

The Strategic Display of Emotions*

Daniel Chen[†]

Astrid Hopfensitz[‡]

Boris van Leeuwen[§]

Jeroen van de Ven[¶]

August 2018

Abstract

The emotion that someone expresses has consequences for how that person is treated. We study whether people strategically adjust their expressed emotions in economic games. In a laboratory experiment, participants play a task-delegation game in which managers assign a task to one of two workers. We vary whether getting the task is desirable or not. Workers are instructed to take pictures expressing happiness and anger, and choose which picture to show to the manager. We find that workers can avoid getting the task by showing the picture on which they express anger and are more likely to show anger when the task is not desirable. When we do not instruct workers to express emotions, we find that they still strategically adjust their facial expressions but fail to exploit the full potential benefits.

*We thank for helpful suggestions and comments.

[†]IAST, [email](#)

[‡]IAST, [email](#)

[§]Department of Economics and CentER, Tilburg University. b.vanleeuwen@uvt.nl

[¶]Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute.
j.vandeven@uva.nl

Emotions are important in many economic interactions. Like verbal communication, they act as social cues and convey intentions. Appearing angry during negotiations leads to more favorable offers (Reed et al., 2014; Andrade and Ho, 2009; Van Kleef et al., 2004), while smiling helps fostering trust (Scharlemann et al., 2001; Centorrino et al., 2015). This opens the possibility to manipulate emotional expressions in order to elicit a more favorable treatment. While the use of emotional expressions is widespread in strategic interactions, there is only very limited research on whether people can and try to use this to their advantage.

We study whether people strategically adjust their expressed emotions in different economic games. In a laboratory experiment, participants play a task delegation game in which 'managers' assign an investment task to one of two 'workers'. The main treatment variation is whether getting the task is desirable or not. In either treatment, the manager benefits from a worker's investment in the task, while investing is always costly to the worker. It is therefore in the manager's interest to assign the task to the worker that is most likely to make the investment. Many tasks in organizational environments fit this setting. For instance, some tasks are perceived as pleasurable to do for workers (teach a course on game theory), while others are best avoided (administrative duties).

Our main interest is in whether the worker's emotional expression affects the likelihood of getting the task and whether workers exploit this. In the experiment, workers were instructed to take pictures expressing happiness and anger, and could then choose which picture to show to the manager. We hypothesized that anger is associated with a lower trustworthiness in that person making an investment. If so, workers can avoid getting the task by showing anger, and we predicted that they would show anger more frequently when the task is undesirable to get. Managers were not made aware of the fact that workers were instructed to express emotions. We did this so that the managers would not immediately question the sincerity of the expressed emotions.

The main findings are that workers can indeed avoid getting the task by showing the picture on which they express anger, and that they are more likely to show anger when the task is undesirable. Managers assign the task more often to the worker that expresses happiness. We also provide an insight to the mechanism: angry looking workers are perceived as less trustworthy than happy looking workers.

In accordance with this, we find that workers are substantially more likely to show anger when the task is not desirable. When the task is desirable, workers send their angry picture 22 percent of the time. This doubles when the task is undesirable. We also exploit a design feature that allows us to show that workers that do not send their angry picture when the task is undesirable, would in fact have been better off sending their angry picture.

Subjects in our experiment understand the strategic benefits of adapting their expressed emotion to the situation. However, the extent to which they do might be exaggerated by our experimental design. Within this framework, participants can only choose between showing happiness and anger. Furthermore, the request to express different emotions may have made them aware of the potential benefits of using them strategically. In a follow-up experiment, we study whether subjects spontaneously adapt their expressions if no such instructions are given. We find that they still strategically adjust their facial expressions, but they only exploit the potential benefits to a modest extent. Using valence as a measure of emotional intensity on the pictures, we find that the majority of subjects display higher valence (more positive emotions) when the task is desirable compared to when the task is not desirable. However, the difference in mean intensity between the types of tasks is only 41 percent of that in the baseline experiment.

