European Journal of Social Psychology Vol. 24 367–381 (1994)

The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions

URSULA HESS
University of Quebec at Montreal

and

ROBERT E. KLECK
Dartmouth College

Abstract

Dynamic facial expressions, either posed or elicited by affectively evocative materials, were objectively scored to determine the movement cues and temporal parameters associated with the two types of expression. Subjects viewed these expressive episodes and rated each of them on a number of scales intended to assess perceived spontaneousness and deliberateness. Subsequent to viewing all stimuli, subjects reported the specific cues that they felt they had used to discriminate spontaneous from deliberate expressions. The results reveal that (a) subjects were able to accurately report the cues they employed in the rating task and that (b) these cues were not always valid discriminators of posed and spontaneous expressions. Subjects were in fact relatively poor at identifying expressions of the two types and this low discrimination accuracy was found to be a function of the consistent use of these invalid cues. A measure of the level of perceived 'honest demeanour of the stimulus persons based on their neutral expressions was found to relate to perceivers accuracy in discriminating posed and spontaneous expressions.

INTRODUCTION

Facial misrepresentation of our experienced emotions, whether intentional or not, appears to be a central aspect of our social lives. The happy smile shown when an unappreciated gift has been received from a friend or relative, the neutral face

This paper is based on the thesis research conducted by the first author. We would like to thank Arvid Kappas for his support during this project. The research was supported in part by funds associated with the John Sloan Dickey Third Century Professorship and with the William H. Spoor Dialogues on Leadership Program at Dartmouth (Kleck).

Correspondence concerning this article should be addressed to Ursula Hess, Department of Psychology, University of Quebec at Montreal, C.P 8888, succ. A, Montreal (Quebec) H3C 3P8, Canada.

CCC 0046–2772/94/030367–15

Received 24 April 1992
Accepted 15 July 1993

shown by a young adult male about to receive a highly feared injection, or the 'poker' face of a card player who has just drawn a winning hand are common examples of circumstances where facial displays do not map onto the underlying affect of the encoder. Indeed, it has been argued that one of the socialization outcomes sought in all cultures is to educate individuals in the modulation of emotion-relevant facial behaviour to render it appropriate to given interaction contexts or to individual social motives (e.g. Ekman, 1973; Wundt, 1903).

In many social contexts there is probably little if any motivation to detect these facial 'misrepresentations', either because they are relatively trivial or the overt detection of them would result in social discomfort and embarrassment (e.g. Kraut, 1980). At the same time, in dealing with our children or with intimates we may be highly motivated to distinguish between their expressions of truly felt emotions and expressions that have been intentionally displayed in the interests of controlling our behaviour or feelings.

What we first have to know is whether facial expressions that reliably reflect an underlying affect and those that do not are different in a way that *could* be detected by a perceiver. Ekman and Friesen (1982) have argued that smiles that occur in conjunction with positive affect (spontaneous smiles) differ from deliberate smiles (smiles that occur in the absence of the appropriate affect) in both their temporal and topographical features. Specifically, they expect that spontaneous smiles should be accompanied by wrinkles around the eye due to the activity of the orbicularis oculi muscle (Ekman, Davidson and Friesen, 1990), while deliberate or deceptive smiles will be more asymmetrical and will have shorter onset times and more irregular offset times than spontaneous smiles. Finally, they argue that deliberate smiles should be shorter than 2/3 second or longer than 4 seconds.

Recent studies provide empirical support for some of these assertions. Hager and Ekman (1985), for example, found false or deliberate smiles to be more asymmetrical than 'felt' ones. The hypothesis regarding differences in onset time has been supported by Weiss, Blum and Gleberman (1987) for posed smiles and by Hess and Kleck (1990) for 'masking deceptions', that is, for facial expressions employed to cover a different underlying emotional state. In the latter two studies, posed smiles as well as masking deceptions were also found to be more irregular than were spontaneous smiles. Empirical support for the notion that spontaneous (or felt) smiles should be accompanied by crow's feet wrinkles around the eyes is provided by Ekman, Friesen and O'Sullivan (1988).

As a group these studies suggest that cues are available to permit perceivers to distinguish between posed and spontaneous expressions at least for some affects. It is plausible, of course, that decoders cannot detect these differences or that they rely upon other, non-valid cues when attempting to make this discrimination. Judging from the deception literature in general, it appears to be the case that not all cues assumed by perceivers to discriminate deceptive and non-deceptive behaviours do so in a reliable way (see DePaulo, Stone and Lassiter (1985) for a review). Further, in this broader literature, evidence has been found for both a perceiver and a demeanour bias. The first describes the tendency of perceivers to label any given message as honest or non-deceptive; the latter the tendency for some stimulus persons to be seen as relatively honest or dishonest independent of their actual veracity (Zuckerman, DeFrank, Hall, Larrence and Rosenthal, 1979).

The present study is an attempt to investigate systematically (a) which cues per-

ceivers report to be valid cues for the discrimination of spontaneous and deliberate emotional facial expressions, (b) whether they use these cues in deciding if an expression is spontaneous or deliberate and, (c) whether the cues employed by the perceivers are actually valid cues for this discrimination task.

It was hypothesized that subjects could discriminate between spontaneous and deliberate emotional facial expressions, but with a small margin of accuracy A path analysis approach was used to relate nonverbal cues that the subjects believe to be markers of deliberateness, to their actual judgments of the degree of deliberateness of the expression. We expected that subjects could identify the cues they were using to make the discrimination. Further, we expected that some cues believed to be important by the subjects would not be valid cues for the discrimination of spontaneous from deliberate emotional facial expressions. That is, a model was hypothesized, in which the cues mentioned by the subjects as important would be systematically related to their actual judgments, but not to the elicitation conditions.

