

Do we know what we show? Individuals' perceptions of their own emotional reactions

Ursula Hess¹, Sacha Sénécal², and Pascal Thibault¹

1. *University of Québec at Montreal, Canada*
2. *Health Canada*

Abstract

The present research investigates the intrapersonal perception of emotions. For this, we assessed how well individuals' own perceptions of their emotional expressivity corresponds to (a) their own evaluation when they actually see themselves on video, and (b) evaluations made by naive observers. The results show high correlations between self-judgments; however, correlations between self-judgments and observer judgments were much lower, suggesting that people's own judgments may not be shared by others due to the use of idiosyncratic modes of expression. Further, the data show that individuals tend to overestimate their sending accuracy and that gender stereotypes influence observer ratings of anger and happiness.

Key words: emotion, self-awareness, nonverbal communication.

Correspondence should be sent to Ursula Hess, Department of Psychology, University of Quebec at Montreal, CP 8888, station A, Montréal, QC, H3C 3P8, Canada.
(email: hess.ursula@uqam.ca)

INTRODUCTION

The study of the *interpersonal* perception of emotions has a long tradition in psychology; however, much less interest has been focused on the study of *intrapersonal* perceptions. Yet, in everyday life we often have to rely on self-perceptions of our behavior. In fact, the ability to accurately perceive one's emotions is an important facet of the concept of emotional intelligence (Salovey & Mayer, 1989). Further, emotional behavior is subject to a large number of social norms. Display rules (Ekman & Friesen, 1971) and feeling rules (Hochschild, 1979) prescribe emotional displays and reactions for many social situations (see also Gallois, 1994; Kirouac & Hess, 1999). Developmental research shows that children as young as 5 years old are able to act in accordance with some of these rules (see Saarni & Weber, 1999, for a review). In order to be able to display emotions in accordance with social norms individuals need a certain amount of emotional self-awareness. That is, they need to know what they express during a specific interaction. However, it does not suffice that their expressions match their intentions, e.g., to show an empathic sad expression when told about a sad event, but their expressions must also be recognized by others. This implies two criteria for emotional self-awareness. First, individuals' ratings of recordings of their own emotional behavior should be congruent with their retrospective assessment of their behavior following an interaction. Second, this assessment of their behavior should be congruent with ratings by naïve observers. The first criterion relates to individuals being able to show what they intend to show, whereas the second relates to their communicative ability.

Surprisingly, the research interest in both issues has been muted and the very few studies in this domain tend to focus on the second issue, that is, on the question to what degree peoples' perception of their own emotional reactions is congruent with the external frame of reference of the observer. The first issue, whether individuals' "online" perceptions are congruent with their own frame of reference has not been studied.

The extant research focuses on the congruence between self-ratings and observed behavior with regard to personality constructs and general levels of emotional expressivity as well as the ability to label one's own facial behavior (e.g., Cooper, 1986; Shields & Padawer, 1983; Riggio, Widaman, & Friedman, 1985). The general conclusions drawn from these diverse studies are quite similar – intrapersonal accuracy tends to

be low to not more than moderate. On the other hand, when individuals' ratings of their own video-taped behaviors are compared to ratings by strangers congruence can be quite high (e.g., Ansfield, DePaulo, & Bell, 1995). That is, when watching their own behaviors, individuals may use similar cues to those employed by strangers. That individuals who rate their own behaviors online – while, or just after it occurs – may base their judgment on factors other than the observable behavior is also suggested by a study by Barr and Kleck (1995). Barr and Kleck compared participants' self-reports immediately after an emotional event with self-ratings based on the video-taped behavior. When participants were asked to what degree their taped expression corresponded to their expectations, the majority was surprised by their lack of expressivity.

The present study aims to investigate to what degree individuals are able to accurately perceive their emotional reactions. In this context, it is important to note that the literature on sex differences in emotion generally assumes that women are more aware of their emotions (see, e.g., Fischer, 1993). A secondary aim of the study is therefore to investigate sex differences.

The studies cited above often did not clearly distinguish between self-ratings of emotional states such as “I felt happy” versus self-ratings of emotional behaviors such as “I smiled a lot”. However, these two types of judgments are not the same and one might assume that congruence between participants and observers on behavioral measures might well be superior than on measures of the internal emotional state of the participant. Further, these two measures need not yield congruent results. As mentioned above, Barr and Kleck's (1995) participants were frequently surprised at their lack of overt expressivity. In contrast, a person may well be aware of experiencing, for example, anger, but assumes that the anger was not expressed overtly, when in fact it was. Therefore, in the present study, two sets of scales were used. First, an emotional reaction scale that assessed generalized emotional states and, second, an emotional behavior scale that asked about specific emotional expressive behaviors. A third aim of the study is therefore to compare emotional self-awareness regarding generalized emotional states and specific emotional behaviors.

