# Facial Appearance, Gender, and Emotion Expression

Ursula Hess
University of Quebec at Montreal

Reginald B. Adams, Jr.
Harvard University

Robert E. Kleck
Dartmouth College

Western gender stereotypes describe women as affiliative and more likely to show happiness and men as dominant and more likely to show anger. The authors assessed the hypothesis that the gender-stereotypic effects on perceptions of anger and happiness are partially mediated by facial appearance markers of dominance and affiliation by equating men's and women's faces for these cues. In 2 studies, women were rated as more angry and men as more happy—a reversal of the stereotype. Ratings of sadness, however, were not systematically affected. It is posited that markers of affiliation and dominance, themselves confounded with gender, interact with the expressive cues for anger and happiness to produce emotional perceptions that have been viewed as simple gender stereotypes.

That men's and women's emotion displays differ in many ways is well established. Two differences in particular have been noted. First, women report smiling more and are considered by others to smile more than men, and second, men's displays of anger have been reported to be both more pervasive and are generally more acceptable (see Brody & Hall, 2000; Fischer, 1993). Even stronger than observed differences in emotional expressivity between men and women are the stereotypic expectations that individuals hold regarding such differences. These expectations are socialized early and can have dramatic consequences for the perception of emotion in others. For example, even children as young as 5 years tend to consider a crying baby as "mad" when the baby is

purported to be a boy but not when it is purported to be a girl (Haugh, Hoffman, & Cowan, 1980). Also, Gaelick, Bodenhausen, and Wyer (1985) found that husbands tend to interpret the simple absence of smiling during a marital dispute as a sign of hostility on the part of their wives, whereas wives tend to interpret the simple absence of hostility displays by their husbands in such disputes as a sign of love. In this context, Hess, Blairy, and Kleck (1997) found that women's expressions of happiness were perceived as more intense and their expressions of anger and disgust as less intense than expressions of the same physical intensity shown by men. Thus, there is evidence from different domains suggesting that the stereotypical view of men's and women's emotional dispositions translates also into a bias in the perception of their emotions. That is, when men and women or boys and girls show the same expression, this may not lead to the attribution of the same emotion or to the same intensity of emotion.

Although both certain differences in actual behaviors and differences in stereotypic expectations regarding male and female expressivity are well documented, the reasons for these differences remain unclear. Henley (1977, 1995) as well as LaFrance and Hecht (1999, 2000) traced both to a power inequity between men and women. They suggest that women smile more and are expected to smile more because across time and cultures, women have had less power than men, and low social power is thought to be associated with more smiling. This notion is in part

based on the observation that the primate homologue to the human smile, the silent bared teeth display, is generally a sign of appeasement (but also occurs in play situations, just like the human smile; Preuschoft & van Hooff, 1997). Yet, although ample evidence exists that women do smile more, and that social norms and roles influence the size of this difference (for a review, see LaFrance, Hecht, & Paluck, 2003), the evidence that smiling behavior is linked to low social power is somewhat less clear (see Hall & Halberstadt, 1994, for a discussion).

A related explanation for the stereotypical expectations regarding appropriate emotion displays by men and women focuses on their respective social roles (e.g., Brody & Hall, 2000; Shields, 2000). Specifically, it is generally assumed that women's nurturing role favors the acquisition of superior interpersonal skills and the ability to communicate nonverbally, whereas men's roles are seen as more agentic and hence may foster more goal-directed displays (Eagly, 1987; Eagly & Steffen, 1984). Hence, on the one hand, women may be expected to be more affiliative in general in order to sustain nurturing relationships. On the other hand, it may be more acceptable for men to show anger as a goal-directed behavior to address obstacles. For example, male managers whose voice expresses anger when learning about a problem are perceived as more competent than those whose voice expresses sadness, or remains neutral, whereas the same is not true for female managers (Lewis, 2000).

Given these considerations, the gender of an expressor may be considered a reasonable cue for social orientations such as dominance and affiliation, and thus expectations of anger proneness versus proneness to show happiness may be simply based on stereotypical expectations regarding men's and women's roles. Thus, women are presumed to be more affiliative, and hence, they are expected to be more likely to smile, whereas men are presumed to be more dominant and hence may be given more leeway in their expression of anger.

However, gender per se is not the only cue on which people may base such expectations. One important cue in this context is facial appearance. In fact, physical appearance is, in and of itself, an important part of the gender stereotype (Deaux & Lewis, 1984). Thus, Friedman and Zebrowitz (1992) posited that manipulation of physical appearance should influence perceptions of gender-stereotypic behaviors. Specifically, they found that facial maturity drives the perception of gender-stereotypic traits and that the gender stereotype is weakened when comparing males and females of equal facial maturity.

