

EPSA09: Genuine versus Deceptive Emotional Displays

This paper contributes to the explanation of human cooperative behaviour, examining the implications of Brian Skyrms' modelling of the prisoner's dilemma (PD). Augmenting a PD with signalling strategies promotes cooperation, but a challenge that must be addressed is what prevents signals being subverted by deceptive behaviour. Empirical results suggest that emotional displays can play a signalling role and, to some extent, are secure from subversion. I examine proximate explanations and then offer an evolutionary explanation for the translucency of emotional displays. Selection acts on the basis of *lifetime* fitness consequences and, crucially for my argument, the intensity of selection decreases over the course of a lifetime. Hence we tend to possess traits that promote survival when young and, with regard to emotional displays, translucency allows successful maturation over our protracted period of nurturing by close kin. This is due to the vital role played by emotional interactions in the normal cognitive and social development of *Homo sapiens*.

The prisoner's dilemma, positive assortment and signalling.

Evolutionary models of one-shot PDs clearly demonstrate that where cooperators and defectors are paired randomly cooperation is driven to extinction (Alexander 2007). Skyrms' conclusion on the PD is that positive assortment of cooperators can promote the evolution of cooperation (Skyrms 1996). That is, if cooperators interact with each other at high enough frequency above the random pairing rate, 100% cooperation is an equilibrium outcome, a claim that has been established by evolutionary game theorists across other disciplines. Indeed, it has been argued that many models of the evolution of altruism share an underlying mathematical structure- that of Hamilton's Price equation formulation of inclusive fitness theory. Hamilton's relatedness coefficient between individuals can be interpreted at a general level as the degree of positive assortment of types and need make no reference to common descent (McElreath & Boyd 2007). Interesting though this common mathematical form is, it is impossible to test the relationship of a model to actual strategic interactions without considering the mechanism that generates positive assortment. However, Skyrms is clear that his focus is only on the theoretical possibility of cooperation.

Suppose that nature has somehow- I don't care how- arranged high correlation between like strategies among individuals playing the Prisoner's Dilemma.
(Skyrms 2000: 280)

In light of Skyrms' project to explain the evolution of various features of the social contract, it is not enough to offer only a proof of the possibility of cooperative behaviour. Proposals are also required for an actual mechanism that yields positive assortment. Even without a formal model one can see that if only cooperators could get together, and exclude defectors, they would do better.¹ The taxing problem is how assortment is achieved.

One mechanism is population viscosity, where players interact only with their neighbours (Alexander 2007). Another involves cooperators signalling their type and conditionalising their behaviour on the signal or its absence. It is the second of these to which Skyrms alludes in his model of correlated interaction. Implicit in the model is the assumption that cooperators can recognise each other, something the following remark makes explicit.

Correlation may be established by some sort of sensory detection. For instance, cooperators and defectors might emit different chemical markers. (Skyrms 1996: 56)

Various features of the model might be questioned but, for my purposes, I concentrate on the signalling assumption. The question to be asked is what prevents defectors emitting the cooperative signal but defecting in the PD or, following Skyrms' example, what prevents a defector developing the capability to release the cooperators' chemical marker? Arthur Robson's "secret handshake" model further illustrates this concern (Robson 1990). Here cooperators use an unsubverted signal of their cooperative nature and, unsurprisingly, such "handshakers" can invade defectors and 100% cooperation is an ESS of the model. Skyrms cites Robson's paper as an example of a model of the evolution of cooperation via assortment. Unfortunately, from the point of view of cooperators, whilst cooperation is an ESS of the model, it is not resistant to invasion by a novel strategy and Robson is explicit on this.

The difficulty is just that such an [cooperative] ESS is a "sitting duck" for the introduction of still another mutant, one which would prey on the first [signalling] mutant. The second mutant should carry the signal, but play "u" [defect] against the first mutant as well as against the non-signaling population. (Robson 1990: 387. Parentheses added)

To reiterate, a strategy that emits a "fake" signal of cooperation and then defects in the PD can invade the population. Such a faking defector could occur in two ways. Firstly, an ordinary defector could mutate and gain the signal or, secondly, a handshaking cooperator could mutate and switch PD behaviour. Unless the handshake remains secret, cooperation is unstable.