METHOD

Overview

The data for the present study was collected as part of a larger study that focused on couple communication (Senécal, Murard, & Hess, 2003). The study was conducted in three phases. First, participants were asked to fill out a series of self-report questionnaires. The second phase consisted of a role-playing task. For this, participants played out their most likely reaction to a series of emotion vignettes while being videotaped. Directly following the role-play, participants rated their performance in terms of both their perceived emotional states and their emotional behaviors. The third phase of the experiment was conducted as a separate session. During this session, which took place two to four weeks after the initial session, participants rated their own performance based on the video recording. In addition, each participant rated the expressions of 5 participants of the opposite sex.¹

Participants

Thirty-six men and 36 women participated. Participants were recruited as couples and had been together for at least six months. Participants' age ranged from 18 to 34 years with an average of 24.1 years. For one couple, only session 1 was completed.

Procedure

Session 1. Participants completed a series of questionnaires and took part in a role-playing session. The order of the two tasks was counter-balanced across couples.

For the role-playing task, participants were conducted to a room that was designed to resemble a living room. The participants sat in a comfortable chair, facing an empty chair with a small VCR camera placed

1. Each participant also rated the video-taped expressions of their spouse. This data will not be presented in the present context.

behind the empty chair and aimed towards the participant. The experimenter gave all instructions face to face but stood behind a partition while the participant was engaged in the actual role-play so as not to intimidate or distract the participant. The participants were told that the experimenter would read scenarios aloud and that their task was to try and imagine their spouse sitting in the chair in front of them. They were instructed to react as naturally as possible by saying or doing whatever they would say or do in such a situation. Participants knew that they would be filmed but were assured that sound would not be recorded to respect the privacy of their couple relationship. Each participant completed eight role-play segments. The first two were practice trials. The practice trials were intended to familiarize the participants with the procedure and the experimental situation. The six other role-play scenarios were situations taken from a set of vignettes aimed at eliciting six emotional reactions (happiness, fear, anger, sadness, shame, and guilt).

Session 2. The second session was held 14 to 32 days after the first, with an average of 20 days. Participants watched individually the videotaped segments from the first part of the experiment and rated the intensity of the emotional expressions shown during the role-play. They were instructed to base their judgments on the nonverbal behavior of the person on the tape. Each participant first rated two practice segments while the experimenter was available to answer procedural questions. Participants were then shown a recording of their own role-play and were asked to rate their own emotional behavior. After each segment, participants paused the VCR, rated the intensity of the expression, and then moved on to the next segment. Participants also rated the role-play segments of five individuals of the opposite sex whom they did not know.

Vignettes

The emotion vignettes were created based on a large database of descriptions of emotion eliciting events from 2921 participants from 37 countries (ISEAR, International Survey on Emotion Antecedents and Reactions). The specific vignettes used in this study were selected to represent the core relational themes (Lazarus, 1991) for six emotions: happiness, fear, anger, sadness, shame, and guilt. The final set consisted

of three vignettes per emotion.² The following is an example for a guilt vignette: “You lie to your partner about the reasons why you were late”. An example for an anger vignette is: “Your partner accuses you of something that you have not done”.

Dependent measures

Post role-play ratings. Following each role-play segment, participants were asked to describe their emotional state as well as their expressive behaviors during the role-play. The scales were anchored with 0% (not at all) and 100% (as intense as possible). The nine scales aimed at assessing the participants’ emotional state were: happiness, serenity, anger, contempt, disgust, sadness, fear, shame, and guilt. The nine corresponding emotional behaviors were: to laugh/to get excited (happiness), to smile/to relax/to contemplate (serenity), to hit/to insult/to criticize (anger), to stare/to look hard at (contempt), to grimace/to vomit (disgust), to tremble/to freeze (fear), to cry/to isolate oneself (sadness), to withdraw/to lower one’s head/to blush (shame), and to justify one self/to apologize (guilt). The behaviors were selected based on a series of pre-tests. The first pre-test consisted of asking male and female participants to report the behaviors they would be likely to show when experiencing each of a series of emotions. The second phase consisted of presenting a second group of male and female participants with the behaviors and asking them to indicate which emotion each behavior reflects. Behaviors for which more than 60% of participants agreed regarding their significance and which were not selected as indicative of any other emotion by more than 10% of participants were retained. For ease of reference, we will refer to the emotion described by the behaviors rather than to the behaviors throughout the manuscript.³

Video ratings. To judge the videotaped role-play segments, participants used the same scale as for the post role-play self-ratings.

