Psicológica (2013), 34, 179-198.

Interpersonal effects of emotion in a multi-round Trust Game

María I. Tortosa, Tatiana Strizhko, Mariagrazia Capizzi & María Ruz*

University of Granada (Spain)

Emotions displayed by others are pivotal ingredients of the decisions we make in social contexts. However, most of the research to date has focused on the subjective emotion of the decider rather than on the emotional expressions of the partners in the interaction. The present investigation was designed to explore how happy and angry facial expressions modulate cooperative responses in multi-round Trust Games. Our results show that happy partners generate higher levels of trust than angry partners even after repeated experience in a context in which emotional displays are not predictive of the partners' cooperation rates. This effect disappears once the social meaning of emotional displays is removed from the game. An additional study shows that participants are able to learn specific associations between discrete emotions and positive or negative cooperative tendencies, although they need more evidence when the associations counteract prior expectations. Overall, our results stress the reliability of discrete emotions as cues in interpersonal interactions and the resilience of the effect of these positive and negative cues in contexts in which they lack real predictive power.

Decision-making in social contexts is guided not only by our knowledge of the purpose and rules of the interactions between people, but also by our beliefs and expectations about those with whom we interact (e.g. Ruz, Moser & Webster, 2011). One important factor influencing social decision-making is emotion (Adolphs, 2003; Olsson & Ochsner, 2008). A

* Acknowledgments: Financial support to this research came from the Spanish Ministry of Science and Innovation through a 'Ramón y Cajal' research fellowship (RYC-2008-03008) and grant PSI2010-16421 to MR, grant P09.HUM.5422 to Pío Tudela and MR, plus the Andalusian Autonomous Government through grant SEJ2007.63247 to Pío Tudela and MR. Correspondence should be addressed to María Ruz. Mind, Brain and Behaviour Research Centre. Dept. of Experimental Psychology, Campus de Cartuja s/n, 18071 Granada, Spain. E-mail: mruz@ugr.es

branch of research has focused on the *intrapersonal* effects of emotion, or how moods or emotional states of the decider affect judgment and decision-making in several situations (see Angie, Connelly, Waples & Kligyte, 2011, for a recent review). For example, the induction of positive or negative emotional states on the decider increases or decreases trusting behaviors respectively (Dunn & Schweitzer, 2005; see also Harle & Sanfey, 2007). However, the interpersonal effects of emotions, or how the emotions expressed by other people could be used as cues to predict their most likely behavior, have been much less explored to date (Van Kleef, de Dreu, & Manstead, 2010).

Given their functional role in communicating intentions (Darwin, 1872; Fridlund, 1995; Keltner & Haidt, 1999), facial expressions of emotion are especially strong candidates to influence trust decisions in social encounters, as they play a major role in indicating when a person is willing to be cooperative and trustworthy and when a person is not (Buck, 1984; Boone and Buck, 2003). Along evolution, and sustained by the basics of classical conditioning, we have learned to associate different emotions with specific meanings. In most cases, positive emotions such as happiness predict positive consequences whereas negative emotions such as anger indicate that bad things may happen (Darwin, 1872; see Ruz & Tudela, 2011; Ruz, Madrid & Tudela, *in press*). These associations, crafted along the years, help us to adjust our behavior in light of the predispositions of others.

The link between emotions and social decision-making also emerges in research conducted in the field of social neuroeconomics (Fehr & Camerer, 2007; Sanfey, 2007), which widely employs experimental economic games to study the patterns of social behaviors in interactive situations. This research indicates that decision makers do not always behave "rationally", or follow the strategy of strict self-interest and individual maximization (Camerer, 2003). Some of the most common games are the Ultimatum and Dictator Games, the Prisoner´s Dilemma or the Trust Game. The last two have a similar structure, in the sense that mutual cooperation provides better payoffs tan mutual defection. The experimental Trust Game (Berg, Dickhaut, & McCabe, 1995; Camerer & Weigelt, 1988) involves a minimum of two players, a *trustor* and a *trustee*. The trustor is endowed with a sum of money and has to decide whether or not to share it with her/his game partner. If s/he keeps the money for her/himself, the trustee gets nothing. If s/he decides to share, the trustee receives the initial endowment multiplied by an amount. If the trustee then reciprocates the trust, the sum is divided between the two players; otherwise the trustor obtains nothing. In this game, the decision of the trustor is

hazardous because the trustee's reciprocation is not enforced by the rules. Still, substantial amounts of trust are observed across studies (Berg et al., 1995). These effects are attributed to altruism and reciprocation, which activate reward brain circuits (Fehr & Camerer, 2007).

Social preferences for trust are not unconditional but depend on the belief that the partner is likely to reciprocate the trust (Camerer, 2003), combined with the general tendency of people to trust others (Berg et al., 1995). Two different studies have shown that happy facial expressions –either schematic line drawings (Eckel & Wilson, 2003) or photographs of people (Scharlemann, Eckel, Kacelnik, & Wilson, 2001; see also Averbeck & Duchaine, 2009)- generate higher levels of *initial* trust in one-shot games. Moreover, the choices seem to be influenced by the facial dynamics that distinguish between genuine and fake smiles (Krumhuber, Manstead, Cosker, Marshall, Rosin, & Kappas, 2007; Niedenthal, Mermillod, Maringer, & Hess, 2010), with authentic smiles generating higher cooperation rates.

