Smiling is a Costly Signal of Cooperation Opportunities:

Experimental Evidence from a Trust Game*

Samuele Centorrino,[1] Elodie Djemai,[1] Astrid Hopfensitz, [1] Manfred Milinski,[2] Paul Seabright[1]

This version: 11th April 2011

Abstract

We test the hypothesis that "genuine" or "convincing" smiling is a costly signal that has evolved to induce cooperation in situations requiring mutual trust. Potential trustees in a trust game made video clips for viewing by potential trusters before the latter decided whether to send them money. Ratings of the genuineness of smiles vary across clips; it is difficult to make convincing smiles to order. We argue that smiling convincingly is costly, because smiles from trustees playing for higher stakes are rated as significantly more convincing, so that rewards appear to induce effort. We show that it induces cooperation: smiles rated as more convincing strongly predict judgments about the trustworthiness of trustees, and willingness to send them money. Finally, we show that it is a honest signal: those smiling convincingly return more money on average to senders. Convincing smiles are to some extent a signal of the intrinsic character of trustees: less honest individuals find smiling convincingly more difficult. They are also informative about the greater amounts that trustees playing for higher stakes have available to share: it is harder to smile convincingly if you have less to offer.

---

* Any correspondence should be addressed: Paul.Seabright@tse-fr.eu. The authors are grateful to Lecio Costa-Simoes for excellent and very professional filming in Toulouse; to Marie-Claire Villeval and her team at GATE at the University of Lyon, especially Sylvain Ferriol, for their hospitality, encouragement and technical support; and the Max Planck Institute for Evolutionary Biology in Plön for financial support. We thank Ruxanda Berlinschi, Julia Bird, Claudia Capozza, Marie Lalanne, Maxime Marty, Andrea Matranga, Racha Ramadan, Karine Van Der Straeten and Stéphane Peysson for participation in pilot versions, and Jean-Louis Rullière, Alice Seabright and James Tremewan for valuable comments on preliminary results. The usual disclaimer applies.

[1] Toulouse School of Economics, Manufacture des Tabacs, 21 allée de Brienne, 31042 Toulouse, France.

[2] Max-Planck-Institute for Evolutionary Biology, August-Thienemann-Str. 2 24306 Plön, Germany.

I. Introduction

The man who indulges us in this natural passion, who invites us into his heart, who, as it were, sets open the gates of his breast to us, seems to exercise a species of hospitality more delightful than any other. No man, who is in ordinary good temper, can fail of pleasing, if he has the courage to utter his real sentiments as he feels them, and because he feels them.

Adam Smith – The Theory of Moral Sentiments[1]

Smiling is a form of behavior that is found in all human societies[2]. It appears to be more elaborate and more central to communication in humans than in any other species, and to play an important part in judgments of individuals about the character and general trustworthiness of others. There is no scientific consensus as to why it has evolved to be like this, nor about what it is in smiling that makes it an appropriate basis for judgments of others. There is consensus, however, about a number of its characteristics. First, viewers perceive smiles as varying in their degree of "genuineness" or "convincingness". Since the work of Duchenne (1862) and Darwin (1872) in the 19th century it has been known that smiles perceived as genuine (known as enjoyment or "Duchenne" smiles) are characterized by use of the orbicularis oculi (which surround the eyes) in combination with the zygomatic major (which raises the corners of the mouth); symmetry is also an important characteristic of Duchenne smiles. More recent research focuses on the importance of temporal dynamics such as smile onset, apex, and offset durations for perceived genuineness[3]. Second, Duchenne smiles are not under straightforward voluntary control. Some individuals can make them more often and more easily than others, and all individuals find them easier to make when in certain affective states. Such states include a relaxed mood in

[1] Smith (2000), p.497.
[2] See Darwin (1872), Ekman (1982), Niedenthal et al. (2010).
[3] See Krumhuber et al. (2007).

general, and feeling well disposed to a communication partner in particular. Third, smiles induce mimicry, both in the sense that individuals viewing smiles by others have an increased tendency to smile themselves[4], and in the sense that individuals trying to make a good impression on others (as when posing for photographs) make an effort to smile well. Although individuals can smile when alone, smiling behavior seems to be a form of communication. But if so, what is it communicating, and why have we evolved a form of communication behavior that is under such imperfect conscious control?

In this paper we test the hypothesis that smiling is a form of costly communication (costly in a sense we make precise below) that has evolved to induce cooperation between individuals in situations requiring mutual trust. According to this view, the necessary costliness of smiling is precisely the reason why it is under such imperfect conscious control. This hypothesis is not original to us[5], but to our knowledge it has not previously been subjected to a comprehensive experimental test. By a "comprehensive" test we mean one that tests separately the three component hypotheses that smiling "genuinely" or "convincingly" is:

a) costly to the smiler in terms of effort,

b) causally effective in inducing the target of the smile to cooperate with the smiler, and

c) a reliable signal of the likely benefits to the target of cooperating with the smiler.

[4] See Niedenthal et al. (2010).
[5] See Owren & Bacharowski (2001).

Component a) is important in distinguishing this hypothesis from two alternative views: first, that smiling is a form of costless communication that solves pure coordination problems (like "cheap talk"), and secondly, that it is not communication at all but merely an outward sign of an inner emotional state (like blushing, say). Component b) is important in explaining why human beings should have evolved the habit of communicating in this costly way. Component c) is important to explain why human beings should also have evolved the tendency to be influenced by the smiles of others. There exists some corroborating evidence for components b) and c) in the literature, though in neither case have the components in question been rigorously tested. Shug et.al. (2010) demonstrate that individuals who display relatively cooperative tendencies as proposers in an ultimatum game are more emotionally expressive in the face of unfair treatment by others than those who do not, including in the tendency to emit Duchenne as opposed to non-Duchenne smiles, which is consistent with component c). However, there is no test of any association between their emission of Duchenne smiles and their gestures of cooperation, and the sample is small (only 20 participants). Component b) is the only one of the three to be tested directly, and has received significant support (Scharlemann et al. 2001; Johnston et al. 2010). However, the first of these studies uses still pictures, a methodology that captures only a small part of the complex interactions involved in a smile. The second study indeed uses video clips but tests cooperation in a prisoners' dilemma (where non-cooperation is a dominant strategy, unlike in the trust game); furthermore it does so on the basis of comparison of only two clips and cannot control for other differences between clips.

Costly signaling has been extensively studied both in economics since the work of Spence (1974), and independently in biology since the work of Zahavi

(1975). A signal is any observable trait that imposes a cost on its bearer (a pecuniary or non-pecuniary effort cost in economics, a fitness cost in biology) but which reliably indicates the presence of some advantageous hidden trait because the signal is *more* costly for those individuals that do not possess the trait than for those who do[6]. The benefit from signaling the hidden trait is that it attracts partners in mating or in some other mutually beneficial cooperative activity, and the benefit to the signaler of doing so must exceed the cost of the signal. So what is the hidden trait that is signaled by smiling? In economic exchange the hidden trait could be an intrinsic characteristic of the smiler (such as their degree of altruism or tendency to display reciprocity[7]), or a characteristic of the situation in which the smiler finds herself (such as the size of the pie she is proposing to share with the smilee).

How can smiling convincingly be the result of costly effort? Emitting some kind of smile requires rather little effort. One might imagine that making the smile appear genuine was something that people could either do or not do, depending both on their personal characteristics and on features of the situation outside their control, but in any case outside the reach of their conscious will and therefore requiring no more effort than for a rudimentary smile. However, consider what happens in job interviews. Candidates can rarely produce convincing smiles to order, but instead they typically make a strong effort to be agreeable, to pay close attention to their interviewers, and interact, both conversationally and through body language in such a way as to make smiling "come naturally"; there is also some evidence that smiling can itself alter felt emotional states[8]. Robert Trivers (2000) has hypothesized that self-deception may be a means of reducing the cognitive load required to deceive others

---

[6] See Grafen (1990).
[7] See Gintis et al. (2003).
[8] See Strack et al. (1988).

successfully, a cognitive load that has been described in detail by Vrij et.al. (1996). Our hypothesis, therefore, is that non-Duchenne smiles require negligible effort but Duchenne smiles require significant cognitive effort, to a degree that probably varies between individuals, unless the smiler is already in a state that makes them more likely to be cooperative towards the target of the smile.