To our best knowledge, ours is the first paper to study the strategic display of emotions. The most closely related paper is that by Andrade and Ho (2009). In their experiment, receivers in an ultimatum game can send a verbal message stating their level of anger. They find evidence that people strategically over-report their level of anger. Our focus is on nonverbal, facial cues. While there are several parallels, facial expressions are different in that they are more hard-wired and perceived as harder to mimic. Furthermore, when both verbal and nonverbal cues are present, the nonverbal cues can be of first-order importance: when the facial expression of a person does not match his or her verbal statements, people frequently rely more on the former. The importance of having facial cues over (only) verbal cues is illustrated by the high value that is placed on face-to-face interactions (see later).

Another closely related paper is that by Gneezy and Imas (2014). They find that people strategically manipulate the emotions in others. When they have the option to anger their opponent, they do so in environments in which this pays off to do.

Our study also relates to a growing literature showing the importance of facial cues more generally. In many economic interactions, people highly value facial cues. For many jobs, face-to-face interviews are a key part of the recruitment procedure, and most people post pictures of themselves on their professional websites. Online marketplaces such as Airbnb and Uber make it possible for buyers and sellers to see each other on photos before trading. Businesspeople spend large amounts of time and money to meet with their business partners before signing a contract (Forbes 2009). In trust games, many people are willing to pay to see a photo of the other (Eckel and Petrie 2011) or to show their own photo to others (Heyes and List 2016). Facial cues are predictive of behavior, which justifies the value that is placed on them. Based on seeing pictures of unknown others, observers can predict behavior in ultimatum games (Van Leeuwen et al. 2017), trust games (Bonnefon et al. 2013; Vogt et al. 2013), dictator games (Fetchenhauer et al. 2010) and prisoner dilemma or public good games (Verplaetse et al. 2007; Tognetti et al. 2013) above chance levels. Pupil dilation is related to deceptive behavior (Wang et al., 2010).

1 Experiment I

1.1 Experimental Design and Hypotheses

Our main setup is a task delegation game, inspired by that of Babcock et al. (2017). In all treatments, we divided subjects into 'managers' and 'workers' (we used neutral labels in the experiment). Each manager was paired with two workers. The manager saw pictures of the workers and then had to allocate an investment task to one of them. The chosen worker became the 'designated' worker. The designated worker could then accept or refuse to invest. Investing is costly to the worker but beneficial to the manager.

Our interest is in the emotional expression of the workers on their pictures. Prior to receiving the instructions, each worker was asked to take two photos: one with a happy expression and one showing anger. They could then choose which picture to show to the manager. The main purpose is to examine whether workers adapt their choice to the strategic environment. Managers were not informed that workers were asked to express emotions and workers knew this. We did this so that the managers

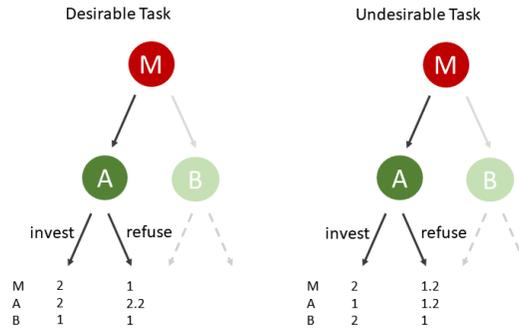


Figure 1: Payoffs when the task is assigned to worker A. Left: desirable task. Right: undesirable task. The case in which the task is assigned to worker B is symmetric.

would not immediately question the sincerity of the expressed emotions.

The treatment variation was whether or not it was desirable for the workers to get the task. In [treatment 'Desirable'](#), the worker was always better off by getting the task than not getting the task, independent of his investment decision. The designated worker earned 2 if he invested, and 2.2 if he did not invest. The other worker always earned 1. The manager earned 2 if the designated worker invested, and 1 if the designated worker did not invest. The left part of [Figure 1](#) summarizes the payoffs for the case in which the task is assigned to worker A (the case in which the task is assigned to worker B is symmetric). In [treatment 'Undesirable'](#), a worker was better off if the task was assigned to the other worker and the other worker invested. In this version, if the designated worker refused to invest, the other worker had to invest. The worker that ended up investing always earned 1. The worker not investing earned 2 if the designated worker invested and 1.2 if the designated worker refused. The right part of [Figure 1](#) summarizes the payoffs for this case.

