms that our group manipulations were successful, and that participants in the *group identification* and *group categorization* conditions identified more with their group than the participants in the *control* condition.[7]

Figure 1 plots the empirical cumulative distribution functions of prior beliefs (μ_0) by condition. The *group identification* condition presents a flatter distribution of priors than the *control* condition, which means that mass shifts towards higher priors (i.e. the average prior belief is higher in the *group identification* condition than in the *control* condition). In particular, the median in the *control* condition coincides with the unbiased prior of 50%: the cumulative distribution function crosses 0.5 at a prior of 50%. The median prior in the *group identification* condition is 60%. Interestingly, the empirical distribution of prior beliefs in the *group categorization* condition is very similar to the one in the *group identification* condition, which suggests that random assignment to groups and pure categorization effects are already enough to put participants into a group mindset favorable to the distortion of prior beliefs. Taken together, this is first evidence that beliefs are distorted in an in-group enhancing way.

Figure 1: Empirical cumulative distribution functions of prior beliefs by condition

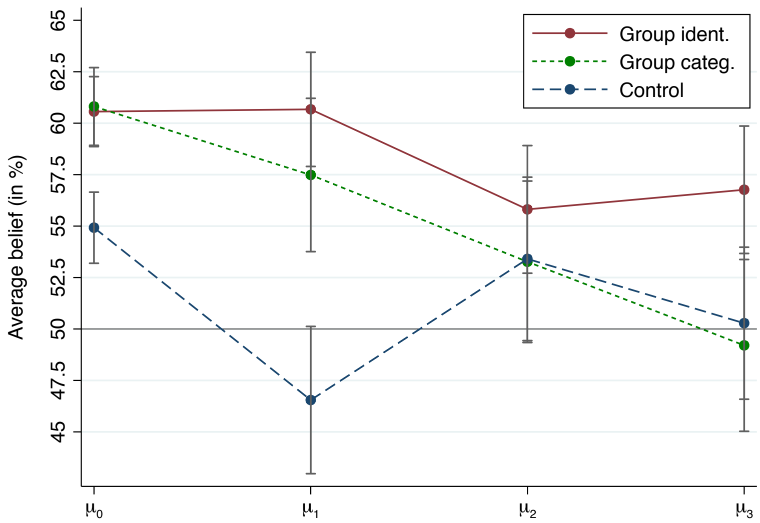


[7]We used a French translation of the manipulation check used by Chen & Li (2009). The mean identification (measured on a 10-point Likert scale from 1 to 10) in the *group identification* condition was 5.45, whereas in the *control* condition it was 3.47, which is significantly lower ($t(126) = 4.13$, $p < .001$). In the *group categorization* condition the average identification was 4.80, which is also significantly higher than in the *control* condition ($t(126) = 2.81$, $p = .006$), but not significantly lower than in the *group identification* condition ($t(126) = 1.41$, $p = .161$; all p-values reported are for two-tailed tests).

Figure 2 shows the dynamics of average beliefs (first the prior belief μ_0, and then posteriors μ_1 to μ_3, that were elicited after participants had received a signal about their group's relative performance) by condition. The capped-spikes show the standard errors. We can see that the prior and the first posterior are both significantly higher in the *group identification* condition than in the *control*. Posteriors μ_2 and μ_3, that were elicited after the second and the third signal, are on average still higher in the *group identification* condition, but the difference between the two conditions becomes statistically insignificant. Nevertheless, the average belief in the *group identification* condition is always clearly above 50% (despite the fact that the number of positive and negative signals was equalized by experimental design), whereas it fluctuates somewhat randomly around the average rational benchmark of 50% in the *control* condition.

The beliefs of participants in the *group categorization* condition converge almost linearly to the posteriors of control participants. The effect of the minimal group framing, even if effective in generating a biased prior, seems to disappear faster than the effect of the group manipulation tasks.

Figure 2: Dynamics of beliefs by condition



4.2 Estimating In-group Overconfidence

Participants in the experiment do not receive any information regarding relative performance of other group members before the elicitation of prior beliefs. Prior beliefs of participants in the *group identification* condition and in the *group categorization* condition may, however, differ from those of control participants due to the effect of the group framing and/or the group identification manipulation.[8] We can thus model priors as follows.

$$\mu_{i0} = \gamma_0 + \gamma_1 I_i + \gamma_2 C_i + \epsilon_{i0} \tag{1}$$

where $\mu \in [0, 100]$ is the subjective probability that the in-group scored higher than the out-group in the IQ quiz, I is a condition-dummy that takes value 1 for participants in the *group identification* condition and 0 otherwise, and C is a condition-dummy that takes value 1 for participants in the *group categorization* condition and 0 otherwise.

Standard OLS estimates of the parameters in model (1) are $\hat{\gamma}_1 = 5.641$ (robust standard error $SE = 2.424$, $p = 0.021$) and $\hat{\gamma}_2 = 5.891$ ($SE = 2.560$, $p = 0.022$). The parameter γ_0 measures the average prior belief in the control condition and is equal to 54.922 ($SE = 1.728$, $p < 0.001$). In line with the evidence presented in Figure 1, prior beliefs are significantly higher in both the *group identification* and *group categorization* conditions, compared to prior beliefs of control participants.

After the elicitation of prior beliefs, participants receive three binary signals about the relative performance of their in-group. A signal in period t is good $s_t = G$ if the in-group scored higher in 3 randomly selected questions of the IQ quiz (drawn without replacement), otherwise the signal is bad $s_t = B$. Ties were broken randomly and posterior beliefs were elicited directly after each of the three signals was communicated to participants.

Since the posterior beliefs in the *group categorization* condition converge fast to beliefs in the *control* condition, and for ease of presentation, in the remainder of this section we will compare the belief updating process of participants in the *group identification* condition to beliefs of control participants.[9]

To identify a potential in-group bias in the way participants process information, we begin by considering the Bayesian posterior belief

$$\mu_t \equiv P(W|s_t, ..., s_1) \quad = \quad \frac{P(s_t|W, s_{t-1}, ..., s_1)}{P(s_t|s_{t-1}, ..., s_1)} \mu_{t-1} \tag{2}$$

where W (for "winning") is the event "the in-group scored higher than the out-group in the IQ quiz" and s is the signal received about relative performance of the in-group. Let p_t denote the informativeness of a good signal, i.e. $p_t = P(s_t = G|W, s_{t-1}, ..., s_1) = 1 - P(s_t = G|L, s_{t-1}, ..., s_1)$,[10] where L (for "losing") is the event "the in-group scored lower than the out-group in the IQ quiz." For good and bad signals, the likelihood ratios of posteriors are:

$$\frac{P(W|s_t = G, s_{t-1}, ..., s_1)}{P(L|s_t = G, s_{t-1}, ..., s_1)} = \frac{p_t}{(1 - p_t)} \frac{\mu_{t-1}}{1 - \mu_{t-1}}$$

$$\frac{P(W|s_t = B, s_{t-1}, ..., s_1)}{P(L|s_t = B, s_{t-1}, ..., s_1)} = \frac{(1 - p_t)}{p_t} \frac{\mu_{t-1}}{1 - \mu_{t-1}}$$

[8] Section SM1 of the Supplementary Material shows evidence that participants in the *group identification* condition could not infer relative performance of their in-group in the IQ test from the tasks they performed in the group manipulation stage. Section SM2 shows that beliefs about group performance are not driven by beliefs about individual performance.