To test our hypotheses, and to help determine the kind of hidden trait signaled by smiling, we observe non-verbal behavior in an economic experiment involving trust. In a trust game[9] subjects (called "senders") each decide whether to send a sum of money to a second player, called a trustee. If they do so the sum is tripled, and the trustee may choose to keep the money, or to return some part of it to the sender. In particular, trustees face a choice between returning the sender's original stake, which corresponds to a form of reciprocating behavior, and returning an additional 50% of the stake so as to share the total surplus equally, which is a considerably more generous form of behavior.

The innovation in our version of the trust game is that we asked trustees to make short video clips to be shown to senders before the senders took their decision; 84 subjects produced a total of 168 clips. Earlier studies have observed the impact of seeing still pictures of partners in trust games and results indicate that pictures of smiling partners induce more trust[10]. Whether trustworthy partners can be detected from still pictures is controversial and might depend on the moment when the picture was taken[11]. Dynamic pictures might in this respect be better[12]. However no earlier study observes reactions to dynamic pictures of trustees in a situation where video

[9] See Berg et al. (1995).
[10] Scharlemann et al. (2001).
[11] Yamagishi et al. (2003), Verplaetse et al. (2007).
[12] Brown et al. (2003).

messages were obtained from participants knowing that this was their only means to convince their partner to cooperate.

We said nothing to trustees about smiling, but asked them to present themselves in a simple common format to the senders in such a way as to convince the latter to send the money. We showed the clips to the senders and asked them to take a decision about sending the money, then to rate the clips along a number of dimensions, one of which was the perceived genuineness of the smile. As our results indicate, this turned out to play a crucial role in the interaction between the sender and the trustee.

The outline of our paper is as follows. In section II we review the literature on smiling and on costly signaling to situate our paper in the context of prior work. In section III we set out a simple formal model of a trust game in which trustees can send a costly signal prior to the decision taken by trusters. We use this model to derive testable empirical predictions, and in particular the three component hypotheses described above. In section IV we describe our experimental method. In section V we present the results. Section VI concludes.

II. Literature Review

Even though our investigation of smiles as a costly signaling device is novel, a large number of studies in economics and psychology have in recent years investigated the importance of emotions in games. Inspired by results from affective sciences that emotions are not just some random noise but an essential part of the decision making mechanism (Damasio, 1994), theoretical and experimental work has turned to investigate the effect of different emotions and other visceral factors on decision making (Elster, 1998; Loewenstein, 2000; Kahneman, 2003; Frijda et al., 2004). While the focus has been mostly on such negative social emotions as anger and guilt (Bosman and van Winden, 2002; Sanfey et al., 2003; de Quervain et al., 2004; Hopfensitz and Reuben, 2009), increasing attention has been given to the use of rewards and the experience of happiness (e.g. Kahneman et al., 1999; Frey, 2008; Frey and Neckermann, 2009). Smiles are on the one hand an expression of experienced happiness and might be used as a coordination device (Manzini et al., 2009), but smiling is also an important component in social exchange (Owren & Bacharowski, 2001).

To detect whether an interaction partner can be trusted we can either rely on third party information regarding the target individual's reputation (Sommerfeld et al., 2008) or use visual signals concerning the individual's character (Frank, 1988). Indeed it has been observed that players in a trust game are willing to spend money on visual information of their partner (Eckel and Petrie, 2008). Which visual information is used is however not clear, but honest smiles could play a crucial role in that respect (Cohn and Smith, 2004). Because activation of the orbicularis oculi (one of the main markers of Duchenne smiles) is believed to be under emotional and involuntary control (Ekman and Friesen, 1982), it might therefore be an informative signal.

Whether smiles are indeed perceived as a signal of trust has been subject of a number of recent experimental studies. Scharlemann et al. (2001) presented participants in a modified trust game with static pictures of sixty photographic models, which were portrayed smiling and with a neutral expression. The results show that smiling pictures are more often trusted than their non-smiling counterpart (68.3% versus 55%). Similarly Mehu et al. (2007b) assess which characteristics are associated with honest smiles. Fifty faces (half male and half female) were rated concerning ten different attributes: attractiveness, generosity, trustworthiness, competitiveness, health, agreeableness, conscientiousness, extroversion, neuroticism, and openness to experience. The findings show that Duchenne smiles played a significant role in the assessment of generosity and extroversion. Further, van 't Wout and Sanfey (2008) observe that judgments of facial trustworthiness is related to sending money in a trust game. In this study, ratings of player trustworthiness (based on showing a photograph of each player) were a significant predictor of how much money these players were given in a standard one-shot trust game. In a follow-up study, Chang et al. (2010) replicate these findings in a repeated trust game.

One drawback of such studies is the use of static pictures of smiles. Indeed Scharlemann et al. conclude that “the use of still photograph biases the perception of the facial expressions. [...] Dynamic footage of facial expression may be a more appropriate mean of eliciting naturalistic viewing patterns”. Such facial dynamics have been studied by Krumhuber et al. (2007). In their study participants played a trust game against partners that were presented with short video clips. These video clips were manipulated so as to obtain a true smile, a fake smile and a neutral expression for the same face. Indeed partners displaying an authentic smile were rated higher on perceived trustworthiness, than their fake smiling or non-expressive counterparts.

Another drawback of the above studies is that they can only be used to assess that smiles elicit trust in others. Whether smiles are truthful indicator of a person's intentions is essential for the importance of smiles as a social signal.

Altruism and cheater detection in social dilemmas has received considerable attention in economics and biology (Cosmides and Tooby, 1992; Gintis et. al., 2001). It is evident that signals that can be used to identify altruists might quickly be imitated by non-altruists and would thus not be reliable (Fehr and Fischbacher, 2005). One suggestion is that altruism as such can serve as a reliable signal of trustworthiness (Smith and Bliege Bird, 2000; Gintis et al., 2001; Lotem et al., 2003). However in many situations behavior of the interaction partner cannot be observed. In order to detect trustworthy partners reliably in one-shot interactions, it is therefore necessary to base decisions on verbal or non-verbal signals sent by the partner.

Brown and Moore (2002) stress that honest signals with reliable emotional basis may be needed to guarantee positive intentions of a counterpart. This leads to the importance of 'emotional expressivity' i.e. the ability to accurately communicate your internal feeling state (Boone and Buck, 2003). However to be reliable, these signals must be costly and therefore difficult to mimic. Smiles and especially honest smiles might be just that. Brown et al. (2003) were the first to observe that videos from self reported altruists are rated differently by neutral observers than videos of non-altruists. Further an analysis of video recordings from altruists and non-altruists showed that self reported altruists showed more orbicularis oculi activity and more symmetric smiles (see also Oda et al.; 2009).

To test whether smiles are indeed related to actions in a game requires thus the observation of both behavior and facial expression. Mehu et al. (2007a) suggest that human smiles are more prevalent in situations which involve sharing or exploitations of resources. By filming sixty pairs of friends during a neutral and a sharing decision they observe that significantly more Duchenne smiles are produced during sharing situations. Thus situations requiring sharing elicit smiles and laughter (Mehu and Dunbar, 2008). Whether smiles are also predictive of a specific sharing decision has so far not been studied and is subject of our work.