The goal of the present research is to investigate the cues decoders use to differentiate spontaneous and voluntary emotional facial expressions in general, that is, cues not linked to specific emotions. We therefore employed expressions of happiness and disgust, since one is representative for a positive emotion and the other for a negative emotion (e.g. Hess and Kleck, 1990). While the use of a wider range would have been preferable we felt constrained to two emotions by practical reasons. Specifically, it is difficult to elicit recognizable spontaneous displays of emotional states such as fear or anger in adults in a laboratory setting without violating ethical standards of laboratory research or without using a different emotion elicitation method for each emotion (which in the context of this research would not have been acceptable).

METHOD

Overview

Posed and spontaneous facial expressions of happiness and disgust were shown to male and female judges who rated each expression on a number of scales designed to assess the perceived degree of voluntary control or deliberateness in the facial behaviour. They were also asked to indicate the specific cues they had employed in making these judgments across the stimulus set as a whole. Honest demeanour judgments of the stimulus persons were acquired from an independent group of judges. Finally, the expressive episodes were objectively scored on a number of temporal and muscle movement parameters.

Stimuli

The facial expressive displays employed as stimuli were drawn from an earlier study (Hess and Kleck, 1990). In that study 18 males and 17 females were candidly videotaped, using an arrangement detailed in Kappas, Hess and Kleck (1990), while posing happy and disgust expressions and while responding to visual or cognitive stimuli designed to elicit spontaneous expressions of these affects. From the stimulus subject's point of view, the experiment was concerned with the psychophysiological responses taken during the experimental session and the posing instruction was introduced

as a 'control condition' used to examine the impact of facial movements on the physiological measures. Subjects were asked to pose realistic expressions of both happiness and disgust and were given three minutes to achieve each of these. Two elicitation procedures were used with all subjects to obtain spontaneous expressions of affect. In one they were asked to relive a previous situation in which they had experienced disgust or happiness. In the second they were shown video materials previously demonstrated to elicit reliably happy and disgust responses in individuals drawn from the same subject population.

The three elicitation tasks (pose, watch elicitors, and relive experience) were counterbalanced across subjects. After each task subjects reported the degree of emotional arousal they had experienced regarding several specific affects (e.g. fear, disgust, happiness, etc.). They were videotaped without their knowledge by a hidden camera. Each subject's permission to use his or her video records for rating purposes was obtained at the completion of the experimental session.

The expressive episodes employed in the present study were selected based on the presence of discernible facial behaviour as well as the self-reports of emotional arousal provided by the expressors. For the spontaneous elicitation conditions an episode was retained only if it was accompanied by a self-report of the relevant emotion (happiness or disgust) at a moderate level or above (a rating of 3 or above on the 5-point scale). For the posed facial expressions, episodes accompanied by an arousal level greater than 0 or 1 on the rating scales were discarded. This was done to assure that the spontaneous expressions of happiness and disgust were occurring in the presence of appropriate emotional arousal while the posed expressions occurred in the absence of emotional arousal. While this procedure assures that the voluntary expressions were produced while subjects did not feel the relevant emotional state it is possible that expressions were included that do not portray the intended emotion due to a lack of expressive skills on the part of the expressor. As a manipulation check judges were therefore asked to also rate the emotion portrayed. From the total pool of episodes meeting these criteria (140) posed and spontaneous expressions within each affect were roughly matched regarding length[1] This resulted in the identification of 17 posed and 17 spontaneous happy expressions and 13 posed and 13 spontaneous disgust expressions[2] These 60 expressive episodes were then randomly edited into two different sequences for presentation to the subjects in the present study The average length of the episodes was 15.2 seconds. The expressive behaviours were shown to the judges on a 19″ black and white monitor placed approximately 3 feet in front of the judge.

On the basis of recent research it is plausible to argue that decoders may differentially attend to or process whatever aspects of an image are presented to their right or left visual fields (e.g. Borod, St. Clair, Koff and Alpert, 1990; Rhodes, Ronke and Tan, 1990). To control for possible confounding effects of the placement of the encoders' two hemifaces within the video image, one half of the decoders saw the stimuli horizontally reversed, that is, with the encoder's left hemiface to their

[1] We chose not to include length of expression as a cue, even though it is a valid cue (Hess and Kleck, 1990), since we wanted to closer examine the role of the dynamic aspects of the expression (e.g. onset, offset, etc.) without the confounding effects of length.

[2] Due to the selection procedure chosen, only two subjects contributed expressive episodes for all four conditions. That is, the discrimination of spontaneous and voluntary episodes was performed across different models.

right and *vice versa*. The reversal was accomplished electronically and did not alter the image in any other way nor was it detectable by the decoders.

Subjects

Twenty male and 20 female undergraduates participated for extra course credit. They were run individually in sessions of approximately 1 ½ hours by a female experimenter. To enhance the subjects' motivation to persist in a task of this length a prize was awarded to the person with the highest accuracy in identifying posed and spontaneous expressions.