[9] In section SM3 of the Supplementary Material of this paper, we compare the updating process of participants in the *group categorization* condition with that of participants in the *control* condition. The results confirm the evidence in Figure 2: participants in the *group categorization* condition do not update their beliefs differently than participants in the *control* condition.

[10] Notice that this equality is given by experimental design.

Table 2: Interpretation of parameters

	Control cond.	Group identification cond.
Perfect Bayesian updating	$\alpha_0^G = \alpha_0^B = 1$	$\alpha_0^G + \alpha_1^G = \alpha_0^B + \alpha_1^B = 1$
Conservative updating	$\alpha_0^G, \alpha_0^B < 1$	$(\alpha_0^G + \alpha_1^G), (\alpha_0^B + \alpha_1^B) < 1$
Asymmetric updating	$\alpha_0^G > \alpha_0^B$	$\alpha_0^G + \alpha_1^G > \alpha_0^B + \alpha_1^B$
No in-group bias in updating		$\alpha_1^G, \alpha_1^B = 0$

Taking logs and combining these two equations we can write the posterior odds as

$$\ln\left(\frac{\mu_t}{1-\mu_t}\right) = \ln\left(\frac{p_t}{1-p_t}\right)\mathbb{1}(s_t = G) + \ln\left(\frac{1-p_t}{p_t}\right)\mathbb{1}(s_t = B) + \ln\left(\frac{\mu_{t-1}}{1-\mu_{t-1}}\right)$$

with $\mathbb{1}(s_t = S)$ an indicator that takes value one if the signal is S. Allowing all components to differ between conditions according to the indicator $I_i = 1$ if participant i is in the *group identification* condition and $I_i = 0$ if she is in the *control* condition, we arrive at the following regression model:

$$\begin{aligned} \tilde{\mu}_{it} &= \alpha_0^B s_{it}^B + \alpha_1^B [s_{it}^B \times I_i] + \alpha_0^G s_{it}^G + \alpha_1^G [s_{it}^G \times I_i] + \\ &\quad \beta_0 \tilde{\mu}_{i,t-1} + \beta_1 [\tilde{\mu}_{i,t-1} \times I_i] + \epsilon_{it} \end{aligned} \tag{3}$$

for $t \geq 1$, where $\tilde{\mu}_{it} \equiv \ln[\mu_t/(1-\mu_t)]$ (this term is called "logit beliefs" by Moebius et al. (2014)), $s_{it}^G \equiv \ln[p_{it}/(1-p_{it})]\mathbb{1}(s_{it} = G)$ and $s_{it}^B \equiv \ln[(1-p_{it})/p_{it}]\mathbb{1}(s_{it} = B)$. The effect of the group identification manipulation on beliefs is given by

$$E[\tilde{\mu}_{it}|I_i = 1] - E[\tilde{\mu}_{it}|I_i = 0] = \alpha_1^B s_{it}^B + \alpha_1^G s_{it}^G + \beta_1 \tilde{\mu}_{i,t-1} \tag{4}$$

We are interested in the difference in information processing between conditions, rather than deviations from Bayesian updating. Thus, our main focus is on the significance of the coefficients with subscript one.[11] Nevertheless, model (3) has the advantage of allowing us to say something about asymmetry and conservatism of belief updating in general: if $\alpha < 1$ participants are updating conservatively with respect to the rational Bayesian prescription, whereas if $\alpha^G \neq \alpha^B$ they give different weight to good and bad signals and update asymmetrically (Moebius et al., 2014). Table 2 provides guidance to interpret the parameters in model (3).

Before estimating the parameters in model (3) we computed the empirical informativeness (p_t) of the signals received by participants in the experiment. For the first period, we drew 2024 signals by selecting each possible triplet out of the 24 questions in the IQ test and comparing the scores of each group; p_1 is equal to the proportion of positive signals among the "winning" groups (i.e. groups that scored higher overall than their out-group).

[11] Notice that the presence of individual effects in the errors will render the OLS estimated coefficient β_0 biased and inconsistent. No inference will be made about this coefficient. Nevertheless, our parameters of interest are consistent due to exogenous assignment to conditions and exogeneity of signals.

For period two, we eliminated the actual questions that were drawn in the first signal and considered the remaining 21 questions in the test; we computed p_2 as the proportion of positive signals among the "winning" groups *conditional* on the realized first signal. Finally, we eliminated the questions drawn in the second signal and drew all possible third signals from the remaining 18 questions and computed p_3, again *conditional* on the history of signals.

Column 1 of Table 3 presents the OLS estimates of a pooled version of model (3) that does not separate the effects by condition. Overall, participants are in line with Bayesian prescriptions and incorporate information correctly: both good and bad signals are considered with equal strength and the estimated α_0-coefficients are close to one. The picture is more interesting when looking at column 2, which reports estimates of the interacted model (3). Participants in the *group identification* condition update differently than control participants: α_1^G and α_1^B are statistically significant. In particular, participants that have been exposed to the group identification manipulation react more to positive information than control participants, that is $\alpha_1^G > 0$. Moreover, these participants discount negative information more strongly, and do not update downwards as much as control participants after having received a bad signal, $\alpha_1^B < 0$. Because participants in the *group identification* condition interpret good signals as being more informative about in-group performance and bad signals as being less informative about bad performance, they persist in their upward-biased prior belief. As a result, the updating process of participants in the *group identification* condition displays only slow convergence to the beliefs of participants in the *control* condition.

5 Conclusion and Discussion

In a controlled laboratory experiment, we test the hypothesis that overconfidence exists with regard to group-level attributes by observing beliefs and information-processing about group-level ability, while explicitly ruling out confounding effects of inflated individual self-confidence. Our design identifies in-group overconfidence when the subjective beliefs of participants in the *group identification* condition are higher than the beliefs of participants in the *control* condition. We find that: a) participants are more confident about the relative intelligence of their in-group than they are about the intelligence of a group of random others, and b) the intial overconfidence fades out only slowly when participants receive relevant information about actual performance.