The model we develop hypothesizes that individuals may be motivated, to a greater or lesser degree, both by reciprocity and by altruism. There is a large literature addressing ways of incorporating social preferences in individual utility functions (see Sobel, 2005, for a survey). It is safe to say that there is no consensus as to the appropriate way of modelling such motivations, and it is emphatically not our intention to propose a general theory here. For instance, in many models of behavior in public goods games, individuals are considered to be motivated either by reciprocity or by altruism but not both (Fehr, Fischbacher & Gaechter, 2003); this is a useful device for focusing on the distinction between unconditional contributors and conditional contributors. Other papers (Hwang & Bowles, 2010; Brülhart & Usunier, 2004) hypothesize that individuals may have both motivations simultaneously to different degrees, and that is the approach we adopt here. This is a plausible and parsimonious way to capture the phenomenon, clearly present in our data, that some trustees are willing to return the same amount that was sent to them, while others adopt a more generous approach. We model this more generous approach as indicating relatively high altruism, but it could equally well arise from inequality aversion and in our experiments the two interpretations would be observationally equivalent; the choice between them is not what interests us in this work.

III. A model of costly signaling prior to a trust game

a. Outline

There are two players, A and B. To avoid confusion we shall refer to A as "he" and to B as "she", though in the experiment individuals of both sexes take each role.

Player A receives a stake of value s and must decide whether or not to send it to player B (we consider s to be greater than or equal to 1, without loss of generality). If it is sent it is multiplied by three, and player B may choose to send some part of the new enlarged stake back to A.

Player A's decision will be influenced by his beliefs about Player B along two dimensions – how much Player B cares about strong reciprocity, and how altruistic she is (we make these terms precise below). With respect to strong reciprocity, Player B may be one of two types $\theta \in (L,H)$; for simplicity we assume there are equal proportions of the two types in the population, though nothing of importance turns on this. H-types have stronger preferences for reciprocity than L-types (we can call these High Reciprocators and Low Reciprocators respectively). With respect to altruism, Player B has a component of her utility that is a stochastic function of the amount she sends back to A. Player B knows her own type at the start of the game, and notably when she makes a video clip in order to persuade player A to send her his stake.

If player A sends the stake, player B must decide to send back to player A a multiple *m* of the original stake, where $m \in (0,1,1.5)$. Since the stake has been multiplied by 3, this means that she has a choice between keeping half the stake, keeping two-thirds of it, and keeping all of it.

The order of moves is therefore as follows: first B makes a clip which involves sending a costly signal. Then A views the clip and forms a belief about B's type based on the evidence from the clip. If A chooses not to send the stake the game ends, A keeps the stake and B receives a zero monetary payoff (and a total payoff that may include a cost of effort involved in sending the signal). If A chooses to send the stake then B finally chooses what multiple of the stake to return to A, and the game ends.

As is standard we solve the game backwards from the end.

b. Player B's move

We model player B's motivation for returning a multiple of A's original stake using a random utility function. It is separable in money and in two types of social preference. The first social preference is for strong reciprocity, which we model as a fixed utility derived from sending back at least the original stake to player A, but not otherwise varying according to the amount sent. This utility, which differs between types, is given by α_θ, where $1 > \alpha_H > \alpha_L > 0.5$.

The second motivation is altruism, which is increasing in the amount sent back by B to A (it can be thought of as reflecting B's pleasure at knowing that she is increasing A's payoff). We model this as a utility that is a multiple β of the amount returned, plus a random error term ε. The coefficient β is itself random and may be greater or less than one (capturing the fact that, of players who return at least some money, some return only the original stake while others return a larger amount). Specifically, $\beta \in \{0.5,1.5\}$ with probability $(1 - p_\theta, p_\theta)$. We assume that $p_H > p_L$ to reflect the fact that individuals with a greater propensity for reciprocity are also likely to be more altruistic.

We therefore model player B's utility function as follows:

(1) $$U_B = 3s - ms + \alpha_\theta + \beta ms + \varepsilon \qquad \text{if } m>0$$

(2) $$U_B = 3s \qquad \text{if } m=0$$

where the error term ε has a zero mean, and is uniformly distributed between -0.5 and +0.5.

It is straightforward to see that if β=1.5, player B will always choose *m*=1.5, since his utility is always strictly increasing in *m*. Thus either type of player will choose *m*=1.5 with probability p_θ.

If β=0.5 on the other hand, player B's utility is strictly decreasing in *m* once *m* is positive. Thus B will either choose *m*=0 or *m*=1. The probability of choosing *m*=1 is therefore the probability that:

(3) $$\alpha_\theta > \frac{s}{2} - \varepsilon$$

If *s*=1, $\frac{s}{2} - \varepsilon$ is distributed uniformly on [0,1], so the probability that *m*=1 is just $(1 - p_\theta)\alpha_\theta$.

If *s*=2, $\frac{s}{2} - \varepsilon$ is distributed uniformly on [0.5,1.5], so the probability that *m*=1 is just $(1 - p_\theta)(\alpha_\theta - 0.5)$.

We can write this probability as a function of *s*, namely as

$(1 - p_\theta)(\alpha_\theta + \frac{(1-s)}{2})$.

We therefore summarize in Table 1 the probabilities of choosing different values of *m* according to whether the player is of high or low type and whether the stakes are high or low, as follows:

[Table 1 here]

c. Player A's move

Player A will send the money if the expected value of doing so is greater than the sure value of keeping it.

We also model player A's decision using a random utility function. We ignore altruism on the part of player A[13] and consider his utility as given by his expected payoff plus an error term η which is uniformly distributed between –e and 0 (we can consider this as a way of allowing for risk aversion while keeping the advantages of linear utility: η=0 corresponds to risk neutrality, while η=-e is the highest risk aversion in the population).

Player A's decision then depends on γ, his subjective probability of facing a High Reciprocator type. He will send the money if the gain from receiving a net profit of half the original stake, multiplied by the probability that B chooses *m*=1.5, exceeds the loss of the whole original stake, multiplied by the probability that B chooses *m*=0. Formally, A sends the money iff:

$$(4) \qquad 0.5(\gamma p_H + (1-\gamma)p_L) + \eta \quad >$$

$$\gamma(1-p_H)\left(\frac{(1+s)}{2} - \alpha_H\right) \quad + \quad (1-\gamma)(1-p_L)\left(\frac{(1+s)}{2} - \alpha_L\right)$$

[13] One reason for doing so is that it is plausible that A players would be less likely to feel altruism towards those B players they believed were likely not to return them any money, and therefore the calculation how likely the B player is to return money precedes and predetermines the effect of altruism on player A's decision. This is a hard phenomenon to analyze, and we have chosen to ignore it to focus on the issues more central to this paper.

Notice that the right hand side of equation (4) is strictly increasing in s. This means that, for given γ, player A is less likely to send the money when the stakes are high than when they are low. Thus if we observe a higher probability of sending the money when the stakes are high, this must mean that A players have higher levels of γ.

Because of the uniform distribution of η, we can write the probability that an A player sends the money, given the value of γ, as q_γ, where

$$(5)\ q_\gamma = \frac{[0.5(\gamma p_H + (1-\gamma)p_L) - \gamma(1-p_H)((1+s)/2 - \alpha_H) - (1-\gamma)(1-p_L)((1+s)/2 - \alpha_L)\ ]}{e}$$

Differentiating (5) with respect to γ yields:

$$(6)\quad \frac{\partial q_\gamma}{\partial \gamma} = \frac{[0.5(p_H - p_L) - (1-p_H)((1+s)/2 - \alpha_H) + (1-p_L)((1+s)/2 - \alpha_L)\ ]}{e} > 0$$

Differentiating (6) with respect to s yields

$$(7)\qquad \frac{\partial^2 q_\gamma}{\partial \gamma \partial s} = \frac{(p_H - p_L)}{2e} > 0$$

which shows that a given increase in γ will result in a larger increase in q_γ when *s*=2 than when *s*=1. So higher stakes make the probability of sending the money more sensitive to player A's subjective probability that player B is the High Reciprocator type.

d. The signal

Now consider the making of the video clip. Player B invests effort e, which has an increasing convex cost $c_\theta(e)$, where $c_H(e) < c_L(e)$ for all positive values of e.

This effort produces a smile whose quality is related to the effort exerted via an increasing function $g(e,\tau)$, where τ is a random variable, and the probability distribution function $f(g|e)$ has the Monotone Likelihood Ratio Property.