Procedure

When the subject arrived for the session he/she was seated at a table in the stimulus viewing room and asked to read written instructions describing the general nature of the task. The experiment was presented as one in which they would be making judgments regarding whether a facial expressive episode was spontaneous (i.e. caused by an emotion with little or no intentional control by the person expressing it) or was intentional (i.e. did not reflect an underlying emotion parallel to the expression being shown). They were verbally instructed in the use of the video player and the computer program controlling the presentation of the rating scales. These scales were displayed on a computer screen placed to their right. The subjects then rated two practice episodes to familiarize themselves with the procedures and the experimenter answered any questions they had and left them to complete the task.

Subjects could proceed at their own pace in rating the 60 episodes, but were not allowed to view any particular stimulus more than once. Following the completion of the rating task, subjects filled out a form asking for the decision rules they would pass along to a potential judge for discriminating whether an expression was spontaneous or voluntary Following this, the purpose of the experiment was fully described to the subjects and any questions regarding the study were answered in detail by the experimenter.

Dependent measures

Judge ratings

For each expressive episode subjects were instructed to indicate on seven-point Likert-scales (1—not at all; 7—completely) their answers to the following questions: (a) to what degree is the expression you just saw spontaneous, (b) to what degree is the expression you just saw voluntary/controlled, and (c) to what degree does the expression reflect an underlying emotional state. Furthermore, they were asked to make a categorical judgment regarding whether the expression was posed or spontaneous. This judgment was compared to the actual elicitation condition under which the facial display was produced. If the judges' decision was accurate it was coded as 1, if it was inaccurate as 0.

Lastly, subjects indicated on eight-point Likert-scales how intensely each of several

emotional states (happiness, anger, fear, sadness, disgust, and other) were reflected by the facial expression. All questions were presented on an Apple Macintosh SE using a custom Hypercard program.

Objective coding

As noted above, after subjects had responded to the set of 60 stimuli they completed a form asking them to describe to another potential subject the important cues to look for when trying to discriminate between spontaneous and voluntary expressions. A content analysis of these reports revealed that subjects focused primarily upon the dynamic temporal aspects of the expression (e.g. expression held to long, quick repetition) and upon the eyes and the eye region (e.g. person looks away, expression of the eyes) in attempting to discriminate between posed and spontaneous expressions. For the purpose of attempting to use a path analytic approach to relate the reported cues and the perceivers' actual judgments, behaviours corresponding to these cues were objectively scored[3]

Temporal variables. All expressive episodes were scored using Ekman and Friesen's Facial Affect Coding System (1978)[4] This approach yields precise measurements of the speed of onset, time at apex, and speed of offset of an expressive episode. It is important to note, of course, that an expression can be characterized by several phases. It can, for example, become visible and intensify to a certain point, be held at that level or decrease a bit, intensify again, drop slightly, again intensify, and then return to baseline. In the present case an expressive episode was defined as lasting from the first appearance of an action unit also present in the expression during an apex period to the disappearance of the last action unit. Since new action units were usually not added to the expression during apex phases, phases were scored for the complete display and not for each action unit separately The frequency of phasic changes was scored by counting the number of such episodes from the moment when the expression became first visible until it faced back to baseline. This measure will subsequently be referred to as the 'irregularity' of the expression. Since an expression could have several onset or offset phases the speed of onset and offset was calculated as the average length of the onset or offset phases within a particular episode. A separate coder examined 20 randomly chosen episodes for the number of phases and the length of each phase. The intercoder reliability was 0.88 (Cohen's kappa) for the first measure and 0.92 (Pearson correlation) for the second.

Eye variables. Two coders independently recorded the number of eye blinks and gaze aversions for the total set of expressive episodes. The measure used for both variables was frequency divided by the length of the episode in seconds. Reliability for the gaze aversion measure was 0.65 (Cohen's kappa) and for the eye blink measure was 0.98 (Cohen's kappa). The FACS code AU6 provided a measure of the activity of the orbicularis oculi — the muscle surrounding the eye that creates 'crow feet wrinkles' The involvement of this action has been implicated in previous thinking

[3] Another frequently mentioned cue was 'exaggerated expression'· this cue was not followed up since the phrasing was too vague to allow finding a cut-off point of intensity for the differentiation of 'exaggerated' versus normal expressions.

[4] The episodes were coded by the first author, who is a certified FACS coder. The coder was blind to condition at the time of coding.

about spontaneous and deliberate expressions and it has been argued to produce a specific 'look in the eye' (Ekman, 1985). Since subjects mentioned the 'look in the eye' as a possible cue, the percentage of time that AU6 was present during the expressive episode was included as an objective measure.

Honest demeanour

As noted in the Introduction, the deception literature has suggested that whether a particular behaviour is seen as deceptive or not depends in part on the target person's 'demeanour' To assess this possibility, a still photograph was made of the video record of each stimulus person who provided expressive episodes for the present study These black and white photographs showed the person face on, from the shoulders up, and were taken when the individual showed a neutral facial expression as determined by a FACS score of 0. These photos were presented in two different orders to small groups (2–4) of males and females drawn from the same population as the judges employed in the present study A total of 32 males and 32 females saw the set of photographs one at a time and rated them on a number of scales. Embedded within these bipolar adjective pairs were the focal dimensions of 'honest–dishonest' and 'untrustworthy–trustworthy' Since across stimulus persons ratings on these scales were highly correlated ($r = 0.95$) only the mean ratings on the honest–dishonest dimensions were retained for the analysis presented below

Data analysis

The stimuli used in this study were 60 facially expressive episodes that were intended to reflect two qualitatively different emotions (happiness and disgust). These had been displayed by our stimulus persons under conditions where they either had been instructed to pose the expressions or been exposed to emotion-eliciting stimuli, which in turn evoked spontaneous emotion displays.

