

Running head: FACIAL DYNAMICS

Facial dynamics as indicators of trustworthiness and cooperative behavior

Eva Krumhuber, Antony S. R. Manstead

School of Psychology, Cardiff University, U.K.

Darren Cosker, Dave Marshall, Paul L. Rosin

School of Computer Science, Cardiff University, U.K.

Arvid Kappas

School of Humanities and Social Sciences, Jacobs University Bremen, Germany

Word count: 3814

Abstract

Detecting cooperative partners in situations that have financial stakes is crucial to successful social exchange. We tested whether humans are sensitive to subtle facial dynamics of counterparts when deciding whether to trust and cooperate. Participants played a 2-person trust game prior to which the facial dynamics of the other player were manipulated using brief (< 6 seconds) but highly realistic facial animations. Results showed that facial dynamics significantly influenced participants' (a) choice of whom to play the game with and (b) decisions to cooperate. We also found that inferences about the other player's trustworthiness mediated these effects of facial dynamics on cooperative behavior.

Keywords: emotions; facial expressions; dynamics; trust; cooperation.

Facial dynamics as indicators of trustworthiness and cooperative behavior

In many situations we face a choice of whether to pursue our own short-term interests, or to rely on another person to maximize collective interests (De Cremer, 1999). There are potential gains if cooperation is achieved, but a risk that one might be exploited by cheaters who take advantage of one's own cooperation. It is therefore valuable to be able to spot interaction partners who are likely to be cooperative (Frank, 2004).

Trust has been shown to be an important precursor in the development of cooperation (Ross & LaCroix, 1996). We only make risky choices that render us vulnerable to others if we trust those others. When this trust is absent we are less likely to expose ourselves to risk of exploitation. But how can we tell whether or not another person can be trusted? Are there some signals that indicate trustworthiness and future cooperative behavior?

Clearly, it would be efficient in evolutionary terms to be able to detect trustworthiness in others quickly and on the basis of nonverbal cues. There is recent evidence showing that facial appearance influences attributions of competence and might affect voting behavior (Todorov, Mandisodza, Goren, & Hall, 2005). Such effects are, at least partly, linked to automatic attributions of trustworthiness based on the shape of the face and can occur in less than a second (Willis & Todorov, 2006).

However, the face is more than a static appearance cue with specific physical features. Facial expressions have been shown to provide behavioral and situational information in trust related contexts (Boone & Buck, 2003). They are important signals of emotional states (Ekman, 1982) and communicate our intentions to others (Keltner & Haidt, 1999). For example, someone who smiles may appear to be happy and approachable, and therefore likely to engage in cooperative behavior. Detecting such expressions would therefore seem to be a route to successful social exchange (Scharlemann, Eckel, Kacelnik, Wilson, 2001).

Yet not all facial displays are genuine signals of underlying emotions and intentions (Ekman, 1985). We may fake expressions in order to appease and to appear trustworthy without really meaning it. In this sense, a smile can easily be ‘put on’ in order to give an impression of cooperativeness and trustworthiness. Such false expressions may allow access to resources that would otherwise be denied. In any culture where individuals and groups can make gains by exploiting others’ trust there will therefore be cheaters who try to simulate moral emotions (Frank, 1988).

A central question is how humans distinguish between genuine emotions and fake or dishonest ones. There is evidence that facial motion conveys useful information for emotion and face perception (Ambadar, Schooler, Cohn, 2005; Christie & Bruce, 1998). For example, the speed with which an emotion unfolds has been found to affect the perception of the expression (Kamachi, Bruce, Mukaida, Gyoba, Yoshikawa, & Akamatsu, 2001). Given the fleeting nature of facial expressions, these differences in expression dynamics take place within a matter of milliseconds. Do these facial dynamics also provide clues about who is trustworthy?