We begin by assuming that this smile has a predictable positive effect on player A's subjective probability γ that player B is the High Reciprocator Type. Without such an effect neither player would have any incentive to exert any effort at all. This effect can be represented by the "smile function" $\gamma = \gamma(g)$, where $\gamma' > 0$. The function $\gamma(g)$ need not be concave but if not $c_\theta(e)$ must be sufficiently convex to yield a unique interior solution.

We next go on to show that if player B knows this, and if the quality of the smile responds to her effort, she has reason to invest effort in smiling in such a way that the smile will indeed be a positive signal not just of her effort but also of the probability that she is the High Reciprocator type. Thus A's tendency to display greater trust in individuals who have more convincing smiles is one that could be expected to evolve under natural selection since it would correspond to a real empirical regularity.

To see this, write $V_{s\theta}$ for the expected utility B will receive if player A sends the money and note that $V_{sH} \geq V_{sL}$.[14] Writing $e^*_{s\theta}$ for the optimal choice of effort by a player B who is playing for stake s and is of type θ, since $c_H(e) < c_L(e)$ it follows that

(8) $$e^*_{sH} > e^*_{sL}$$

It is also straightforward that $V_{2\theta} > V_{1\theta}$, and therefore that

(9) $$e^*_{2\theta} > e^*_{1\theta}$$

Any function $g(e,\tau)$ that has the Monotone Likelihood Ratio Property will imply that the conditional probability that Player B is the High Reciprocator Type is increasing in the value of $g(e,\tau)$. To see this note that Bayes' Law with a uniform prior implies that

[14] The reason why the expected utility for B players of type H is higher than the utility for those that are L is that they have more altruism payoff than L players do. They could choose to return the same amount as L players do and would get at least as much utility as L players from doing so. In fact they choose to return more (in expected terms) than L players do, so their expected utility must be higher.

$$(10) \qquad prob(\theta = H|g(e,\tau)) = \frac{1}{1 + f(g|e_{sL}^*)/f(g|e_{sH}^*)}$$

which is monotonically increasing in g by equation (8) and the Monotone Likelihood Ratio Property. This means that an increasing smile function $\gamma(g)$ is indeed consistent with natural selection and therefore we can predict, substituting the smile function into equation (6), that

$$(11) \qquad \frac{\partial q_\gamma}{\partial g} > 0$$

Finally, given that the convincingness of smiles is the result of effort in the way described in equation (10), we can calculate how the expected gain to A from sending money is related to smile quality. We write the expected gain to A from sending the money, conditional on smile quality as follows

$$(12) \qquad E\ (U_A|g,s)$$

$$= \left(pr(\theta = H|g)\right)\left[1.5s.p_H + s(1-p_H)(\alpha_H + \frac{(1-s)}{2})\right]$$

$$+\left(1 - pr(\theta = H|g)\right)\left[1.5s.p_L + s(1-p_L)(\alpha_L + \frac{(1-s)}{2})\right] - s$$

We can rewrite (12) as

(13) $E\ (U_A|g,s)$

$$= \left(pr(\theta = H|g)\right)\left[1.5s.(p_H - p_L) + s(\alpha_H - \alpha_L) - \frac{s(1-s)}{2}(p_H - p_L)\right.$$

$$\left. + s(p_L\alpha_L - p_H\alpha_H)\right] + \left[1.5s.p_L + s(1-p_L)(\alpha_L + \frac{(1-s)}{2})\right]$$

and therefore we can write the derivative of $E\ (U_A|g,s)$ with respect to $pr(\theta = H|g)$ as

(14) $$\frac{\partial E\ (U_A|g,s)}{\partial pr(\theta = H|g)} = s[(s+0.5)(p_H - p_L) + (\alpha_H - \alpha_L) + (p_L\alpha_L - p_H\alpha_H)\ ]$$

which is strictly positive because

(15) $[(s+0.5)(p_H - p_L) + (\alpha_H - \alpha_L) + (p_L\alpha_L - p_H\alpha_H)\ ] >$

$(s+0.5)(p_H - p_L) + \alpha_H(1-p_H) - \alpha_H(1-p_L) =$

$[(s+0.5)(p_H - p_L) - \alpha_H(p_H - p_L)\ ]$

and the expression on the RHS is positive for any $s \geq 0.5$.

From this it follows, given (10), that

(16) $$\frac{\partial E\ (U_A|g,s)}{\partial g} > 0 \qquad \forall g,s$$

which is just the statement that the expected gain to player A from sending the stake to player B is increasing in the perceived convincingness of player B's signal.

e. Testable implications

Our test of the hypothesis that smiling convincingly is a costly signal, therefore, amounts to testing the following:

The costliness of the signal: inequality (9);

The effectiveness of the signal in inducing cooperation: inequality (11);

The informativeness of the signal: inequality (16).

IV. Experimental Methods and Data Description

We use a simplified version of the original trust game, proposing senders a binary choice of trust or no trust and trustees three different return options[15]. Sessions for trustees were conducted first, to allow them to record their video messages. Their actions were obtained by the strategy method at the same moment. Decisions for senders and trustees were incentivized and earnings were according to their partners' decisions. Therefore payout for trustees did not take place until after senders had made their respective trust decisions.

Video messages were produced by eighty-four volunteers aged between 18 and 35 years recruited from the general population in Toulouse, France. We told trustees they would face two different but unknown partners who might be persuaded to send them a sum that would be tripled if they did so. The difference between the two games was that in the first, trustees recorded their video message before being

[15] See Hopfensitz & Reuben (2009).

informed about the precise payoffs and thus before taking their decision, while in the second game they recorded their message after taking their return decision[16]. The trustee had to choose between returning nothing, the sender's original stake or 1.5 times the sender's original stake (thus half the total amount). Trustees were randomly split into two payoff treatments: stakes in the Low treatment were 4 euros, and stakes in the High treatment were 8 euros. Video clips were made on a professional TV platform, and a practice clip helped participants to become familiar with the environment. A total of 168 video clips was obtained, two for each trustee. Verbal messages during clips were standardized by giving a predetermined sentence to trustees that had to be memorized. To make the message natural for trustees they included in this sentence their name, age and occupation and were reassured that the precise wording did not matter. Video clips lasted around fifteen seconds on average, with the fastest at around ten and the slowest at around twenty seconds.

Senders' behavior and evaluation of video clips were obtained in two waves of experiments conducted in a different experimental laboratory to minimize the risk that senders might recognize trustees. A total of 198 student participants were recruited at the University of Lyon; 84 senders participated in the first wave and 114 in the second. The difference between the two waves was that participants in the first wave were matched with trustees and their decisions determined trustees' payoffs. Participants in the second wave made trust decisions and were paid according to

[16] These two treatments were always presented in the same order, due to the impossibility of having participants first play a game where they are informed about payoffs and then a game where they do not know about the different payoff options.

trustees' initial decision. Observations from the first wave therefore concern games where sender and trustee faced the same stake size. In the second wave senders were informed only that trustees might return nothing, their initial endowment or 1.5 times their endowment. The actual stake size of their partner was varied: half of each sender's video messages came from participants in the High stake treatment and half came from the Low stake treatment.

In the first and second waves each sender viewed a total of 42 (respectively 28) clips in two series of 21 (resp. 14). For each clip the sender was asked to decide whether to send money to the trustee, and then to rank the trustee on an 8-point scale along a number of dimensions including how much they smiled, how genuine were their smiles, their attractiveness, their trustworthiness, their intelligence and their self-confidence. Senders were matched at random with one of the trustees from each series and received payoffs determined by the actual decision of this real partner. To ensure anonymity for trustees it was not revealed which of the clips viewed had been selected to determine senders' payoffs. From the first wave a total of 21 decisions and ratings concerning each of the 168 clips was obtained. In the second wave senders were presented with one of three different payoff treatments (Low: 4 Euro; High: 8 Euro; Super High: 12 Euro) and made half of their decisions concerning partners from the Low and half concerning partners from the High trustee treatment. From the second wave an average of 19 decisions and ratings for each clip were obtained[17]. Figure 1 summarizes the choices made by the participants. [Figure 1 here]

All results reported use the pooled data from the two waves of the experiment. A dummy variable distinguishing the two waves was never significant in any specification, indicating that the two waves were conducted under indistinguishable conditions.

