COGNITION AND EMOTION, 2013
Vol. 27, No. 4, 603–620, http://dx.doi.org/10.1080/02699931.2012.726608

Routledge
Taylor & Francis Group

Who do you trust? The impact of facial emotion and behaviour on decision making

Timothy R. Campellone and Ann M. Kring

Department of Psychology, University of California, Berkeley, CA, USA

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

During social interactions, we use available information to guide our decisions, including behaviour and emotional displays. In some situations, behaviour and emotional displays may be incongruent, complicating decision making. This study had two main aims: first, to investigate the independent contributions of behaviour and facial displays of emotion on decisions to trust, and, second, to examine what happens when the information being signalled by a facial display is incongruent with behaviour. Participants played a modified version of the Trust Game in which they learned simulated players' behaviour with or without concurrent displays of facial emotion. Results indicated that displays of anger, but not happiness, influenced decisions to trust during initial encounters. Over the course of repeated interactions, however, emotional displays consistent with an established pattern of behaviour made independent contributions to decision making, strengthening decisions to trust. When facial display and behaviour were incongruent, participants used current behaviour to inform decision making.

Keywords: Facial emotion; Decision making; Behaviour; Trust Game.

Consider the following scenario: you walk on to a used car lot and are immediately greeted by a smiling salesperson. The two of you walk around the lot, making friendly small talk while he continues to be all smiles. After a few minutes, the salesman excuses himself to take a phone call. As you continue to browse the lot alone, you overhear him talking about how he plans to mark up the price on whatever car you choose to make a greater profit. Do you act on the salesman's untrustworthy behaviour and choose a different lot, or do you trust his smile and otherwise pleasant demeanour? Depending on what you use to inform your decision, you could end up being duped into overpaying for a car. In this example, the salesman's smiling facial expression is incongruent with his behavioural intention of selling a car for more than it is worth. Decisions regarding whether a person is trustworthy, regardless of whether he or she is smiling, can be accurately made based on facial features within 100 ms (Todorov, Pakrashi, & Oosterhof, 2009). However, the information we rely on to facilitate social decision making is often in conflict, such as when a person's facial display of emotion is incongruent with their behaviour.

Researchers investigating social decision making have utilised behavioural economic paradigms

Correspondence should be addressed to: Timothy R. Campellone, 2205 Tolman Hall, University of California, Berkeley, Berkeley, CA 94720-1690, USA. E-mail: tcampellone@berkeley.edu

© 2013 Taylor & Francis

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

as a way to study human interactions while maximising experimental control over variables of interest (see Rilling & Sanfey, 2011, for a review). One paradigm in particular, the Trust Game (Berg, Dickhaut, & McCabe 1995), is well suited for investigating how people make decisions about whether to trust a social partner.[1] During the original Trust Game, two people were seated in different rooms. The person in the role of Player A was given $10 and the option to invest any amount of the $10 in Player B. The experimenter multiplied the amount given to Player B, and Player B was then instructed to return any amount of the multiplied sum to Player A. The amount invested by Player A represents how much that person trusts Player B to reciprocate the investment, with more invested money indicating greater trust. Player B's decision to return money to Player A reflects trustworthy behaviour, with lesser returns indicating untrustworthy behaviour. Diminished trust, therefore, is reflected by a small (or no) investment by Player A accompanied by a lack of reciprocation by Player B. The conservative or "safe" choice would be for Player A to keep the $10 and invest nothing. This way, Player A is guaranteed to leave the interaction with a positive gain. However, 30 of the 32 participants in the position of Player A gave some amount of money to Player B (average Player A investment was $5.16). Thus, without knowing anything about the other person with whom they were playing the game, most participants in the role of Player A were willing to invest money (i.e., trust the other player) in hopes that Player B would reciprocate their trust.

Why do people decide to trust? Indeed, the decision to trust is a potentially risky choice. Participants' decision to invest an average of half the allotted money defies theoretically predicted behaviour of a zero investment (Camerer, 2003). While previous studies have examined single interactions between interaction partners, more recent studies have examined repeated interactions with the same partners. Indeed, most social encounters involve an ongoing series of exchanges that provide new information that in turn shapes impressions and decisions (Keltner & Kring, 1998). When deciding whether to trust a person, research suggests that we continue to modify our decision making over the course of repeated interactions, tracking changes in a social partner's behaviour. For example, King-Casas and colleagues (2005) had participants play an iterated version of the Trust Game with another person, with the amounts invested and returned by each player presented in real time. The amount invested for each trial was dependent on the previous trial's return, with larger returns leading to greater investment on the next trial. Thus, we tend to learn about a social partner's behaviour, trustworthy or not, over the course of repeated interactions with that person.

In addition to examining the influence of behaviour across repeated interactions, other studies have examined how social factors may influence decisions to trust, including the perceived trustworthiness of interaction partners. For example, participants in one study engaged in a series of single Trust Game interactions with different simulated partners represented by black and white emotionally neutral faces. Participants invested more in faces that were rated as trustworthy looking (van t' Wout & Sanfey, 2008). Indeed, judgements of trustworthiness based on facial features can be made very quickly (Todorov et al., 2009), are highly reliable across raters (e.g., Brownlow, 1992), and are argued to play an essential role in human survival (e.g., Cosmides & Toobey, 2000). However, we know less about how the information signalled by emotional displays influences decisions of trust.

As the used car salesman example illustrates, facial displays of emotion can also be potent signals during social interactions, communicating both feelings and intentions (e.g., Keltner & Haidt, 1999) that can potentially influence

[1] In the original study conducted by Berg and colleagues, the game was called the Investment Game. However, since other studies employing this paradigm have referred to the Investment Game as the Trust Game, we will henceforth refer to the Investment Game as the Trust Game for the purposes of clarity.

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

decisions of trust. As with trustworthy judgements of neutral faces, affective judgements (e.g., like/dislike) about faces also occur as quickly as 160 ms after presentation (e.g., Pizzagalli et al., 2002). Moreover, emotional facial displays convey information that can influence behaviour. For example, facial displays of happiness, such as smiling and laughter, have been shown to promote affiliative tendencies in observers (e.g., Keltner & Bonanno, 1997) and are associated with approach-related behaviour (e.g., Knutson, 1996). Displays of happiness have also been shown to promote decisions to trust during social interactions. For example, social partners displaying a smile were co-operated with and trusted more than non-smiling social partners during a single interaction version of the Trust Game (Scharlemann, Eckel, Kacelnik, & Wilson, 2001). Facial displays of anger, on the other hand, signal to observers to keep their distance (Marsh, Ambady, & Kleck, 2005). People who exhibit angry displays are rated as being less trustworthy (Dunn & Schweitzer, 2005), and are perceived as being less likely to engage in affiliative behaviour (Montepare & Dobish, 2003).

Although emotional facial displays and changes in real-time trustworthy behaviour constitute important types of information used in decision making, other characteristics of a social partner's behaviour are also influential. For example, knowing that a social partner has a history of immoral behaviour, regardless of how trustworthy their current behaviour may be, was associated with less participant trust during an iterated version of the Trust Game (Delgado, Frank, & Phelps, 2005). These studies highlight the conundrum presented by our example of the smiling used-car salesperson. Is his untrustworthy behaviour salient enough to help you look past that smiling display? Or, is his happy expression enough to override your knowledge of the salesperson's ill intentions?