2. The complete set of 18 vignettes used in this study is available upon request from the authors.

3. Due to a clerical error, the first six couples who participated in the study did not fill out these scales and were thus assigned missing data for this part of the study.

Observer ratings. Observers used the same scales as the actors to judge the video-taped expressions of other participants.

RESULTS

Data treatment

As mentioned above, two criteria for emotional self-awareness were used: congruence of post role-play ratings with video self-ratings and congruence of video self-ratings with observer video ratings. The first assesses whether participants actually expressed what they had intended to express during the role-play according to their own frame of reference. The second assesses whether they successfully communicated the intended emotion to others.

To assess correspondence between the participants' post role-play ratings and their self-ratings of the video, we correlated the ratings for each pair of judgments separately across the 9 items of the emotion and the 9 items of the behavior scales. To assess the correspondence between participants' post role-play ratings and observer ratings of the video, two additional sets of correlations were obtained for each participant: the correlations between post role-play ratings and the observer ratings and the correlations between the video self-judgments and the observer ratings. As correlations can not be submitted to parametric statistics, the r values were transformed into z -scores using the Fischer r to z transform for the purpose of the analyses. For the correlations reported below, the z -scores were transformed back into correlations.

Correlations are insensitive to differences in level. To assess whether participants overestimated their level of expressivity following the role-play, as suggested by Barr and Kleck (1995), profile analyses were employed to compare the post-role-play ratings with the self- and observer ratings of the video.

Intrapersonal accuracy

Based on the relevant t -statistic, all correlations were significantly different from zero ($p < .05$, detailed tables available from first author).

To assess the level of congruency as a function of sex, type of rating scale, and emotion situation, an analysis of variance with the between-subject factor Sex of participants and the within-subjects factors Emotion (happiness, fear, anger, sadness, shame, and guilt), and Type of scale (emotion, behavior) was conducted on the z-transformed correlations. A main effect of Sex, $F(1, 58) = 6.34, p = .015$, as well as a Sex \times Type of scale, $F(1, 58) = 3.87, p = .054$, and a Type of scale \times Emotion situation, $F(5, 54) = 4.41, p = .002$, interaction emerged. Although participants were overall more congruent when rating emotions than behaviors for happiness ($r = .66$ versus $r = .60$), anger ($r = .66$ versus $r = .46$), fear ($r = .69$ versus $r = .61$), and sadness ($r = .70$ versus $r = .52$) situations, the reverse pattern was found for shame ($r = .47$ versus $r = .69$) and guilt ($r = .50$ versus $r = .68$) situations. As regards sex differences, women's ratings were overall more congruent than men's; however, this contrast was significant for the emotion ratings only. That is, women were more congruent in their perception of their emotional *feelings* ($r = .51$, for men and $r = .71$ for women) than were men but not in their perception of their emotional expressive *behaviors* ($r = .57$, for men and $r = .63$ for women).

Actor – observer correspondence

Overall, correspondence levels were considerably lower for actor-observer correspondence than for intrapersonal correspondence ($r = .34$ versus $r = .62, z = 2.08, p = .038$ for emotional reactions, and $r = .38$ versus $r = .60, z = 1.63, p = .106$ for emotional behaviors). In fact, only for shame and guilt situations and for men's ratings of happiness situations were the correlations significantly different from 0 for both scales ($p < .05$). For the emotion scales only, women's ratings of the anger situation were also above chance congruent with observer ratings. The analysis of variance revealed a main effect of Emotion situation, $F(5, 52) = 14.03, p < .001$, as well as an Emotion situation \times Type of scale interaction, $F(5, 52) = 5.58, p < .001$. Overall, ratings for fear and sadness situations were least congruent (fear: $r = .12$ and $r = .02$; sadness: $r = .18$ and $r = .09$, for emotional reactions and behaviors, respectively), those for happiness and anger showed middling congruence (happiness: $r = .38$ and $r = .30$; anger: $r = .36$ and $r = .38$, for emotional reactions and behaviors, respectively), and ratings for

shame and guilt situations showed the highest level of congruence (shame: $r = .49$ and $r = .71$; guilt: $r = .46$ and $r = .68$, for emotional reactions and behaviors, respectively).