Consistent with evidence about the origins of individual overconfidence (Eil & Rao, 2011; Moebius et al., 2014), our participants display a directional bias in updating when processing information about their in-group compared to information about random others. They give more weight to positive information than to negative information about the in-group. This asymmetry slows down the speed of convergence in beliefs, counteracting the effect of the two dynamic forces that should reduce overconfidence in the in-group over time: namely, i) learning, and ii) the deterioration in group attachment as decisions take place at later times after the group manipulations tasks.[12]

Given that we report results from a laboratory experiment using random assignment to

[12] For supporting evidence see Chen & Li (2009). We thank Yan Chen for suggesting this channel.

Table 3: OLS regressions for $\tilde{\mu}_t$
Group identification vs. Control condition

	Pooled	Interacted
	(1)	(2)
α_0^G	0.953***	0.455
	(0.215)	(0.331)
α_0^B	0.986***	1.409***
	(0.218)	(0.404)
β_0	0.410***	0.372***
	(0.112)	(0.112)
α_1^G		0.918**
		(0.443)
α_1^B		-0.827*
		(0.477)
β_1		0.021
		(0.236)
Observations	336	336
Participants	118	118
R^2	0.224	0.245
$\alpha_0^G = 1$	0.829	0.102
$\alpha_0^B = 1$	0.948	0.313
$\alpha_0^B = \alpha_0^G$	0.915	0.049
$\alpha_0^G + \alpha_1^G = 1$		0.209
$\alpha_0^B + \alpha_1^B = 1$		0.103
$\alpha_0^G + \alpha_1^G = \alpha_0^B + \alpha_1^B$		0.089

Notes: *** p<.01, ** p<.05, * p<.10. SE clustered by participant in parentheses.
The bottom part of the table reports p-values of the corresponding tests.
Extreme-valued prior beliefs are excluded from the sample.

artificial, minimal groups that are very short-lived, we suspect that our results demonstrate only a lower-bound of the effects we identify. With real-existing groups—that are likely to be much more relevant to people's identities—, it seems safe to expect that overconfidence in the own group and the bias in updating are actually stronger.

Our results are of particular interest when considering phenomena like intergroup discrimination and spatial segregation between communities. Initial prejudice, stereotypes, or certain socioeconomic forces, can lead people of the same group to cluster together, and to segregate from other groups. Moreover, the lack of interaction between communities can prevent individuals from collecting relevant information about the characteristics of the other groups, and so prejudice and segregation are perpetuated. The results of this study suggest that more interaction between groups may be an appropriate means for reducing prejudice and stereotypes, as people do seem to be able to incorporate relevant information into their intergroup-judgments. The idea that fostering intergroup-contact may be a fruitful way to counter prejudice and discrimination is also in line with perspectives from social psychology (e.g., Pettigrew, 1998). However, the result that updating is asymmetric means that convergence to unbiased beliefs may take a long time, and shows the limits of such an approach.

Biased beliefs about group-level abilities can also have important consequences when statistical discrimination based on group membership occurs, and beliefs or forecasts about an individual's performance are formed using the individual's membership to a certain group as additional information. Economic models of statistical discrimination (Phelps, 1972; Arrow, 1973; Aigner & Cain, 1977; Coate & Loury, 1993) usually assume that beliefs about group-level ability are rationally formed to maximize their instrumental value, and that they are therefore on average correct. By showing that such beliefs about group-level performance may be systematically distorted in an in-group favoring way, our results shed some doubt on this assumption. From a theoretical perspective, this finding could provide a building stone for a more nuanced economic model of discrimination that integrates the taste-based approach (Becker, 1957) with the statistical discrimination literature. Accounting for the fact that beliefs about groups' skills may actually be distorted provides a natural link between these two perspectives.

References

Aberson, C. L., Healy, M., & Romero, V. (2000). Ingroup bias and self-esteem: A meta-analysis. *Personality and Social Psychology Review*, *4*(2), 157 – 173.

Aigner, D. J. & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 175–187.

Akerlof, G. A. & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*(3), 715–753.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S.

(1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*(5), 804–825.

Alport, G. W. (1954). *The Nature of Prejudice*. Oxford: Addison-Wesley.

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets* (pp. 3–33). Princeton University Press.

Babad, E. & Katz, Y. (1991). Wishful thinking—against all odds. *Journal of Applied Social Psychology*, *21*(23), 1921–1938.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.

Bénabou, R. & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics*, *117*(3), 871–915.

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, *100*, 1913–1928.

Benoît, J. & Dubra, J. (2011). Apparent overconfidence. *Econometrica*, *79*(5), 1591–1625.

Bigler, R., Jones, L., & Lobliner, D. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*(3), 530–543.

Blake, R. R. & Mouton, J. S. (1961). Reactions to intergroup competition under win-lose conditions. *Management Science*, *7*(4), 420–435.

Brookins, P., Lucas, A., & Ryvkin, D. (2014). Reducing within-group overconfidence through group identity and between-group confidence judgments. *Journal of Economic Psychology*, *44*, 1–12.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2013). Overconfidence and social signalling. *Review of Economic Studies*, *80*(3), 949–983.

Cattell, R. B. (1940). A culture-free intelligence test. *Journal of Educational Psychology*, *31*(3), 161–179.

Chen, Y. & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431 – 457.

Coate, S. & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *American Economic Review*, *83*(5), 1220–1240.

Crocker, J. & Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology*, *58*(1), 60.

Eil, D. & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–138.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*(1), 155–194.

Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, *96*(2), 212 – 216.

Goette, L., Huffman, D., & Meier, S. (2012). The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. *American Economic Journal - Microeconomics*, *4*(1), 101–115.

Greiner, B. (2003). An online recruitment system for economic experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003 GWDG Bericht*, volume 63 (pp. 79–93). Goettingen: Ges. f. Wiss. Datenverarbeitung.

Healy, A. & Offenberg, J. (2007). Overconfidence, social groups, and gender: Evidence from the lab and field. Loyola Marymount University Working Paper.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*(1), 575–604.

Howard, J. W. & Rothbart, M. (1980). Social categorization and memory for in-group and out-group behavior. *Journal of Personality and Social Psychology*, *38*(2), 301.

Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica*, *77*(2), 603–606.

Koszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, *4*(4), 673 – 707.

Krizan, Z. & Windschitl, P. D. (2007a). The influence of outcome desirability on optimism. *Psychological Bulletin*, *133*(1), 95.

Krizan, Z. & Windschitl, P. D. (2007b). Team allegiance can lead to both optimistic and pessimistic predictions. *Journal of Experimental Social Psychology*, *43*(2), 327–333.

Moebius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). Managing self-confidence.

Moore, D. & Healy, P. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65–85.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, *62*, 659–661.

Price, P. C. (2000). Wishful thinking in the prediction of competitive outcomes. *Thinking & Reasoning*, *6*(2), 161–172.

Schaller, M. (1992). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of Personality and Social Psychology*, *63*(1), 61.

Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 33–48). Chicago: Nelson-Hall.