In the present context, it is important to note that a number of aspects of facial appearance that lead to perceptions of dominance and affiliation are highly confounded with gender. Thus, a high forehead, a square jaw, and thicker eyebrows have been linked to perceptions of dominance and are typical for men's faces (Senior, Phillips, Barnes, & David, 1999), whereas a rounded baby face is both feminine and perceived as more approachable (Berry & Brownlow, 1989) and warm (Berry & McArthur, 1986), central aspects of an affiliative or nurturing orientation. This leads to the hypothesis that—regardless of gender—individuals who appear to be more affiliative are expected to show more happiness, and individuals who appear to be more dominant are seen as more anger prone. As cues to gender and cues to affiliation/dominance are highly confounded, this would lead to more women being expected to be happiness prone and more men to be anger prone.

This hypothesis is supported by a recent study (Hess, Adams, & Kleck, in press) showing that the perceived emotionality of men and women is partially mediated by the extent to which their faces are perceived to reflect dominance or affiliation dispositions. Specifically, the tendency to perceive women as more likely to show happiness, surprise, sadness, and fear was mediated by their higher perceived affiliation and lower perceived dominance, respectively. In the same vein, the tendency to perceive men as more likely to show anger, disgust, and contempt was partially mediated by both their higher level of perceived dominance and their lower level of perceived affiliation. This finding suggests that the tendency to rate men's anger expressions as more angry and women's happiness expressions as more happy, observed by Hess et al. (1997), may also be, at least partially, explained by differences in perceived dominance and affiliation of men's and women's faces rather than by gender per se.

Two studies were conducted to test this hypothesis. Faces were manipulated to be equal in terms of the presence and absence of physical cues to dominance and affiliation but to differ with regard to a salient gender cue. If gender category membership per se underlies stereotypical expectations, then male anger faces should be rated as angrier than female faces, and, conversely, female happiness expressions should be rated as happier even if there is no difference in the perceived dominance and affiliation of the faces. In contrast, if perceived dominance and affiliation un-

derlies the gender stereotype, then this bias should disappear when faces are equated for these characteristics.

Specifically, we predicted that if gender stereotypes for emotionality are mediated by gender per se, equating faces on the physical characteristics that drive perceptions of dominance and affiliation would not change the stereotypical differences typically reported in the literature. In contrast, if perceptions of emotionality are driven primarily by perceptions of dominance and affiliation, then equating faces on the pertinent physical characteristics would result in no differences in the emotional attributions made to men and women.

Equating actual men's and women's faces for the physical characteristics associated with perceived dominance and affiliation is not, however, a straightforward proposition. Many of the facial cues that lead to attributions of dominance or affiliation, such as jaw and face shape, eyebrow thickness, and the like, also signal male or female gender (see Keating, 1985; Keating, Mazur, & Segall, 1981; Senior et al., 1999). Thus, for men rated high in dominance, no matching women can be found and vice versa for women rated high in affiliation. Furthermore, the use of actual men's and women's faces makes it difficult to assure that the expressive behavior is equivalent across the sexes. To create male and female stimuli that were both equivalent with regard to physical cues to dominance and affiliation and identical with regard to expressive displays, we used the fact that facial dominance and affiliativeness cues (i.e., eyebrow thickness, height of forehead and jaw form, facial rounding) are located in the interior of the face. In contrast, hairstyle is a powerful cue for gender identification that does not affect expressive behavior or other facial cues. Thus, in our first study, the stimuli were drawings of an expressive androgynous face to which was added a male or female hairstyle. Both the facial cues that convey dominance and affiliation and the expressive behaviors signaling emotion were therefore held constant while varying the apparent gender of the stimuli.

## Study 1

### Method

*Participants.* Twenty-three women and 21 men participated individually. All were students at the University of Quebec at Montreal.

*Stimulus material.* Black-and-white drawings of facial expressions of happiness, sadness, anger, and disgust were created on the basis of an existing set of photographs of emotional displays (Japanese and Caucasian Facial Expressions of Emotion [JACFEE]; Matsumoto & Ekman, 1988). A drawing of a neutral facial expression was also created. On the basis of the neutral (0%) and the relatively intense emotional facial expression (100%), intermediate expressions differing in physical intensity by 20% steps were constructed using Morph 2.5 software for Macintosh. Morphing to create stimuli of varying intensity has been shown to be a valid procedure by a number of investigators (Etcoff & Magee, 1992; Hess et al., 1997; Young et al., 1997). These expressions were then combined with both male and female hairstyles. The resulting set of 48 stimuli (6 intensity steps × 4 emotions × 2 sexes) was presented in a different random order for each participant using an Apple Macintosh with a 14-in. (35-cm) screen. Participants saw each facial expression twice, once with a female hairstyle and once with a male hairstyle. Figure 1 shows examples for happiness and anger.