As regards the mean correlations for the correspondence between participants' ratings of their video-recording and observer ratings of the same recording for the emotion scale, correspondence levels were again found to be significantly lower than for intrapersonal correspondence ($r = .33$, $z = 2.15$, $p = .031$). For the behavior scale, the congruence level was also lower but the difference was again less pronounced, ($r = .45$, $z = 1.11$, $p = .266$). In fact, only ratings of emotional reactions ($r = .30$, for men and women combined) and expressive behaviors ($r = .61$, for men and women combined) for shame and guilt situations (emotional reactions: $r = .37$; emotional behaviors: $r = .61$, for men and women combined) correlated significantly above chance. In addition, correlations for men's emotional reactions to happiness situations ($r = .42$) and for women's behavioral reactions to anger situations ($r = .50$) reached significance. The analysis of variance revealed a main effect of Emotion situation, $F(5, 52) = 2.99$, $p = .019$, which was qualified by a Type of scale \times Emotion situation interaction, $F(5, 52) = 3.10$, $p = .016$. Whereas no significant differences in correspondence between self video-ratings and observer video-ratings emerged across emotions for the emotion scale (happiness $r = .35$, fear $r = .27$, anger $r = .38$, sadness $r = .30$, shame $r = .30$, and guilt $r = .37$), correspondence was significantly higher for shame ($r = .61$) and guilt ($r = .61$) situations than for fear ($r = .26$, $z = 2.78$, $p = .006$), sadness ($r = .32$, $z = 2.61$, $p = .010$), and happiness ($r = .35$, $z = 2.06$, $p = .040$) situations for the behavior scale.

These findings replicate previous findings of low levels of accord between self-ratings and observer ratings (Barr & Kleck, 1995), as well as the observation that self-ratings and observer ratings of the same video-taped emotional behaviors ($r = .45$ versus $r = .33$) are relatively more congruent than ratings of emotional state (Ansfield, DePaulo, & Bell, 1995). However, the question arises how observer and self-ratings differ. To answer this question, profile analyses were conducted. These analyses allow us to directly compare the pattern of judgments from all three types of ratings: post role-play ratings, self video-ratings, and observer video-ratings. The analyses were conducted separately for each scale type.

Profile analyses. Profile analysis is a special application of multivariate repeated measures analysis of variance designed to assess whether profiles from different groups differ on a set of measures. The analysis comprises three tests. The first test assesses whether the ratings across the nine scales follow profiles, that is, whether some emotions were rated as more intense whereas other were rated as less intense. The second test, "level", assesses whether one group's scores were on average higher on the collected set of measures. For example, were self-ratings generally more intense than ratings by observers? Finally, the test of parallelism assesses whether different groups have different profiles. In the present context, the question is whether different judgments result in the same pattern of highs and lows across emotions (for more details regarding this procedure see, e.g., Tabachnick & Fidell, 1996).

Table 1
F-values for profile analyses of emotion and behavior ratings as a function of type of judgment

Situation	Emotional reactions			Emotional behaviors		
	Flat <i>F</i> (8, 1578)	Level <i>F</i> (2, 197)	Parallel <i>F</i> (16, 1578)	Flat <i>F</i> (8, 1578)	Level <i>F</i> (2, 197)	Parallel <i>F</i> (16, 1578)
Happy	58.92 ^{***}	3.37 [*]	9.70 ^{***}	54.33 ^{***}	5.05 ^{**}	6.81 ^{***}
Fear	61.91 ^{***}	11.72 ^{***}	18.30 ^{***}	14.28 ^{***}	.75	2.79 ^{***}
Anger	51.04 ^{***}	.79	1.03	46.50 ^{***}	.26	1.21
Sad	30.38 ^{***}	.51	3.38 ^{***}	24.87 ^{***}	.40	2.68 ^{***}
Shame	51.97 ^{***}	3.25 [*]	3.74 ^{***}	119.51 ^{***}	3.18 [*]	4.02 ^{***}
Guilt	69.80 ^{***}	13.37 ^{***}	9.03 ^{***}	146.18 ^{***}	6.14 ^{**}	9.44 ^{***}

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 1 shows the *F*-values and significance levels for both ratings of emotional state and behavior ratings. As expected, across all emotion situations and both types of scales the profiles were significantly non-flat. More pertinent in the present context is whether profiles differ in overall level and whether they are non parallel. Differences in level

reflect differences in the intensity of the ratings across the emotion profile. Profiles are non-parallel when the peaks and troughs of the different types of ratings do not line up, reflecting an emotion by judgment type interaction.

Overall the same general pattern emerged for the emotional reaction and the emotional behavior scales (detailed tables are available from the first author). A series of post-hoc tests ($p < .05$) revealed that in general self video-ratings and observer video-ratings did not differ with regard to the level of intensity of the ratings. However, there are significant differences in intensity ratings between post role-play ratings on one hand, and both self- and observer ratings of the videotape on the other hand. Specifically, for core emotions, that is, the emotions they intended to portray, the participants' post role-play ratings overestimate expressiveness ($M = 38.28$, $SD = 14.79$ versus $M = 22.36$, $SD = 4.67$, and $M = 16.19$, $SD = 5.56$, for self and observer video-ratings, respectively), whereas for secondary emotions, significant differences favor higher ratings by observers. This suggests that participants considered themselves not simply more expressive overall, but rather as expressing the intended emotional state more accurately than was actually the case. In other words, they overestimated their sending accuracy.