Note that in both treatments, the manager earns more money if the designated worker invests. This creates incentives to allocate the task to the worker that appears most likely to invest. On their turn, workers like to appear trustworthy in Desirable. In Undesirable, they may wish to appear untrustworthy so as to avoid getting the task.¹

¹There is a caveat. If a worker is very pessimistic about the likelihood that the other worker will invest, he is better off by getting the task himself and not invest.

We conjectured that anger is associated with lower trustworthiness, and therefore that workers would be more likely to send the angry picture in Undesirable than in Desirable.

Hypothesis 1. *Anger is associated with lower levels of trustworthiness. In both treatments, managers are less likely to assign the task to workers that look angry.*

Hypothesis 2. *Workers are more likely to display anger in treatment Undesirable compared to treatment Desirable.*

Even if Hypothesis 1 is supported, if a worker were to select the happy picture in Undesirable or the angry picture in Desirable, this is not evidence per se that they are making a wrong decision. It is well possible that for a particular player it is not always optimal to display the emotion that is payoff maximizing for the ‘average’ player. For instance, some players may be unconvincing in pretending to be angry, and therefore be better off looking happy in either version of the game. In order to deal with this, we did not always implement the player’s choice. Instead, we told them that for one of the workers the chosen picture would be implemented, and for the other one it would be randomly determined. This creates a counterfactual which we use as a control.

1.2 Experimental Procedures

To preserve anonymity, we used two labs in two different cities. Workers were in Amsterdam (CREED lab) and managers were in Tilburg (CentER lab). A total of 272 subjects participated, 136 in each role (51 percent female, mean age 22). There were 10 sessions. Each session had 24, 28, or 32 participants, depending on the show-up. Sessions were gender-balanced, with a fraction of females that was always between 0.42 and 0.60.²

At the start of a session, workers were first asked to take one picture expressing happiness and one picture expressing anger. We asked them to make it look natural. They could retake a picture until they were satisfied with it. Only then did they

²Before running this experiment, we ran one pilot session in which we used different games (a trust game and ultimatum game). We decided to change the design to make the games more comparable and to create a situation in which looking angry has clear potential benefits.

receive the instructions for the remainder of the experiment. The same pictures were used throughout the experiment.

All players then received the instructions on their screen. The instructions are in Appendix D. To verify their understanding, they always had to answer a series of test questions correctly before they could continue. They also received a hard-copy with a summary of the instructions.

Each subject played a total of 12 rounds and was paid for every round. Subjects were rematched every round. Managers never saw the same worker more than once. At the end of each round, they received feedback about their earnings, which picture was shown, and whether or not the (other) worker invested. Average earnings were XX and each session lasted about 90 minutes. Each session ended with a survey in which we collected additional information. We describe this in section 1.4.4 when we analyze heterogeneity.

Workers had to give consent for the use of their pictures in the experiment if they wanted to participate. They were seated in closed, sound-proof cubicles. To ensure that the software package FaceReader could detect expressions, subjects were always told to capture their entire face on the picture and to look into the camera. Subjects could take pictures themselves using a high-quality webcam (Logitech HD 1080p).

1.3 Measurements

Intensity of expressed emotions. The expressed intensity of different emotions was measured using FaceReader 7.1. This software package classifies expressions based on distances between 538 points on the face as well as muscle tightness. It is trained on an artificial neural network using over 10,000 images, and has been validated in several studies (see Van Leeuwen et al., 2017 for references). As output, it provides for each emotion a number between 0 and 1, reflecting the intensity of the expressed emotion on that picture. The software could successfully capture the face for all but one picture.