In the research reported below we explore whether humans are sensitive to small differences in the facial dynamics of expressive behavior when choosing whether or not to trust and cooperate in situations that have financial stakes. The smile is one of the most common and effective signals in human communication. On the other hand, it is also one of the easiest to fake (Ekman & Friesen, 1982; Ekman, Friesen, & O’Sullivan, 1988). We therefore focus on the smile as an expression with a dual nature and explore whether facial dynamics provide important information about its quality to perceivers¹. If facial dynamics tell us something about the genuineness of an emotion, we should be less likely to trust and cooperate when a smile seems to be fake; but if a smile expression appears to be genuine, there should be higher rates of trust and cooperation.

To test these hypotheses, we presented participants with short video clips of counterparts with whom they would play games that had financial stakes. The facial expression of the counterpart was manipulated and consisted either of a neutral expression or of one of two dynamic smile expressions. Using realistic facial animations we created subtle differences in the dynamic nature of these smiles that affected the onset, apex and offset durations².

Past research has shown that posed expressions have shorter onset and offset durations compared to emotion-elicited expressions of felt joy (Cohn, & Schmidt, 2003; Ekman & Friesen, 1982; Hess & Kleck, 1990). Moreover, in two previous studies we found that smiles with longer onset and offset durations were judged to be significantly more genuine than their shorter counterparts, whereas authenticity ratings decreased the longer the smile was held at the apex (Krumhuber & Kappas, 2005; Krumhuber, Manstead, & Kappas, in press).

We therefore generated smiles that were either more ‘authentic’ (longer onset and offset durations, shorter apex duration) or ‘fake’ in their dynamic pattern (shorter onset and offset durations, longer apex duration). If facial dynamics act as an indicator of someone’s trustworthiness and cooperativeness in a given situation, we predicted that a) the selection of a counterpart and b) decisions about whether to cooperate would be shaped by the dynamic quality of the smile expression. These were the questions we examined in the two experiments described below.

Experiment 1

Method

Participants and Design.

Forty-eight female students (18-27 years, $M = 19.6$) at Cardiff University participated in the study for course credit or monetary compensation of £1.50. The counterpart’s facial dynamics

(fake smile vs. authentic smile vs. neutral expression) was the within-subjects variable and choice of fellow-player was the key dependent variable. Additional dependent variables were impressions of the other, expected cooperativeness of the other, and decisions to engage in the game.

Procedure and Materials.

When participants arrived individually at the laboratory they were informed, via a computer screen, that they would take part in a game in which they and another person (their counterpart) would make a monetary choice.

Trust game. The game was structured in such a way that the participant and counterpart were each endowed with £5. The participant had to decide whether to keep the money or to pass on the entire £5 to the counterpart. If the money was passed on, the amount was doubled by the experimenter. The counterpart could then decide whether to return £7.50 to the participant or keep the £10, leaving the participant with no money. Thus there were potential gains for both players if they cooperated, but risks for the participant if the counterpart decided not to cooperate.

Counterpart's facial dynamics. To make a choice of fellow-player, participants were shown short video clips of each of three possible counterparts. It was explained that not all counterparts would cooperate and that some would try to appear to be cooperative without returning anything. One of three counterparts displayed a neutral expression, one a fake smile, and one an authentic smile. The sequence of facial expressions was counterbalanced across counterparts. To rule out possible effects of attractiveness and honest demeanor, the three female counterparts chosen for this research were matched for trustworthiness and attractiveness, as determined in a pilot study ($n = 16$).

Using realistic facial animations we constructed dynamic smile expressions with standardized timing parameters for each counterpart's face. The facial model was restricted to the

lower face and was shown against a neutral background movie of the same person. Thus, only the mouth region was animated, thereby allowing us to study the effects of facial dynamics independently of other morphological features (such as the ‘Duchenne marker’, which contracts the muscle around the eye socket and produces ‘crow’s feet’ at the corner of the eye).