Gender differences. As mentioned above, there is some reason to believe that women might be more emotionally aware than men. Table 2 shows the results of a series of profile analyses to assess gender differences in ratings. The profile analyses were conducted separately for each type of judgment (detailed tables are available from the first author). All profiles were significantly non-flat. For the post role-play ratings, profiles for emotional reactions were significantly non-parallel for happy, fear, and anger role-plays, that is, men and women rated their reactions differently across the profiles. Profiles for emotional expressive behaviors were significantly non-parallel for fear and sadness situations. The main difference between men and women was that women (reactions: $M = 45.37$, $SD = 36.12$; behaviors: $M = 31.94$, $SD = 35.38$) reported more intense emotional reactions as well as more emotional expressive behaviors than did men (reactions: $M = 31.20$, $SD = 30.05$; behaviors: $M = 22.35$, $SD = 29.08$) for theme congruent emotions, as well as across all emotion scales (reactions: $M = 10.59$, $SD = 7.85$ versus $M = 13.18$, $SD = 8.29$; behaviors: $M = 9.04$, $SD = 5.03$ versus $M = 11.72$, $SD = 8.75$, for men and women, respectively).

Table 2
F-values for profile analyses of emotion and behavior ratings as a function of gender

Situation	Emotional reactions			Emotional behaviors		
	Flat <i>F</i> (8, 1578)	Level <i>F</i> (2, 197)	Parallel <i>F</i> (16, 1578)	Flat <i>F</i> (8, 1578)	Level <i>F</i> (2, 197)	Parallel <i>F</i> (16, 1578)
<i>Post role-play self-ratings</i>						
Happy	47.07***	4.49*	2.69**	29.78***	5.20*	1.33
Fear	55.32***	7.82**	4.59***	3.43***	7.92**	3.44***
Anger	18.42***	2.64	3.28***	11.01***	0.01	1.31
Sad	17.11***	0.47	1.71	5.66***	1.54	4.61***
Shame	26.53***	0.03	0.48	46.41***	1.21	1.15
Guilt	44.87***	0.87	0.96	75.67***	1.16	0.85
<i>Video self-rating</i>						
Happy	8.36***	0.01	2.16*	10.70***	0.00	2.59**
Fear	12.46***	3.74	3.47***	7.69***	0.10	1.60
Anger	14.82***	0.76	2.06*	15.22***	1.65	1.35
Sad	7.62***	0.55	1.21	7.89***	0.02	1.09
Shame	7.77***	0.05	1.18	25.72***	0.07	0.98
Guilt	13.11***	1.70	1.91	30.60***	3.97*	1.96*
<i>Video other-rating</i>						
Happy	14.17***	0.81	0.86	24.22***	0.51	1.36
Fear	15.07***	4.58*	1.46	16.07***	3.37	2.67**
Anger	25.91***	3.27	1.91	34.73***	3.25	4.88***
Sad	12.88***	1.08	0.32	37.71***	4.51	1.16
Shame	28.86***	0.57	1.49	67.51***	1.93	0.14
Guilt	17.39***	2.78	1.37	57.11***	2.83	1.19

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

Further, women tended to report more feelings ($M = 28.42$, $SD = 34.60$ versus $M = 18.72$, $SD = 24.80$) and behaviors ($M = 13.96$, $SD = 26.39$ versus $M = 2.75$, $SD = 7.75$) of sadness across all emotion situations. In turn, men reported more anger in happiness situations (reactions: $M = 5.83$, $SD = 12.74$ versus $M = 0.33$, $SD = 1.27$; behaviors: $M = 2.83$, $SD = 8.48$ versus $M = 1.67$, $SD = 9.12$), but less anger in anger situations (reactions: $M = 18.57$, $SD = 19.30$ versus $M = 36.00$, $SD = 29.89$; behaviors: $M = 19.17$, $SD = 24.88$ versus $M = 25.67$, $SD = 28.37$).