The first question we sought to answer was whether the judges viewing these episodes recognized them as expressions of happiness and disgust respectively This can be viewed as a manipulation check in the sense that if perceivers cannot discern the affective quality of the expression, there would be less interest in the issue of perceived spontaneity and deliberateness.

The second concern was to determine to what degree subjects could accurately discriminate the posed expressions from the spontaneously elicited ones. Four different ratings scales were employed to assess these perceptions and in this exploratory study responses to each scale were analysed separately

Finally, and most importantly, an attempt was made to relate objectively scored cues to the judges' ratings to determine if these judgments were based on valid or invalid cues. For this we employed a path analytic approach using LISREL VII (Jöreskog and Sörbom, 1988, see Ch. 6 for this approach). For this purpose, the perceived degree of 'voluntariness', the accuracy of the judgments, the objective measures described above, and the honest demeanour measure obtained from a separate set of judges were entered as directly observed variables. To assess the goodness of fit of the path analytical model the maximum likelihood chi square statistic, the rho index of practical fit (Bentler and Bonett, 1980), and Bollen's IFI index

(Bollen, 1989) were considered[5] The goal of this study was to investigate the judges' use of general cues discriminating spontaneous from voluntary expressions, i.e. cues not necessarily linked to a specific emotion. Therefore, the decision was made to include the happy and spontaneous expressive episodes in one model. Even with that the total number of cases employed here (60) is probably too low to allow a confident generalization of the findings. It is, however, large enough to permit the attempt to model the specific data set under consideration[6]

RESULTS AND DISCUSSION

Manipulation check

To assess whether the stimulus persons' facial displays were recognized as happiness and disgust, subjects' ratings on these two 8-point scales were analysed using a multivariate mixed design analysis of variance. The between-subjects factors were sex of rater and orientation of image (normal/reversed); within-subjects factors were emotion (happy/disgust) and condition (posed/spontaneous). Significant main effects for emotion and condition emerged for happiness (($F(4,33) = 327\ 49$, $p < 0.001$ and $F(4,33) = 459-66$, $p < 0.001$ respectively) and disgust ratings ($F(4,33) = 102.14$, $p < 0.001$ and $F(4,33) = 54.46$, $p < 0.001$ respectively), an emotion × condition interaction ($F(4,33) = 24.25$, $p < 0.001$) was found for ratings of disgust. Inspection of the means showed that the spontaneous happiness expressions were rated as showing more happiness than the posed happiness expressions (mean of 3.44, $S.D = 0.97$ and 2.86, $S.D = 0.88$ respectively, $t(39) = 9.56$ $p < 0.001$). The reverse was the case for the posed and spontaneous disgust expressions, here, the posed expressions were rated as expressing more disgust than the spontaneous expressions (mean of 3.76, $S.D = .0$ and 2.90, $S.D = 0.78$ respectively, $t(39) = 7.34$ $p < 0.001$).

For both the happiness and the disgust expressions the ratings were highest for the target emotion category In comparison, ratings on all other specific emotion scales were low For the spontaneous and voluntary happy expressions the next highest rating was for sadness: (spontaneous $m = 0.52$; voluntary $m = 0.64$) while for the spontaneous and voluntary disgust expressions it was fear (spontaneous $m = 1.16$; voluntary $m = 1.14$). These results indicate that the expressive episodes on average conveyed the intended emotional state to the raters. Both the spontaneous happy expressions and the spontaneous disgust expressions were rated as expressing the underlying emotional state to a higher degree than was the case for posed expressions ($t(39) = 3.65$, $p < 0.001$ and $t(39) = 2.56$, $p < 0.05$ respectively).

Perceived spontaneity/control of the expressions

A 2 (sex of rater) × 2 (orientation of image) × 2 (emotion) × 2 (condition) mixed design analysis of variance was conducted on the ratings of perceived spontaneity and perceived control. Again, the between-subjects factors were sex of rater and

[5] The χ^2 should be nonsignificant and the Bentler-Bonett index and the IFI larger than 0.90 for a model to be an adequate representation of the data.
[6] Anderson and Gerbing (1988) recommend between 50 and 150 cases.

orientation of image (normal/mirror); the within-subjects factors were emotion (happy/disgust) and condition (posed/spontaneous). No significant main effects for sex of rater or orientation were revealed. For both perceived spontaneity and perceived control significant main effects for condition emerged $F(1,36) = 67\ 41$, $p < 0.001$, $F(1,36) = 93.71$, $p < 0.001$, respectively); for ratings of perceived spontaneity an emotion × condition interaction was found ($F(1,36) = 5.53, p < 0.05$).

As expected, spontaneous expressions were rated as more spontaneous and less controlled than were controlled expressions. However, a comparison of these ratings shows that subjects exhibit a bias toward attributing spontaneousness to all expressions (see Table 1). This is consistent with findings by DePaulo *et al.* (1985) that while deceptive messages are perceived as less honest than honest messages, subjects tend to assume most messages are honest.