Perceived trustworthiness and attractiveness. An independent group of observers rated pictures on trustworthiness and attractiveness on a 7-point scale. Each picture was

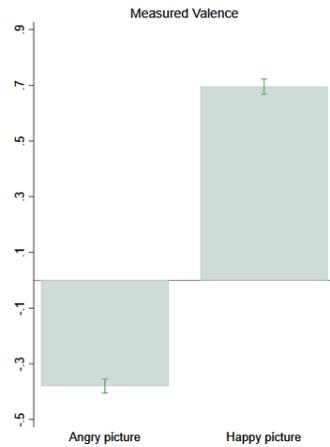


Figure 2: Mean valence by expression on the picture. Error bars indicate +/- 1 s.e.

rated 8 times (4 males, 4 females) on each dimension. Each observer rated pictures on only one dimension and saw only one picture of each person. Pictures were sorted by gender. Raters received a flat payment of XXX. Cronbach’s alpha, a measure of interrater reliability, is around 0.8, which is usually considered to be a good reliability (0.809 and 0.814 for the trustworthiness and attractiveness ratings, respectively).

1.4 Results

1.4.1 Anger, Happiness, and Perceived Trustworthiness

We first verify that the happy and angry pictures indeed express different emotions, and that this affects trustworthiness ratings. We use valence as a composite measure of the intensity of expressed emotions. Valence is defined as the difference in positive and negative emotions, and can vary from -1 (very negative) to +1 (very positive).³ Figure 2 shows a substantial difference in valence between the pictures. As expected, mean valence is negative on the angry picture and positive on the happy picture.

³Valence is calculated as the level of ‘happy’ minus the maximum level of any negative emotions (anger, sad, scared, and disgust). Subjects in our experiment express very little negative emotions besides anger, so that valence is essentially the level of ‘happy’ minus the level of ‘anger’.

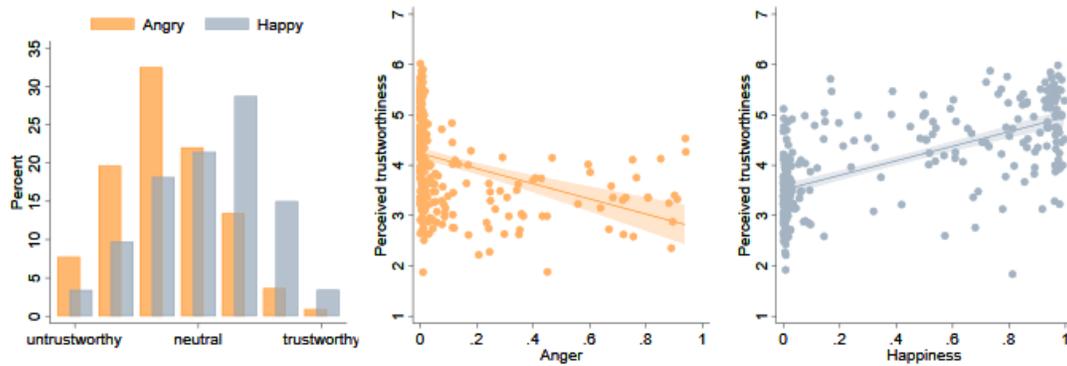


Figure 3: ... Error bars indicate +/- 1 s.e.

Furthermore, the perceived trustworthiness is substantially higher on the happy picture (4.67) compared to the angry picture (3.43). Telling is that almost all subjects (124 out of 135, or 92 percent) look more trustworthy on their happy picture than on their angry picture. Figure 3 shows that there is a clear shift in the distribution of ratings (signed-rank test, $p < 0.001$). The modal rating is below ‘average trustworthiness’ for the angry picture and above ‘average trustworthiness’ for the happy picture. Perceived trustworthiness is strongly negatively associated with the degree of anger ($\rho = -0.484$, $p < 0.001$) and positively associated with the degree of happiness ($\rho = 0.645$, $p < 0.001$).