The studies that have explored the evolution of trust in an iterated exchange presented either no photos, thus no information regarding the partners' displayed emotions (King-Casas, Tomlin, Anen, Camerer, Quartz, & Montague, 2005), or 'expressionless' neutral photos of the game partners (Delgado, Frank, & Phelps, 2005). Other recent studies that have used multi-round exchanges have focused on the effects of negative emotion as anger on future interactions in computer-mediated negotiations (Van Kleef et al. 2010) and disappointment vs. anger with the tit-for-tat strategy (Wubben, De Cremer, & van Dijk, 2009). They did not present faces though, but statements expressing the emotions in question.

Our study explored how the behavior elicited by emotional facial expressions of happiness and anger is maintained over the course of an extended social interaction. Along three experiments, we investigated: (1) whether happy and angry facial expressions of partners in a Trust Game modulate trust decisions even after several rounds in which such emotional expressions are not predictive of the partners' cooperation rates; (2) whether these long-lasting effects remain once emotional displays are not linked to the partner in the game; and (3) whether people are able to use happy and angry facial expressions as cues that predict their natural or their unnatural consequences in inter-personal situations (i.e. non-cooperative and cooperative tendencies, respectively). Our hypotheses predict: (1) a higher rate of cooperative responses after happy than after angry emotional expressions; (2) no effect of emotional expressions when they are displays presented at random by the computer, without social significance; and (3) a

rapid association of happy and angry emotional expressions with cooperative and non-cooperative purpose/intent, while a delayed learning of associations that are not consistent with initial priors linking emotions and their most likely consequences.

METHODS

EXPERIMENT 1

Participants. 32 students (8 males, mean age of 21 years) from the University of Oxford (12) or the University of Granada (18) participated in exchange for course credits. They all signed a consent form approved by the local Ethics committees and received a chocolate token for their participation. They all had normal or corrected-to-normal vision. Two of them were excluded from analyses because they did not have enough observations in all conditions.

Stimuli and procedure. At the beginning of the session, participants were instructed that the experiment explored the cooperation patterns that emerge between people during the so-called Trust Game. They were going to play multiple rounds with three different players over the course of the game. At the beginning of every round, participants were presented with a symbolic Pound/Euro and had to decide whether to keep it (by pressing the *k* on the keyboard) or share it with their partner (by pressing *s* on the keyboard). The keep decision would yield no earnings for the partner at the end of the trial. The share decision would result in £5/5 Euro given to the partner who, in turn, would decide whether (1) to reciprocate the cooperation, in which case each of them would receive £2.5/2.5 Euro; or (2) not to reciprocate, in which case the participant would receive nothing because the partner kept the £5/5 Eur. The participants' goal was to maximize their payoffs in the game, and they were told that *mutual* cooperation was the best strategy for reaching this goal. They could, however, to the best of their knowledge freely decide on every trial whether to trust their partner or not. The prize for maximizing their payoffs was a chocolate bar once the game was finished. Participants were also told that even though partners were represented by photos on the computer, their behavior mimicked normal patterns of play by real people. Therefore, they did know that they were not playing against real people on-line, but the instructions stressed that the face photos represented the choices of normal partners. Participants were not told about the different emotions that

partners would be displaying, and thus they were unaware of the main goal of the study, which was to explore how emotions influenced cooperation rates. At the end of the game, all participants received a chocolate bar.

The game was presented on a PC running E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Frontal photographs of three female (see Aguiar, Brañas-Garza, Cobo-Reyes, Jimenez & Miller, 2009) white faces displaying happy, neutral, or angry emotional expressions were selected from the NimStim face stimulus set (Tottenham, Tanaka, Leon, McCarry, Nurse, Hare et al., 2009) to represent the participants' partners. All stimuli were presented against a grey background (see Fig. 1). Every trial began with a 1-sec presentation of a Pound or Euro symbol, replaced by a fixation point (+) for another 1 sec, and was followed by the picture of the partner for that trial. After 1.5 sec, the picture was replaced by the fixation point for 1 sec. Next, a question mark (?) was displayed for 2.5 sec; this served as a prompt for the participants to make their decision. Finally, participants received feedback about the payoffs in the trial (see Fig. 1), which was displayed for 2 sec. If they decided to keep the money, the message was 'You have decided not to cooperate. You add £1/1 Euro and your partner adds £0/0 Euro'. If they decided to share, either 'You have decided to cooperate. Your partner receives £5/Eur and decides to correspond. You add £2.5/2.5 Euro and your partner adds £2.5/2.5 Euro' or 'You have decided to cooperate. Your partner receives £5/5 Euro and decides not to correspond. You add £0/0 Euro and your partner adds £5/5 Euro' were presented. Participants played this game 84 times with each of the three partners (for a total of 252 trials), who on a trial-by-trial basis displayed random happy, neutral, or angry emotions with equal probability and reciprocated at the constant rate of 50% regardless of their emotional expression. The session lasted about 30 minutes.

We analyzed the percentage of participants' cooperation rates across conditions. Our main variable was the emotion displayed by the partner, which was manipulated at 3 levels: happy, neutral, and angry. In addition, we included two more variables in the design. First, responses were divided in blocks of 50 trials to examine the effect of practice with the task. Second, we included the feedback that participants received from their partner in the previous trial to explore how this may have affected their subsequent decision (3: non-cooperation, non-reciprocated cooperation and reciprocated cooperation). Thus, the average levels of acceptance rates per participant and condition were analyzed with repeated-measures ANOVAS with the factors Emotion (3) x Block (5) x Feedback (3).

The Greenhouse-Geisser correction for non-sphericity was applied where appropriate (Jennings and Wood, 1976) to adjust the degrees of freedom. Only corrected probability (as well as epsilon) values are reported.