Fake smiles and authentic smiles differed solely in their onset, apex, and offset durations and were synthesized at a frame rate of 25 images per second. Authentic smiles had a relatively long onset duration of 20 frames (total 0.8 s), a relatively long offset duration of 53 frames (or 2.12 s), and a relatively short apex duration of 47 frames (or 1.88 s). Fake smiles were characterised by a relatively short onset duration of 9 frames (or 0.36 s), a relatively short offset duration of 10 frames (or 0.4 s), and a relatively long apex duration of 101 frames (or 4.14 s).

The smile expression was operationally defined as an upper smile (lip corner pull) with mouth opening, and synthesized at a medium level of intensity. All smile stimuli lasted 120 frames (i.e., 4.8 seconds) and were preceded by one second in which the counterpart’s neutral face was shown. This was done to familiarize participants with the counterpart’s face before the expression was shown. The three female counterparts showing three different facial expressions were displayed in random order as movie-clips (451 x 361 pixels, 5.8 sec) and presented in Director MX 2004 (Macromedia).

Dependent measures. Participants rated each video excerpt on a 7-point scale with respect to how likeable, attractive, and trustworthy they perceived the counterpart to be, and how cooperative they expected her to be (0 = *not at all*, 6 = *very*). Participants then chose a counterpart with whom to play the trust game. For the chosen counterpart, participants also indicated whether they would ‘engage’ by passing on their £5 to the counterpart, or ‘exit,’ thereby keeping the endowment (0 = *exit*, 1 = *engage*).

Results and Discussion

Preliminary analyses showed that there was no significant effect of counterpart identity, $F(8, 40) = 0.99, p > .05, \eta^2 = .16$. The data were therefore collapsed across the three counterparts and all dependent variables were entered into a MANOVA with a single repeated measures factor (facial dynamics). As shown in Figure 1, the facial dynamics of the counterpart had a powerful effect on participants' impression ratings and expectations concerning cooperation, multivariate $F(8, 40) = 24.66, p < 0.001, \eta^2 = .831$. Counterparts who showed an authentic smile were perceived as more likeable, $F(2, 94) = 108.83, p < .001, \eta^2 = .698$, attractive, $F(2, 94) = 20.81, p < .001, \eta^2 = .307$, and trustworthy, $F(2, 94) = 51.35, p < .001, \eta^2 = .522$, than those who showed a fake smile or a neutral expression. Participants also expected counterparts with an authentic smile to be more cooperative than fake smiling or non-expressive counterparts, $F(2, 94) = 76.29, p < .001, \eta^2 = .619$ (all means differ at $p < .05$ or better).

For the trust game participants were most likely to choose to play with authentically smiling counterparts, $\chi^2(2) = 21.12, p < 0.001$. Just over 60% of participants chose a counterpart displaying an authentic smile, whereas 33.3% of participants chose a fake smiling counterpart, and 6.25% of participants chose a non-expressive counterpart. It is clear, then, that participants made use of facial dynamics in making their choice of counterpart.

There was also a significant association between choice of counterpart and subsequent decisions to engage in the game, $\chi^2(2) = 6.28, p < 0.05$. Participants who chose a counterpart with an authentic smile were more likely than other participants to engage. More than 89% of participants with an authentically smiling counterpart decided to engage, compared with 75% of participants with a fake smiling counterpart, and 33.3% of participants with a non-expressive counterpart (Figure 2).

These findings show that facial dynamics significantly influence the selection of a counterpart and decisions about whether to cooperate. In a situation involving trust, one's fate is at the mercy of another person. Choosing a trustworthy counterpart is likely to be key to successful exchange. Our results show that facial dynamics help to convey relevant information about this personal quality to perceivers when they have to choose whom to trust.

In real life, decisions often involve actual consequences for an individual. Moreover, there is sometimes no opportunity to choose with whom we deal. In a second experiment, we wanted to test whether humans are sensitive to the facial dynamics of another person when they participate in a trust game that is played for real money and when the fellow-player is assigned rather than chosen. Furthermore, we aimed to explore the means by which facial dynamics influence cooperation. If facial dynamics serve as an index of trustworthiness, we predicted that the effects of facial dynamics on participants' cooperative behavior would be mediated by the perceived trustworthiness of the fellow-player.