To date, most studies have focused on manipulating either social partner facial displays *or* trustworthy behaviour, not both, in order to better understand the relative influence of each factor in decision making. However, to illuminate the mechanisms underlying decisions to trust, we need to understand how changes in *both* facial displays and behaviour converge to shape social decision making. By simultaneously investigating changes in facial display and behaviour, we can examine what happens when these two factors are congruent (e.g., smiling and trustworthy behaviour) or incongruent(e.g., angry and trustworthy behaviour), thus helping to unpack how behaviour learned across repeated interactions and changes in facial display influence how we reach decisions about whom to trust.

Only one study that we are aware of has investigated facial features and trustworthy behaviour simultaneously. Chang, Doll, van t' Wout, Frank, and Sanfey (2010) asked participants to play an iterated version of the Trust Game with four simulated players represented by pictures of neutral faces that had been previously rated as being high or low in trustworthiness. Player behaviour was manipulated so that simulated players were either trustworthy (i.e., they returned 80% of the invested amount during every trial) or untrustworthy (i.e., they returned 20% of the invested amount during every trial). Results indicated that while participant decisions of trust were initially associated with player facial features, subsequent decisions were associated with the player's behaviour. That is, greater investments were made in players that reciprocated participant trust, regardless of whether that player's face had been deemed high or low in trustworthiness.

Chang et al.'s findings suggest that trustworthy behaviour may be more influential than trustworthy facial features in deciding whom to trust over repeated interactions. However, there are two key areas in need of further investigation. First, the amount returned by each simulated player was constant from trial to trial, making player behaviour predictable. In the present study, we varied player behaviour from trial to trial in order to more closely approximate the variability in behaviour typically found in real-life social exchanges. It is rarely the case that someone behaves constantly from interaction to interaction, and

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

thus we sought to manipulate the variability in player behaviour.

Second, the pictures used to represent the players in the Chang et al. (2010) study were emotionally neutral. While these faces were rated as having features that were higher and lower in trustworthiness, emotional expressions can signal additional information to help inform decision making. For example, Van Kleef, De Dreu, and Manstead (2010) argued that emotions are integral for navigating and making decisions during social interactions. While previous studies have shown that facial features can reliably signal *behavioural or personality* information, such as trustworthiness, emotional expressions convey additional *affective* information, such as the sender's feelings (e.g., Keltner & Haidt, 1999). Signalled affective information by a partner can influence the emotional state of an observer, and subsequently their decisions and behaviour (Keltner & Kring, 1998; Van Kleef, 2009). For example, seeing a happy person smiling at you can in turn make you feel happy, which may increase your likelihood of engaging in affiliative behaviour.

Van Kleef and colleagues also discussed the importance of the interaction context for determining the meaning of an emotional expression. Although their focus was on the distinction between competitive and co-operative contexts, learning patterns of behaviour over time can also create a context for decision making. For example, discussions with a friend will likely occur within a context of trustworthiness, the result of a history of trustworthy behaviour with that friend built up over time. Conversely, a conversation with your unreliable landlord, who has a history of negligence when it comes to making repairs, is likely to occur in an untrustworthy context. Thus, the information conveyed by an emotional expression, such as a smile on the face of your friend or landlord, can change given the context created by their past behaviour (Van Kleef et al., 2010).

Given the evidence for the influence of emotional displays, such as happiness and anger, on decision making (e.g., Loewenstein & Lerner, 2003; Scharelmann et al., 2001; Van Kleef et al., 2010), the additional affective information signalled by emotional displays is likely to impact decisions to trust. While Chang and colleagues found that trustworthy neutral expressions did not influence decisions to trust above and beyond player behaviour, it remains unclear whether emotional displays will be salient enough to override tendencies to make trust decisions based on behaviour such as whether to trust a smiling, but untrustworthy player.

The present study

In the present study, we sought to investigate how emotion and behaviour shape decision making. Participants played an iterated version of the Trust Game during which they interacted with four simulated players. There were two study versions, each with two blocks. The study versions varied in the amount and timing of information provided from the simulated players across the repeated interactions. Across the repeated interactions in the first block of the Behaviour First (BF) version, participants learned each player's behaviour, defined as the degree to which players reciprocated participant trust. Importantly, we varied the amount each player returned from trial to trial, making player behaviour less predictable and thus more akin to actual social interactions. In the second block of the BF version, we added emotional displays of happiness or anger to the simulated players, creating conditions where player behaviour and emotional display were congruent (e.g., trustworthy and happy) and incongruent (e.g., trustworthy and angry). The addition of emotional rather than neutral displays allowed us to assess whether the information signalled by emotional displays can override learned player behaviour from the previous block of repeated interactions that contained only player behaviour.

Across repeated interactions during the first block of the Face First (FF) version, participants learned the behaviour of each player while simultaneously seeing a picture with an emotional display. During the second block of the FF version, we removed the concurrent emotional

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

displays and changed the behaviour of two of the players to again create congruent and incongruent conditions. Creating these conditions through changes in player behaviour, rather than changes in emotional displays, allowed us to examine if behaviour changes can override learned player behaviour alongside a congruent emotional display. In other words, after repeatedly interacting with a smiling trustworthy player, we assessed whether decisions to trust would change if that player no longer exhibited a smile yet remained trustworthy or if the player no longer smiled and began acting untrustworthy. In other words, we assessed whether changes in player behaviour are salient enough to override learned displays of emotion.

Using this design allowed us to investigate how learning player behaviour with or without concurrent emotional displays influenced participant decision making. Further, by having two study versions, we were able to investigate the distinct as well as combined contributions of emotional displays and behaviour over the course of repeated interactions in making decisions to trust.

We tested several predictions. First, we expected that participants would invest more in trustworthy players and less in untrustworthy players, with player trustworthiness indicated by the amount each simulated player returned. Second, we expected that facial displays would influence decisions to trust in initial encounters. Specifically, we predicted that participants assigned to the FF version would invest more in trustworthy players with happy emotional displays compared to the amount invested by participants assigned to the BF version in trustworthy players with no accompanying facial display. In addition, we expected that participants assigned to the FF version would invest less in untrustworthy players with angry displays compared to participants assigned to the BF version who interacted with untrustworthy players showing no concurrent displays of anger. In other words, the information signalled by emotional displays would be associated with trust (i.e., investment behaviour) above and beyond player behaviour during these initial encounters.

Following repeated interactions with simulated players, however, we predicted that when emotional display and player behaviour were incongruent (i.e., happiness with untrustworthy behaviour and anger with trustworthy behaviour), participants' decisions to trust would be more associated with the simulated player's behaviour. Specifically, following repeated interactions with players in each Trust Game version, we expected that participants would invest more in trustworthy and less in untrustworthy players, regardless of whether their current emotional display or behaviour was incongruent with their established pattern of reciprocation. Finally, we examined whether men and women would differ in their decisions to trust. Recent evidence suggests that women are more risk averse than men (see Croson & Gneezy, 2009, for a review; Byrnes, Miller, & Schafer, 1999, for a meta-analysis), particularly in domains such as ethics and finances (Figner & Weber, 2011), suggesting that women may be less likely than men to trust others in this type of paradigm. However, studies using behavioural economic paradigms to investigate social decision making have typically not found, or at least not reported, gender differences in risk preference.