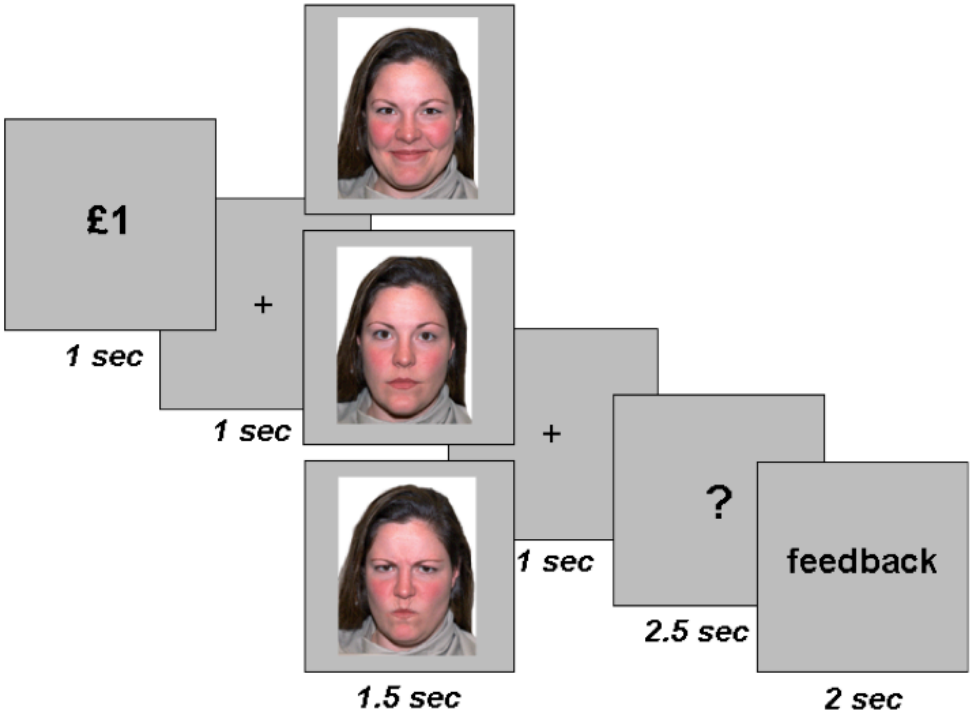

Figure 1. Sequence of events in a trial.

RESULTS

Participants cooperated on 62% of the trials (SD=33.7). Results showed a main effect of Emotion, $F_{2,58}$=12.53, p<0.001, ε=0.710, ηp^2=0.302, as participants cooperated more with happy than with angry partners (70.4% vs. 49.4%), $F_{1,29}$=14.83, p<0.001. The difference between angry and neutral (65.1%) was also significant, $F_{1,29}$=13.19, p=0.001, but between happy and neutral it only reached marginal significance, $F_{1,29}$=3.66, p=0.066 (see Fig. 2). The effect of Block of trials was not significant, F<1. Nevertheless, to test whether the factor Emotion was still significant after repeated experience with the game, we evaluated the effect of the partners' emotional expression in the last block of trials (5). In this block, cooperation rates were still lower for angry (44%) than for happy (67%), $F_{1,29}$=11.8, p=0.001 or neutral partners (64%), $F_{1,29}$=9.41, p<0.01. The

difference between happy and neutral partners was not significant, $F<1$. The variable Feedback did not modulate cooperation rates, $F<1$, and no interaction reached significance levels (all $ps>0.1$).

An additional analysis was performed to test for potential differences between participants from Oxford and Granada universities. However, this between-subject variable did not produce any reliable effect (all relevant $Fs<1$).

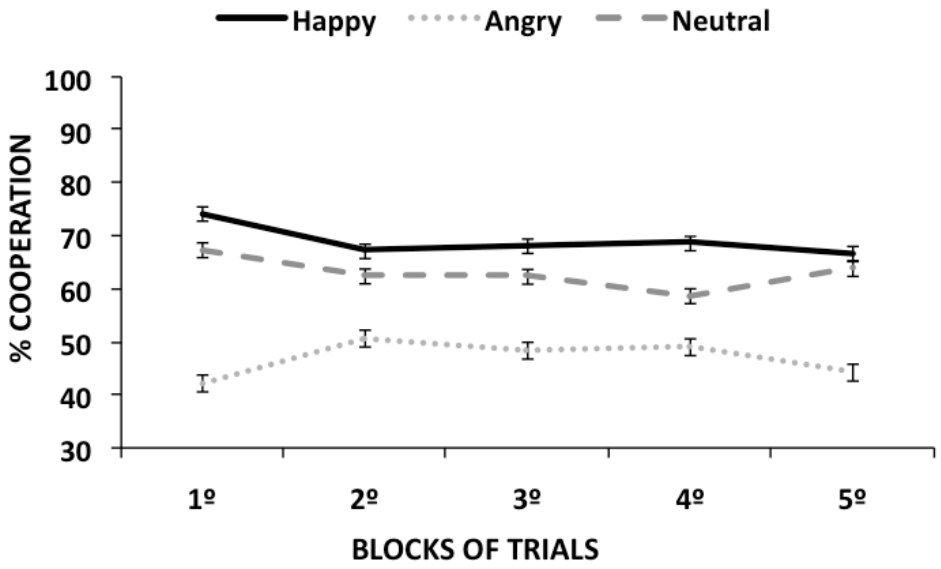

Figure 2. Participants' cooperation rates with partners displaying happy, angry or neutral facial expressions along five blocks of trials in Experiment 1.