Wilson, W. & Miller, N. (1961). Shifts in evaluations of participants following intergroup competition. *Journal of Abnormal and Social Psychology*, *63*, 428–431.

Supplementary Material

SM1 Checking for private information before the elicitation of priors

To ensure that no inference about group-intelligence can be made before measuring the prior beliefs, and that decisions are independent (not correlated within groups) also in the *group identification* condition, we check that participants in this condition cannot infer their in-group's intelligence by seeing the in- and the out-group's flag on their screens or by observing the size of their own group's circles at the end of the circle tasks in stage A.

Figure 3 shows the group flags constructed by the participants in the *group identification* condition. The flags right next to each other are the flags of the groups that were in competition with each other in stage A of the experiment. In order to test whether it is possible to infer a group's intelligence from these flags, we presented the eight pairs of flags to an additional sample of 45 student participants, and asked them, based on the flags only, to guess which of the two groups they believed scored higher on an intelligence test. Participants were not incentivized in this follow-up study. On average, these 45 additional participants guessed correctly 55 percent of the time. This is only marginally significantly different from a chance level of 50 percent ($p = .060$).[13] On the level of the individual comparisons, participants guessed significantly correctly for comparisons 2.b, 3.a, and 3.b. They were significantly wrong for comparison 1.b.[14] Overall, we interpret this as evidence that it is not possible to reliably infer group intelligence from seeing the flags that the two groups constructed in stage A.

Figure 4 plots the in-group's score in the IQ test (excluding the self) against the size of the in-group's circle, for the two clicking games of stage A. In none of the two games there seems to be a correlation between the size of the circles and the in-group's score in the IQ test. If any, there is a positive one in the circle-reduce game, though the slope coefficient is clearly not significant ($t = 0.51$, $p > .10$).

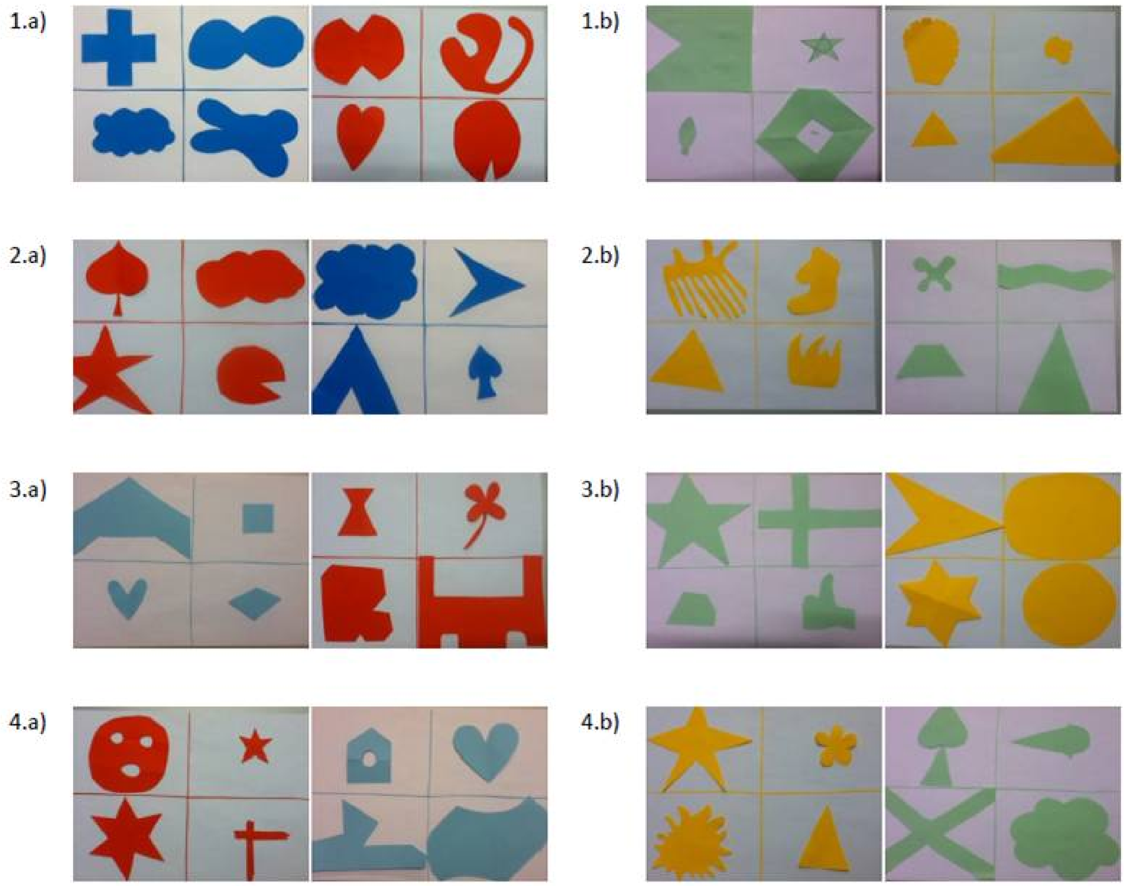
The lack of correlation between the outcome of the circle tasks and the groups' scores on the IQ test, plus the fact that intelligence cannot be reliably inferred from observing the flags, allow us to rule out within-group correlation in the elicitation of beliefs as a result of the group manipulation tasks in stage A. Moreover, participants do not get any feedback about the outcomes of the games of stage A: they do not know if they won against the other group in the circle tasks and they do not know how they scored in the IQ test individually nor at the group level. Hence, when we elicit priors, their decisions can be considered as independent. When eliciting posteriors, they are only informed of whether their group scored better or not than the other group at three randomly drawn questions. They do not see the other participants' beliefs, nor their scores, nor their signals. Thus, also for the *group identification* condition, we can consider the individual participant as

[13] p-value from OLS regression clustered by individual. A non-clustered binomial probability test yields a p-value of $p = .065$.