1.4.2 Behavior of managers

Our main research question is whether subjects strategically adapt their expression to the situation. This presumes that there is some benefit of choosing a certain expression. We find clear evidence that the expression matters for the manager’s decision. For both task types, managers are more likely to assign the task to the worker that looks happy. When the two workers express different emotions, the one that looks happy is 17 percentage points more often chosen when the task is desirable ($p = 0.019$) and 29 percentage points more often when the task is undesirable ($p < .001$). These differences are still present, and even somewhat larger, if the two workers are of the same gender.

Table 1 shows the estimates of the likelihood that the task is assigned to a worker. The reported coefficients are in terms of odds ratios. In both task types, a happy

Table 1: Assignment of task.

Task:	(1) Desirable	(2) Desirable	(3) Undesirable	(4) Undesirable
Happy expression	1.405** (0.207)	1.251 (0.251)	1.818*** (0.270)	1.382 (0.293)
Perceived trustworthiness		1.102 (0.139)		1.291** (0.154)
Female		1.171 (0.183)		0.731** (0.111)
N	840	840	792	792

Conditional logit estimates (odds ratios). Dep. var.: being the designated player. Robust s.e.in parentheses clustered at the manager level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

expression significantly increases the likelihood of becoming the designated worker (columns 1 and 3). The coefficients get closer to 1 and become insignificant after controlling for trustworthiness and gender. Perceived trustworthiness is above 1 and significant in the case of an undesirable task, supporting the idea that managers select happy looking workers because they look more trustworthy (columns 2 and 4).⁴ We also note that, compared to males, females are less likely to be chosen when the task is undesirable. When one of the workers is male and the other is female, the female is 9 percentage points less likely to be chosen. This is counter to the findings of Babcock et al. (2017). However, the task we use is not identical.

1.4.3 Behavior of workers

Our results indicate that there are clear benefits for workers of sending the happy expression when the task is desirable and the angry expression when the task is undesirable. We find that workers act in accordance, and are roughly twice as likely to send their angry picture when the task is undesirable compared to when the task is desirable (44 percent versus 22 percent), a difference that is highly significant (signed-rank test, $p < 0.001$). Figure 4 shows that this difference already exists in the

⁴Attractiveness is not included as a control because of multicollinearity problems. It is highly correlated with trustworthiness (correlation coefficient 0.562, $p < 0.001$).

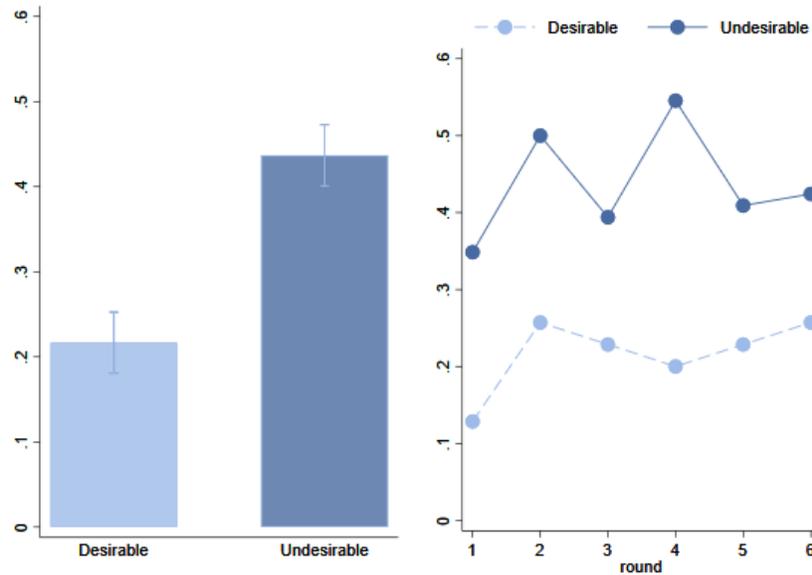


Figure 4: Percentage of workers choosing the angry picture. Error bars indicate +/- 1 s.e.

first round and persists over the rounds. A very similar picture arises when we look at it from the perspective of trustworthiness. When the task is desirable, workers send their least trustworthy picture relatively rarely (20 percent of the time). This happens more frequently when the task is undesirable (45 percent of the time).