Experiment 2

Method

Participants and Design

Ninety students (18-30 years, $M = 20.84$, male = 40, female = 50) at Cardiff University participated in the study for course credit or a snack (soft drink and potato chips). The 3 x 2 factorial design included the counterpart's facial dynamics (fake smile vs. authentic smile vs. neutral expression) and the sex of the participant (male vs. female) as between-subjects variables and cooperative behavior as the main dependent variable. Additional dependent variables were perceptions of the counterpart, emotions, and behavioral intentions.

Procedure and Materials

On arrival at the laboratory participants were instructed via computer that they would play a game for real money with another person (their counterpart). The money was paid in cash to them after the experiment. Participants were led to believe that their counterpart was seated at another computer in a separate room of the laboratory (although the ‘other player’ was in fact a pre-programmed strategy and always reciprocated). Prior to the actual game two trial tasks were conducted to allow participants to familiarize themselves with the structure of the game.

Trust game. The trust game was similar to used by Scharlemann, Eckel, Kacelnik, and Wilson (2001) and consisted of the following predetermined payoff structure. The participant and the counterpart were first endowed with £1.00 and £0.50, respectively. The participant had to decide whether to keep the £1.00 or pass the choice to the counterpart. If the participant chose to pass, the counterpart could end the game, keeping £1.25 and leaving the participant with £0.80, or pass the move back to the participant. If the counterpart chose to pass, the participant could then decide between £1.20 and £1.00 payoffs for each of the players. Again, the participant’s initial “trusting” move was crucial because the counterpart then had an incentive to quit, leaving the participant worse off; but if the initial trusting move were reciprocated both players would be better off.

Counterpart’s facial dynamics. Just before participants began to play the game they were shown one of several short video sequences of their ostensible counterpart. It was explained that the person had been pretested on cooperativeness and trustworthiness and was videotaped while being asked about her likely strategy in the trust game. The person shown in the video was one of the three different counterparts used in the first study and displayed an authentic smile, a fake smile, or a neutral expression. The facial stimuli were exactly the same as in the first study and presented in Director MX 2004 (Macromedia).

Dependent measures. Before participants made their decisions about whether or not to cooperate (0 = not cooperate, 1 = cooperate), they rated on 7-point scales how genuine the counterpart appeared to be in the video they had seen, and how cooperative they expected the counterpart to be (0 = *not at all*, 6 = *very*). After the trust game, further impressions of the counterpart were measured by five items pertaining to the counterpart's trustworthiness, hostility, insincerity, cooperativeness, and likeability (0 = *not at all*, 6 = *very*). In addition, participants rated how trusting, uncertain and relaxed they felt during the game (0 = *not at all*, 6 = *very*). Negatively framed items were reverse scored. With the exemption of the item "relaxed," which loaded on a separate factor, all nine judgment ratings were combined into a single index of the counterpart's perceived trustworthiness ($\alpha = .89$).

At the end of the task, participants also rated their counterpart with respect to 6 emotions: happy, afraid (reverse scored), sad (reverse scored), relaxed, angry (reverse scored), and surprised (0 = not at all, 6 = very). Except for "surprised," which loaded on a separate factor, all items were averaged into a single index of the counterpart's perceived positive emotionality ($\alpha = .75$).

Participants were then asked to indicate on a 7-point scale their behavioral intentions with respect to a) how much they would like to be paired with the *same counterpart* or *a different one* if they were to play the online game again (0 = prefer other, 6 = prefer same); b) how likely it was that they would make the *same decision* again (0 = very unlikely, 6 = very likely); and c) how much they would like to *meet* the counterpart outside the context of this research (0 = not at all, 6 = very much).