Hence, models that postulate cooperative signals must offer reasons to believe that the signals are secure from subversion or commit what Ken Binmore calls the "transparent disposition fallacy" (Binmore 1994). David Gauthier exemplifies the fallacy in his work on "constrained maximization".

Cooperators are “translucent” but no account is given of a mechanism underpinning translucency and, hence, the question of subversion cannot even be raised (Gauthier 1986). In this paper I examine empirical and theoretical support for emotional displays as a candidate signalling mechanism.

Emotional displays as signals.

First, let me stress that I am not examining the potential for emotions to act as commitment devices that bind individuals to cooperation. Several authors make this claim and I set that debate to one side, focusing only on the possible signalling role of emotions (Elster 2000). Robert Frank has a combined commitment plus signal theory of emotion but on the signal side restricts his empirical evidence to work on deception detection (Frank 1988). The next section considers that evidence but later I add considerations from the neurobiology of emotion and, more important with respect to evolutionary considerations, work in theoretical biology. Like Skyrms, I take it that cultural evolution is the mechanism to examine when it comes to PD behaviour (Skyrms 1996). However, my interest is in signalling and I assume that emotional displays are, to a great extent, the product of biological evolution. I am not suggesting that the *origin* of emotional displays has its roots in their signal value in PD-like interactions. Rather, once evolved, they are available resources for a process of cultural evolution. An initial question to ask is whether or not players in PD-like interactions track emotional display signals. Scharlemann et al. (2001) investigate whether or not smiles increase trusting behaviour and find conclusively that they do. It might be argued that this result is unsurprising but the authors claim that theirs is the first controlled test with financial stakes of the relationship between smiles and trust.

My aim is to offer theoretical support for an evolutionary explanation for the translucency of emotions. Before doing so I turn to evidence on deception detection in general and the detection of cooperative intent in PD-like interactions. By translucency I mean that visible displays are well, but imperfectly, correlated with genuine emotional episodes.

Detection of deception and cooperation.

The literature on deception detection is intimately tied to that on emotion detection precisely because emotional displays are difficult to control. Emotion and deception can be connected in a number of ways. Firstly, an emotional response can play the dual action-causing and signalling roles

proposed by Frank. Secondly, the act of deception itself can trigger an emotional response.² The phenomena of detection apprehension, where deceivers fears being caught, and deception guilt are well known. False “promises” in PDs are likely to cause both, typically meeting many of their elicitation criteria. These include; target suspicion, target unwillingness to be fooled, target suffering, the deceit being selfish and, often, the deceiver and target sharing social values (Ekman 1992).

In a comprehensive meta-analysis, Bond & DePaulo (2006) find deception detection accuracy of 53.5%. Such a margin above random performance (50%) is, on the face of it, unimpressive. However, the result is significant at a level of 0.01% and the confidence interval excludes 50%. Also, the effect size is larger than 60% of a sample of 474 reported effects in the social psychological literature. The phenomenon of deception detection is well confirmed as real.

Also relevant are careful studies that demonstrate “leakage” of emotional displays, sometimes via what Paul Ekman calls “microexpressions” (Ekman 2009). It is these studies, referred to in the next section, that form evidence for the involuntary nature of some muscle movements and also reflect the other proximate explanations for emotional translucency. Ekman criticises many deception studies on the basis that they involve trivial and ecologically unrealistic situations in which emotional responses to committing an act of deception are slight (Ekman 2003: 206).