[14] For comparison 1.b 82.2 percent guessed falsely that the green group was more intelligent; for 2.b 71.1 percent guessed correctly that the yellow group was more intelligent; for 3.a 82.2 percent guessed correctly that the red group was more intelligent, and for 3.b 71.1 percent guessed correctly that the green group was more intelligent.

the independent level of observation, and we do therefore not cluster standard errors on the group-level in all regression analyses provided in the paper.[15]

Figure 3: Flags constructed by participants in the *group identification* condition

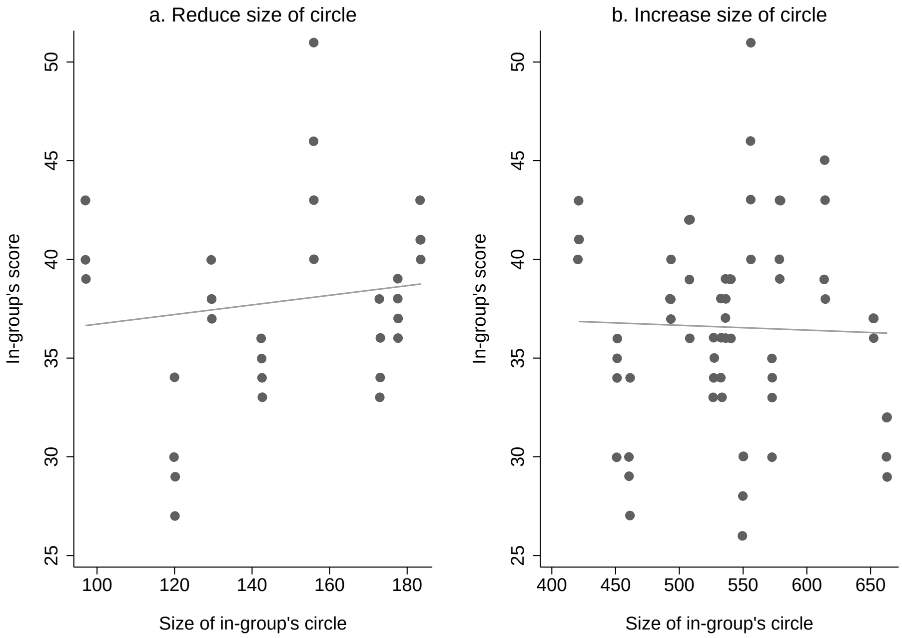


SM2 In-group overconfidence vs. individual overconfidence

Figure 5 plots the subjective prior belief that the in-group scored higher than the out-group in the IQ test, i.e. μ_0 (elicited in stage C of the experiment) against the measure of individual overconfidence (elicited in stage E), by condition. Albeit positive, the relationship between these two measures is not significant in the *group identification* condition ($t = 1.34$, $p > 0.1$). It is significant in the other conditions ($t = 2.01$ $p = 0.049$ in the *control* condition and $t = 2.12$, $p = 0.038$ in the *group categorization* condition). The fact that the correlation is significant in the *control* condition, but non-significant in our main treatment condition (*group identification*) means that individual self-confidence cannot be the driver behind the treatment differences we find (specifically that prior beliefs about

[15] For the *control* and *group categorization* condition clustering is evidently not necessary, as participants did not engage in any of the interactive group tasks of stage A.

Figure 4: In-group score in the IQ test by size of circles



group performance are higher in the *group identification* condition than in the *control* condition). It does not seem to be the case that overconfidence in groups and overconfidence in the self are strongly related. The small positive correlations we find in two of our three treatments may be driven by anchoring effects (when entering their choice regarding confidence in the self in stage E, participants may remember their earlier decisions about confidence in the group in stages C and D of the experiment, and use them as an anchor) or also by some individuals consistently misinterpreting the belief-elicitation mechanism. By comparing the results from our group identification condition with the *control* condition, our design allows us to take such potential misunderstandings of the belief elicitation mechanism into account and prevents them from biasing our results.

SM3 Estimating in-group overconfidence in the *group categorization* condition

Column 1 of Table 4 presents the OLS estimates of the model that pools together participants in both the *group categorization* and *control* conditions. Again, both good and bad signals are considered with equal strength and the estimated α_0-coefficients are close to one. In column 2 we report estimates of the interacted model. The results suggest that there is a slight asymmetry in the updating process as participants in the *group categorization* condition tend to react more to good signals than to bad signals, but these effects are not statistically significant (α_1^G and α_1^B are statistically zero). All in all, participants in the *group categorization* condition do not update differently than control participants.

Figure 5: Prior beliefs and self-confidence

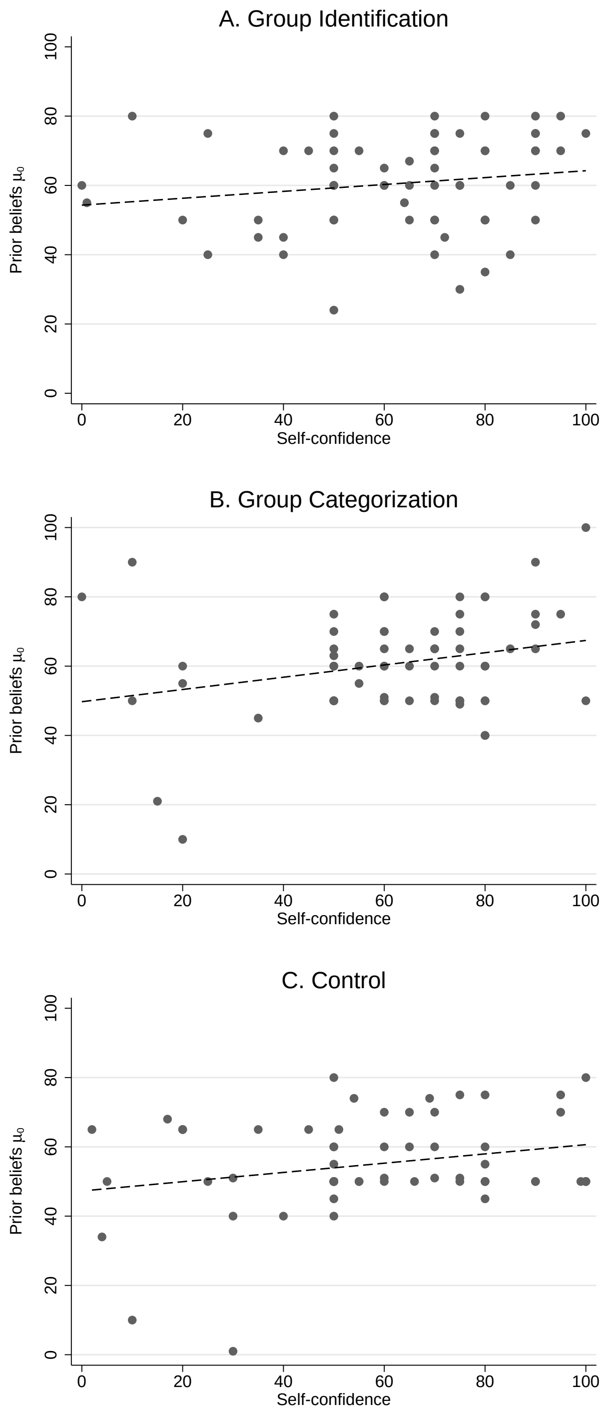


Table 4: OLS regressions for $\tilde{\mu}_t$
Group categorization v Control condition

	Pooled	Interacted
	(1)	(2)
α_0^G	0.836***	0.455
	(0.302)	(0.331)
α_0^B	1.105***	1.409***
	(0.272)	(0.404)
β_0	0.368***	0.372***
	(0.128)	(0.112)
α_1^G		0.854
		(0.650)
α_1^B		-0.593
		(0.541)
β_1		-0.039
		(0.239)
Observations	308	308
Participants	112	112
R^2	0.187	0.198
$\alpha_0^G = 1$	0.589	0.102
$\alpha_0^B = 1$	0.700	0.313
$\alpha_0^B = \alpha_0^G$	0.421	0.049
$\alpha_0^G + \alpha_1^G = 1$		0.582
$\alpha_0^B + \alpha_1^B = 1$		0.611
$\alpha_0^G + \alpha_1^G = \alpha_0^B + \alpha_1^B$		0.355

Notes: *** p<.01, ** p<.05, * p<.10. SE clustered by participant in parentheses.
The bottom part of the table reports p-values of the corresponding tests.
Extreme-valued prior beliefs are excluded from the sample.