For video self-ratings, profiles for emotional reactions were again non-parallel for happiness, anger, and fear situations. For emotional behaviors, profiles were non-parallel for happiness and guilt situations only. Men rated themselves as showing less happiness (reactions: $M = 14.43$, $SD = 25.31$ versus $M = 28.77$, $SD = 36.76$; behaviors: $M = 6.29$, $SD = 14.87$ versus $M = 13.94$, $SD = 23.03$) and serenity (reactions: $M = 13.57$, $SD = 25.08$ versus $M = 14.86$, $SD = 25.93$; behaviors: $M = 12.71$, $SD = 23.87$, versus $M = 26.71$, $SD = 32.40$) during the happiness situation, as well as less sadness (reactions: $M = 16.57$, $SD = 28.30$ versus $M = 29.31$, $SD = 31.80$; behaviors: $M = 5.43$, $SD = 14.87$ versus $M = 12.43$, $SD = 23.12$) and fear (reactions: $M = 7.57$, $SD = 13.03$ versus $M = 14.71$, $SD = 23.98$; behaviors: $M = 4.86$, $SD = 15.02$ versus $M = 7.14$, $SD = 19.68$) during the sadness situation.

For the anger situation men rated themselves as showing more happiness (reactions: $M = 5.57$, $SD = 16.03$ versus $M = 0.29$, $SD = 1.69$; behaviors: $M = 5.43$, $SD = 17.71$ versus $M = 2.00$, $SD = 9.01$) and disgust (reactions: $M = 9.71$, $SD = 18.19$ versus $M = 3.28$, $SD = 6.18$; behaviors: $M = 5.57$, $SD = 17.27$ versus $M = 4.54$, $SD = 13.74$) feelings. Finally, for the guilt situation men rated themselves as showing more anger (reactions: $M = 10.00$, $SD = 23.39$ versus $M = 5.71$, $SD = 14.41$; behaviors: $M = 8.86$, $SD = 15.82$ versus $M = 2.29$, $SD = 7.31$), but less guilt (reactions: $M = 22.43$, $SD = 29.71$ versus $M = 29.71$, $SD = 32.45$; behaviors: $M = 32.43$, $SD = 31.04$ versus $M = 47.57$, $SD = 38.49$) behaviors that did women.

When men's and women's expressions were rated by others, no differences in attributed feeling states emerged. Only the profiles for emotional behaviors during fear and anger role-plays were significantly non-parallel. Men were rated as showing more sadness behaviors during both situations ($M = 9.60$, $SD = 10.22$ versus $M = 1.58$, $SD = 3.27$

and $M = 3.26$, $SD = 5.18$ versus $M = 1.45$, $SD = 2.41$, for fear and anger situations, respectively), as well as showing more anger and contempt behaviors during the anger role-play ($M = 27.55$, $SD = 18.71$ versus $M = 15.85$, $SD = 16.73$ and $M = 18.38$, $SD = 18.08$ versus $M = 7.92$, $SD = 9.11$, for anger and contempt, respectively). Thus, whereas men actually *reported* having felt less anger than women during the anger role-play, and did not rate their own behaviors differently from women, observers nonetheless *perceived* more anger and contempt.

DISCUSSION

Two different facets of emotional self-awareness were assessed. First, we compared participants' self-reports of emotional reactions and expressive behaviors following a role-play situation with their assessment of the same behavior when seen later on video. In general, correlations between these two ratings were significant and ranged from $r = .43$ for men's sadness behaviors to $r = .81$ for women's emotional state of sadness. Women were found to be significantly more congruent with themselves when rating emotional states; however, no sex differences in congruence were found for ratings of emotional expressive behaviors.

Overall, these data seem to suggest relatively high levels of intrapersonal accuracy in emotional self-perception. However, the profile analyses suggest that participants tended to overestimate their sending accuracy; specifically, they overestimated the intensity with which emotion theme congruent emotions (i.e., sadness in response to a sadness vignette) were shown as well as to underestimate the presence of secondary emotions (i.e., emotions other than sadness for a sadness vignette), that is, participants tended to overestimate the clarity of their emotional messages.

The second facet of emotional self-awareness regards whether the participants' self-assessment of the emotional reactions corresponds to observer ratings. That is, whether the emotional reaction was communicated successfully. In this case, both the correlations and the profile analyses revealed relatively lower levels of agreement. The ratings correlated overall only at below chance level and actors again overestimated their sending accuracy. Finally, we compared self-ratings of the videotape with observer ratings of the videotape. Although there was some

overall agreement in terms of the rated intensity, the rating profiles were non-parallel and congruence scores were low to moderate. These findings suggest that actors and observers differed in their qualitative perception, that is, while observers rated some emotions more intensely actors rated others more intensely.

These findings demonstrate that estimates of intrapersonal accuracy depend on the criterion that is used. When the actors' own perceptions of their video-taped behavior were used as criterion to assess the accuracy of their post role-play ratings, individuals showed – contrary to some previous findings – generally high levels of awareness of their emotional reactions.⁴ However, at the same time, the present findings also replicate Barr and Kleck's (1995) observation that individuals overestimate their sending accuracy somewhat. Specifically, they consider themselves as more expressive of the core emotion corresponding to the emotion situation (i.e., anger in an anger situation), when they make the post role-play ratings than they do when making the self video-ratings.