Additional studies test directly whether subjects can predict cooperation in PDs and other interactions that require trust. The evidence here is somewhat mixed. Ockenfels & Selten (2000) find no evidence of involuntary signals in bargaining interactions and Jeannette Brosig (2002) finds better than chance prediction of cooperation but not of fair bargaining in face-to-face games. On the other hand, Pradel et al. (2009) do find accurate prediction of fairness but in a study in which players were know to each other. Frank et al. (1993) find better than chance prediction in a PD in which players spend 30 minutes together beforehand. Yamagishi et al. (2003) find that observers have better recall of photos of defectors faces than of cooperators and Verplaetse et al. (2007) find that observers can predict cooperative behaviour from photos taken at the moment of decision.³ Fetchenhauer et al. (2010) even find that subjects can distinguish between fair and unfair sharers by observing 20s videos recorded *before* players are informed of the task. Finally, Schug et al. (2010) find that FACS analysis (see below) shows cooperators to be more generally expressive than defectors.

It appears that we exhibit a degree of emotional translucency but why is that so? First I deal with proximate explanations.

Proximate mechanisms for securing emotional translucency.

The first explanation is that there are features of emotional facial expression that are not under voluntary control. These have been documented using Ekman & Freisen's "facial action coding system" (FACS) for careful analysis of muscular movement (Ekman et al. 2002). For instance, "Duchenne" smiles involve movement of the difficult to control orbicularis oculi and pars lateralis muscles and such movement is absent in smiles made without the accompanying emotion. Ekman labels these, "reliable facial muscles" and argues that studies that find no anti-correlation between smiling and deception can be explained because they do not distinguish between Duchenne and non-Duchenne smiles (Ekman 2003).⁴ Other identified reliable muscle movements form part of the expressions of fear, anger and sadness and Porter & ten Brinke (2008) find evidence of leakage of negative emotions via these muscles.

The second explanation for translucent emotions is the degree to which they are initiated without input from higher cognition. In the case of fear, Joseph LeDoux demonstrates that the brain region central to fear initiation is the amygdala and two paths can act as triggers. One consists of signals travelling directly from the thalamus while the other sends signals via cortical sensory processing regions. The thalamus carries out much less sophisticated processing than the cortical sensory regions but the direct route to the amygdala is twice as fast as the cortical route (LeDoux 1998). The actions prompted directly from the thalamus are totally involuntary and fire first. That being the case, suppression of emotional displays is difficult to achieve.

The third proximate account follows from work on the neurobiology of the conscious feeling of emotional responses. Antonio Damasio and others demonstrate that the substrate of emotional feelings lies in somatosensory maps of the unfolding responses. Individuals with damage to the relevant sensory areas suffer from subdued emotional experiences (Damasio 2000). Regarding deception, manipulation of bodily responses is difficult, particularly during on-going episodes, if their production *precedes* their conscious perception. All three of these proximate explanations involve evolutionarily ancient structures that are the result of biological evolution. We can therefore ask the question, is there an evolutionary explanation for their presence and persistence?

An objection to the suggestion that producing deceptive emotional displays is difficult is the claim that people, particularly actors, do indeed produce fake emotional displays. At least four replies

can be made to the actor objection. Firstly, the intuition can be challenged. Actors able to give convincing performances are held in high regard, suggesting that they are rare. It is also questionable that, on turning on the television, people typically find it hard to tell whether they are watching a play or a documentary.⁵ Secondly, the argument relies on anecdotal observation whereas researchers commonly hold the view, corroborated by Gosselin et al. (1995), that actors' "posed" emotions deviate from genuine expressions. This leads to a third response which is that the way in which to "fake" an emotion is to elicit a "genuine" display by thinking about, say, a sad event. This familiar Stanislavski (2008) "system" supports the view that emotions are difficult to fake and leads to the final response to the actor objection. The conditions in which actors produce genuinely felt, but deceptive, emotions are atypical of realistic situations in which one wants to control emotional displays. Film actors can make use of repeated takes to generate a display and stage actors inhabit their role within a script for hours at a time. None of this applies to the real-time unscripted interactions of everyday life.