METHODS

Participants

Seventy-two undergraduate students (36 men, 36 women) from the University of California, Berkeley, were randomly assigned to one of two versions of the iterated Trust Game. The mean participant age was 20.51 years ($SD = 2.98$). The sample was predominantly Asian (63.9%), followed by Caucasian (23.6%) and Hispanic (6.9%). Students received partial course credit in a psychology class upon completing the study procedures.

Procedure

After providing informed consent, participants were told that they would be playing a computer game with other people. Participants were led to

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

believe that the other players were real people when in fact each player was simulated, behaving according to a predetermined pattern. Participants were told they would see pictures of the other players during some of the game trials. To increase believability, each participant had his or her picture taken, ostensibly to be incorporated into the game for the other players to see.

Participants played an iterated version of the Trust Game, created using E-Prime 2.0 software and presented on a Dell desktop computer with an 18″ monitor. On each trial, participants had the option of investing from 0 to 10 points in a player. The invested amount was then quadrupled, and the other "players" would then return an amount of the quadrupled sum to the participant (Figure 1 shows an example trial). Participants played against four simulated players (labelled Player 1, 2, 3, or 4). Player behaviour, similar to a previous study (van t' Wout & Sanfey, 2008), was predetermined so that two players were trustworthy (i.e., their average return was double the initial amount invested), and the other two players were untrustworthy (i.e., their average return was half of the initial amount invested). For example, a trustworthy player given 6 points by a participant would return an average of 12 of the 24 possible points (the 6 points were quadrupled to 24). By contrast, an untrustworthy player would return an average of 3 of the 24 points. The average amount of points participants invested on a trial, ranging from 0 to 10, represented the degree of trust placed in the simulated players. The total amount of points a participant received did not accumulate across trials and was reset after each player interaction.

To further promote optimal performance in the Trust Game, participants were told that they would earn a monetary incentive equal to the sum of six randomly selected trial outcomes. To compute the incentive total, each point above the original ten that the participant was initially given would result in $0.10. Thus, a player return of 23 points would yield a payoff of $1.30 for that trial (13 points above the original 10 multiplied by $0.10). However, in order to make the incentives equivalent for all participants, everyone received $6 regardless of actual trial outcome. After the Trust Game procedures were completed, participants were asked two questions (realness of the interactions, believability of the cover story) to assess whether they believed the cover story using a 1 (*Not at all*) to 5 (*Very much so*) Likert scale. A full debriefing interview explained the rationale for deception in detail.

Trust Game versions

Participants were randomly assigned to one of two versions of the Trust Game, called the Behaviour First (BF)and Face First (FF) versions. As mentioned earlier, each study version contained two blocks. Within each block, participants interacted with each of the four simulated players 16 times, for a total of 64 trials in each block and 128 total study trials.

Figure 1. *Example of study trial. The first screen showed the player number. During some blocks (as shown in this example), players were also represented by an emotional display. The second screen prompted participants to make their investment. The third screen displayed the quadrupled point total. The fourth screen showed the amount returned by the player as well as both participant and player totals for that trial.*

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

Two key variables distinguished the BF and FF versions. The first variable was the *presence of an emotional display* on simulated players, which occurred in the first block of the FF version and the second block of the BF version. Thus, participants first learned player behaviour with concurrent emotional displays in the FF version, but not the BF version. The second variable was the *congruence or incongruence of facial display and player behaviour*, which is described in more detail in the following sections. Trials with facial displays included simulated players displaying either a happy or angry facial expression.

Behaviour First version. During the first block of the Behaviour First (BF) version, participants played with four simulated players labelled by a number and with no concurrent facial display (see Table 1). Two players exhibited trustworthy behaviour (i.e., the average amount returned over the entire first block was double the amount invested by participants); the other two exhibited untrustworthy behaviour (i.e., the average amount returned over the entire first block was half the amount invested by participants). At the end of the first block, participants rated the four simulated players' trustworthiness and likeability using a 1 (*Not at all*) to 7 (*Very much so*) Likert scale (see Table 2). In the second block of the BF version, participants played with the same four players identified by the same number as in the first block. Player behaviour during the second block was the same as in the first block, with the same two players being trustworthy and untrustworthy. However, participants now also saw a picture of each player expressing either a happy or angry facial expression. With the addition of facial displays during the second block of the BF version, we created two conditions: behaviour/display congruence (trustworthy/happy player, untrustworthy/angry player) and behaviour/display incongruence (untrustworthy/happy player, trustworthy/angry player). At the end of the second block participants again rated player trustworthiness and likeability.

Face First version. During the first block of the Face First (FF) version, participants played with four simulated players depicted by an emotional facial display and a number (also in Table 1). The pairing of emotional display and player behaviour was congruent such that two trustworthy players were paired with happy faces and two untrustworthy players were paired with angry faces. Each player was paired with the same picture throughout the first block. At the end of the first block, participants rated player trustworthiness and likeability (see Table 2). In the second block of the FF version, participants played with the same four simulated players, but they no longer saw the pictures that were displayed in the first block. Instead, each player was represented by number only. In addition, player behaviour changed in the second block such that two players' behaviour was congruent with behaviour in the first block and two players' behaviour was incongruent with behaviour in the first block. Specifically, one player who was trustworthy and represented by a happy display during the first block behaved in an untrustworthy manner in Block 2 and another player who was previously untrustworthy and paired with an angry face behaved in a trustworthy manner in Block 2. This manipulation allowed us to investigate how incongruence created by behaviour change, rather than by the addition of an emotional display, influenced decisions of trust. Further, by changing emotional displays during the second block of the BF version but changing behaviour in the second block of the FF version, we could investigate whether different types of

Table 1. *Overview of study design with player behaviour and emotional displays for each block and version*

Behaviour First version		*Face First version*	
Block 1	*Block 2*	*Block 1*	*Block 2*
TR	TR + HAP	TR + HAP	TR
UN	UN + HAP*	UN + ANG	TR*
TR	TR + ANG*	TR + HAP	UN*
UN	UN + ANG	UN + ANG	UN

Notes: TR = Trustworthy behaviour; UN = Untrustworthy behaviour; HAP = Happy emotional display; ANG = Angry emotional display. *Incongruent conditions.