SM4 Instructions used in the experiment

(The following instructions were originally written in French.)
Welcome to this experiment! You will make decisions that will affect your earnings. Although we express all earnings in terms of coins, these coins will be exchanged for swiss francs at the end of the experiment using the following exchange rate:

5 coins = CHF 1.-

Independently from your decisions during the experiment, you will receive a fix amount of 8 CHF for your participation. The final amounts will be rounded to the nearest integer.

It is **strictly forbidden to talk with other participants**. It is important that you respect this rule for the experiment to run smoothly. If you have any questions, please raise your hand to contact the assistants. If you do not follow this rule, we will have to exclude you from the experiment.

SM4.1 Control condition

The following part was contained in the instructions distributed to participants in the Control condition.

This experiment is divided in three parts (A, B and C). We will now explain what you will do in the first part (part A).

PART A

What is it about? You will do a test of logic based on images. Your objective is to answer correctly as many questions as possible.

What will you do? The test is composed of 3 sections. Each section of the test will be described in detail directly on the screen at the beginning of each section. We will also provide you with examples of the type of questions contained in the section. Each section is composed of 8 questions. There are thus a total of 24 questions in the test. You have 1 minute 30 seconds per section to answer as many questions as possible. For every correct answer, one point will be added to your score. If your score places you among the 50% best participants in the room, you earn a prize of 10 coins. In case of a tie, the computer will determine randomly who earns the prize.

Be aware that certain questions, towards the end of each section, are very difficult and almost nobody can expect to answer them on time. Even if you ignore the right answer, try and guess, because an empty answer is considered as false.

You will be informed of your earnings in this part at the end of the experiment.

Did you understand? At the beginning of each section, you will be able to practice with two trial questions. The answer to these questions will not influence your score.

PART B

What is it about? In this 4-section part, we propose you two games, game A and game B. Your single task is to fix a rule that determines which game you will play.

Game A is a **comparison** game, that consists in comparing the score in the image test of part A ("the test") obtained by three randomly-selected participants in the room with the score of three other randomly-selected participants in the room. For ease of explanation, we will refer to the first three participants as "persons X,Y,Z" and the other three participants as "persons A,B,C." Notice that you are not included among these six persons.

Game B is a **lottery** game, whose rules are explained below.

What will you do?

Section 1: The two games are explained in the table below:

Game A: comparison game	Game B: lottery game
If the persons XYZ have obtained MORE points in total in the image test of part A than persons ABC, you earn 10 coins. If the persons XYZ have obtained LESS points in total in the image test of part A than persons ABC, you earn 0 coins. If XYZ and ABC have obtained the same number of points in the test, the computer will toss a coin to determine who has "obtained" more points.	A random device in the computer will determine if you earn the 10 coins or if you earn 0 coins. The probability of earning the 10 coins is X% (see explanation below).

Game B offers the possibility of earning 10 coins with a probability already determined by the computer (a value between 0 and 100%, that we call here X%). What does this mean? If X is equal to 0%, for exemple, it means that you have no chance of winning in game B. On the contrary, if X is equal to 100%,you will systematically win in game B.

You don't know what is exactly the probability X of winning the 10 coins in the lottery game B, so how to chose which game you prefer to play? Actually, you will fix the rule that determines which game you will play, by telling us a value between 0 and 100%. The rule is:

- If the probability X determined by the computer is lower or equal to the value that you tell us, you play game A.

- If the probability X determined by the computer is greater than the value that you tell us, you play game B.
- To sum up, the rule is:

$$\begin{array}{lll} X & \leq & \text{value that you tell us} \quad \rightarrow \text{game A} \\ X & > & \text{value that you tell us} \quad \rightarrow \text{game B} \end{array}$$

In other words, **you will tell us from which value of the probability X of winning in game B, you prefer to play game B rather than game A.**

Please, look at the examples in the table below to better understand the consequences of your decision.

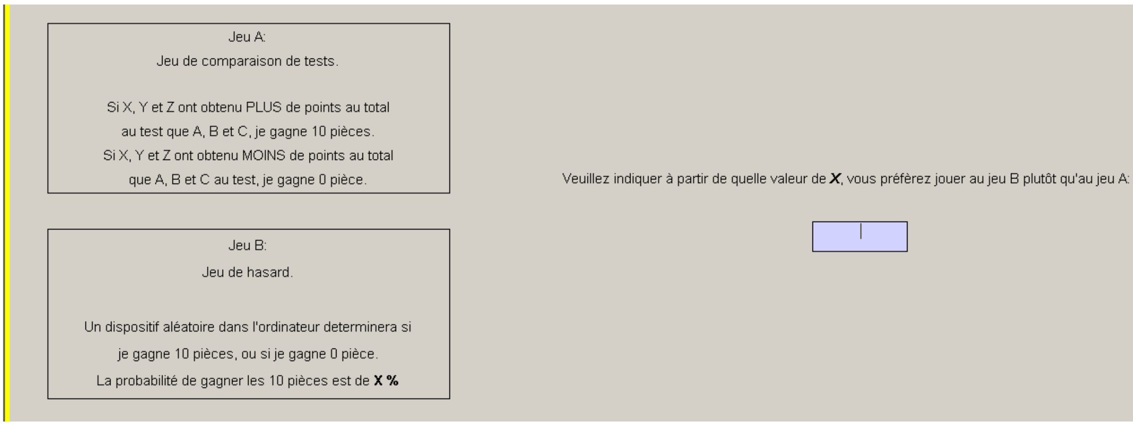
Probability X (determined by the computer)	Value that you tell us	You play
25%	20%	Game B: you earn 10 coins with a 25% probability
25%	30%	Game A: you earn 10 coins if the score of XYZ in the images test is higher than the score of ABC
75%	30%	Game B: you earn 10 coins with a 75% probability
75%	40%	Game B: you earn 10 coins with a 75% probability
75%	90%	Game A: you earn 10 coins if the score of XYZ in the images test is higher than the score of ABC

These examples show you that, if you want to maximize your chances of earning the 10 coins, you have to think carefully and tell us from which value of the probability X of winning in the lottery game you really prefer to play game B rather than game A.