Having presented proximate explanations of emotional translucency I lay the groundwork for my evolutionary explanation by looking at situations in which we would expect displays to be transparent.

Emotions and common-interest interactions.

In addition to PDs, there are other conflict-of-interest interactions. In these cases too, signals suffer from pressures in favour of deception. For instance, in chicken games, there is a pressure to fake an emotional display of commitment to defection where no commitment exists. Contrast this with coordination games in which players' interests coincide. Here, emotional displays can play a signalling role that is free from the danger of subversion. These games can possess multiple equilibria and signals allow players to coordinate on optimal outcomes and, since interests coincide, there is no pressure in favour of deceptive signalling (Robson 1990).

Rather than talking of coordination games and equilibrium selection, social and developmental psychologists study the phenomenon of "joint attention". Indeed, connecting the game theoretic and psychological terms of art, joint attention is defined as "coordinated joint engagement" and plays an important role in infants' development of language and theory of mind (Bakeman & Adamson 1984). Developmental research demonstrates that both children and chimpanzees rely on emotional displays to facilitate joint attention and social learning (Russell et al. 1997). There is also

evidence of crucial role of emotional interaction in generating secure attachment in infants (Hrdy 2009).

The fact that emotions play such vital roles early in life explains, in part, the involuntary nature of emotional responses in social situations. It is not open to infants and young children to produce voluntary signals via deliberation. Even in older children and adults, in real-time, common-interest interactions, it is very often important to respond quickly in social situations. Therefore there are active pressures in favour of involuntariness as well as the pressure in favour of deception in conflicts. Hence it is not just that there is pressure for us not to be actively deceptive, there is also pressure in favour of the involuntariness that makes deception intrinsically difficult.⁶

To summarize, emotional displays appear to play a signalling role in both common and conflict-of-interest interactions. In the former there is pressure against deceptive capacity whilst, in the later, there is pressure in favour. The question to which I now turn is, do we have reason to believe that either pressure is stronger?

Balancing pressures: Age-dependent intensity of selection.

It is mean *lifetime* reproductive consequences of behaviours that count when it comes to assessing selection pressures (West et al. 2007). If the situations in which a trait is beneficial outweigh those where it is costly then it is selected. With regard to our focus, if all interactions during a lifetime were cases in which interests coincide there would be no pressure in favour of a capacity for deception. Given that we actually encounter a variety of strategic encounters, pressures for and against deceptive capability must be weighed against each other. The most simple-minded way to do so would be to count how many of each type occur on average across the population. That cannot be correct since we must weight interactions by what is at stake in each case. High stakes interactions cause a greater pressure than where the stakes are low. In spite of the enormous complexity of counting interactions we can say that there is a threshold in the rates of interaction types beyond which the pressure due to common-interest interactions outweighs pressure due to conflicts. However, the weighted comparison remains flawed. Where evolution is genetic, and I am interested in the *biological* evolution of emotions, an inclusive fitness perspective must be taken. Taking account of kin directed behavioural traits pushes down the threshold at which conflicts are outweighed by shared interest interactions because relatedness moves interests into closer alignment.

Finally, and this is the point on which my evolutionary argument hinges, the most sophisticated treatment of weighting interactions considers the point during a lifetime at which interactions occur. As we will see, this consideration pushes the threshold down even further, sealing a good evolutionary explanation for why it is fitness maximising for human emotional displays to be towards the transparent end of the opaque→transparent spectrum. Selection is most intense on traits expressed early in an organism's life and gets progressively weaker as the trait is expressed up to and beyond the mean age of final reproduction. To illustrate, Huntington's chorea is caused by a lethal dominant allele. If the condition were manifested at a pre-pubescent age there would be extremely intense selection against the allele since all of its carriers would die before reproducing. In fact, the average age of onset of Huntington's is 35½ and therefore many carriers successfully pass on their genes (Rose 1991). It is this age-dependent variation that underwrites theories of antagonistic pleiotropy in which an allele is selected that has negative effects later in life but beneficial ones earlier. William Hamilton's standard measure of selection intensity is the partial derivative of the "Malthusian parameter" of population growth w.r.t. age-specific survival probability. It is at its maximum throughout the pre-reproductive phase and then decreases with age (Charlesworth 1980).