DISCUSSION

In Experiment 1, participants cooperated more with happy and neutral partners than with those displaying angry facial expressions. This result fits with previous findings that relate happiness expressions with trustworthiness (Eckel & Wilson, 2003) or signals of cooperative intents (Fridlund, 1995). In addition, it extends the results of experiments that showed that happy expressions relate to initial levels of high cooperation

rates in single-round Trust Games (Eckel & Wilson, 2003; Scharlemann et al., 2001).

The cooperation rate for neutral was closer to happy than to angry expressions. The fact that no-emotion had a similar meaning of happiness could be explained within the frame of the positive-negative asymmetry effect (Baumeister, Bratslavsky, Finkenauer & Vohs, 2001). According to this approach, negative events, emotions, and information have more impact than positive ones and contribute more strongly to form an impression. Thus, the effect of positive interactions (i.e., with happy expressions) would be closer to the effects of neutral interactions, whereas bad ones (angry expressions in our case) would be clearly different and distinct. In addition, people are more concerned with avoiding bad feedback than with maximizing positive feedback (Tice, 1991). If, as predicted by evolutionary and classical conditioning theories, angry facial expressions are associated to negative consequences in contrast to positive and neutral expressions, angry expressions would have stronger effects than positive or neutral ones.

The lack of interaction between the factors of emotion and block of trials, and the significant effect of emotional expressions on participants' cooperation rates during the last block of trials, strongly suggests that this emotion-related bias persists even after repeated experience in a game setting in which emotional expressions are not predictive of the partner's reciprocation of cooperation. In addition, the feedback that participants received on the previous trial seemed to have no effect on subsequent trust decisions, as cooperation rates were not altered across feedback conditions.

However, at this point there is a potential and less appealing alternative explanation to our results. The effect that emotional expressions had on the cooperation rates of participants may derive from general mood effects or affective priming (i.e. cooperation may be enhanced by the mere display of positive stimuli) rather than by the social meaning of emotions. From this perspective, the presentation of any emotional material unrelated to social interactions may have produced the same results due to their condition of primes with evaluative value.

To test this alternative hypothesis, we set up Experiment 2. Instead of introducing additional confounds by changing the nature of the affective stimuli, we repeated the exact same procedure using emotional photos of the partners, but modified the instructions that participants received to devoid the emotional expressions of their social meaning.

EXPERIMENT 2

Participants. 32 students (1 male, mean age of 21.9 years) from the University of Granada participated in exchange for course credits. They all had normal or corrected-to-normal vision.

Stimuli and procedure. The experiment was the same as the previous one, except for the modified instructions. The background story remained the same. They were going to play a Trust Game with several different partners, who were represented by photos on the computer but had a behavior that mimicked normal patterns of play by real people. Their goal, again, was to maximize pay-offs and they were told that *mutual* cooperation was the best strategy for reaching this goal. However, whereas participants in Experiment 1 were not told about the emotions that the partners would display, in the current one they were informed that the computer *randomly* assigned different emotions to the partners. As in the previous experiment, though, they were reminded that the identity of the partner was relevant for the game.

The design included the factors Emotion (happy, angry, neutral), Block (5) and Feedback (noncooperation, non-reciprocated cooperation and reciprocated cooperation) as within-subject variables.

RESULTS

Participants cooperated on 61.9% of the trials (SD=12.8). There was a main effect of the Feedback, $F_{2,34}$=4.79, p=0.01, ηp^2=0.220, as participants tended to cooperate more after not cooperating in the previous trial (66.2%) compared to when the feedback was both a non-reciprocated, $F_{1,17}$=4.77, p<0.05; 59.6%), or a reciprocated cooperation, $F_{1,17}$=8.26, p=0.01; 56.5%). In contrast to the previous experiment, there was no effect of Emotion, (F<1; 59.3% vs. 59.8% vs. 63.1%, for angry, happy and neutral, respectively). No interaction reached significance levels (all ps>0.1).

DISCUSSION

Results from Experiment 2 indicate that the biasing effect of emotion on cooperation rates in a Trust Game was not due to automatic mood or affective priming effects. If this was the case, in the current experiment we should have found the same effects as in Experiment 1, because everything remained the same except for the association of the partners to their emotional expression in the context of a social interaction. Instead, we

found that the cooperation rate of participants was equal across happy, neutral and angry partners. This suggests that the link between the partners and their expression in a social context drove the effect of emotions observed in Experiment 1, and not mere mood or affective priming effects (see also Ruz, Moser and Webster, 2011; Gaertig, Moser and Ruz, 2012).

Together, experiments 1 and 2 suggest that the priors that we hold relating happy emotional expressions with cooperative consequences are strong (see Averbeck & Duchaine, 2009; Ruz & Tudela, 2011) and resistant to evidence that regards them as not informative (as emotional expressions did not predict the partners' cooperation intents). If this were so, we would expect that in game context in which emotional displays predicted the cooperative consequences that they are naturally associated to, participants would adjust their cooperative behavior in a fast way. On the contrary, if the game context associated emotions with the opposite of their natural associations, people would need more evidence (i.e. trials) to adjust their behavior. These were the manipulations in Experiments 3.

EXPERIMENT 3

Participants. 26 students from the University of Oxford (13 females, mean age 21.5) participated in the experiments. They all signed a consent form approved by the local Ethics Committee. All of them had normal or corrected-to-normal vision.