The following figure shows you the screen where you will tell us this value. After having written your value click on the button "continue".

Section 2: This section is almost identical to the previous one. However, in the beginning of this section, the computer will inform you of the outcome of a "match" between ABC and XYZ.

To determine the outcome of the "match," the computer will randomly draw three out of the 24 questions of the image test of part B and will compute the number of points obtained in total at these three questions by persons ABC and persons XYZ. The three persons that have obtained the most points in these three questions win the "match." If ABC and XYZ have obtained the same number of points, the computer will toss a coin to determine who wins the "match." You will see whether XYZ have WON or LOST the match on top of your screen. After having learned the outcome of the match, you will tell us, just like in Section 1, from which value of the probability X, you prefer to play game B rather than game A.



Sections 3 and 4: These sections are identical to Section 2. Just like in Section 2, in the beginning of theses sections the computer will inform you of the outcome of a "match" between ABC and XYZ. For every match the computer will consider the answers to three out of the 24 questions of the test. However, the three questions drawn randomly by the computer are different from the ones drawn in the previous matchs. The answers to a question are thus never part of more than one match.

Also in sections 3 and 4, after having learned the outcome of the match, you will tell us from which value of the probability X, you prefer to play game B rather than game A.

You will be informed of your earnings in this part at the end of the experiment.

Did you understand? Before starting, we want to make sure that you and everybody else have understood correctly what you will do. For this purpose, please answer the control questions that will appear on your screen. These questions have no influence on your earnings.

PART C

What is it about? In this part, we will propose you two games: game A and game B. Just like in part B, you will fix the rule that determines which game you will play.

Game A is a **ranking** game, where you earn coins if the total number of points that you have obtained in the image test of part A (your individual score) places you among the best 50% of participants in the room in terms of this score.

Game B is a **lottery** game, identical to the one in part B.

What will you do? We will classify the individual scores of each of the 16 participants in this room in two halfs: the top 50% (the 8 highest scores) and the bottom 50% (the 8 lowest scores).

In the table below we explain the two games:

Just like before, you will tell us from which value of the probability X of winning at the lottery game, you prefer to play game B, rather than game A. Once you have entered your value, click on the button "continue."

Game A: ranking game	Game B: lottery game
If your individual score places you among the top 50% in the room, you earn 10 coins. If your individual score places you among the bottom 50% in the room, you earn 0 coins. If your score is right in the middle, the computer will toss a coin to determine if your score is among the top or among the bottom 50%.	A random device in the computer will determine if you earn the 10 coins or if you earn 0 coins. The probability of earning the 10 coins is X% (see explanation below).

Your earnings will be computed in the same way as in part B. You will be informed of these earnings at the end of the experiment.

Did you understand? If you have finished reading these instructions and you have no questions, click on "start". If you have questions, raise your hand to call one of the assistants.

SM4.2 Group categorization condition

The following part was contained in the instructions distributed to participants in the Group Categorization condition.

In this study, each participant is part of a group composed of four participants. Your group will always be the same throughout the entire experiment.

This experiment is divided in three parts (A, B and C). We will now explain what you will do in the first part (part A).

PART A

What is it about? Your group is related to another group in the room. You will do a test of logic based on images. The objective of your group is to obtain a score in the test higher than the other group.

What will you do? The test is composed of 3 sections. Each section of the test will be described in detail directly on the screen at the beginning of each section. We will also provide you with examples of the type of questions contained in the section. Each section is composed of 8 questions. There are thus a total of 24 questions in the test. You have 1 minute 30 seconds per section to answer as many questions as possible. For every correct answer, one point will be added to your group's score. The group with the highest score in the test earns a prize of 10 coins for each group member. The members of the losing group do not earn anything. In case of a tie, the computer will determine randomly who earns the prize.

Be aware that certain questions, towards the end of each section, are very difficult and almost nobody can expect to answer them on time. Even if you ignore the right answer, try and guess, because an empty answer is considered as false.

You will be informed of your earnings in this part at the end of the experiment.

Did you understand? At the beginning of each section, you will be able to practice with two trial questions. The answer to these questions will not influence your score.

PART B

What is it about? In this 4-section part, we propose you two games, game A and game B. Your single task is to fix a rule that determines which game you will play.

Game A is a **comparison** game, that consists in comparing the score in the image test of part A ("the test") obtained by the three other members of your group ("your group") with the total score of three members of the other group ("the other group"). The three members of the other group that will be considered for the score comparison have been randomly chosen among the four members of the other group. These three members of the other group will always be the same persons until the end of the experiment.

Game B is a **lottery** game, whose rules are explained below.

What will you do?

Section 1: The two games are explained in the table below:

Game B offers the possibility of earning 10 coins with a probability already determined by the computer (a value between 0 and 100%, that we call here X%). What does it mean? If X is equal to 0%, for exemple, it means that you have no chance of winning in game B. On the contrary, if X is equal to 100%,you will systematically win in game B.

You don't know what is exactly the probability X of winning the 10 coins in the lottery game B, so how to chose which game you prefer to play? Actually, you will fix the rule that determines which game you will play, by telling us a value between 0 and 100%. The rule is:

- If the probability X determined by the computer is lower or equal to the value that you tell us, you play game A.
- If the probability X determined by the computer is greater than the value that you tell us, you play game B.
- To sum up, the rule is:

$$\begin{array}{lll} X & \leq & \text{value that you tell us} \quad \rightarrow \text{game A} \\ X & > & \text{value that you tell us} \quad \rightarrow \text{game B} \end{array}$$

Game A: comparison game	Game B: lottery game
If **your group** has obtained MORE points in total in the image test of part A than the other group, you earn 10 coins. If **your group** has obtained LESS points in total in the image test of part A than the other group, you earn 0 coins. If your group and the other group have obtained the same number of points in the test, the computer will randomly determine which group has "obtained" more points.	A random device in the computer will determine if you earn the 10 coins or if you earn 0 coins. The probability of earning the 10 coins is X% (see explanation below).