Table 2. *Trustworthiness and likeability ratings in each version by block*

Behaviour First				*Face First*			
Block 1		*Block 2*		*Block 1*		*Block 2*	
Player	*M (SD)*	*Player*	*M (SD)*	*Player*	*M (SD)*	*Player*	*M (SD)*
TR	T = 4.91 (1.5)	TR + HAP	T = 5.74 (1.1)	TR + HAP	T = 4.37 (2.0)	TR	T = 5.63 (1.2)
	L = 5.11 (1.5)		L = 5.80 (1.0)		L = 4.71 (2.0)		L = 5.23 (1.6)
UN	T = 2.76 (1.2)	UN + HAP	T = 2.23 (1.1)	UN + ANG	T = 3.00 (1.6)	UN	T = 5.69 (1.0)
	L = 3.18 (1.5)		L = 3.63 (2.0)		L = 3.34 (1.8)		L = 5.31 (1.4)
TR	T = 5.14 (1.4)	TR + ANG	T = 5.29 (1.5)	TR + HAP	T = 4.46 (1.7)	TR	T = 2.34 (1.4)
	L = 5.09 (1.5)		L = 5.11 (1.2)		L = 4.20 (1.6)		L = 2.66 (1.8)
UN	T = 2.89 (1.6)	UN + HAP	T = 2.00 (1.1)	UN + ANG	T = 3.26 (1.2)	UN	T = 2.17 (1.4)
	L = 2.97 (1.5)		L = 2.26 (1.7)		L = 3.40 (1.5)		L = 2.51 (1.8)

Notes: All ratings were made using a 1 (*Not at all*) to 7 (*Very much so*) Likert scale. TR = Trustworthy behaviour; UN = Untrustworthy behaviour; HAP = Happy display; ANG = Angry display; T = Trustworthiness rating; L = Likeability rating; *M* = Mean; *SD* = Standard Deviation.

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

congruency and incongruency differentially influenced participant decision making.

During blocks that included pictures (i.e., first block of the FF version and the second block of the BF version), the four players displayed the same emotion throughout the block. However, facial displays were randomly assigned to player across participants such that the same face was paired with different predetermined behaviours (i.e., trustworthy, untrustworthy) across participants. We pseudo-randomised player gender into the congruent and incongruent study conditions. Thus, half of participants completed the task with male players in the congruent condition and female players in the incongruent condition, while the other half completed the task with male players in the incongruent condition and female players in the congruent condition. Finally, to control for any effect of order, the sequence in which players appeared was pseudo-randomised across participants so that the same player never appeared twice in a row.

Facial stimuli

For both versions of the game, we used four pictures (two men, two women) displaying either happiness or anger from the NimStim Facial Stimulus Set (Tottenham et al., 2009). Each picture was rated by an independent sample ($n = 99$) to ensure equivalence in intensity, attractiveness, and trustworthiness within valence. That is, pictures were selected such that the two angry pictures were similar in these domains and the two happy pictures were similar in these domains. The mean ratings were as follows, with each rating being made using a 1 (*Not at all*) to 7 (*Very much so*) Likert scale: trustworthiness (angry: 2.01/happy: 5.04), attractiveness (angry: 1.93/happy: 4.15), and intensity (angry: 5.5/happy: 3.2).

Data analysis plan

We first examined whether participants assigned to the different Trust Game versions differed on any demographic characteristics, such as age, race, or education. In addition, we examined whether player gender and order interacted with any other study variables. We tested our main study hypotheses using multiple measures, including participant investment *behaviour* and participant self-reported *experience* of their interactions with simulated players. Participant gender was included in all analyses to investigate whether trust (i.e., investment behaviour) was different for men and women. To test our first hypothesis, we conducted a paired-samples *t*-test to examine the difference between the average amount invested in trustworthy and untrustworthy players during the first block of the BF version. We chose this approach since this block did not include emotional displays, and thus differences in participant investment

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

would reflect an understanding of player behavioural patterns. We also tested this hypothesis by examining whether participants' ratings of player trustworthiness and likeability for trustworthy and untrustworthy players were consistent with participants' investment behaviour.

To test our second hypothesis about greater trust being given to smiling trustworthy players compared to trustworthy players with no smile, and less trust being given to angry untrustworthy players compared to untrustworthy players with no scowl, we conducted a 4 (Player) × 2 (Version) × 2 (Gender) multivariate analysis of variance (MANOVA) for the first block only. By comparing the first blocks of the two versions, we were able to examine the relative contributions of each emotional display (FF version, Block 1) above and beyond player behaviour (BF version, Block 1). Similarly, we examined whether participants in the BF versus FF versions differed in their ratings of player likeability and trustworthiness following the first block of interactions.

To test our third hypothesis, we conducted separate 4 (Player) × 2 (Block) × 2 (Gender) MANOVAs for each Trust Game version. Analysing each version independently allowed us to investigate whether the addition of congruent or incongruent emotional displays (BF version) or the change in behaviour (FF version) differentially impacted participant's decisions to trust. We also examined whether participant's ratings of trustworthiness and likeability would be associated with current player behaviour. For all MANOVAs, Greenhouse–Geisser corrections were used when assumptions for sphericity were not met, and corrected p-values are reported. Corrected t-values and two-tailed p-values are reported when Levine's test for equality of variance was not met. Effect sizes are reported as partial eta squared (η_p^2).

RESULTS

There were no differences between the participants assigned to the BF or FF versions in education, $t(70) = 0.36$, $p = .72$, or age, $t(70) = 0.67$, $p = .51$. In addition, participant race, $F(6, 63) = 0.52$, $p = .80$, the order of player interactions, $F(8, 61) = 0.52$, $p = .84$, and player gender, $F(1, 68) = 1.62$, $p = .19$, had no effect on investment behaviour, thus they were excluded from further analyses.

Previous studies that have used simulated players often assume (or at least do not report whether) participants believed the cover story. Ratings of how real the interaction felt (average of 2.47 on a 5-point scale) and how believable the cover story about the simulated players was (average of 2.92 on a 5-point scale) suggest that participants did not fully believe nor disbelieve that the interactions were real and with actual players. Of importance, these ratings did not differ between study versions and were not related to task performance.

Our first hypothesis, that participants would invest more in trustworthy compared to untrustworthy players during the first block of the BF version, was supported, $t(35) = 9.57$, $p < .001$. That is, participants invested more in trustworthy ($M = 6.94$, $SD = 1.9$) compared to untrustworthy players ($M = 4.58$, $SD = 1.7$), $t(35) = 9.57$, $p < .001$, indicating that participants varied their investment behaviour depending upon whom they were interacting with. This finding was corroborated by participant ratings of trustworthiness and likeability of the simulated players at the end of the first block (see Table 2). That is, participants rated trustworthy players as being more trustworthy, $t(35) = 8.33$, $p < .001$, and likeable, $t(35) = 6.40$, $p < .001$, than the untrustworthy players (i.e., players 2 and 4).

Next, we tested our second hypothesis that happy trustworthy players would be trusted more than trustworthy players with no concurrent facial display, and that angry untrustworthy players would be trusted less than untrustworthy players with no concurrent facial display during initial encounters in the first block. Examining only the first block of both Trust Game versions, a 4 (Player) × 2 (Version) × 2 (Gender) repeated-measures MANOVA revealed a main effect for Player, $F(3, 66) = 107.79$, $p < .001$, $\eta_p^2 = .83$, and a main effect for Gender, $F(1, 68) = 3.97$, $p = .05$, $\eta_p^2 = .06$. Overall, participants trusted (i.e., invested

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

more in) trustworthy players more than untrustworthy players, and women trusted less (i.e., invested less) than men. Key to our hypothesis, the Player × Version interaction was also significant, $F(3, 66) = 6.83$, $p < .001$, $\eta_p^2 = .24$. As shown in Figure 2, independent sample t-tests indicated that participants in the FF version invested significantly less in untrustworthy players than participants in the BF version, $t(70) = 3.43$, $p < .001$, indicating that player displays of anger alongside untrustworthy behaviour were associated with lower investment amounts than untrustworthy behaviour alone. However, participants in the FF version did not invest significantly more in trustworthy players than those in the BF version, $t(70) = -0.60$, $p = .55$, suggesting that the addition of happy displays was not associated with greater trust. Thus, we found only partial support for this hypothesis. Further, these observed differences in investment behaviour between the BF and FF versions were not reflected in the trustworthiness and likeability ratings made at the end of the first block in each version. That is, trustworthy and untrustworthy players were rated similarly, regardless of whether they were accompanied by an emotional display.