[Table 2 here]

Table 2 summarizes the characteristics of our participants for each group of players. The first column reports the means and standard deviations for the senders. This subsample is exactly gender-balanced, the average age is 22 years old and 92% are currently students. For a given sender, the proportion of decisions for which she decided to send money ranges from 0 to 1. Of the 198 participants, 3 participants decided to send money to every partner, and 20 decided never to send money. On average, each sender decides to cooperate with 37% of their partners. Column (2) concerns the trustees, who are 25 years old on average. The sample of trustees comes from a less homogeneous population than the senders since only 46% are students. The sample of trustees is fairly balanced with 55% of women. As described above, the trustees recorded two video clips and made a decision about how to share the pie for each clip. In the first decision (made after the first clip), 55% of the participants

---

[17] Owing to some volunteers' not turning up for their session of the experiment there was a slight variation in numbers of ratings per clip: the clips in our dataset have an average of 40 ratings each, with a minimum of 38 and a maximum of 45.

decided to share equally the amount received multiplied by 3, 31% decided to send back the original stake to the sender and to keep 2/3 of the total pie, and the remaining 14% decided to return nothing. In the second decision (made before the second clip), the distribution suggests a slight shift from equal sharing to sending nothing.

V. Results

Figure 2 shows a comparison of High and Low treatments of trustees in terms of the average ratings by senders of the genuineness of their smiles, their average trustworthiness rating, the proportion of senders who decided to send money, and the proportion of trustees who chose to return at least some of the money they received. The latter is a measure of their being motivated by social preferences, since trustees had nothing financially to gain from returning the money to senders.

[Figure 2 here]

Figure 2 indicates that trustees under the High treatment (playing for twice the stakes of those under the Low treatment) are perceived as having more genuine smiles and as being more trustworthy, and are associated with a higher percentage of senders sending money, although a smaller percentage of trustees under the High treatment actually return any money to the senders. It might be thought that the positive correlation of the High treatment with smile genuineness and with selfish behavior would imply that genuine smiles are positively correlated with selfish behavior. However, Figure 3 indicates that this is not so. The explanation is that those in the

High group who *succeeded* in making genuine smiles were not a random subset of those in the High treatment; they were a more unselfish group than those who did not succeed.

In Figure 3 clips are divided into those whose smiles were given average ratings above 5 (46% of the clips) and the rest. Clips with smiles perceived as genuine were given higher ratings for trustworthiness, attractiveness and intelligence, and were associated with a higher willingness to send money, but were also associated with a higher willingness to return some money to senders. So overall it appears that the High treatment created both a higher incentive to smile in a way perceived as genuine, and a higher incentive to be selfish instead of returning money to the sender, but that those who succeeded in smiling genuinely were those who were most likely to return the money. As Figures 2 and 3 indicate, all of these comparisons were significant at 5% levels, and all but two were significant at well under 1%. We report one-tailed p-values as the theory yields unambiguous predictions as to the sign of the difference between the groups.

[Figure 3 here]

These comparisons by treatment cannot tell us whether there is a genuine causal effect, for two reasons. First, because of the experimental design and the need to match senders with trustees in the same treatment in the first wave, there was some correlation in the overall data between the treatment of trustees and the treatment of senders: these two treatments need to be separately controlled. Secondly, it is important to control for a number of other dimensions along which clips differed, such as age, gender and so forth. For these reasons multivariate regression analysis is required. We consider our three component hypotheses in turn. In all cases, in order to avoid possible "justification effects" in which users' ratings are influenced by the decisions they have already taken whether to send money to the trustees, we use as measures of smile quality, trustworthiness and attractiveness the average rating of each clip across all viewers, rather than the rating given by the individuals themselves; this requires, however, that standard errors be calculated clustering by clip.

Table 3 reports our tests of these three hypotheses. Equation A tests the hypothesis that smiling convincingly is costly, and therefore is more likely to happen under the High treatment when the returns to smiling convincingly are higher. The treatment effect is significant at under 2%: 0.12 points, which is about 36% of one standard deviation of the distribution of mean ratings by clip[18].

[Table 3 here]

Other notable features of Equation A are that more convincing smiles are associated with trustees who are rated as more intelligent, and also with older trustees. Trustees with beards are rated as having less convincing smiles, and women with a significant décolleté are rated as having significantly more convincing smiles (it is not

[18] Though this is not a large effect, our treatment was itself not very large. Although the difference between a 24 Euro maximum gain and a 48 Euro maximum gain may seem important for an experimental session that lasted only around half an hour, the mean gain to trustees in the High treatment (which was of course reduced if senders did not send money, and if trustees returned part of the money) was just over 12 Euros, compared to a little over 7 Euros for those in the Low treatment.

clear whether the causal mechanism is via the psychology of the smiler or of the viewer). There is no significant effect of gender of either the sender or the trustee, nor of perceived attractiveness of the trustee, and no effect of whether the clip is filmed before or after the decision to return the money has been sent. These coefficients (like those of other controls) are not reported, though the full specification is available from the authors.

Equations B and C test the hypothesis that smiling convincingly induces cooperation, the first by looking at what subjects say and the second by looking at what subjects do. Equation B considers whether convincing smiles are associated with judgments of greater trustworthiness. There is a massively significant correlation (t-ratio of over 8); a one-point deviation increase in smile quality is associated with slightly more than a half point increase in perceived trustworthiness. Perceived intelligence is also positively and very significantly correlated with perceived trustworthiness, and there is a significant effect of the High treatment independently of smile quality, suggesting that trustees are putting effort into other dimensions of non-verbal communication as well. Equation C examines whether convincing smiles lead to an increased probability of sending money to the trustees. Once again there is a massively significant association: a one point increase in smile quality leads to a 21% increase in the probability of sending the money, which is equivalent to a 7% increase per standard deviation of smile quality. Perceived intelligence is again a very important factor in the decision.

Equation D tests the hypothesis that convincing smiles are an informative signal, that is, whether they are associated with higher gains to the sender from cooperating with the trustee. We therefore regress gains from sending money on average smile quality, without other controls. The effect is small but positive (0.9 Euros per one point increase, or 0.3 Euros per standard deviation) and significant at under 1%[19].

Finally, we investigate whether convincing smiles are associated more closely with the amount of money the trustee has available to offer the sender or with the intrinsic character of the trustee. We therefore regress a dummy variable for the High treatment on the mean smile quality, and in a separate equation we regress on the same measure of smile quality a dummy variable indicating that the trustee takes an unselfish decision. The purpose is to see whether the smile quality is a reliable informative of the amount available to share, and of the character of the trustee. As reported in Table 4, the smile quality is positively related to both the size of the pie to be shared and the unselfish behavior by the trustee. Two caveats are in order, however. First, the coefficient on unselfish behavior by the trustee has a large standard error so we cannot be confident in its measurement, which is not statistically significantly different from zero at conventional levels (unlike the coefficient on the High treatment). Secondly, as we saw in Figure 2, unselfish behavior itself appears to be influenced by the treatment, so we cannot be confident in treating it as a measure of the intrinsic character of the trustee. So we should conclude that smile quality is

[19] In the decision equation (though not in the others) the treatment effect for senders is large and highly significant, and follows an inverted-U shape: a High treatment makes senders more likely to send money, but a Superhigh treatment makes them much less likely to do so.

definitely informative about high cooperation opportunities, and *may* also signal the intrinsic character of the smiler.

VI. Conclusions

We have tested three component hypotheses of the theory that smiling convincingly is a costly signal that has evolved to induce cooperation in situations requiring mutual trust. All three components are supported by the evidence. First, smiling convincingly is costly, because smiles from trustees playing for higher stakes are rated as significantly more convincing, by over a third of one standard deviation; this strongly suggests that they are produced when there are rewards to the additional effort required. Secondly, we show that smiling convincingly induces cooperation: smiles rated as more convincing are strong predictors of judgments about the trustworthiness of trustees, and of the revealed willingness to send them money. Finally, we show that smiling convincingly is an informative signal: those smiling convincingly return more money on average to senders. It is clearly informative of the amount the trustee has available to share with the truster; there is weaker evidence that it may also be a signal of the intrinsic trustworthiness of the trustee independently of the amount at stake.