The reluctance by some workers to send their angry picture when the task is undesirable is not per se evidence that they are making the wrong choice. Some workers may be better off (in monetary terms) sending their happy picture even when the task is undesirable. This, however, is not the case. To show this, we exploit our design feature that it is sometimes randomly determined which of the worker's pictures is sent. Even when workers prefer to show a happy expression, their angry picture is sometimes shown instead. This allows us to estimate the effect of sending an angry picture for all cases in which a worker preferred to send a happy picture.

Column 1 of Table 2 first replicates the finding that angry-looking workers are less likely to get selected for the task when the task is undesirable; the odds-ratio is significantly lower than 1. Column 2 estimates the effect among the group of workers that *intended* to send the happy picture. We find that the odds-ratio of being selected is even lower. Thus, if anything, the benefits of sending the angry

Table 2: Assignment of task.

	(1)	(2)
Task:	Undesirable	Undesirable
Sample:	All	Send Happy
Angry expression	0.550*** (0.082)	0.486** (0.141)
Observations	792	248

Conditional logit estimates (odds ratios). Dep. var.: being the designated player. Robust s.e.in parentheses clustered at the manager level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

picture are even larger among this group of workers.

Another way to see that those workers are making suboptimal choices is by comparing the trustworthiness on their pictures. Among workers that prefer to send a happy picture when the task is undesirable, the mean trustworthiness on their happy picture is higher than on their angry picture (4.67 against 3.40). Consequently, they could have decreased their perceived trustworthiness by sending their angry picture. Possibly, they nevertheless send their happy happy picture because that is the one on which they look more attractive.

@@Should we say something about investments? fraction investing is about .35 (.4 for desirable, .29 for undesirable). No correlation with perceived trustworthiness, but this could mean they successfully use facial expressions strategically. also, not a random sample, since most trustworthy looking are asked to invest.

1.4.4 Heterogeneity

We next examine if there is heterogeneity in how strategic people are in their emotional expressions. We collected information on gender, the level of strategic reasoning, and emotional intelligence. The level of strategic reasoning is measured using an adapted version of the racing game (see Gneezy and Rustichini, @@@). This captures a subject's ability to perform backward induction. Emotional intelligence is measured in two ways. The first is performance on the 'Reading the Eyes in the

Mind’ test (Baron-Cohen, @@@), a standard test in psychology in which subjects are asked to describe the emotions displayed in pictures. The second is a subject’s ability to predict which people would reject a low offer in an ultimatum game, based on facial expressions (see the task developed in Van Leeuwen et al., @@@). All these measures were collected in a survey following the main part of the experiment and were not incentivized. More details about the tasks and the determination of scores can be found in Appendix A.

We measure subjects’ responsiveness to the situation by comparing the fraction of times that they sent the picture that matches the situation; the happy picture when the task is desirable and the angry picture when the task is undesirable. Column 1 in Table 3 shows the results. Females appear less responsive than males but the difference is not significant. Subjects who score high on the eyes test are more responsive, while we do not find a significant impact of the level of strategic reasoning on responsiveness. One might expect that subjects with a large difference in perceived trustworthiness between the two pictures will be more responsive, as for them the impact of sending the right picture is larger. The coefficient is indeed positive but small and not significant. Of course, these correlations should be interpreted with caution. For instance, a low score on the eyes test may be indicative of low effort rather than a true absence of emotional intelligence.

2 Experiment II

Subjects in our experiment understand the strategic benefits of adapting their expressed emotion to the situation. Arguably, the request to express different emotions on the pictures made them aware of the potential benefits. In a follow-up experiment, we study whether subjects spontaneously adapt their expressions.