A different picture of participants' emotional self-awareness emerged when observer judgments were used as the criterion. In this case, congruence scores were much lower, suggesting low levels of emotional self-awareness. That is, participants did indeed convey the intended emotions according to their own frame of reference, but were not necessarily able to communicate these emotions to others. This is especially interesting given the specific circumstances of the present experiment. Participants were asked to role-play emotions in a situation which invited them to communicate their feelings clearly. The explicit requirement to communicate clearly is not part of everyday situations and hence if anything the present study overestimates the success of emotional communications in everyday life. This has clear implications for interpersonal communication in general. It is a frequent complaint that an interaction partner should have known how someone felt about one issue or the other. This complaint about interaction partners' perceived lack of empathy is easy to understand when we consider individuals' tendency to overestimate their encoding accuracy as shown in the present study.

4. Given that the video ratings sessions were conducted at least 2 weeks after the recording and that the segments were shown in a random order it is very unlikely that the participants remembered their previous ratings and were able to reproduce them for the video-rating.

Although people may well feel that they communicated an emotion clearly enough and may also have been able to observe their own emotional behavior as congruent with the intended communication – they may not realize that for an outside observer this communicative intent was less obvious.

However, this overall conclusion does not apply to shame and guilt situations for which congruence levels were high regardless of the criterion used. It is interesting to speculate on the reasons for the high level of congruence for these two emotions. It is likely that participants have a different expressive repertoire for different emotions. For emotions that are frequently communicated, the expressive repertoire may include fairly idiosyncratic modes of expression (as suggested by Fernandez-Dols & Carroll-James, 1997), which would tend to reduce sending accuracy when observers who did not know the participant were used as a criterion. In contrast, it is possible that for somewhat rarer emotions such as shame and guilt, participants have not developed these idiosyncratic modes of expression thus making it easier for the observers to decode their displays.

It may be argued that participants' high levels of intrapersonal congruence can be explained by the specific circumstances of the experiment; specifically, the fact that participants were facing a camera during the role-play. Cameras have been used to induce increased self-consciousness and it has been argued that increased self-consciousness leads to higher levels of intrapersonal accuracy (see Silvia & Gendolla, 2001, for a review). Specifically, in the present case one may argue that participants treated themselves as objects of their own perception already during the post role-play ratings and hence it is not surprising that when they actually see themselves in the video they come to the same conclusions regarding their emotional state. However, two issues speak against that interpretation. First, while participants made their ratings they were not facing a camera – and the induced self-consciousness effects of a camera tend not to be stable over time. Secondly, and more importantly, induced self-consciousness has not been shown to actually improve awareness of internal states such as physiological symptoms or emotional reactions (Silvia & Gendolla, 2001). This observation favors our original interpretation, that is, individuals express emotions often by using an idiosyncratic expressive style, which may not be shared by others. This effect would be more pronounced for frequently expressed

emotions such as anger and happiness and less so for emotions that are presumably rarer such as shame and guilt.

A third aim of the study was to assess whether there are differences in congruence of judgments of emotional reactions versus behaviors. Interestingly, overall very few differences were found. In fact, only the congruence of self video-ratings with observer video-ratings was found to be somewhat higher for behaviors ($r = .45$) than emotional reactions ($r = .33$) but this difference does not reach significance. The pattern of results for the profile analyses also does not differ between the two measures. It is interesting to note that those differences between scales that did emerge related to gender differences.

Overall, gender differences emerged mainly for anger, fear, and happiness role-plays. Generally, women reported feeling and showing more emotions during the role-play. Women also tended to report more sadness and fear. Interestingly, although men and women did not rate their anger behaviors differently, observers attributed more anger behaviors to men. The attribution of anger to man is part of the emotional gender stereotype (see, e.g., Fischer, 1993). This tendency for observers to rate the actors' behaviors in congruence with the emotional gender stereotype and differently from the actors own evaluations supports Hess and Kirouac's (2000) notion that gender stereotypes can influence the decoding of emotions.

In sum, these results suggest that men and women show high levels of emotional self-awareness when their own judgments are used as criterion but not when observer' ratings are used as a criterion. This finding applied to both ratings of emotional reactions and ratings of emotional behaviors. This implies that although people may think that they successfully communicated an intended emotional message as judged by their own frame of reference, this message may not be understood in the intended way by others. This finding is especially interesting as in the context of the role-play participants are more likely to attempt to make an effort to communicate their emotional reactions clearly. The lack of congruence with naïve observers suggests that individuals may use fairly idiosyncratic means of emotion expression, which are not necessarily shared by others. This is supported by the finding that for the - presumably rarer shown - emotions of shame and guilt considerably higher levels of congruence between self- and other-ratings were found.