Table 1. Means and standard deviations for dependent variables as a function of emotion and elicitation condition

Condition	Happy		Disgust	
Rating	Spontaneous	Deliberate	Spontaneous	Deliberate
Spontaneous	4.71 (0.58)	3.88 (0.52)	4.55 (0.70)	4.08 (0.77)
Controlled	3.54 (0.74)	4.35 (0.59)	3.64 (0.78)	4.21 (0.83)

One goal of the present study was to assess the degree to which subjects can discriminate between posed and emotion-elicited expressions. For this, we tabulated the percentage of accurately decoded expressions for each category using the accuracy measure described previously The results are shown in Figure 1, the dotted lines indicate the limits above and below which the accuracy rates are significantly different from chance at the $p < 0.05$ level[7]

Subjects distinguished, on average, between posed and spontaneous expressions at an above-chance level. However, only in the case of spontaneous expressions of both happiness and disgust were these discriminations statistically significant. Raters were unable to identify the posed expressions at better than chance level. Overall, happiness expressions were more accurately decoded than were disgust expressions.

It is noteworthy, that the subjects could accurately identify spontaneous expressions at above-chance level based on information from the face alone. That spontaneous expressions were better identified than posed expressions is consistent with the 'perceiver bias' found in other studies on decoding accuracy (see Zuckerman, DePaulo and Rosenthal, 1981). However, since only 32.5 per cent of the perceivers showed a marked preference for the label 'spontaneous' (that is, applied the label 'spontaneous' to significantly more than 50 per cent of the stimuli) the results rather seem to indicate that subjects were somewhat more accurate in decoding spontaneous expressions than in decoding deliberate expressions. The data, however, cannot be viewed as providing strong support for this assertion.

[7] The limits were calculated by using the t-test formula for point estimation, choosing the appropriate $t(39)$ at the $p = 0.05$ level, and solving the equation for the sample mean.

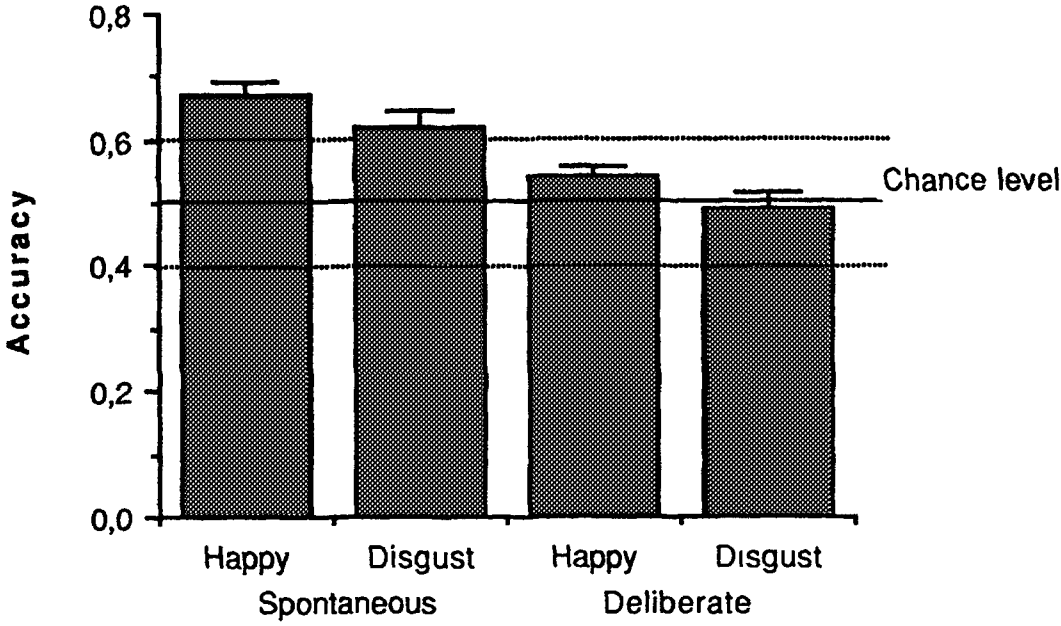

Figure 1. Mean accuracy score as a function of elicitation condition and emotional expression

Structural equation modelling

The results described above suggest that subjects recognized spontaneous expressions at above-chance level, but not posed expressions. Yet, even in the former case, the level of decoding accuracy evidenced is not very striking and must be qualified by the perceivers' preferential use of the 'spontaneous' label. This relatively unimpressive ability to discriminate the two types of expressions stands in stark contrast to the subjects' willingness to advise others (via their decision rules) regarding what cues signal when an expression is deliberate and when it is spontaneous and to their own level of confidence in their judgments. The possibility thus exists that perceivers as a group are responding to very similar cues in attempting to discern the degree of voluntary control present in an expression but that the cues they are focused on are less than optimal.

To assess whether the observers *agreed* at above-chance level in their judgments (regardless of whether these judgments were correct or not), a binomial test was conducted. For 28 episodes an interrater agreement greater than chance was found, a significantly larger number of episodes than would be expected by chance ($z = 14.81$, $p < 0.001$). These results suggest that the observers based their overall rating on shared rules when decoding the expressive episodes for approximately half the episodes. However, these findings relate only agreement regarding the overall judgement (which is presumably based on the use of shared rules), to investigate whether subjects agree regarding specific cues a structural equation modelling approach is appropriate. Given that the above-mentioned results provide support for the notion that the subjects' ratings do not simply reflect idiosyncratic rules or a random choice it is justifiable to further analyse the data using this approach. Therefore, a path analysis was performed (using LISREL VII, Jöreskog and Sörbom, 1988).

As noted earlier, the cues selected for inclusion in this model were based in part

on the decision rules suggested by the subjects as a group (see above). The independent ratings of 'honest' demeanour were included since this variable was expected to influence the accuracy of the subjects' judgments, in the sense that honest appearing people may be expected to not dissimulate their emotional expressions. First, a null model assuming that the relations among all variables are zero was evaluated. As Table 2 shows this model can be easily rejected.