References:

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. “Trust, reciprocity, and social history.” *Games and Economic Behavior*, 10: 122-142.

Boone, R. Thomas, and Ross Buck. 2003. “Emotional Expressivity and Trustworthiness: the role of nonverbal behavior in the evolution of cooperation.” *Journal of Nonverbal Behavior*, 27(3): 163-182.

Bosman, Ronald, and Frans van Winden. 2002. “Emotional hazard in a power to take experiment.” *The Economic Journal*, 112:147-169.

Brown, William Michel, and Chris Moore. 2002. “Smile Asymmetries and Reputation as Reliable Indicators of Likelihood to Cooperate: An Evolutionary Analysis.” In *Advances in Psychology Research*, ed. S.P. Shobov, vol. 11, 59-78. New York: Nova Science Publishers.

Brown, William Michel, Boris Palameta, and Chris Moore. 2003. “Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm.” *Evolutionary Psychology*, 1: 42-69.

Brülhart, Marcus, and Jean-Claude Usunier. 2004. “Verified Trust: Reciprocity, Altruism, and Noise in Trust games”, Working paper 04.15, Université de Lausanne.

Chang, Luke J., Bradley B. Doll, Mascha van 't Wout, Michael J. Frank, and Alan G. Sanfey. 2010. “Seeing is believing: Trustworthiness as a dynamic belief.” *Cognitive Psychology*, 61: 87–105.

Cohn, Jeffrey F., and Karen L. Schmidt. 2004. “The timing of facial motion in posed and spontaneous smiles.” *International Journal of Wavelets, Multiresolution and Information Processing*, 2 : 1-12.

Cosmides, Leda, and John Tooby, J. 1992. “Cognitive adaptations for social exchange.” In *The adapted mind: Evolutionary psychology and the generation of culture,* ed. Jerome H. Barkow, Leda Cosmides, and John Tooby. New York: Oxford University Press.

Damasio, Antonio. 1994. *Descartes’ error: Emotion, reason and the human brain.* New York: Avon Books.

Darwin, Charles R.. 1872. *The Expression of the Emotions in Man and Animals.* London: John Murray, London.

de Quervain, Dominique, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. “The neural basis of altruistic punishment.” *Science*, 305 (5688): 1254-1258.

Duchenne de Boulogne, Guillaume. 1862. *The Mechanism of Human Facial Expression.* Paris: Jules Renard.

Eckel, Catherine C., and Ragan Petrie. 2008. “Face Value”, Experimental Economics Center Working Paper Series 2008-11, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Ekman, Paul. 1982. *Emotion in the Human Face*. Cambridge: Cambridge University Press.

Ekman, Paul, and Wallace V. Friesen. 1982. “Felt, false and miserable smiles”. *Journal of Nonverbal Behavior*, 6: 238-252.

Elster, Jon. 1998. “Emotions and economic theory.” *Journal of Economic Literature*, 36 (1): 47–74.

Fehr, Ernst, and Fischbacher, Urs. 2005. “Altruists with Green Beards.” *Analyse & Kritik*, 27/2005, 73–84.

Fehr, Ernst, Urs Fischbacher, and Simon Gächter, S. 2003. “Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms.” *Human Nature*, 13 (1):1-25.

Frank, Robert H. 1988. *Passions within Reason. The Strategic Role of the Emotions*. New York: Norton.

Frey, Bruno. 2008. *Happiness: A Revolution in Economics*. Cambridge, MA and London, UK: MIT Press.

Frey, Bruno, and Suzanne Neckermann. 2009. “Awards: A Disregarded Source of Motivation.” In *Perspectives in Moral Science – Contributions from philosophy, economics, and politics in honour of Hartmut Kliemt*, ed. Michael Baurmann and Bernd Lahno, 177-182. Frankfurt: Frankfurt School Verlag.

Frijda, Nico, Antony Manstead, and Agneta Fischer. 2004. “Feelings and emotions: Where do we stand?” In *Feelings and Emotions: The Amsterdam Symposium*, ed. Antony Manstead, Nico Frijda, and Agneta Fischer. Cambridge: Cambridge University Press.

Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr. 2003. “Explaining altruistic behavior in Humans.” *Evolution and Human Behavior*, 24: 153-172.

Gintis, Herbert, Eric Alden Smith, and Samuel Bowles. 2001. “Costly Signaling and Cooperation”. *Journal of Theoretical Biology*, 213(1): 103-119.

Grafen, Alan. 1990. “Biological signals as handicaps.” *Journal of Theoretical Biology*, 144: 517-546.

Hopfensitz Astrid, Ernesto Reuben. 2009. “The importance of emotions for the effectiveness of social punishment.” *The Economic Journal*, 119: 1534-1559.

Hwang, Sung-Ha, and Samuel Bowles. 2010. “Is Altruism Bad for Cooperation?” Working paper, University of Massachusetts.

Johnston, Lucy, Lynden Miles, and C. Neil Macrae. 2010. “Why Are You Smiling at Me? Social functions of enjoyment and non-enjoyment smiles.” *British Journal of Social Psychology*, 49: 107-127.

Kahneman, Daniel, Ed Diener, and Norbert Schwarz (Eds.). 1999. *Well-being: The foundations of hedonic psychology*. New York: Russell Sage Foundation.

Kahneman, Daniel. 2003. “A psychological perspective on economics.” *The American Economic Review*, 93 (2): 162–168.

Krumhuber, Eva, Anthony S.R. Manstead, Darren Cosker, Dave Marshall, Paul L. Rosin, and Arvid Kappas. 2007. “Facial dynamics as indicators of trustworthiness and cooperative behaviour.” *Emotion*, 7: 730-735.

Loewenstein, George. 2000. "Emotions in economic theory and economic behaviour." *American Economic Review*, 90(2): 426–432.

Lotem, Arnon, Michael A. Fishman, and Lewi Stone. 2003. "From reciprocity to unconditional altruism through signalling benefits." *Proceedings of the Royal Society B,* 270: 199-205.

Manzini, Paola, Abdolkarim Sadrieh, and Nicolaas J. Vriend. 2009. "On Smiles, Winks and Handshakes as Coordination Devices." *The Economic Journal*, 119 (537): 826–854.

Mehu, Marc, Karl Grammer, and Robin Dunbar. 2007a. "Smiles when Sharing." *Evolution and Human Behavior*, 28: 415–422.

Mehu, Marc, Anthony C. Little, and Robin Dunbar. 2007b. "Duchenne smiles and the Perception of Generosity and Sociability in the face." *Journal of Evolutionary Psychology*, 5, 1–4, 183–196.

Mehu, Marc, and Robin Dunbar. 2008. "Naturalistic Observation of Smiling and Laughter in Human Group Interactions." *Behaviour*, 145(12): 1747-1780.

Oda, Ryo, Noriko Yamagata, Yuki Yabiku, and Akiko Matsumoto-Oda. 2009. "Altruism Can Be Assessed Correctly Based on Impression." *Human Nature,* 20 (3): 331-341.

Owren, Michael J, Jo-Anne Bachorowski. 2001. "The evolution of emotional expression: A "selfish-gene" account of smiling and laughter in early hominids and humans." In *Emotions: Current Issues and Future Directions*, ed. Tracy J. Mayne, and George A. Bonnano, 152-191. New York: Guilford Press.

Niedenthal, Paula, Martial Mermillod, Marcus Maringer, and Ursula Hess. 2010. “The Simulation of Smiles (SIMS) Model: Embodied Simulation and the Meaning of Facial Expression.” *Behavioral and Brain Sciences,* 33 (06): 417-433.