Stimuli and procedure. The current experiment followed the same structure as Experiment 1 but added a contingency between facial expressions and reciprocity rates. For one group of participants (bias-consistent), emotions in the game were associated to their natural consequences. All partners displayed happy, angry and neutral expressions with equal probability (33.3%), but their reciprocation rate depended on their emotional display. Each reciprocated on 80% of the trials in which their facial display was happy, on 20% when it was angry, and on 50% when they had a neutral expression. For another group of participants (bias-inconsistent), the contingency between facial expressions and reciprocity rates was reversed, and now the emotion displayed by the partner predicted the opposite of their natural consequences. Thus, partners reciprocated in 20% of the trials when their facial expression was happy, 80% when it was angry, and 50% when they had a neutral display. Only partners' expressions and not identities were predictive of their trustworthiness. The design was multifactorial with Emotion (Happy, Angry, Neutral), Block (5) and

Feedback (non-cooperation, non-reciprocated cooperation and reciprocated cooperation) as within-subject variables, and Group (bias-consistent, bias-inconsistent) as between-subject factor.

RESULTS

Mean cooperation rate was 63.7% (SD=33.2). The ANOVA showed an interaction between Experimental Group, Emotion, and Block, $F_{8,192}$=5.34, p<0.001, ε=0.675, ηp^2=0.182, due to differences between groups in cooperation rates across the blocks.

In the bias-consistent group, where contingencies between emotion and cooperation rates were the expected, there was a main effect of Emotion, $F_{2,24}$=31.85, p<0.001, ε=0.701, ηp^2=0.726. Participants cooperated more with happy (85.7%) than with neutral (67.9%), $F_{1,12}$=18.47, p=0.001 or angry partners (35.9%), $F_{1,12}$=40.42, p<0.001 (see Fig. 3). The difference between neutral and angry partners was also significant, $F_{1,12}$=24.65, p<0.001. Crucially, there was no interaction with Block, F<1, as the effect of Emotion was constant along the task. In this group, the variable Feedback modulated cooperation rates, $F_{2,24}$=4.27, p<0.05, ηp^2=0.263. Participants cooperated more after a reciprocated cooperation feedback (68.2%) than after a non-reciprocated one (58.9%), $F_{1,12}$=5.49, p<0.05, and more than after a non-cooperative decision (63.2%), $F_{1,12}$=5.75, p<0.05. There was no difference in participants' cooperative behavior between non-cooperation and non-reciprocated feedbacks, $F_{1,12}$=1.12, p>0.3.

In the bias-inconsistent group, where contingencies were reversed, there was a main effect of Emotion, $F_{2,24}$=5.00, p<0.05, ηp^2=0.294. Overall, cooperation rates were higher for angry (76.8%) than for happy (51.0%) partners, $F_{1,12}$=6.38, p<0.05, and higher for neutral (65.4%) than for happy partners, $F_{1,12}$=6.05, p<0.05, with non-significant differences between angry and neutral, $F_{1,12}$=2.12, p>0.1 There was also an interaction between Emotion and Block, $F_{8,96}$=7.37, p<0.001, ηp^2=0.381, which showed that the effect of Emotion appeared as participants acquired practice with the game contingencies. In the first block, cooperation rates tended to be equal for happy (65.9%), angry (61.4%) and neutral partners (67.5%), all Fs<1. Block 2 showed a main effect of Emotion, $F_{2,24}$=3.53, p<0.05, ηp^2=0.227, that increased until Block 5, $F_{2,24}$=22.91, p<0.001, ηp^2=0.656. In this last block, cooperation rates were higher for angry (93.1%) than for happy (43.8%), $F_{1,12}$=42.53, p<0.001, and neutral partners (56.6%), $F_{1,12}$=27.21, p<0.001, although there were no differences between neutral and happy

partners, $F_{1,12}$=2.52, p>0.1 (see Fig.4). In this bias-inconsistent group, the variable Feedback did not modulate cooperation rates, F<1.[1]

DISCUSSION

Participants learned the contingency between a facial expression and the reciprocity rate, as their cooperation mimicked the rates assigned to each type of emotions. They were able to learn such associations when they were consistent with the expectations set by the emotional expressions, but also when they were inconsistent with the links that we hold naturally between valence of emotions and trustworthiness. Participants from the bias-inconsistent group, however, arrived at the 'correct' contingencies later than those in the bias-consistent group, as evidenced by the significant interaction between Group, Emotion and Block. Thus, whereas there were no discrepancies between the expectations generated by the emotional expressions and the behavior of the partners in the bias-consistent group, the initial predictions failed in the bias-inconsistent group, and thus they needed more time to learn the correct cooperation rates.

In addition, and in contrast to Experiment 1, participants' trust decisions were influenced by the previous cooperation feedback from their partners, but only in the bias-consistent group, as they cooperated more after their partners in the previous trial reciprocated the cooperation than when they did not. However, the reciprocity feedback obtained from the previous trial did not modulate cooperation rates in the bias-inconsistent group. We will turn to why this may be so in the General Discussion section.

In sum, these results bear on the bias that people have to hold positive expectations when confronted with an expression of happiness, and negative expectations when confronted with an expression of anger. Most important, our results show that people are able to learn associations between emotions and cooperation tendencies that mismatch their priors (Averbeck & Duchaine, 2009; Ruz & Tudela, 2011) and adapt their

[1] We performed an additional meta-analysis to measure the extent of the association between emotional displays and cooperation behaviour of participants in experiments 1 and 3 together. For every participant, we obtained an index of the differences in percentage of cooperation with happy vs. neutral partners, and correlated this value with the specific contingency between happiness and partner's reciprocation rates (50% in Experiment 1, 80% in Experiment 2a and 20% in Experiment 2b). Results show a strong correlation between these two variables, $r=.53$, $p<.001$.

behavior to such unnatural associations, even though this takes an extended experience.