Table 2. Indices of fit

Model	X^2	*df*	*p*	Δ	IFI
0. Null model	126.09	36	0.001	—	—
1. Model 1 assuming that all cues are valid and used by the observers	30.04	19	0.051	0.76	0.90
2. Model 2, respecified model 1, with nonsignificant paths dropped and path from crow's feet wrinkles to blink frequency added	8.32	15	0.910	0.93	1.06
3. Model 3, model 2 with three nonsignificant paths dropped	11.67	18	0.864	0.91	1.06

The first attempt to model the observed data assumed an 'ideal decoder', specifically, a situation where all (see Figure 2) the objectively scored cues are both employed by the subjects and are valid cues. That is, all cues are influenced by the elicitation condition (i.e. the type of expression) and in turn influence the ratings of the degree to which the expression is spontaneous or deliberate[8] The accuracy of the ratings in turn should be influenced by the variables used as well as by the apparent honesty of the stimulus persons, since voluntary expressions of honest appearing stimulus persons should be rated as less controlled or more spontaneous.

Again, as expected, the indices of goodness of fit (see Table 2) show that this model does not provide a good fit of the data. Several paths were not significant and the modification indices suggested the addition of other paths. Specifically, crow's feet wrinkles and gaze aversion were found to not be influenced by elicitation condition. Further, the paths from speed of onset to both control and accuracy, as well as the paths from blink frequency and gaze aversion to accuracy were nonsignificant. In addition, the modification indices suggested the inclusion of a path from blink frequency to crow's feet wrinkles and from elicitation condition to both accuracy and control.

Therefore, the model was respecified as shown in Figure 3 by dropping nonsignificant paths and adding the paths suggested by the modification indices. This resulted in a significant improvement of the model ($\text{Chi}^2_{(1)}$ 21.72; $p < 0.001$). As Table 2 shows this model (model 2) has very adequate levels of fit. Figure 3 shows three nonsignificant paths in the model, but dropping these paths did not result in a significant improvement of the model (see Table 2). Model 2 was therefore considered to provide an adequate description of the relationships between the observed variables.

[8] Degree of control and not degree of spontaneousness was selected for the analyses, even though the two measures were highly correlated ($r = -0.9678$) because a trend towards a bias of the observers' to attribute spontaneousness to all episodes was found in the previous analyses.

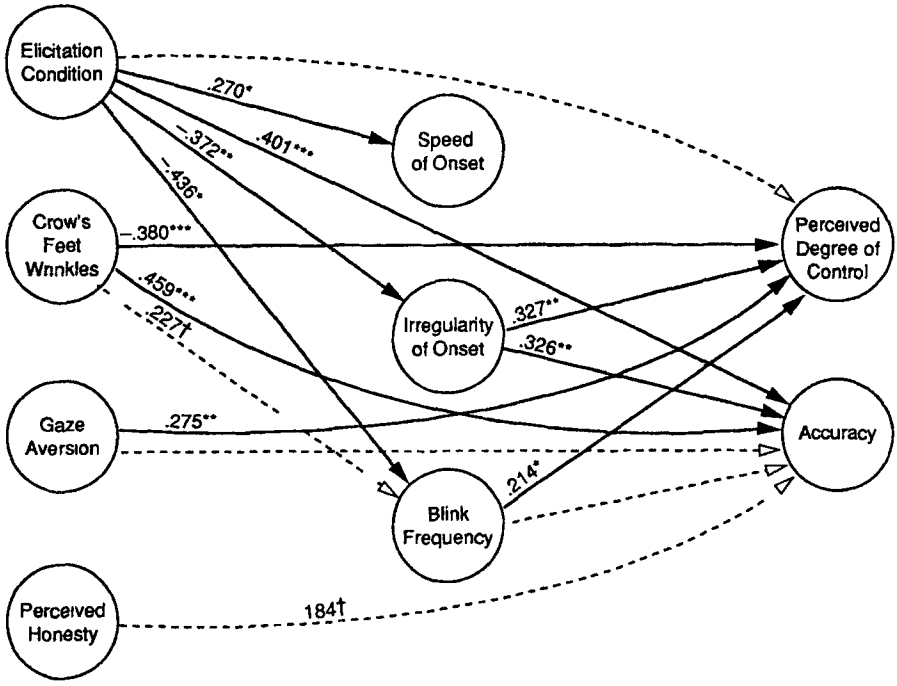

Figure 2. Standardized estimates for the relationships between nonverbal cues and observers' ratings: model 1. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

One should note that the model shown in Figure 3 provides only a description of the data from the current study and cannot be seen as general. Yet, it can be used to evaluate our hypotheses regarding the present study Specifically, we expected that subjects in this study (a) used the nonverbal cues they reported using and (b) that these cues would not constitute an optimal set.

Further, one should note, that the model contained only those cues mentioned by the observers. It is therefore possible that the observers may have used additional cues but did not mention them. For instance, DePaulo *et al.* (1985) found posture shifts to be reliably related to perceived deception, but this behaviour was not mentioned by the observers. The squared multiple correlations (perceived degree of control: $R^2 = 0.55$; accuracy· $R^2 = 0.44$), however, do not suggest that a large part of the variance might be explained by the consistent use by the observers of such an additional cue. The weak, nonsignificant path between elicitation condition and control is consistent with the notion that there are no additional valid cues that influenced the observers' ratings.