Having established that pressures due to behavioural traits are most intense at younger ages I move on to the questions of when, during a human lifetime, interests tend to conflict or coincide.

Conflicting and common-interests across a lifetime.

Since my focus is the biological evolution of social emotional displays, these "interests" are inclusive fitness consequences for alleles. Hence the effect of an interaction on related individuals, weighted by relatedness, r , is added to the fitness effect on the focal individual (Frank 1998). Where r is positive, inclusive fitness considerations push interests into closer alignment than the interaction's direct fitness payoffs. Therefore kin effects can transform a conflict of interest into a case where interests coincide (Binmore 2005). Note that I am not suggesting that *all* interactions with kin are common-interest, as evidenced by the substantial literature on parent-offspring conflict and sibling rivalry (Salmon 2005). The identification of these phenomena was important in countering a naïve view of kin groups according to which interests *always* coincide. The typical examples of both involve what in game theoretic terms are zero-sum interactions, such as food provisioning, in which interests

are perfectly opposed. Even in these cases relatedness pushes inclusive fitness values towards common-interests compared to direct fitness consequences.

Considering the pattern of human parental care, offspring experience an extremely lengthy period of intensive support and evidence suggests that this has been the case for up to 800 000 years (Flinn et al. 2007). Additionally, our nearest ancestors exhibit briefer but still protracted periods of maternal care, suggesting that the phenomenon is many millions of years old (Boesch & Boesch-Achermann 2000). An objection at this point might be that current Western periods of parental care are longer than during our evolutionary past. It is true that human offspring were not always reliant on their parents for 16 years and the period during which the vast majority of interactions were with kin was briefer. Balancing this change though, it is likely that the average age of first reproduction is older now than it was for a large proportion of our evolutionary past. Children may not have been raised beyond puberty in the past but they were likely to have children themselves not long after and it is at that point that selection intensity begins to drop. Additionally, mean global life expectancy has probably only risen above 30 in the last 200 years (Riley 2001). Consequently, for most of our evolutionary past, the period during which humans interacted principally with close kin formed a significant proportion of their lives and coincided with the period of most intense selection.

Summarizing the two preceding sections, we tend to inherit traits that make us successful in our youth, what tend to be important when we are young are common-interest interactions and in those cases there is pressure in favour of making transparent signals. Notice that my account can be interpreted as an example of antagonistic pleiotropy with translucency being the result of the trade-off between transparency and deceptive capacity.

Plasticity of displays.

Before concluding, I stress an important assumption of the preceding sections. Variation in selection intensity with age is a factor potentially affecting all traits, but its effect depends on the plasticity over time of the trait in question. The trade-off between beneficial and negative consequences across a whole lifetime need not be considered if the trait is not present from birth until death. To illustrate, one might argue that *hardwired* cooperative behaviour in PD interactions can be selected due to its benefits early in life when inclusive fitness effects typically skew payoffs. The harm suffered during adult interactions with non-kin occurs when selection is less intense and so does not

weigh as heavily as the positive effects early in life. The problem here is that there is no reason to believe that specific cooperative behaviours are inflexible over the course of a lifetime and so the trade-off argument cannot go through. In contrast, in the case of emotional displays, I *do* assume that they are relatively developmentally fixed. This assumption is based primarily on the proximate mechanisms that explain the translucency of displays. Involuntary facial muscles, automatic sub-personal emotion pathways and somatosensory feedback that produces feelings are built into the morphology of the adult face and brain rather than being surface behavioural routines and so seem less amenable to variation over a lifetime. There is some evidence comparing deception in children and adults although it does not specifically examine emotional displays. It also focuses on whether observers can detect deception rather than whether there are differences between honest and deceptive displays. These results suggest that distinguishing between lies and truths in adults and children is of similar difficulty (Crossman & Lewis 2006).