The setup of the new treatment is very similar to the baseline experiment. The main difference is that this time we did not instruct them to express emotions on the pictures. Subjects in the role of workers took a selfie at the start of each round. We predicted that the expressed emotions would be less extreme in this setup, making it harder to detect any differences. We therefore used a within-subject design to increase statistical power. Each subject played six rounds in each version of the game, and we varied the order of the games across sessions. They received the

Table 3: Responsiveness and background characteristics

Sample Dep. var.	(1) Experiment 1 Matching picture	(2) Experiment 2 Increase in valence
Female	-0.093 (0.058)	-0.062 (0.060)
Eyes score ^a	0.109*** (0.029)	-0.014 (0.030)
Strategic reasoning score ^a	0.043 (0.028)	0.041 (0.030)
Δ trustworthiness ^b	0.028 (0.039)	
Constant	0.627*** (0.060)	0.122*** (0.043)
Observations	136	74
R-squared	0.124	0.047

OLS estimates. ^a: standardized scores. ^b: difference in trustworthiness between the happy and angry picture. Robust s.e. in parentheses clustered at the worker level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

instructions for the second part only after completing the first part. 148 subjects participated in the experiment, 74 in each role (54 percent female, mean age 22). FaceReader could successfully read emotions on 866 (97.5 percent) of the pictures. More details about the procedures can be found in Appendix B.

2.1 Results

We again find evidence that subjects strategically express emotions. The mean valence when the task is desirable is 0.24, against a mean valence of 0.15 when the task is undesirable, a difference that is statistically highly significant (signed-rank test, $p < 0.001$). The majority of subjects (73 percent) displays a higher valence in the desirable task. In qualitative terms, this is consistent with the results from the previous experiment.

In terms of magnitude, the difference between tasks is much more modest. Fig-

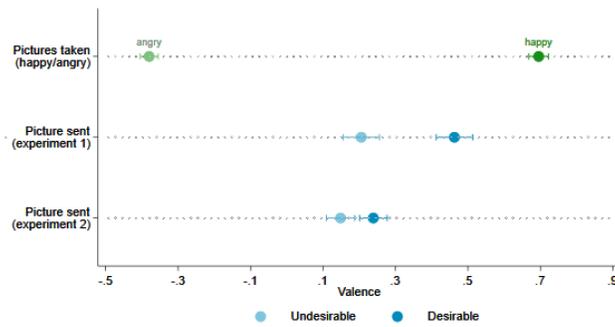


Figure 5: Valence on pictures taken and pictures sent. Error bars indicate ± 1 s.e.

Figure 5 shows the mean valence on the happy and angry pictures (green dots) from experiment 1. The difference of 1.07 gives an indication of what subjects can achieve in terms of expressing different emotions. The actual difference in valence between the tasks is lower, because subjects do not always send the happy picture when the task is desirable and the angry picture when the task is undesirable. The difference in valence on the pictures that are actually sent is about 0.22 in experiment 1 and 0.09 in experiment 2 (blue dots). Thus, the variation in intensity of expressed emotions between tasks in experiment 2 is only 41 percent of that in experiment 1 and only 8 percent of the potential difference.

The kernel estimates of the distribution of valence on the pictures sent by subjects provide additional support, see Figure 6. In both experiments we see a shift in the distributions across the different tasks, so subjects do respond to the situation they face. In experiment 1, the valence has substantial mass almost everywhere on the scale. By contrast, in Experiment 2 it is much less dispersed and largely concentrated around 0, with little mass below zero.

While subjects in Experiment 2 vary their expressions to a lesser extent, the potential benefits appear to be of a similar magnitude. Table 4 shows the effect of valence on the likelihood of becoming the designated player. The odds ratios are above one for both tasks and both experiments. The coefficient is not significant for the the undesirable task in Experiment 2 (column 4). This is probably due to a lack of power; few subjects express strong emotions in that treatment, and especially negative emotions are mostly absent. The odds ratio is, however, in the same