Results and Discussion

Trustworthiness scores, positive emotionality scores, and the three behavioral intention measures (would prefer same counterpart, would make same decision, would like to meet) were

entered into a 2 x 3 x 3 (sex of participant x identity of counterpart x facial dynamics) MANOVA. The only significant effect was the main effect of facial dynamics, multivariate $F(10, 136) = 23.09, p < .001, \eta^2 = .629$. Counterparts displaying an authentic smile were rated higher on perceived trustworthiness, $F(2, 72) = 160.75, p < .001, \eta^2 = .817$, and positive emotionality, $F(2, 72) = 85.05, p < .001, \eta^2 = .703$, than their fake smiling or non-expressive counterparts. Participants with authentically smiling counterparts also expressed greater willingness to be paired with the same counterpart again, $F(2, 72) = 37.58, p < .001, \eta^2 = .511$, and to meet outside the context of the research, $F(2, 72) = 29.12, p < .001, \eta^2 = .447$. Overall, the neutral expression was perceived most negatively, with low ratings on all dependent measures (see Figure 3). (All means differ at $p < .05$ or better). So showing some sort of a smile, even it had dynamic properties that led others to see it as less genuine, seems to be more advantageous than a non-expressive ‘poker’ face.

Importantly, participants were more likely to cooperate with counterparts when they displayed an authentic smile than a fake smile or a neutral expression, $\chi^2(2) = 25.62, p < .001$. Just over 93 % of participants with an authentically smiling counterpart trusted her; whereas 63.3% of participants with a fake smiling counterpart and 30% of participants with a non-expressive counterpart did so. Thus the facial dynamics of the counterpart affected the cooperative behavior of participants in this trust game.

To explore the prediction that intentions to cooperate would be mediated by inferences of trustworthiness derived from facial dynamics, we conducted a series of regression analyses. As shown in Figure 4, facial dynamics predicted cooperative behavior, as well as trustworthiness. Similarly, trustworthiness predicted cooperative behavior. However, when controlling for trustworthiness, facial dynamics no longer predicted cooperative behavior. Sobel’s (1982) test was significant, $z(89) = 2.61, p < .01^3$, showing that the perceived trustworthiness of a counterpart

mediated the effect of facial dynamics on cooperative behavior. Facial dynamics therefore seem to provide an important basis for inferring the trustworthiness of an unknown other, and for shaping the decisions we make about trusting this other.

General Discussion

In two experimental studies we demonstrated that humans are sensitive to subtle facial dynamics when choosing whether or not to trust another person in situations that have financial stakes. The facial dynamics of others as seen in video segments lasting less than 6 seconds were sufficient for participants to make inferences about the other's trustworthiness and future cooperative behavior. These findings extend recent evidence that judging trustworthiness based on the static shape of a face can be done on the basis of very short exposure times and is a fast and intuitive process (Willis & Todorov, 2006). Functional imaging studies suggest that decisions about trustworthiness involve brain structures (such as the amygdala) that process emotions (Adolphs, 2002; Winston, Strange, O'Doherty, Dolan, 2002). Humans therefore seem to have evolved special systems that allow them to detect cheaters (Cosmides & Tooby, 2005) who fake emotions in order to appear trustworthy and cooperative. Such a view is also compatible with theories of economists who argue that a cooperative system would not have evolved if cooperators and defectors could not be distinguished (Frank, 1988).

Counterparts showing a neutral expression were rated as least trustworthy. A non-emotional face when seen in a dynamic context may therefore signal different qualities than when it is seen as a static image. Moreover, any mechanism that involves classifying people as trustworthy or untrustworthy on the basis of static cues would limit the chances of making trustworthiness judgments in the course of social interaction. On the basis of the present findings we propose that facial dynamics have the capacity to serve as a behavioral 'fingerprint' of

someone's trustworthiness in a given situation. Fleeting facial movements convey temporal information that has an impact on whether we will trust others and cooperate with them or refrain from doing so.