A second source for the misinterpretation of intended emotional messages was suggested by the analysis of gender differences in observer

ratings. Although the male and female actors did not rate their happiness and anger behaviors differently, the observers attributed more anger behaviors to men and showed a tendency to attribute more happiness behaviors to women. This suggests that gender stereotypes may have functioned as decoding rules and biased the observation of emotional behaviors in others.

ACKNOWLEDGEMENTS

The study was supported in part by a grant from the Fond FCAR to the first author.

RÉSUMÉ

Cette recherche étudie la perception intrapersonnelle des émotions. Pour cela, nous évaluons à quel point la perception que les individus ont de leur expressivité émotionnelle correspond à (a) leur propre évaluation quand ils se voient sur une vidéo et (b) aux évaluations faites par des observateurs naïfs. Les résultats montrent des corrélations élevées entre les auto-évaluations. Toutefois, les corrélations entre les auto-évaluations et les jugements donnés par des observateurs sont beaucoup plus faibles, suggérant que les propres jugements des individus ne seraient pas partagés par les autres en raison de l'utilisation de modes idiosyncrasiques d'expression. De plus, les données montrent que les individus tendent à surestimer l'exactitude de l'émotion transmise et que les stéréotypes de genre influencent les estimations des observateurs concernant la colère.

REFERENCES

- Ansfield, M. E., DePaulo, B. M., & Bell, K. L. (1995). Familiarity effects in nonverbal understanding: Recognizing our own facial expressions and our friends'. *Journal of Nonverbal Behavior*, *19*, 135-149.
- Barr, C. L., & Kleck, R. E. (1995). Self-perception of the intensity of facial expressions of emotion - Do we know what we show? *Journal of Personality and Social Psychology*, *68*, 608-618.
- Cooper, V. W. (1986). Participant and observer attribution of affect in interpersonal conflict: An examination of noncontent verbal behavior. *Journal of Nonverbal Behavior*, *10*, 134-144.

- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*, 124-129.
- Fernandez-Dols, J. M., & Carroll-James, M. (1997). Is the meaning perceived in facial expression independent of its context? In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression. Studies in emotion and social interaction* (2nd edition, pp. 275-294). New York: Cambridge University Press.
- Fischer, A. (1993). Sex differences in emotionality: Fact or stereotype? *Feminism and Psychology*, *3*, 303-318.
- Gallois, C. (1994). Group membership, social rules, and power: A social psychological perspective on emotional communication. *Journal of Pragmatics*, *22*, 301-324.
- Hess, U., & Kirouac, G. (2000). Emotion expression in groups. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions* (2nd edition, pp. 368-381). New York: Guilford Press.
- Hochschild, A. R. (1979). Emotion work, feeling rules, and social structure. *American Journal of Sociology*, *85*, 551-573.
- Kirouac, G., & Hess, U. (1999). Group membership and the decoding of nonverbal behavior. In P. Philippot, R. Feldman, & E. Coats (Eds.), *The social context of nonverbal behavior* (pp. 182-210). Cambridge, UK: Cambridge University Press.
- Lazarus, R. R. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Riggio, R. E., Widaman, K. F., Friedman, H. S. (1985). Actual and perceived emotional sending and personality correlates. *Journal of Nonverbal Behavior*, *9*, 69-83.
- Saarni, C., & Weber, H. (1999). Emotional displays and dissemblance in childhood: Implications for self-presentation. In P. Philippot, R. Feldman, & E. Coats (Eds.), *The social context of nonverbal behavior* (pp. 71-105). Cambridge, UK: Cambridge University Press.
- Salovey, P., & Mayer, J. D. (1989). Emotional intelligence. *Imagination, Cognition, and Personality*, *9*, 185-211.
- Senécal, S., Murard, N., & Hess, U. (2003). Do you know what I feel? Partners' predictions and judgments of each other's emotional reactions to emotion-eliciting situations. *Sex Roles*, *48*, 21-37.
- Shields, A. S., & Padawer, J. L. (1983). Children's standards for judging their own facial expressions of emotion. *Journal of Nonverbal Behavior*, *8*, 109-125.
- Silvia, P. J., & Gendolla, G. H. E. (2001). On introspection and self-perception: Does self-focused attention enable accurate self-knowledge? *Review of General Psychology*, *5*, 241-169.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York: Harper Collins.

Received 23 September, 2003

Accepted 15 January, 2004