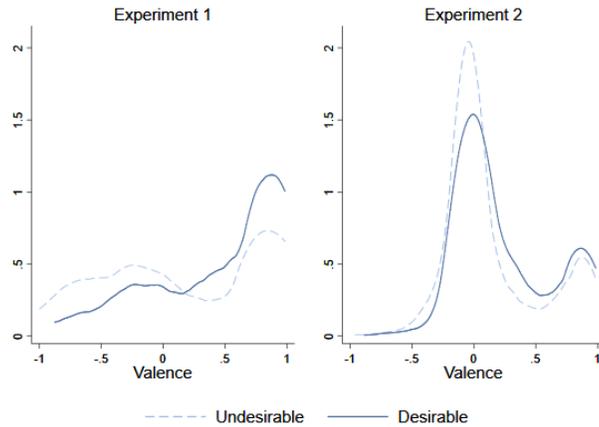


Figure 6: Kernel estimates of the distribution of valence on pictures sent. Error bars indicate ± 1 s.e.

ballpark as that in Experiment 1 (column 2).

We conclude that subjects spontaneously adapt their expressed emotions to the situation, but they only exploit the benefits to a modest extent.

2.2 Heterogeneity

We do not find much evidence of heterogeneous responsiveness depending on background characteristics. We measure responsiveness as the difference in valence between the two types of tasks. Column 2 in Table 3 shows how the increase in valence between the desirable and undesirable task depends on several background characteristics. Females again appear less responsive, but as in the previous experiment the coefficient is not significant. Subjects' scores on the eyes test and strategic reasoning are not significantly related to responsiveness.

2.3 Investments

Waiting for trust measures...

Table 4: Assignment of task.

Task	(1) Desirable Experiment 1	(2) Undesirable Experiment 1	(3) Desirable Experiment 2	(4) Undesirable Experiment 2
Valence	1.231* (0.148)	1.447*** (0.164)	1.898** (0.531)	1.360 (0.341)
Observations	820	792	412	432

Conditional logit estimates (odds ratios). Dep. var.: being the designated player. Robust s.e.in parentheses clustered at the manager level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3 Discussion and Concluding Remarks

-why reluctance to send angry pic? perhaps pessimistic that other invests. perhaps they want to keep control. perhaps they don't like it (a la 'lying cost' or because they like to show nice pic of themselves). perhaps they don't believe it would work.

Appendix A Survey Measures

Level of Strategic Reasoning. The level of strategic reasoning is measured using an adapted version of the racing game (see Gneezy and Rustichini, @@@). In this game, there is a number of chips available. The subject and computer take turns, and can take 1 or more chips each turn. The player to take the last chip wins. We implemented two versions. In the first game, they started with 15 chips and they could take 1,2, or 3 chips each turn. In the second game, they started with 17 chips and they could take 1,2,3, or 4 chips each turn. The subject always started first. We programmed the computer such that at every round, there was a winning strategy for the subject if she took the appropriate number of chips from that round onwards. A subject's score is the number of rounds that she followed a winning strategy. Once the number of chips remaining is less than the maximum number of chips that a subject can take, the solution is trivial and those rounds are not scored. The mean score is 0.497 (median 0.5, stand. dev. of 0.183).

Level of Emotional Intelligence. We have two measures of emotional intelligence. The first one is a subject's score on the 'Reading the Mind in the Eyes Test' (Baron-Cohen, @@@). This is a standard test in psychology, in which pictures of eyes are shown together with a set of words describing emotions. The subject's task is to select the word that best fits the expression on the picture. There are 36 pictures in total and the average score is 25.7 (median 26, standard deviation 4.3).

Our second measure is based on Van Leeuwen et al. (@@@@). We showed ten pictures of participants in a previous experiment who all received a low offer in an ultimatum game. Exactly five of those participants rejected the offer. We asked our subjects to predict which of the participants rejected the low offer. ...@@ to compute...

Appendix B Details Procedures Experiment 2

In the first session of the no-posing treatment, some subjects used hand gestures on the pictures and some considerably slowed down the experiment by taking many selfies before submitting their picture. We therefore slightly modified the instruc-

tions in that treatment, and told participants they were not allowed to use hand gestures, and limited the number of possible selfies to three. An experimenter inspected the submitted pictures. Subjects seldomly violated the instructions. If they did, they were asked to retake the picture.

Appendix C Robustness

perhaps exclude those who took many attempts test questions, non-compliant

Appendix D Instructions