Sanfey, Alan, James Rilling, Jessica Aronson, Leigh Nystrom, and Jonathan J. Cohen. 2003. “The neural basis of economic decision making in the ultimatum game.” *Science*, 300 (5626):1755-1758.

Scharlemann, Jörn PW, Catherine C. Eckel, Alex Kacelnik, and Rick K. Wilson. 2001. “The value of a smile: Game theory with a human face.” *Journal of Economic Psychology,* 22: 617-640.

Schug Joanna, David Matsumoto, Yutaka Horita, Toshio Yamagishi, and Kemberlee Bonnet. 2010. “Emotional expressivity as a signal of cooperation.” *Evolution and Human Behavior*, 31 (2): 87-94.

Smith, Adam. 2000. *The Theory of Moral Sentiments*, New York: Prometheus Books.

Smith, Eric Alden, and Rebecca L. Bliege Bird. 2000. “Turtle hunting and tombstone opening: public generosity as costly signalling.” *Evolution and Human Behavior*, 21(4): 245-261.

Sobel, Joel. 2005. “Interdependent Preferences and Reciprocity.” *Journal of Economic Literature*, Vol. XLIII (June), 392–436.

Sommerfeld, Ralf D., Hans-Jürgen Krambeck, and Manfred Milinski. 2008. “Multiple gossip statements and their effect on reputation and trustworthiness.” *Proceedings of the Royal Society B*, 275: 2529-2536.

Spence, Michael. 1974. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge: Harvard University Press.

Strack, Fritz, Leonard Martin, and Sabine Stepper. 1988. "Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis." *Journal of Personality and Social Psychology,* 54 (5): 768-777.

Trivers, Robert. 2000. "Elements of a scientific theory of self-deception." *Annals of the New York Academy of Sciences*, 907: 114-131.

van 't Wout, Mascha, and Alan Sanfey. 2008. "Friend or foe: The effect of implicit trustworthiness judgments in social decision-making." *Cognition*, 108 (3): 796–803.

Verplaetse, Jan, Sven Vanneste, and Johan Braeckman. 2007. "You can judge a book by its cover: the sequel. A kernel of truth in predictive cheating detection." *Evolution and Human Behavior,* 28: 260-271.

Vrij, Aldert, Gün R. Semin, and Ray Bull. 1996. "Insight into behavior displayed during deception". *Human Communication Research,* 22: 544-562.

Wylie, Douglas R., and Melvyn A. Goodale. 1988. "Left-sided oral asymmetries in spontaneous but not posed smiles." *Neuropsychologia*, 26: 823-832.

Yamagishi, Toshio, Shigehito Tanida, Rie Mashima, Eri Shimoma, and Satoshi Kanazawa. 2003. "You can judge a book by its cover: Evidence that cheaters may look different from cooperators." *Evolution and Human Behavior,* 24: 290-301.

Zahavi, Amotz. 1975. "Mate Selection – A Selection for a Handicap." *Journal of Theoretical Biology*, 53 (1): 205-214.

Table 1: Probabilities that player B chooses various values of *m*

	***m*=0**	***m*=1**	***m*=1.5**
High Reciprocator type (θ=H)	$(1-p_H)(\frac{(1+s)}{2}-\alpha_H)$	$(1-p_H)(\alpha_H+\frac{(1-s)}{2})$	p_H
Low Reciprocator type (θ=L)	$(1-p_L)(\frac{(1+s)}{2}-\alpha_L)$	$(1-p_L)(\alpha_L+\frac{(1-s)}{2})$	p_L

Table 2: Summary statistics

	(1)	(2)
	A Players (trusters)	B players (trustees)
Proportion Male	0.50 (0.50)	0.45 (0.50)
Mean Age	21.58 (4.37)	24.86 (4.73)
Proportion Student	0.92 (0.27)	0.46(0.50)
Proportion Trusting	0.37 (0.23)	
First decision:		
Equally share the triple stake		0.55
Send back the original stake		0.31
Send back nothing		0.14
Second decision:		
Equally share the triple stake		0.49
Send back the original stake		0.32
Send back nothing		0.19
Number of observations	198	84

Note: standard deviations in parenthesis.

Table 3: Tests of components of costly signaling hypothesis

	Equation A	**Equation B**	**Equation C**	**Equation D**
Dependent Variable:	**Smile Quality (scale 1-8)**	**Trustworthiness (scale 1-8)**	**Decision to send money (send=1)**	**Gain (Euros) from sending money**
Ind. variable (characteristics of trustee):				
High treatment	0.124** (0.017)	0.079** (0.026)	-0.011 (0.400)	
Smile quality		0.54*** (0.000)	0.219*** (0.000)	0.911*** (0.005)
Intelligence	0.213** (0.030)	0.310*** (0.000)	0.154** (0.012)	
Age of trustee	0.0149** (0.017)	0.007 (0.139)	0.005 (0.142)	
Beard	-0.262*** (0.003)	-0.006 (0.867)	-0.050 (0.404)	
Décolleté	0.229** (0.011)	-0.008 (0.821)	0.067 (0.188)	

Note: p-values in parentheses, one-tailed values reported for High treatment and Smile quality, two-tailed values for other variables. Standard errors clustered by clip, **=significant at 5%; ***=significant at 1%. Equations A, B and D are estimated by ordinary least squares, Equation C is a probit estimated by maximum likelihood. Other controls include gender of truster and trustee, perceived attractiveness, video sequence, truster's treatment, dummy variables for the trustee being black, having visible piercing and wearing glasses in equations A, in addition to perceived self-confidence, and smile frequency in Equation B. Equation C adds to the previous set of controls a dummy variable for A players' self-reported unselfish behavior from the General Social Survey, income, age of truster and score on a simple intelligence test. The estimated coefficients for the other controls not reported, available from authors on request. Number of observations= 6,720. Variables Smile Quality, Intelligence and Trustworthiness are means by clip.

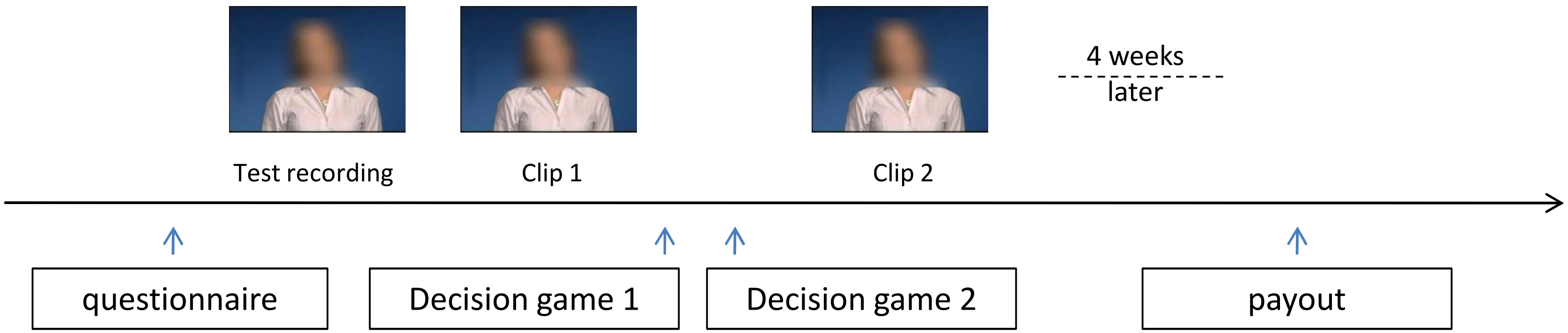
Table 4: Character or Opportunity? Determinants of smile quality

Dependent Variable:	**Trustee is in High treatment (dummy variable)**	**Unselfish behavior by trustee (dummy variable)**
Independent Variable:		
Mean smile quality rating	0.516 (0.038)**	0.403 (0.126)

Note: one-tailed p-values in parentheses, standard errors clustered by clip, **=significant at 5%. Probit equations estimated by maximum likelihood. Number of observations=6,720.