The Player × Gender interaction was also significant, $F(3, 66) = 3.38$, $p = .02$, $\eta_p^2 = .13$. Follow-up t-tests indicated that women assigned to both Trust Game versions invested less in trustworthy players compared to men, $t(70) = 3.14$, $p < .001$, but did not differ from men in their investment in untrustworthy players, $t(70) = 0.42$, $p = .65$. No other main effects or interactions were significant.

Since the creation of incongruent conditions was different in the BF and FF versions, we analysed each version independently to test our third hypothesis. In the BF version, a 4 (Player) × 2 (Block) × 2 (Gender) repeated-measures analysis of variance (ANOVA) revealed a main effect of Player, $F(3, 32) = 82.4$, $p < .001$, $\eta_p^2 = .89$, indicating that participants invested more overall in trustworthy players than untrustworthy players and a main effect of Gender, $F(1, 34) = 7.29$, $p = .01$, $\eta_p^2 = .18$, indicating that women invested less than men. In addition, the Player × Block interaction was significant, $F(3, 32) = 26.78$, $p < .001$,

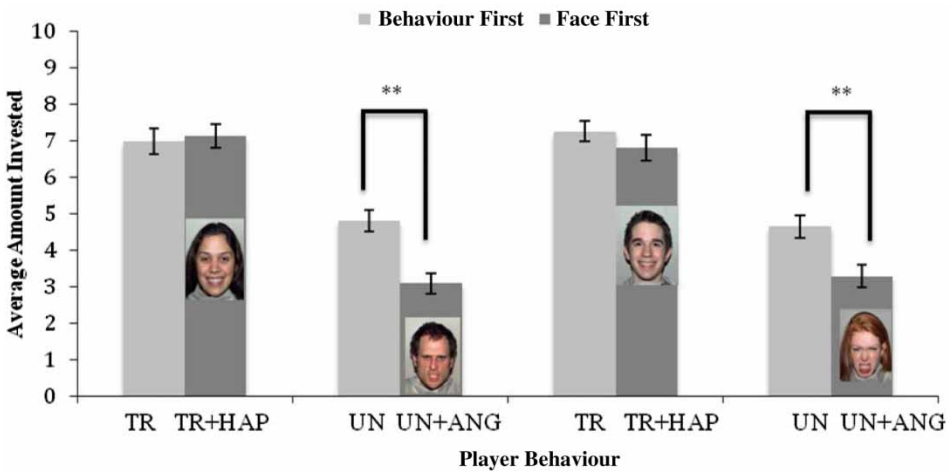

Figure 2. *Average investment during the first block of each version. This figure shows the average amount of points (0–10) invested in each player during the first block of the BF and FF versions. Participants invested significantly fewer points in untrustworthy players with angry emotional displays (UN/ANG) than untrustworthy players with no concurrent angry display (UN). Notes: TR = Trustworthy behaviour; UN = Untrustworthy behaviour; HAP = Happy display; ANG = Angry display. In the BF version, two players were trustworthy (TR) and two were untrustworthy (UN). In the FF version, two players were happy and trustworthy (TR + HAP) and two players were angry and untrustworthy (UN + ANG). The facial stimuli used in this figure are not the faces that appeared in the study. **$p < .01$.*

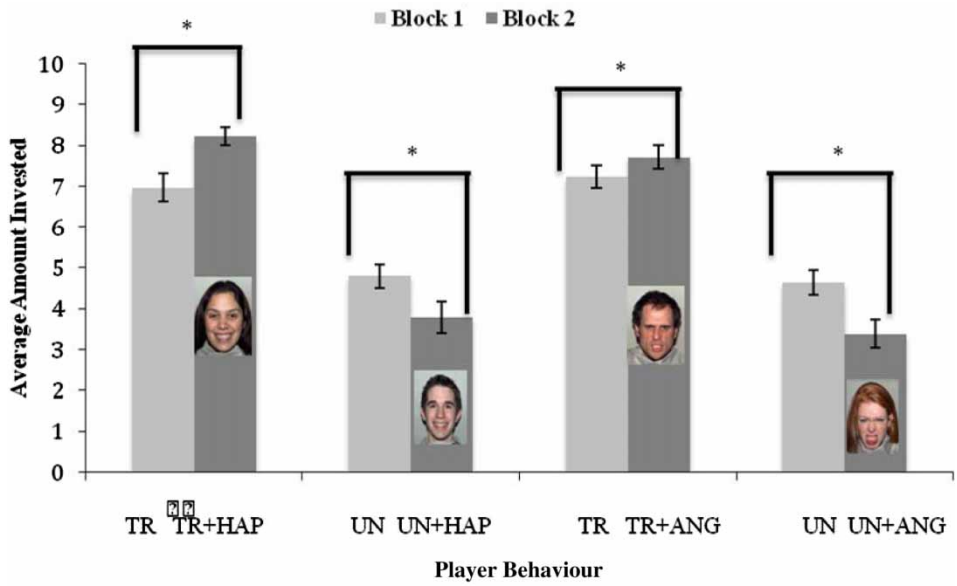

Figure 3. *Average investment during first and second block of the Behaviour First version. This figure shows the average amount of points (0–10) invested in each player during the first and second blocks of the BF version. After learning player behaviour in the first block, facial displays that were congruent or incongruent with player behaviour were added during the second block. Participant investment in the second block reflected player behaviour with significantly more points given to trustworthy players (T/HAP, T/ANG) and fewer points to untrustworthy players (UN/ANG, UN/HAP) compared to the first block. Notes: TR = Trustworthy behaviour; UN = Untrustworthy behaviour; HAP = Happy display; ANG = Angry display. Two players were trustworthy (TR) and two players were untrustworthy (UN) in both blocks. In Block 2, two players were happy (HAP) and two players were angry (ANG). Thus, two players (TR + HAP, UN + ANG) exhibited congruence between behaviour and facial display whereas two players (TR + ANG, UN + HAP) exhibited congruence between behaviour and facial display in Block 2. The facial stimuli used in this figure are not the faces that appeared in the study. *$p < .05$.*

$\eta_p^2 = .72$, and is depicted in Figure 3. Follow-up *t*-tests indicated that participants invested more in trustworthy players (players 1 and 3), $t(35) = -2.33$, $p = .03$, and less in untrustworthy players (2 and 4), $t(35) = 2.81$, $p < .01$, during the second block, regardless of whether they were displaying a happy or angry face. In addition, there were no differences in investment behaviour in the second block between the incongruent and congruent conditions for trustworthy (players 1 vs. 3), $t(35) = 1.23$, $p = .23$, and untrustworthy players (2 vs. 4), $t(35) = 0.90$, $p = .37$. That is, participants did not invest less in a scowling trustworthy player compared to a smiling trustworthy player nor did they invest more in a smiling untrustworthy player compared to a scowling untrustworthy player. Thus, participant' decisions to trust appeared to be influenced more by player behaviour than player facial display. These findings were corroborated by ratings completed after the second block of the BF version. Specifically, participants rated trustworthy players as being more trustworthy and likeable regardless of whether their facial display was congruent with their behaviour—player 1: trustworthiness, $t(35) = 12.62$, $p < .001$; likeability, $t(35) = 10.86$, $p < .001$—or incongruent—player 3: trustworthiness, $t(35) = 9.11$, $p < .001$; likeability, $t(35) = 3.86$, $p < .001$—compared to their untrustworthy counterparts (players 4 and 2, respectively). Interestingly, participants rated the happy untrustworthy player (player 2) as more likeable, $t(35) = 2.91$, $p < .01$, than the angry untrustworthy player (player 4) even though investment behaviour towards these two players did not differ, suggesting that smiles may influence perceptions of liking, but not trust.