Arguing that emotional translucency is not, as a matter of fact, plastic over a lifetime does not prohibit interesting questions regarding pressure for developmental plasticity when age-dependent selection intensity plays a role. One can ask whether lack of plasticity is due to relatively immutable features of the developmental processes involved in building emotional brains and faces, or whether there are selective forces opposing plasticity. In this section I suggested that the former is probably true but there is intriguing evidence that orangutan facial morphology does change during development from adolescence to adulthood (Kuze et al. 2005). Such a phenomenon is not found in humans⁷ but it appears that some of our close relatives exhibit the kind of changes that might be predicted on the basis of variable selection intensity on signalling traits.

Conclusion.

Cooperation in prisoner's dilemmas can evolve via cultural evolution if players can signal their type. However, the problem of deceptive signals must then be tackled and candidate signalling mechanisms proposed and analysed for security from subversion. We have good reason to believe that emotional displays can act as signals of cooperation and that, whilst not transparent, they are translucent. There are proximate mechanisms that explain this but in this paper I have proposed an evolutionary explanation of translucency. Kinship tends to align interests and human development depends crucially on emotional interactions during a protracted period of parental nurturing. This

causes pressure for transparency of emotions and, due to the age-dependent nature of selection intensity, outweighs pressure in favour of deceptive capability later in life.

References.

- Alexander, J. 2007. *The structural evolution of morality*. Cambridge: CUP.
- Bakeman R. & L. Adamson. 1984. Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development* 55: 1278-89.
- Binmore, K. 1994. *Game theory and the social contract: Volume I*. Cambridge, MA: MIT Press.
- Binmore, K. 2005. *Natural justice*. Oxford: OUP.
- Boesch, C. & H. Boesch-Achermann. 2000. *The chimpanzees of the Tai forest*. Oxford: OUP.
- Bond, C. & B. DePaulo. 2006. Accuracy of deception judgements. *Personality and Social Psychology Review* 10: 214-234.
- Brosig, J. 2002. Identifying cooperative behavior. *Journal of Economic Behavior & Organization* 47: 275-290.
- Charlesworth, B. 1980. *Evolution in age-structured populations*. Cambridge: CUP.
- Crossman, A. & M. Lewis. 2006. Adults' ability to detect children's lying. *Behavioral Sciences and the Law* 24: 703-715.
- Damasio, A. 2000. *The feeling of what happens*. London: Vintage
- DePaulo, B., J. Lindsay, B. Malone, L. Muhlenbruck, K. Charlton, & H. Cooper. 2003. Cues to deception. *Psychological Bulletin* 129: 74-118.
- Ekman, P. 1992. *Telling lies. 2nd edition*. New York: Norton.
- Ekman, P. 2003. Darwin, deception and facial expression. *Ann. N.Y. Acad. Sci.* 1000: 205-221.
- Ekman, P. 2009. Lie catching and microexpressions. In *The philosophy of deception*, ed. C. Martin, 118-135. Oxford: OUP.
- Ekman, P., W. Friesen & J. Hager. 2002. *Facial action coding system. 2nd edition*. Salt Lake City: Research Nexus.
- Elster, J. 2000. *Ulysses unbound*. Cambridge: CUP.
- Fetchenhauer, D., T. Groothuis & J. Pradel. 2010. Not only states but traits- Humans can identify permanent altruistic dispositions in 20s. *Evolution and Human Behavior* 31: 80-86
- Flinn, M., R. Quinlan, K. Coe & C. Ward. 2007. Evolution of the human family. In *Family relationships*, eds. C. Salmon & T. Shackelford, 16-38. Oxford: OUP.
- Frank, R. 1988. *Passions within reason*. New York: Norton.
- Frank, R., T. Gilovich & D. Regan. 1993. The evolution of one-shot cooperation. *Ethology and Sociobiology* 14: 247-256.
- Frank, S. 1998. *Foundations of social evolution*. Princeton: Princeton University Press