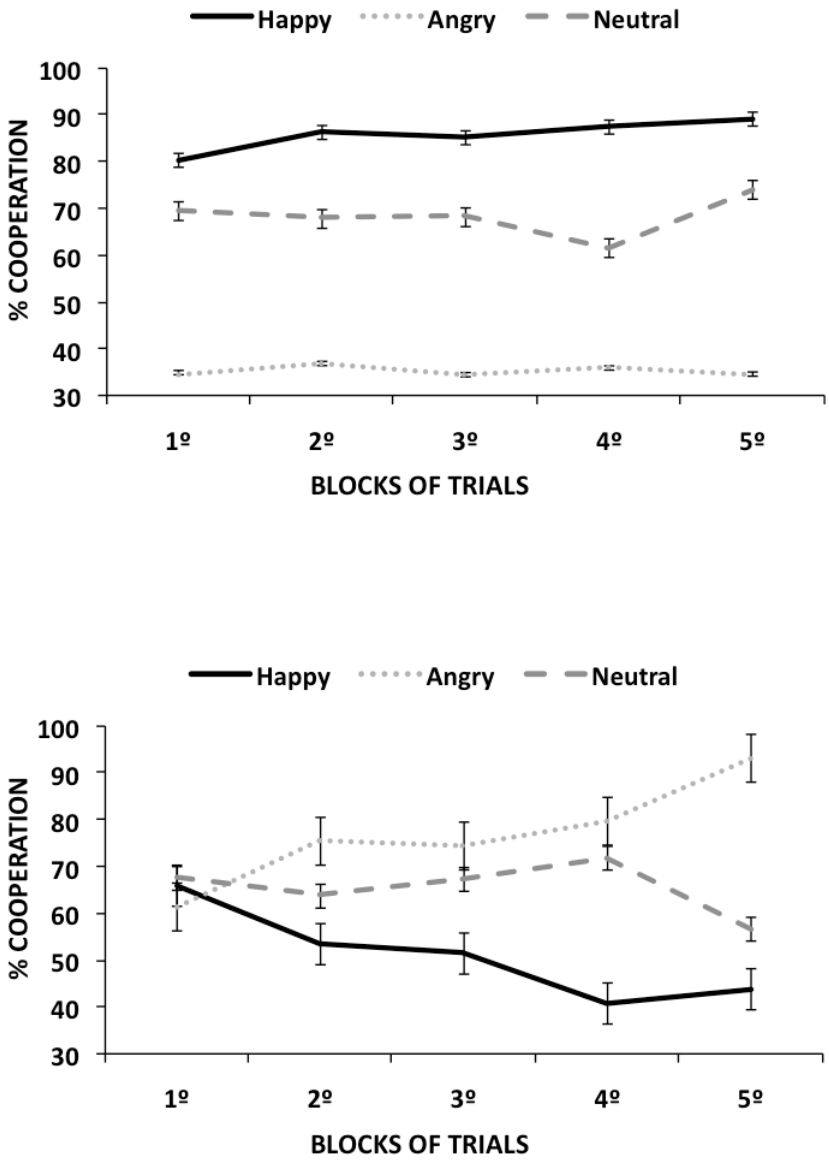

Figure 3 y 4. Participants' cooperation rates with partners displaying happy, angry or neutral facial expressions along five blocks of trials in Experiment 3.

GENERAL DISCUSSION

The present study explored the effect of the emotions displayed by partners in a multi-round trust decision-making setting. Experiment 1 revealed that even after prolonged exposure to a game context in which the emotions displayed by the partners were not predictive of their cooperation rates, participants still cooperated more with happy than with angry partners. Experiment 2 supported the idea that this effect was not due to automatic affective biases generated by the mere presentation of emotional stimuli. In Experiment 3, participants quickly adapted their behavior in a game in which emotions predicted their natural consequences and, although it took them longer, they were able to adapt their cooperation rates to the game contingencies when the association between emotion and cooperation rates was counter-intuitive.

Overall, our results are in line with previous research showing than happy and angry emotional expressions modulate initial levels of trust in interpersonal encounters (Eckel & Wilson, 2003), and support models that posit that facial expressions of emotion are salient cues employed to predict the behavior of others in interpersonal social contexts (Van Kleef, de Dreu & Manstead, 2010). In addition, results from our Experiment 1 extend this literature by showing that the differential levels of trust that happiness and anger generate persist after repeated lack of evidence that emotions are of any use in the game, as they were not associated in any manner to the partners' cooperation rates. Despite this, after more than 200 trials (in the last block of the game), happy partners still generated significantly higher levels of trust than angry ones, and the participants' cooperation rates with partners displaying the latter expression were lower than with those with neutral facial displays. Such bias disappeared completely when the emotions were devoid of their association with the partners by telling participants that they were placed at random by the computer program controlling the game.

The Trust Game we used could be conceived as a cooperative setting, as instructions stressed that *mutual cooperation* between the participant and his/her partners was the best strategy to maximize the payoffs in the game, which was the goal that participants had to fulfill. According to the model proposed by Van Kleef, de Dreu and Manstead (2010; see also Van Kleef, 2009), the effects of the partner's emotions could be mainly driven by affective contagion (e.g. Parkinson & Simons, 2009). Happy partners would engender positive feelings in the participants, who would perceive the exchange as safe and would thus feel that their partner would reciprocate their initial trust, whereas angry partners would generate negative emotions

which would move the participant away from their partner and thus reduce cooperative intents.