Next, we investigated whether the removal of concurrent facial displays during the second block

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

of the FF version influenced decision making. We conducted a 4 (Player) × 2 (Block) × 2 (Gender) repeated-measures MANOVA. The Player main effect was significant, $F(3, 32) = 76.87$, $p < .001$, $\eta_p^2 = .88$, indicating that participants invested more overall in trustworthy players (1 and 3 in Block 1, 1 and 2 in Block 2), $t(35) = 16.01$, $p < .001$, than untrustworthy players (2 and 4 in Block 1, 3 and 4 in Block 2). As predicted, the Player × Block interaction, $F(3, 32) = 64.21$, $p < .001$, $\eta_p^2 = .86$, was also significant, reflecting the change in players 2 and 3's behaviour in the second block compared to the first block. No other main effects or interactions were significant.

As shown in Figure 4, participants invested more in player 2, $t(35) = 10.99$, $p < .001$, and less in player 3, $t(35) = -10.54$, $p < .001$, during the second compared to the first block. In other words, participants seemed to base their decisions to trust on whether a player's current, not past, behaviour was trustworthy. These findings were also corroborated by ratings of player trustworthiness and likeability. The player with consistent trustworthy behaviour across both blocks (player 1) was rated as more trustworthy and likeable than the player who was untrustworthy across blocks (player 4)—trustworthy: $t(35) = 10.86$, $p < .001$; likeable: $t(35) = 7.00$, $p < .001$. Further, the player (player 2) whose behaviour changed from untrustworthy to trustworthy between blocks was rated as more trustworthy and likeable than the player (player 3) who became untrustworthy between blocks—trustworthy: $t(35) = 11.20$, $p < .001$; likeability: $t(35) = 6.90$, $p < .001$. There were no differences in ratings of trustworthiness or likeability for trustworthy (players 1 and 2) or untrustworthy (players 3 and 4) after the second block. Thus, participants' ratings after the second block reflected the changes in player behaviour, not the facial displays from the first block.

DISCUSSION

The present study sought to better understand how people use and integrate the information signalled by facial displays with repeated experience with a person's behaviour to inform decision making during social interactions. Our inclusion of facial displays allowed us to examine how the emotional information signalled by the display was associated with decision making over the course of repeated interactions. Further, the design of this study allowed us to investigate how congruency or incongruency between the information conveyed by a social partner's emotional display and behaviour impacted decisions to trust in an attempt to assess what often happens in real life, as in the enigma posed by the used-car salesman.

The first question we asked was whether participants would be able to identify trustworthy and untrustworthy players based on their behaviour alone. Our findings suggest that participants were able to learn player behaviour, as evidenced by their greater investment in trustworthy compared to untrustworthy players. Similarly, participants rated trustworthy players as more trustworthy and likeable compared to untrustworthy players. Thus, we demonstrated across two modalities (investment behaviour, self-report), that participants identified differences in simulated player behaviour and used this to inform their subsequent decision making over the course of repeated interactions, even as player behaviour varied from trial to trial.

Our second question addressed whether seeing a player's facial display would influence decisions above and beyond a player's reciprocated behaviour. We predicted that happy displays would be associated with greater investment and angry displays associated with lower investment. We found partial support for this hypothesis in that participants invested less in angry, untrustworthy players compared to untrustworthy players with no facial display. This suggests that participants' decisions to trust were influenced by the information being signalled by displays of anger, such as a lower likelihood of affiliative behaviour (e.g., Montepare & Dobish, 2003) and untrustworthiness (e.g., Dunn & Schweitzer, 2005). Interestingly, participants did not rate untrustworthy players with a concomitant angry display as less trustworthy or likable compared to untrustworthy

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

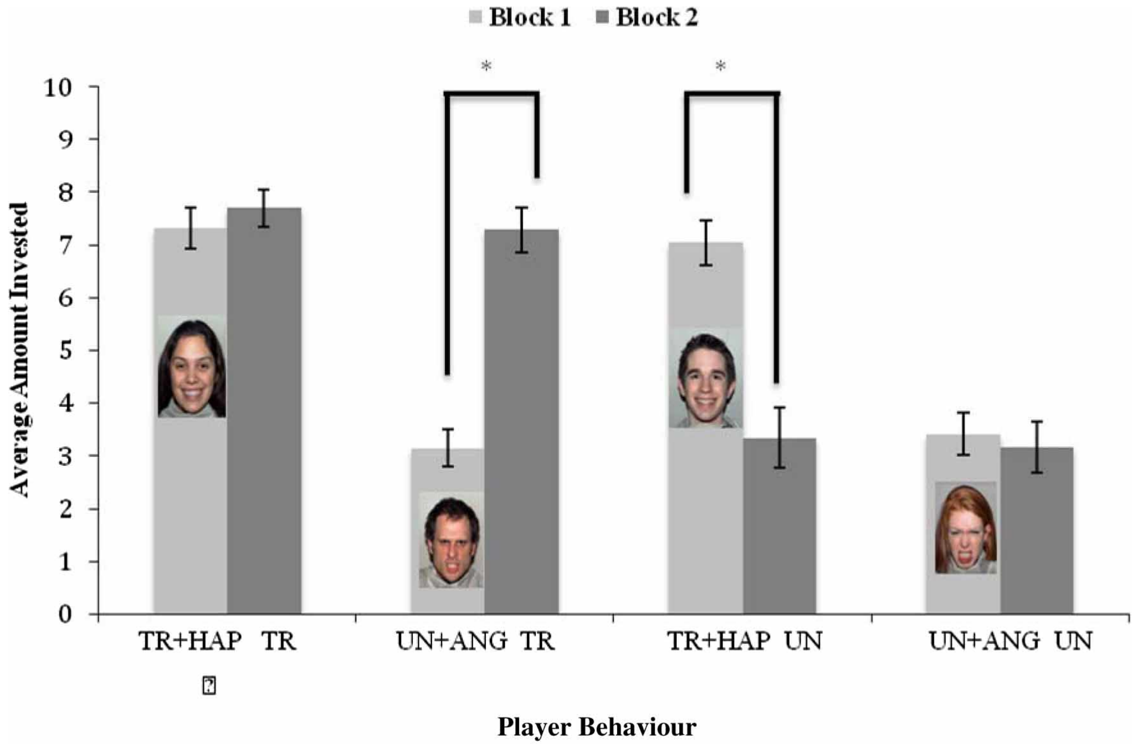

Figure 4. *Average investment during the first and second blocks of the Face First version. This figure shows the average amount of points (0–10) invested in each player during the first and second blocks of the FF version. During the first block, participant's learned player behaviour with concurrent facial displays. In the second block, displays were removed and the behaviour of two players changed. Participants were able to pick up on the changes in the behaviour of these players, but investment in trustworthy and untrustworthy players was not different across blocks. Notes: TR = trustworthy behaviour; UN = untrustworthy behaviour; HAP = happy display; ANG = angry display. In the first block, two players were happy and trustworthy (TR + HAP) and two players were angry and untrustworthy (UN + ANG). In the second block, one trustworthy (TR) player became untrustworthy (UN), and one untrustworthy player became trustworthy. Thus, two players exhibited incongruence between Block 2 behaviour and Block 1 facial display and two players exhibited congruence between current behaviour and Block 1 facial display. The facial stimuli used in this figure are not the faces that appeared in the study.* *$p < .05$.