The description of the data presented by model 2 suggests that the elicitation condition (spontaneous/posed) has an influence on the irregularity of the expression, which in turn has a significant influence on the observer's judgment of the degree of control of the expression. This is in accordance with the notion that subjects based their ratings of the degree of control of emotional facial expressions in part on the irregularity of the expression. The presence of crow's feet wrinkles influenced both degree of control and accuracy, but was not related to elicitation condition. This is contrary to findings by Ekman and Friesen (1982) and by Ekman *et al.*

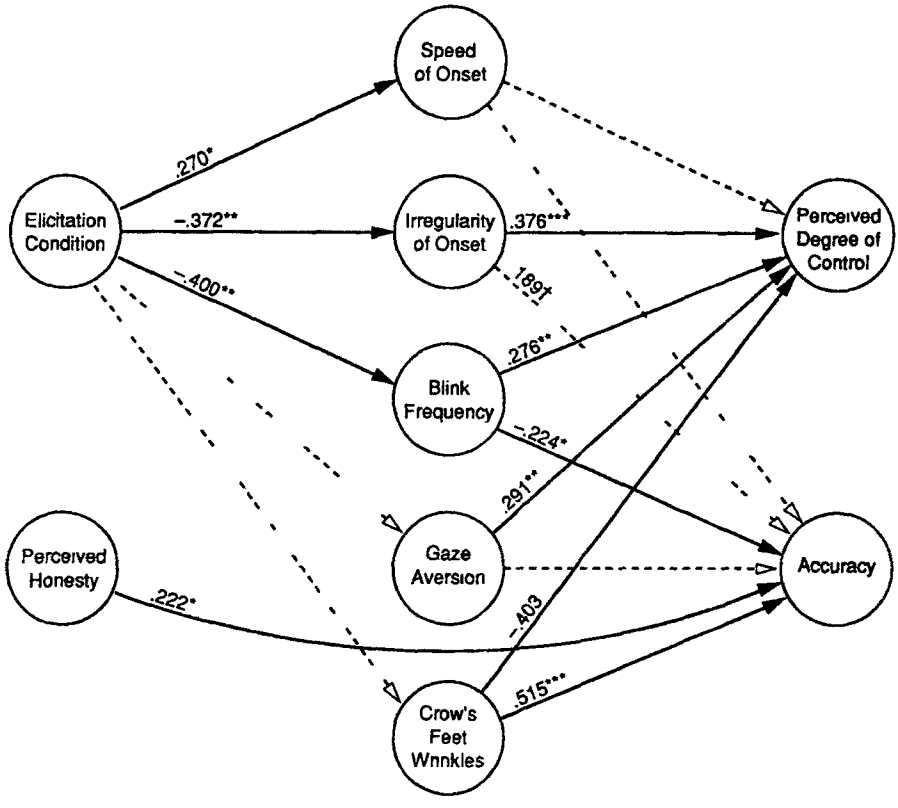

Figure 3. Standardized estimates for the relationships between nonverbal cues and observers' ratings: model 2. († $p < 0.01$, * $p < 0.01$, ** $p < 0.01$, *** $p < 0.001$)

(1988) that crow's feet wrinkles distinguish between spontaneous and deliberate smiles. However, the facial expressions employed here included also expressions of disgust. The strong path from presence of crow's feet wrinkles to the rating of control might, therefore, be seen as an indication that subjects overgeneralized and employed a cue useful for the discrimination of spontaneous and deliberate smiles also for the decoding of the disgust expressions (for which presence of crow's feet wrinkles is not a valid cue)[9]

As expected, the elicitation condition influences eye blink frequency, with posed expressions having a higher eye blink frequency Eye blink frequency and gaze aversion, though the latter was not associated with elicitation condition, influence ratings of the degree of control similarly Observers assessed episodes with higher eye blink frequency and more gaze aversions as more controlled. Further, honest demeanour was found to influence the accuracy scores. While the relationship was only marginally significant, it is interesting to note that even this somewhat crude measure of a 'demeanour bias' could be found to affect the accuracy of the ratings. Since the

[9] One should note that the stimulus pool contained more exemplars of happy than of disgust expressions. The results found regarding the use of the cue crow's feet wrinkles could therefore reflect a bias due to the stimulus material. However, presence of crow's feet wrinkles was not linked to elicitation condition and did not constitute a valid cue. Therefore the conclusion remains that subjects used an invalid cue for their assessment of the spontaneousness of the expression.

assessment of honesty was made based on a photograph of the stimulus person showing a neutral facial expression, this suggests that factors like facial morphology or hairstyle might have influenced the observers' ratings of the facial expressions. One should note that the selection procedure employed here eliminated one potentially valid cue (Hess and Kleck, 1990) — the duration of the expression. The possibility cannot be excluded, therefore, that subjects' accuracy would have been improved if spontaneous and voluntary stimuli had not been roughly matched on length.

In summary, the results show that while on average the stimuli were decoded with above-chance level accuracy, the margin of accuracy was small. The results of the path analysis indicated that the nonverbal cues mentioned in the observers' decision rules did indeed tend to influence the observers' ratings. However, only two of the four cues that influenced the observers' ratings were related to the elicitation condition, that is, were valid cues for the discrimination between spontaneous and posed expressions in this context.