The explanation above does not exclude taking a social functional approach, from which emotional expressions and action tendencies were selected because "they produced consequences that improved the individual's inclusive fitness" (Keltner & Haidt, 1999). Emotions can act as coordinators of social interactions by *informing about others' emotions, beliefs, and intentions;* thus, emotional displays could *also* be used strategically in the current game setting. As in Experiment 1 the association between emotion and reciprocation consequences was random, participants may have disregarded the cooperation behavior of their partners to be guided only by their emotional expressions, as suggested by the lack of effect of feedback information in this experiment.

Also partial reinforcement's theories could help explain the effect of cooperation in Experiment 1. In real life, behaviors learned through partial reinforcement are resistant to extinction. In the current experimental setting it could be the case that the feedback of reciprocity acted as a partial reinforcement for the response of sharing (specially with happy people), which would help maintain responses across the blocks even after repeated lack of evidence that emotions are of any use in the game.

Experiment 3, however, show that once happy and angry emotional displays provide information regarding the partners' behavior, participants are able to use these cues strategically to guide their choices and adapt their trust levels accordingly. Results from Experiment 3 suggest that the initial associations between emotions and their consequences are taken into account in the decision-making process, as evidenced by the slow-down in the group learning bias-inconsistent associations between emotions and cooperation rates. In this line, Ruz & Tudela (2011; see also Ruz, Madrid & Tudela, *in press*) studies suggest that the natural associations between emotions and their consequences are difficult to override and need of additional conflict detection and cognitive control mechanisms, as evidenced by behavioral interference indices, an enhanced frontal N1 potential and neural activation in the anterior cingulate and prefrontal cortices in conflictive situations (Ruz & Tudela, 2011; Ruz, Madrid & Tudela, *in press*). Also, Averbeck & Duchaine (2009) using a non-social reward task, showed that people have a *prior bias* to select happy over angry faces as potential sources of reward. Thus, in our experiments the natural expectations engendered by emotions modulated the amount of evidence needed to combine the emotional expressions with their specific associations in each of the games. Whereas in the bias-consistent group of

Experiment 3 the effect of emotions remained along the blocks (as evidenced by the lack of interaction between these two factors), in the bias-inconsistent group the interaction between emotion and block of trials suggests that participant needed prolonged evidence with the game contingencies to be able to grasp that happy emotions predicted lack of cooperation whereas angry expressions led to higher cooperation rates.

One intriguing aspect of the present set of studies is the lack of evidence that participants used the feedback from the previous trial to influence their current decisions. Whereas in the bias-consistent group of Experiment 3 we observed an effect of feedback in the expected direction (higher cooperation rates after reciprocated than after non-reciprocated trust), there was no hint of such effects in the bias-inconsistent group or in Experiment 1. One possible explanation for this is that our experiments lacked the common association between partners' identity and cooperation tendencies. That is, in none of the experiments the specific identity of the partner predicted cooperation rates (as all partners cooperated at a constant 50% rate). This may have led participants to disregard the feedback from the previous trial, tied to the partner's identity, as relevant information to guide their judgment in the following trial. In contrast, the structure in which the contingencies between emotion and reciprocation rates followed natural expectations, may have made feedback more salient for participants, which may have led them to include it as a relevant factor in their decision. This would fit with the notion that people tend to take more into account the facts (i.e. the feedback) that support their beliefs (for an effect akin to selective exposure, see Kleinhesselink & Edwards, 1975). Future studies associating different identities to differential cooperation rates, combined with emotional displays, should be conducted to help disentangle this matter.

There are some details of the current experiments that limit the scope of our conclusions. First, due to the limited availability of male students at the time data were collected, most of the participants were women, which calls into question the reliability of our results in a male population. To date, however, we have performed some other experiments using similar emotion manipulations with the same Trust Game including equal number of male and female participants, and the factor of gender has never generated a main effect or interacted with any of the variables in the designs (Tortosa, Lupiañez & Ruz, 2013).

Second, the experimental setting is rather artificial, which makes the extrapolation to real-life social situations more difficult and qualifies the scope of the conclusions that may be derived. We tried to minimize this

problem by stressing that the responses of the partners mimicked behavior in real situations. Future studies, however, could be improved on this respect by using videos instead of static pictures of the partners, and/or using actual people as partners. The use of static pictures, however, allowed better experimental control, which could be beneficial, for example, to aid in the adaptation of the paradigm to future electrophysiological experiments.

A drawback related to the artificial nature of the experiments is the extent to which our results can be extrapolated to actual social interactions in which the emotions of the partners are used to infer their internal states in relation to their future cooperative tendencies. It could be argued that emotional faces should be considered as symbolic primes with evaluative value rather than components of actual social interactions. Results from Experiment 2 showing that the effect is lost once emotions are devoid of their social meaning, however, argue against a pure automatic effect from any evaluative prime. In any case, the resolution of this dichotomy falls outside the current study, and should be aided by the use of more naturalistic settings (with real partners or videos of actors, for example).

Overall, our results stress the strong impact of the emotional displays of the people in shaping our decision tendencies, and the resilience of the effect of these positive and negative cues in contexts in which they lack real predictive power. Future research should be aimed at investigating the effect of stable emotional states of others in trust behaviors, and also at exploring the role of other pieces of social information that may be relevant to trust tendencies.