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

players with no facial display. Thus, while anger displays were associated with differential investment behaviour, they were not associated with differences in participants' experience of the simulated players. In other words, the presence of an angry display appeared to facilitate the action of withholding trust, but not the subsequent labelling of that player as being untrustworthy. Thus, despite studies showing how quickly (Todorov et al., 2009) and reliably (e.g., Brownlow, 1992) people can make trustworthiness judgements from facial expressions, participants showed a reluctance to label angry players as untrustworthy, choosing to show their displeasure through action instead. Future studies should further explore this distinction between action and rating.

Contrary to expectations, participants did not invest more in trustworthy players exhibiting a smile compared to trustworthy players with no concomitant facial display. Similarly, participants rated trustworthy players as equally likable and trustworthy, regardless of whether the player exhibited a smile. These findings are inconsistent with a previous study that found smiling displays to be associated with greater trust than neutral displays (Scharelmann et al., 2001). However, that study employed a single interaction Trust Game, with participants seeing each display only one time compared to the sixteen times in our study. Thus, our findings suggest that over the course of repeated interactions, displays of anger, but not happiness, continue to provide information that influences decisions to trust.

Why did we observe differential effects of angry compared to happy facial displays during initial encounters? Brain imaging studies using negatively valenced emotional stimuli, such as angry faces, find activation in the amygdala (e.g., Morris et al., 1998; Yang et al., 2002), a brain area that plays a role in attentional allocation (Phelps, 2006), perceptual saliency (Anderson & Phelps, 2001), and how affective information influences decision making (Ernst & Paulus, 2005; Winkielman, Knutson, Paulus, & Trujillo, 2007). While positively valenced emotions, such as happiness, also activate this region (e.g., Hamman, Ely, Hoffman, & Kilts, 2002), studies have shown greater habituation in this area of the brain to "safety"- versus "threat"-based emotions (e.g., Wright et al., 2000). Although not a focus of this study, our findings showed that participants continued to invest less, on average, across the entire first block when presented with an angry player, but did not show a similar pattern for happy players. Though speculative, decreased saliency coupled with increased habituation to repeated exposures of happy compared to angry displays may have contributed to participants not placing greater trust in smiling trustworthy players. This will be important to disentangle in future research.

Our third question addressed scenarios where a person's behaviour is incongruent with their facial display, as in the example of the used-car salesman. We expected that participants' decisions to trust would be more associated with a player's reciprocated behaviour, even when that behaviour and facial display were incongruent (i.e., happy face but untrustworthy behaviour). This is indeed what we found, as participants invested more in a trustworthy player paired with an angry face and less in an untrustworthy player paired with a happy face. Ratings of trustworthiness and likeability mirrored participants' investment behaviour in that participants rated trustworthy players, regardless of whether they displayed happiness or anger, as more likeable and trustworthy compared to untrustworthy players.

When might a smile influence decision making? The addition of a smiling face to an already trustworthy player contributed to greater investment compared to when that same player was identified by behaviour only. We also found that a smiling player with an established pattern of trustworthy behaviour was trusted more than a smiling player with no behavioural history. This suggests that a smiling display only made independent contributions to decision making when on the face of a player with an established history of trustworthy behaviour. This was also evidenced by significantly greater ratings of trustworthiness for the reliably trustworthy player. Thus, while displays of anger may have made it easier to

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

withhold trust during interactions, regardless of a partner's behaviour history, smiling displays appear to have only influenced decisions to trust when a partner had a history of trustworthy behaviour.

We also investigated what happens when incongruency was created by changes in player behaviour, such as when a formerly trustworthy player became untrustworthy. Again, we expected a player's current rather than their former behaviour would influence that decision making. Consistent with this notion, participants invested more in currently trustworthy and less in currently untrustworthy players, suggesting a reliance on current behaviour for decisions of trust. This was corroborated by ratings of trustworthiness and likeability, which were higher for players with trustworthy players. Support for this hypothesis suggests two things. First, when a social partner's behaviour changes, decision making also updates to reflect the current behaviour. Second, when a social partner's behaviour remains consistent, but their facial display changes, decision making will discount the display in favour of behaviour.

Taken together, these findings point to the dynamic nature of trust. That is, trust develops by incorporating information acquired over the course of repeated interactions to influence current decisions. This study builds upon prior research (Chang et al., 2010; King-Casas et al., 2005) by showing that even when emotional displays are used instead of neutral faces, current player behaviour appears to still be the index for guiding decision making. Changes in the interaction context during the second block of each version appeared to change the significance of player facial displays. That is, a smile added to an untrustworthy player signalled different information than a smile added to a trustworthy player, similar to the example of seeing a smile on your friend versus your landlord. Changes in the meaning of the signalled information likely contributed to the changes in investment behaviour, with a smile only increasing the amount invested in a trustworthy player. A similar pattern emerged with displays of anger, as the addition of a scowling display decreased investment in an untrustworthy player. Thus, emotional displays that are consistent with established behaviour seem to solidify, and even strengthen decisions you already made. Future studies should aim to understand how changes in context modulate the information being signalled during social interactions, and how these changes impact decision making.

Over the course of repeated interactions, our results suggest that trust is placed in someone who exhibits trustworthy behaviour, regardless of how intense or what information is being signalled by their facial display. However, even after repeated interactions with the same person, emotional expressions can make independent contributions to decisions of trust, as evidenced by significant changes in investment behaviour and post-block ratings following the addition of displays congruent with player behaviour. In other words, even with knowledge of someone's behavioural history or pattern, the information signalled by an emotional expression that is congruent with what you already know based on behaviour can still enhance or erode trust.

Consistent with previous findings on gender differences in risk aversion (Croson & Gneezy, 2008), we found that women invested less overall than men. As mentioned earlier, previous research has shown that women, compared to men, are less prone to take risks in certain domains, including finances (Figner & Weber, 2011). In the context of the Trust Game, a larger investment can lead to greater payoffs, but a larger investment is also a riskier decision due to the possibility of receiving fewer points in return. Thus, investing more points in a trustworthy player, despite resulting in a positive return on most trials, can be construed as a riskier decision. Our data showing that women invested less overall, specifically for trustworthy and not untrustworthy players, is consistent with previous literature regarding gender differences in risk taking.