- Gauthier, D. 1986. *Morals by agreement*. Oxford: OUP.
- Gosselin, P., G. Kirouac & F. Dore. 1995. Components and recognition of facial expression in the communication of emotion by actors. *Journal of Personality and Social Psychology* 68: 83-96.
- Gould, S. 1977. *Ontogeny and phylogeny*. Cambridge, MA: Harvard University Press
- Hrdy, S. 2009. *Mothers and others*. Cambridge, MA: MIT Press.
- Kuze, N., T. Malim & S. Kohshima. 2005. Developmental changes in the facial morphology of the Borneo orangutan. *American Journal of Primatology* 65: 353-376.
- Ledoux, J. 1998. *The emotional brain*. London: Weidenfeld & Nicolson.
- McElreath, R. & R. Boyd. 2007. *Mathematical models of social evolution*. Chicago: University of Chicago Press.
- Ockenfels, A. & R. Selten. 2000. An experiment on the hypothesis of involuntary truth-signalling in bargaining. *Games and Economic Behavior* 33: 90-116.
- Porter, S. & L. ten Brinke. 2008. Reading between the lies. *Psychological Science* 19: 508-514.
- Pradel, J., H. Euler & D. Fetchenhauer. 2009. Spotting altruistic dictator game players and mingling with them. *Evolution and Human Behavior* 30: 103-113.
- Riley, J. 2001. *Rising life expectancy*. Cambridge: CUP.
- Robson, A. 1990. Efficiency in evolutionary games. *Journal of Theoretical Biology* 144: 379-396.
- Rose, M. 1991. *The evolutionary biology of aging*. Oxford: OUP.
- Russell, C., L. Adamson & K. Bard. 1997. Social referencing by young chimpanzees. *Journal of Comparative Psychology* 111: 185-193.
- Salmon, C. 2005. Parental investment and parent-offspring conflict. In *The handbook of evolutionary psychology*, ed. D. Buss, 506-527. Hoboken: Wiley.
- Scharlemann, P., C. Eckel, A. Kacelnik & R. Wilson. 2001. The value of a smile. *Journal of Economic Psychology* 22: 617-640.
- Schug, J., D. Matsumoto, Y. Horita, T. Yamagishi & K. Bonnet. 2010. Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior* 31: 87-94.
- Skyrms, B. 1996. *Evolution of the social contract*. Cambridge: CUP.
- Skyrms, B. 2000. Game theory, rationality and evolution of the social contract. In *Evolutionary origins of morality*, ed. L. Katz, 269-284. Thorverton: Imprint Academic.
- Stanislavski, C. 2008. *An actor prepares*. London: Methuen.
- Verplaetse, J., S. Vanneste & J. Braeckman. 2007. You can judge a book by its cover: the sequel. *Evolution and Human Behavior* 28: 260-71.
- West, S., A. Griffin, & A. Gardner. 2007. Social semantics. *Journal of Evolutionary Biology* 20: 415-432.
- Williams, G. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11: 398-411.

Yamagishi, T., S. Tanida, R. Mashima, E. Shimoma & S. Kanazawa. 2003. You can judge a book by its cover. *Evolution and Human Behavior* 24: 290-301.

¹ This is a direct result of the Pareto superiority of mutual cooperation over mutual defection.

² I think that this second role is much less problematic than Frank's "commitment model" of emotions.

³ These studies do not investigate the cues on which judgements are based.

⁴ In a meta-analysis, of 27 studies of smiling during deception, only two did so (DePaulo et al. 2003).

⁵ This could also be explained by differences in the techniques of fiction and documentary.

⁶ Thanks to Mark Bedau (EPSA09) for pressing this point.

⁷ In fact the opposite phenomenon is found. It has been noted that humans exhibit *neoteny*. That is, childhood traits persisting through adulthood to an unexpected degree (Gould 1977).