RESUMEN

Efectos interpersonales de la emoción en un Juego de Confianza. Las emociones que otras personas expresan juegan un papel importante en las decisiones que tomamos en contextos sociales. Sin embargo, la mayoría de la investigación hasta la fecha se ha focalizado en la emoción subjetiva de la persona que toma la decisión, en vez de en la emoción mostrada por los compañeros en una interacción. Nuestro estudio se diseñó para explorar cómo las expresiones de felicidad y enfado de otras personas afectan a las respuestas de cooperación en un Juego de Confianza de interacciones múltiples. Los resultados muestran que los compañeros felices generan niveles de cooperación más altos que los enfadados, incluso después de interacciones repetidas en las que las emociones no predicen la tasa de cooperación de los compañeros. Dicho efecto desaparece cuando el significado social de las emociones se elimina del juego. Otro experimento adicional muestra que los participantes son capaces de aprender asociaciones específicas entre emociones discretas y diferentes tendencias cooperativas,

aunque necesitan más evidencia cuando la asociación es contraria a las expectativas previas. En conjunto, nuestros resultados muestran que las emociones se emplean como señales en las interacciones entre personas, y que su efecto es duradero incluso en contextos en los que carecen de predictividad real.

REFERENCES

Adolphs, R. (2003) Cognitive neuroscience of human social behavior. *Nature Reviews Neuroscience, 4*, 165-178.

Aguiar, F., Brañas-Garza, P., Cobo-Reyes, R., Jimenez, N., & Miller, L.M. (2009) Are women expected to be more generous? *Experimental Economics, 12*, 93-98.

Angie, A.D., Connelly, S., Waples, E.P., & Kligyte, V. (2011) The influence of discrete emotions on judgment and decision-making: A meta-analytic review. *Cognition & Emotion 15*, 1-30.

Averbeck, B.B., & Duchaine, B. (2009) Integration of social and utilitarian factors in decision making. *Emotion, 5*, 599-608.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323-370.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior, 10*, 122-142.

Boone, R. T., & Buck, R. (2003). Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *Journal of Nonverbal Behavior, 27,* 163-182.

Buck, R. (1984). *The communication of emotion*. New York: Guilford Press.

Camerer, C. F., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica, 56*, 1-36.

Camerer, C.F. (2003). *Behavioural game theory: Experiments in strategic interaction*. Princeton: Princeton University Press.

Darwin, C. (1872). *The expression of emotions in man and animals*. London: John Murray.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience, 8*, 1611-1618.

Dunn, J.R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *Journal of Personality and Social Psychology, 88*, 736-748.

Eckel, C.C., & Wilson, R.K. (2003). The human face of game theory: Trust and reciprocity in sequential games. In E. Ostrom & J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 245-274). New York: Russel Sage Foundation.

Fehr, E., & Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences, 11*, 419-427.

Fridlund, A.J. (1995). *Human facial expression: An evolutionary view*. London: AcademicPress.

Gaertig, C., Moser, A.R., Alguacil, S. & Ruz, M. (2012). Social information and economic decision-making in the ultimatum game. *Frontiers in Decision Neuroscience*, 6:103.

Harlé, K.M., & Sanfey, A.G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion, 7*, 876-881.

Jennings, J.R. and Wood, C.C. (1976). The adjustment procedurefor repeated-measures

analyses of variance. *Psychophysiology, 13,* 277-278.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion, 13*, 505-521.

Kleinhesselink, R.R., & Edwards, R.E. (1975). Seeking and avoiding belief- discrepant information as a function of its perceived refutability. *Journal of Personality and Social Psychology, 31,* 787-790.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S. R., & Montague, P.R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. Science, 308(5718), 78-83.

Krumhuber, E., Manstead, A.S., Cosker, D., Marshall, D., Rosin, P.L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*, 730-735.

Niedenthal, P.M., Mermillod, M., Maringer, M., & Hess U. (2010). The simulation of smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral & Brain Sciences*, 33, 417-433.

Olsson, A., & Ochsner, K.N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences, 12*, 65-71.

Parkinson, B., & Simons, G. (2009). Affecting others: social appraisal and emotion contagion in everyday decision-making. *Personality & Social Psychology Bulletin, 35*, 1071-1084.

Ruz, M., & Tudela, P. (2011). Emotional conflict in interpersonal interactions. *Neuroimage, 54*, 1685-91.

Ruz, M., Madrid, E., & Tudela, P. (*in press*) Interactions between emotion and cognitive control in an interpersonal game. *Social, Cognitive and Affective Neuroscience*.

Ruz, M., Moser, A., & Webster, K. (2011) Social expectations bias decision-making in uncertain interpersonal situations. *PLoS ONE* 6(2): e15762. doi:10.1371/journal.pone.0015762.

Sanfey, A.G. (2007). Social decision-making: insights from game theory and neuroscience. *Science, 318*, 598-602.

Scharlemann, J.P.W., Eckel, C.C., Kacelnik, A., & Wilson, R.K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology, 22*, 617-640.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburg: Psychology Software Tools, Inc.

Tice, D.M. (1991). Esteem protection or enhancement? Self-handicapping motives and attributions differ by trait self-esteem. *Journal of Personality and Social Psychology, 60,* 711-725.

Tortosa, M., Lupiañez, J., & Ruz, M. (2013). Race, Emotion and Trust: an ERP study. *Brain Research, 1494*, 44-55.

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T.A., Marcus, D.J., Westerlund, A., Casey, B.J., & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research, 168*, 242-249.

Van Kleef, G.A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science, 18*, 184-188.

Van Kleef, G.A., De Dreu, C.K.W., & Manstead, A.S.R. (2010). An interpersonal approach to emotions in social decision-making: The emotions as social information

model. In M.P. Zanna (Ed.) *Advances in Experimental Social Psychology, 42* (pp. 45-96). Burlington: Academic Press.

Wubben, M.J.J., De Cremer, D., & van Dijk, E. (2009). How emotion communication guides reciprocity: Establishing cooperation through disappointment and anger *Journal of Experimental Social Psychology, 45,* 987-990.

(Manuscript received: 27 April 2012; accepted: 26 October 2012)