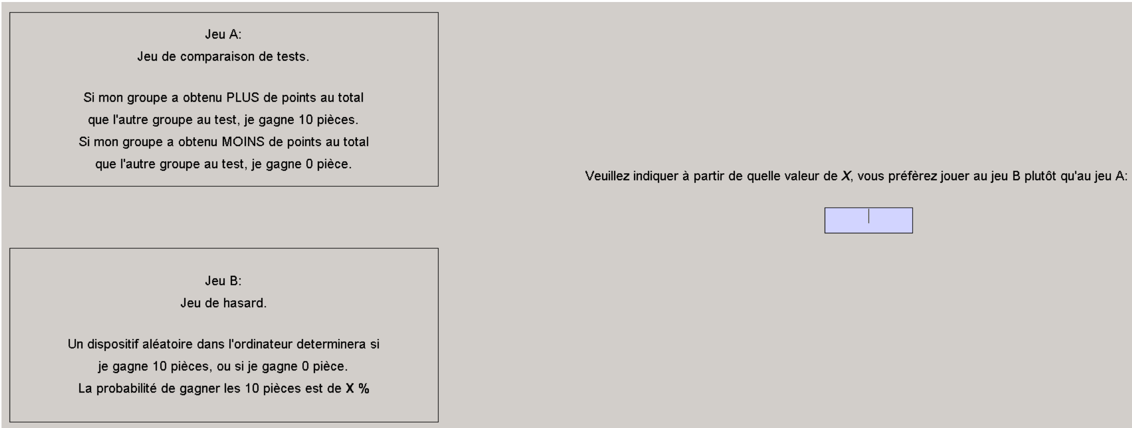
In other words, **you will tell us from which value of the probability X of winning in game B, you prefer to play game B rather than game A.**

Please, look at the examples in the table below to better understand the consequences of your decision.

Probability X (determined by the computer)	Value that you tell us	You play
25%	20%	Game B: you earn 10 coins with a 25% probability
25%	30%	Game A: you earn 10 coins if the score of your group in the images test is higher than the score of the other group
75%	30%	Game B: you earn 10 coins with a 75% probability
75%	40%	Game B: you earn 10 coins with a 75% probability
75%	90%	Game A: you earn 10 coins if the score of your group in the images test is higher than the score of the other group

These examples show you that, if you want to maximize your chances of earning the 10 coins, you have to think carefully and tell us from which value of the probability X of winning in the lottery game you really prefer to play game B rather than game A.

The following figure shows you the screen where you will tell us this value. After having written your value click on the button "continue".



Section 2: This section is almost identical to the previous one. However, in the beginning of this section, the computer will inform you of the outcome of a “match” between the two groups.

To determine the outcome of the “match,” the computer will randomly draw three out of the 24 questions of the image test of part B and will compute the number of points obtained in total at these three questions by the three members of your group and by the three members of the other group. The group that has obtained the most points in these three questions wins the “match.” If the two groups have obtained the same number of points, the computer will toss a coin to determine which group wins the “match.” You will see whether your group has WON or LOST the match on top of your screen. After having learned the outcome of the match, you will tell us, just like in Section 1, from which value of the probability X, you prefer to play game B rather than game A.

Sections 3 and 4: These sections are identical to Section 2. Just like in Section 2, in the beginning of theses sections the computer will inform you of the outcome of a “match” between the two groups. For every match the computer will consider the answers to three out of the 24 questions of the test. However, the three questions drawn randomly by the computer are different from the ones drawn in the previous matchs. The answers to a question are thus never part of more than one match.

Also in sections 3 and 4, after having learned the outcome of the match, you will tell us from which value of the probability X, you prefer to play game B rather than game A.

You will be informed of your earnings in this part at the end of the experiment.

Did you understand? Before starting, we want to make sure that you and everybody else have understood correctly what you will do. For this purpose, please answer the control questions that will appear on your screen. These questions have no influence on your earnings.

PART C

(This part is identical in all conditions. Refer to subsection SM4.1 for instructions.)

SM4.3 Group identification condition

The following part was contained in the instructions distributed to participants in the Group Identification condition.

In this study, each participant is part of a group composed of four participants. Your group will always be the same throughout the entire experiment.

This experiment is divided in seven parts (A to G). We will now explain what you will do in the first part (part A).

PART A

What is it about? In this part, the other members of your group and yourself will construct the flag that will represent you throughout the experiment. For this purpose, each of you will be in charge of decorating one of the fours corners of the flag using colored paper (each group has a different color). In other words, the flag of your group will be composed of a representative element of each one of you.

What will you do? On your table, in front of you, you will find a pair of scissors and an envelope containing colored paper. Start by cutting out a shape. You are free to cut out any shape you want, but you cannot cut out words, nor can you write on the shape. The shape must be composed of a single element (you cannot cut out two separate shapes).

When you finish, write your place number on the back of the shape and put it inside the envelope.

After 5 minutes, the assistants will collect the envelopes containing the shapes. They will then paste the shapes on the corners of the flag. You will see what the flag of your group looks like in the next part of the experiment.

You can now open the envelope and start working.

PART B

What is it about? From now on, your group is related to another group in the room. You will take part in a series of group games. The objective of your group is to win against the other group in each of these games.

What will you do? Firstly, the flag of your group that you have just created in part A will appear on your screen. After having seen your flag, click on "OK" so that the first game can start.

Game 1: reducing the size of circles

You and the other members of your group will see in the middle of your screen a very big circle. The members of the other group will see a circle of the same size on their screens. The aim of the game is to reduce the size of the circle . To reduce the size of the circle you just have to click on it. Every click of a member of your group reduces the size of the circle. The game is over after 30 seconds. The group that finishes with the smallest circle, wins the game, and each member of the winning group earns a reward of 10 coins.

The members of the losing group do not earn anything. In case of a tie, the computer will toss a coin to determine which group earns the reward.

Game 2: image test

The second group game consists of a logic test based on images. The test is composed of 3 sections. Each section of the test will be described in detail directly on the screen at the beginning of each section. We will also provide you with examples of the type of questions contained in the section. Each section is composed of 8 questions. There are thus a total of 24 questions in the test. You have 1 minute 30 seconds per section to answer as many questions as possible. For every correct answer, one point will be added to your group's score. The group with the highest score in the test earns a prize of 10 coins for each group member. The members of the losing group do not earn anything. In case of a tie, the computer will determine randomly who earns the prize.

Be aware that certain questions, towards the end of each section, are very difficult and almost nobody can expect to answer them on time. Even if you ignore the right answer, try and guess, because an empty answer is considered as false.

At the beginning of each section, you will be able to practice with two trial questions. The answer to these questions will not influence your score.

Game 3: enlarging the size of circles

In this game, you and the other members of your group will see in the middle of your screen a small circle. The members of the other group will see a circle of the same size on their screens. The aim of the game is to enlarge the circle. To enlarge the circle you just have to click on it. Every click of a member of your group enlarges the circle. You will notice that the circle will move around the screen. You will have to carefully target the moving circle to be able to enlarge it. The game is over after 30 seconds. The group that finishes with the largest circle, wins the game, and each member of the winning group earns a reward of 10 coins. The members of the losing group do not earn anything. In case of a tie, the computer will toss a coin to determine which group earns the reward.

You will be informed of your earnings in this part at the end of the experiment.

PART C

(This part is identical to part B of the Group Categorization condition. Refer to subsection SM4.2, Part B, for instructions.)

PART F

(This part is identical to part C of the Control and Group Categorization conditions. Refer to subsection SM4.1, Part C, for instructions.)