Taken together, our findings have implications for everyday decision making during social interactions. In the case of the used-car salesman, our data suggest that people are more likely to trust their reaction to overhearing him discussing price

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

gouging than his smiling face, and thus leave the lot, and seek a vehicle elsewhere. Further, as the venue for many social exchanges shifts from in-person to on-line, the quality and quantity of information available to inform our decision making changes as well, with first impressions being made in the absence of actual behaviour. For example, reading a potential dating partner's on-line profile might reveal involvement in on-going volunteer work alongside pictures of a smiling, seemingly trustworthy individual. Initial positive impressions will be associated with greater trusting and decisions to pursue future interactions. However, our data suggest that decisions to trust would be updated based on behaviour that is revealed across repeated interactions, no matter how many smiling photos may be posted.

This study had limitations that should be noted. First, we included only two emotional displays. We chose these emotions because of the evidence indicating their respective influences on perceptions of trust (Dunn & Schweitzer, 2005; Scharelmann et al., 2001) and because these particular emotional displays have been shown to elicit approach or avoidant behaviour in an observer (Knutson, 1996). However, other emotions, even emotions of similar valence, can signal information that may differentially influences decision making during social interactions (Lerner & Keltner, 2000). For example, the degree to which a display is socially engaging (Kitayama, Mesquita, & Karasawa, 2006) or self-conscious (Tracy & Robins, 2004) may provide alternative conceptual frameworks for investigating how emotion influences social decision making. Furthermore, low arousal emotional displays may impact decision making differently. Indeed, both happiness and anger are high arousal emotions(Russell, 1980), and it remains unclear whether lower arousal displays might be differentially associated with decisions to trust.

A second limitation of our study is that the facial stimuli, while balanced within emotion on various attributes, were all Caucasian. Although there was no effect of participant race or interaction between player and participant race, using only Caucasian stimuli limited our ability to investigate the role of race or ethnicity in decision making. Future studies should include a more diverse sample of facial stimuli to further our understanding of how culture shapes decision making during social interactions (Weber & Morris, 2010). Finally, despite our efforts to create a design more akin to actual social interactions, computer interactions lack the complexity and unpredictability found in real-world social exchanges. However, our findings demonstrate that tasks like the Trust Game can be modified to more closely emulate an in-person social interaction, while maintaining the experimental control afforded by such paradigms. One interesting direction would be the inclusion of dynamic versus static stimuli, which may increase the embodiment of the emotional information being signalled in the observer (Sato &Yashikawa, 2006) and increase the social nature of the exchange.

As mentioned earlier, participant decision making was not influenced by whether they believed the study cover story. To our knowledge, only one other study (Singer, Kiebel, Winston, Dolan, & Frith, 2004) has included the results of a believability assessment. However, participants were excluded if they reported not believing that the players were "living persons". Our study shows that the use of deception to increase the social nature of a task like the Trust Game, regardless of whether the cover story is believed, may not adversely impact participant behaviour. Continued investigation of how emotion interacts with social factors and contexts will expand and deepen our understanding of social interactions (Fischer & Van Kleef, 2010).

In summary, the present study sought to investigate how facial emotion and behaviour influence decision making during repeated social interactions. Our findings add to the existing emotion and decision-making literature in three main ways. First, facial displays of anger, but not happiness, can influence decisions of trust during *initial* interactions with a person. Second, with repeated interactions, decision making is continually updated to reflect changes in a social partner's pattern of behaviour, even if this pattern is in conflict with the information being signalled by

their facial display. Finally, even after you have learned a social partner's pattern of behaviour, emotions can make independent contributions to decision making, strengthening choices of whether to trust. These results improve our understanding of how facial displays and behaviour influence decisions in the context of repeated social interactions.

Manuscript received 30 March 2012
Revised manuscript received 27 July 2012
Manuscript accepted 29 August 2012
First published online 27 September 2012

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013

REFERENCES

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*, 305–309.

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

Brownlow, S. (1992). Seeing is believing: Facial appearance, credibility, and attitude change. *Journal of Nonverbal Behavior*, *16*, 101–115.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk-taking: A meta-analysis. *Psychological Bulletin*, *125*, 367–383.

Camerer, C. (2003). *Behavioral game theory*. New York, NY: Russell Sage Foundation.

Chang, L. J., Doll, B. B., van t' Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*, 87–105.

Cosmides, L., & Toobey, J. (2000). The cognitive neuroscience of social reasoning. In M. S. Gazzaniga (Ed.), *The new cognitive neuroscience* (pp. 1259–1276). Cambridge, MA: MIT Press.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*, 448–474.

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.

Ernst, M., & Paulus, M. P. (2005). Neurobiology of decision making: A selective review from a neurocognitive and clinical perspective. *Biological Psychiatry*, *58*, 597–604.

Figner, B., & Weber, E. U. (2011). Who takes risks and why? Determinants of risk taking. *Current Directions in Psychological Science*, *20*, 211–216.

Fischer, A. H., & Van Kleef, G. A. (2010). Where have all the people gone? A plea for including social interaction in emotion research. *Emotion Review*, *2*, 208–211.

Hamman, S. B., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science*, *13*, 135–141.

Keltner, D., & Bonanno, G. A. (1997). A study of laughter and dissociation: Distinct correlates of laughter and smiling during bereavement. *Journal of Personality and Social Psychology*, *73*, 687–702.

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

Keltner, D., & Kring, A. M. (1998). Emotion, social function, and psychopathology. *Review of General Psychology*, *2*, 320–342.

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

Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology*, *91*, 890–903.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*, 165–182.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion*, *14*, 473–493.

Loewenstein, G., & Lerner, J. (2003). The role of affect in decision-making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 619–642). Oxford, UK: Oxford University Press.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach and avoidance related behaviors. *Emotion*, *5*, 119–124.

Montepare, J. M., & Dobish, H. (2003). The contributions of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*, 237–254.

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, A. W., Calder, A. J., et al. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, *121*, 47–57.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Reviews of Psychology*, *57*, 27–53.

Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective judgments of faces modulate early activity (~160 ms) within the fusiform gyri. *NeuroImage*, *16*, 663–677.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Sato, W., & Yoshikawa, S. (2006). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, *104*, 1–18.

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.

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron*, *41*, 653–662.

Todorov, A., Pakrashi, M., & Oosterhof, N. (2009). Evaluating faces of trustworthiness after minimal time exposure. *Social Cognition*, *27*, 813–833.

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249.

Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, *15*, 103–125.

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 emotion in social decision making: The emotions as social information model. *Advances in Experimental Social Psychology*, *42*, 45–96.

van t' Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*, 796–803.

Weber, E. U., & Morris, M. W. (2010). Culture and judgment and decision making: The constructivist turn. *Perspectives on Psychological Science*, *5*, 410–419.

Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective influence on judgments and decisions: Moving towards core mechanisms. *Review of General Psychology*, *11*, 179–192.

Wright, C. I., Fischer, H., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2000). Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. *NeuroReport*, *12*, 379–383.

Yang, T. T., Menon, V., Eliez, S., Blasey, C., White, C. D., Reid, A. J., et al. (2002). Amygdalar activation associated with positive and negative facial expressions. *NeuroReport*, *13*, 1737–1741.

Downloaded by [University of California, Berkeley] at 13:47 05 June 2013