Why are subjects using invalid cues in attempting to make this discrimination? First, it may be that the present task was highly unusual for the subjects, since they could use only visual cues. While they did have some information regarding the context in which the stimulus persons were filmed, this information is clearly less rich than the information typically available in an interactive setting. However, it is not unusual for people to observe others without having access to vocal information (e.g. if someone is in a noisy room, or the person is some distance away) and to make judgments about their emotional states.

Second, it might be that common lore regarding the indicators of voluntarily employed facial expressions is faulty After all, most people do not have the opportunity to receive feedback on misrepresentations they do not detect.

Third, it could be the case that subjects' strategies are adapted to the decoding of only some, very specific, instances of deliberate expressions, like the 'fake smile' This would be supported by the fact that subjects were better in the decoding of the happy expressions. Thus, the decoders might simply have lacked the experience to make the discrimination demanded by the experimental task. Yet, even for the happy expressions, with which the subjects' should be more experienced, the margin of accuracy is low

Fourth, paralleling Kraut's (1980) argument regarding the detection of deception, there may be no social rewards for (or even social sanctions against) detecting dissimulations in the nonverbal behaviour of others. In a Goffman-like manner we may take others' pretences at face value so they will accept whatever face we present to them. If we want to keep smooth well-regulated interactions, the discrimination required by the experimental task is not one we are normally motivated to develop competence in. A well-developed skill in detecting voluntary facial expressions, which are presumably employed to a purpose, is socially not useful.

In this context it is noteworthy, that a relatively crude measure of apparent honesty of the stimulus person, judged based on still photographs, was found to influence decoding accuracy This suggests that the demeanour bias found for the deciding of lies, might also operate in the decoding of spontaneous and posed emotional facial expressions. Overall, the results suggest some parallels between the decoding of deceptive messages and the decoding of deliberate emotional facial expressions exist. However, some of the cues employed by perceivers in this study (e.g. crows' feet wrinkles, and the irregularity of the expression) seem specific to the decoding

of emotional facial expression and suggest that subjects have separate decision rules for the two types of task. A cautionary note is in order insofar as the conclusions are based on a model directly fitted to the observed data, but the results suggest that the decoding of spontaneous and deliberate emotional facial expressions might be a process different from the decoding of lies and thus worthy of investigation in its own right.

REFERENCES

Anderson, J. C. and Gerbing, D. W (1988). 'Structural equation modelling in practice: A review and recommended two-step approach' *Psychological Bulletin*, **103**(3): 411–423.

Bentler, P M. and Bonett, D. G. (1980). 'Significance tests and goodness-of-fit in the analysis of covariance structures' *Psychological Bulletin*, **88**: 588–606.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models' *Sociological Methods and Research*, **17**· 303–316.

Borod, J. C., St. Clair, J., Koff, E. and Alpert, M. (1990). 'Perceiver and poser asymmetries in processing facial emotion' *Brain and Cognition*, **13**: 167–177

DePaulo, B. M., Stone, J. I. and Lassiter, G. D. (1985). 'Deceiving and detecting deceit' In: Schlenker, B. R. (Ed.) *The Self and Social Life*, McGraw-Hill, New York, pp. 323–370.

Ekman, P (1973). *Darwin and Facial Expression: A Century of Research in Review*, Academic Press, New York.

Ekman, P (1985). *Telling Lies*, Norton, New York, NY

Ekman, P., Davidson, R. J. and Friesen, W V (1990). 'The Duchenne smile: Emotional expression and brain physiology II' *Journal of Personality and Social Psychology*, **58**(2): 342–353.

Ekman, P and Friesen, W V (1978). *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto, CA.

Ekman, P and Friesen, W V (1982). 'Felt, false, and miserable smiles' *Journal of Nonverbal Behavior* **6**(4): 238–252.

Ekman, P., Friesen, W V and O'Sullivan, M. (1988). 'Smiles when lying' *Journal of Personality and Social Psychology*, **54**(3): 414–420.

Hager, J. C. and Ekman, P (1985). 'The asymmetry of facial actions is inconsistent with models of hemispheric specialization' *Psychophysiology*, **22**(3): 307–317

Hess, U. and Kleck, R. E. (1990). 'Differentiating emotion elicited and deliberate emotional facial expressions' *European Journal of Social Psychology* **20**: 369–385.

Jöreskog, K. G. and Sörbom, D. (1988). *LISREL VII: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Square Methods*, Scientific Software, Mooresville, IN.

Kappas, A., Hess, U. and Kleck, R. E. (1990). 'The periscope box: A nonobstrusive method of providing an eye-to-eye video perspective', *Behavior Research Methods, Instruments, and Computers*, **22**: 375–376.

Kraut, R. E. (1980). 'Humans as lie detectors: Some second thoughts' *Journal of Communication*, **30**: 209–216.

Rhodes, G., Ronke, K. and Tan, S. (1990). Asymmetries in face perception: Component processes, face specificity and expertise effects' *Cortex*, **26**: 13–32.

Weiss, F., Blum, G. S. and Gleberman, L. (1987). Anatomically based measurement of facial expressions in simulated versus hypnotically induced affect' *Motivation and Emotion*, **11**: 67–81.

Wundt, W (1903). *Grundzuge der physiologischen Psychologie*, Wilhelm Engelmann, Leipzig.

Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T. and Rosenthal, R. (1979). 'Facial and vocal cues of deception and honesty' *Journal of Experimental Social Psychology* **15**: 378–396.

Zuckerman, M., DePaulo, B. M. and Rosenthal, R. (1981). 'Verbal and nonverbal communication of deception' In: Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology*, Academic Press, New York, pp. 1–59