References

- Adolphs, R. (2002). Trust in the brain. *Nature Neuroscience*, 5, 192-193.
- Ambadar, Z., Schooler, J. W., Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, 16, 403-410.
- 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.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, 26, 780-790.
- Cohn, J. F., & Schmidt, K. L. (2003). The timing of facial motion in posed and spontaneous smiles. *Proceedings of the 2nd International Conference on Active Media Technology (ICMAT 2003)* (pp. 57-72). Chongqing, China.
- Cosmides, L., & Tooby, J. (2005). In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 584-627). Hoboken, NJ: Wiley.
- De Cremer, D. (1999). Trust and fear of exploitation in a public goods dilemma. *Current Psychology*, 18, 153-163.
- Ekman, P. (1982). *Emotion in the human face* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Ekman, P. (1985). *Telling lies*. New York: W. W. Norton.
- Ekman, P., & Davidson, R., & Friesen, W. V. (1990). The Duchenne-smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, 58, 342-353.

- Ekman, P., & Friesen, W. V. (1982). Felt, false and miserable smiles. *Journal of Nonverbal Behavior*, 6, 238-252.
- Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology*, 54, 414-420.
- Frank, R. H. (1988). *Passions within reason*. New York: W. W. Norton.
- Frank, R. H. (2004). Introducing moral emotions into models of rational choice. In A.S.R. Manstead, N. Frijda, & A. Fischer (Eds.) (pp. 422-440), *Feelings and Emotions*. New York: Cambridge University Press.
- Hess, U., & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology*, 20, 369-385.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30, 875-887.
- Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13, 505-521.
- Krumhuber, E., & Kappas, A. (2005). Moving smiles: The role of dynamic components for the perception of the genuineness of smiles. *Journal of Nonverbal Behavior*, 29, 3-24.
- Krumhuber, E., Manstead, A.S.R, & Kappas, A. (in press). Temporal aspects of facial displays in person and expression perception. The effects of smile dynamics, head-tilt and gender. *Journal of Nonverbal Behavior*.
- Ross, W., & LaCroix, J. (1996). Multiple meanings of trust in negotiation theory and research: A literature review and integrative model. *International Journal of Conflict Management*, 7, 314-360.
- 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.

- Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290-312). San Francisco: Jossey-Bass.
- Todorov, A., Mandisodza, A. N., Goren, A., Hall, C. C. (2005). *Science*, 308, 1623–1626.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592-598.
- Winston, J., Strange, B., O’Doherty, J., & Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, 5, 277-283.

Author Note

We would like to thank Vicki Bruce and Geoff Thomas for their comments on previous drafts of this paper.

Correspondence concerning this article should be addressed to Eva Krumhuber, School of Psychology, P.O. Box 901, Cardiff University, Cardiff CF10 3AT, United Kingdom; email: KrumhuberE@cardiff.ac.uk.

Footnotes

¹In smile research, there has been recent evidence for a “Duchenne marker” that involves movement in the eye and cheek region in genuine smiles (Ekman, Davidson, & Friesen, 1990). This morphological feature is a perceptible signal in social interaction over and above the effect of dynamic features. However, we argue that facial dynamics may themselves be sufficient to shape perceptions and strategic decisions, independent of this morphological marker.

²Onset duration refers to the length of time from the start of the smile until its maximum intensity, apex duration to the length of time before this maximum smile starts to decrease, and offset duration to the length of time from the end of the apex until the smile disappears.

³Similar results were obtained when combined ratings of genuineness and expected cooperativeness (as measured before the trust game) were taken as mediators, $z(89) = 2.08$, $p < .05$. However, regression analyses testing for mediation by likeability were not significant, suggesting that this process was not driven by how likeable participants perceived the counterpart to be.

Figure Captions

Figure 1. Impression ratings and expected cooperativeness of counterparts as a function of facial dynamics in Experiment 1.

Figure 2. Percentage of chosen counterparts and decisions to engage/exit in the investment game as a function of facial dynamics in Experiment 1.

Figure 3. Perceptions of the counterpart and behavioral intentions as a function of facial dynamics in Experiment 2.

Figure 4. Regression analyses testing for mediation by trustworthiness of facial dynamics on cooperative behavior in Experiment 2.