Figure 1: Order of choices for Second and First Movers in the Experiment

Second mover:



Note: faces are blurred in the Figure to preserve subject anonymity in publication; actual clips were not blurred

First mover:

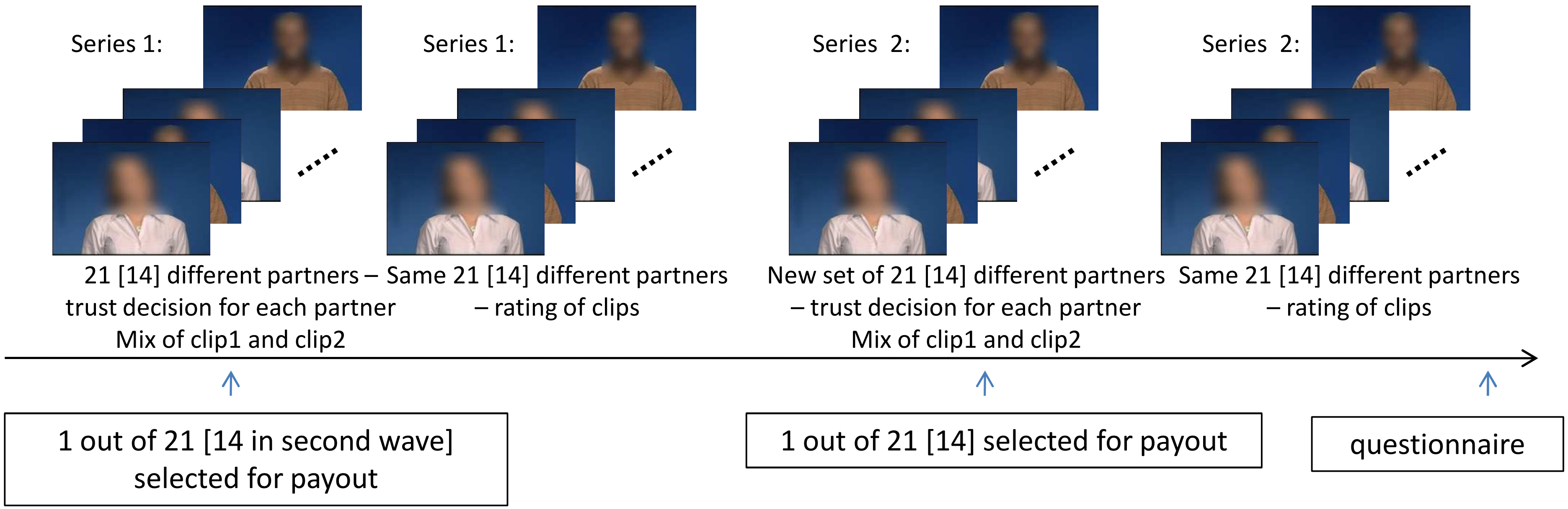
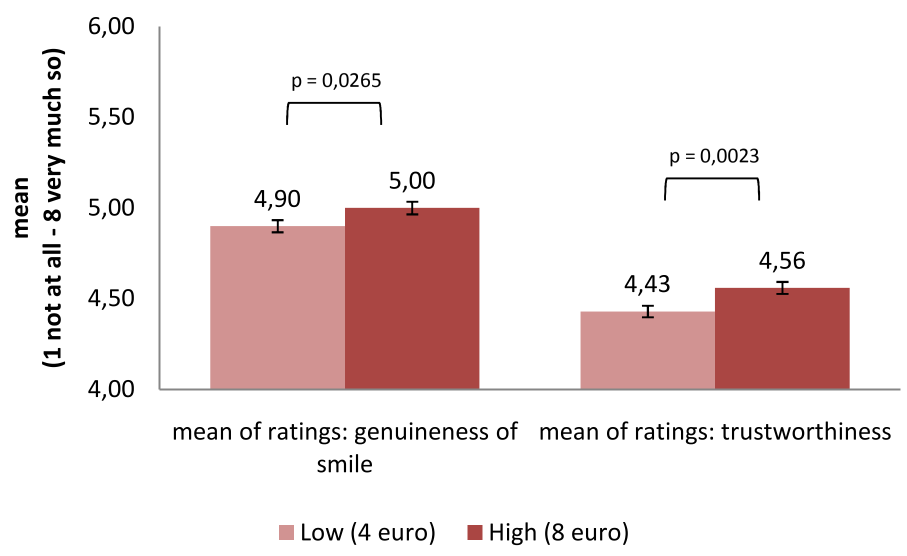


Figure 2: Differences in (a) smile quality and (b) trust and trustworthiness by treatment

(a)



(b)

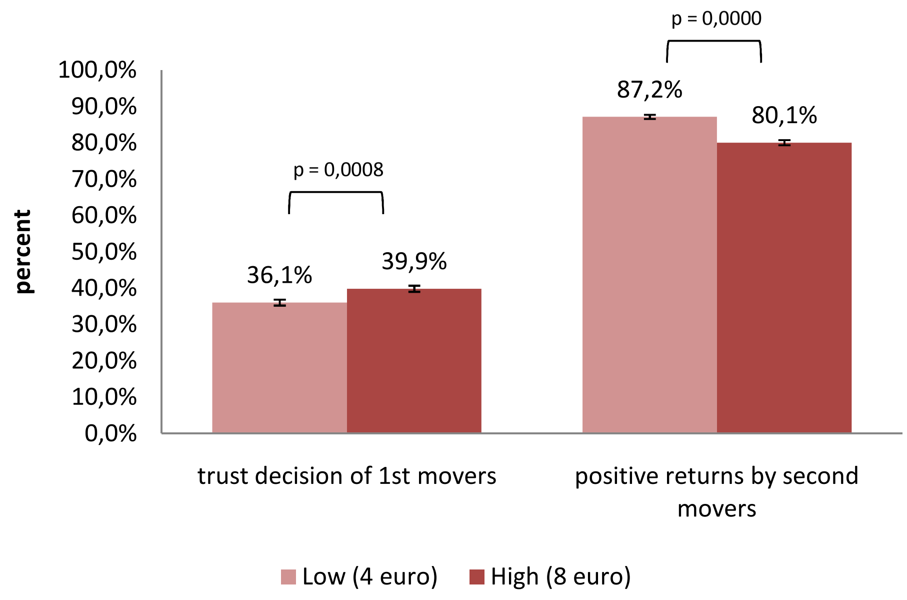
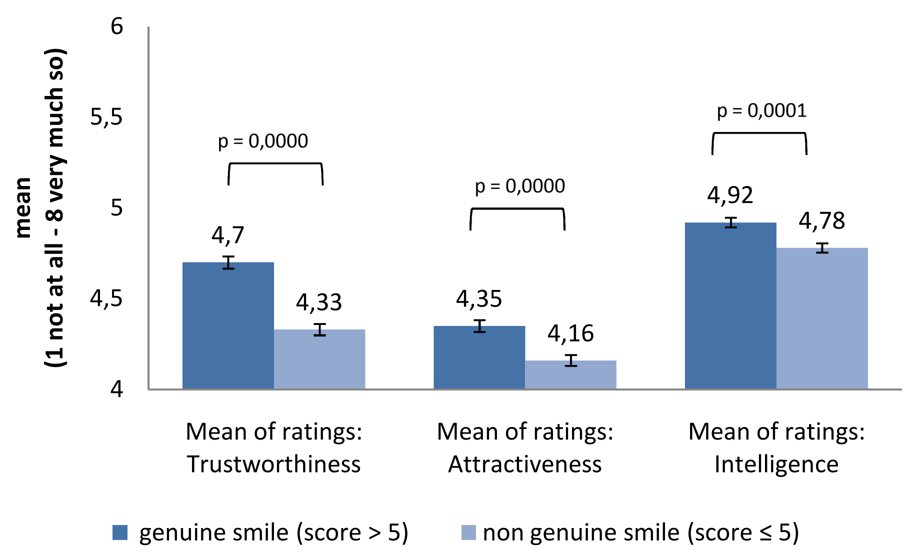


Figure 3: Differences in (a) ratings of trustees and (b) trust and trustworthiness by smile quality

(a)



(b)