*Procedure.* Participants came to the laboratory individually. The experimenter explained that their task would be to rate a series of facial expressions regarding the emotion expressed. More detailed instructions regarding the task were presented on
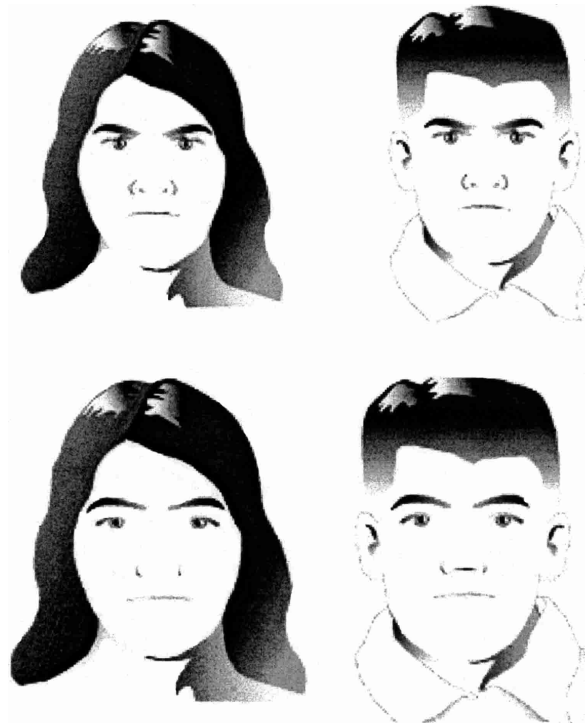


*Figure 1.* Examples of drawings used in Study 1.

Table 1
*Analysis of Variance for the Intensity Ratings on the Target Scale for Study 1*

| Source | df | Anger | Disgust | Sadness | Happiness |
|---|---|---|---|---|---|
| | | | | *F* | |
| Intensity | 5, 39 | 405.83*** | 129.87*** | 54.45*** | 121.46*** |
| Sex | 1, 43 | 13.51*** | ns | 12.43*** | 13.69*** |
| Intensity × Sex | 5, 39 | ns | ns | 2.77* | ns |

* $p < .05$.  ***$p < .001$.

screen. Following the instructions, two practice trials were presented. The emotion ratings were made by clicking on scales presented on the computer screen. Each drawing was shown for 5 s and was then replaced by the rating scales. As soon as the participant had completed all ratings for one stimulus, the next drawing appeared.

*Dependent variables.* Participants were asked to indicate for each drawing the intensity with which it expressed each of the following emotions: anger, contempt, disgust, fear, happiness, sadness, and surprise. The scales were represented by a 200-pixels long, bounded rectangle on the screen, the first 30 pixels of which were white and indicated a judgment of 0.[1] The remaining 170 pixels were graded in color from light gray to dark gray, with the darker end of the scale indicating greater intensity of the emotion. Each scale contained an emotion label and was anchored with the phrases "not at all" and "very intensely."

## Results

We conducted a 6 (intensity levels) × 2 (sex of expressor) multivariate analysis of variance (MANOVA) separately for the rated intensity on the target emotion scale for the expressions of anger, happiness, sadness, and disgust, respectively. All analyses were conducted with an alpha of .05.

We found a main effect of intensity for all four emotions (see Table 1). This main effect suggests that the graded intensity of the expressions was interpreted by the participants as graded intensity in emotional experience. Furthermore, we found a main effect of sex of expressor for all emotions except disgust. We found A Sex of Expressor × Intensity interaction for ratings of the sad expressions only. The means are shown in Figure 2.

Inspection of the means indicated that the same anger expression when shown by an apparent man was rated as less angry than when shown by an apparent woman. This difference was significant for expressions morphed at 80% and 100% intensity. Similarly, sadness expressions were rated as more sad

when shown by an apparent woman than when shown by an apparent man. This difference was nonsignificant for the 40% and 60% expressions and significant for 100% expressions. Finally, expressions of happiness shown by an apparent woman were rated as less intense than the same expressions displayed by an apparent man. This difference was significant for expressions morphed at 20%, 40%, and 60%.

## Discussion

We predicted that controlling for physical cues of dominance and affiliation by using the same faces as apparent men and women would expunge the sex-of-expressor effect in ratings of emotional expressions. This was indeed the case for disgust expressions. However, the present findings for anger and happiness are the opposite of the findings by Hess et al. (1997) using "real" male and female faces. In fact, Hess et al. found that expressions of anger and disgust were rated as more intense when shown by men, and that expressions of happiness were rated as more intense when shown by women.

This suggests that controlling for facial appearance, as was done in the present study, not only reduces gender stereotype effects as predicted, but actually reverses them in some cases. That is, once their higher level of affiliation and lower level of dominance is controlled for, women are actually perceived as expressing more intense anger. Conversely, once their higher level of dominance and lower level of affiliation is controlled for, men are actually perceived as expressing more intense happiness. In contrast, the case for sadness is somewhat less consistent. Hess et al. (1997) found that lower intensity expressions of sadness were rated more intense when shown by men, whereas the 100% expression was rated more intense

---

[1] This scale format was chosen because pretests had shown that participants could not quickly and reliably use the mouse to make a judgment of exactly 0 on the continuous 200-pixel scale.
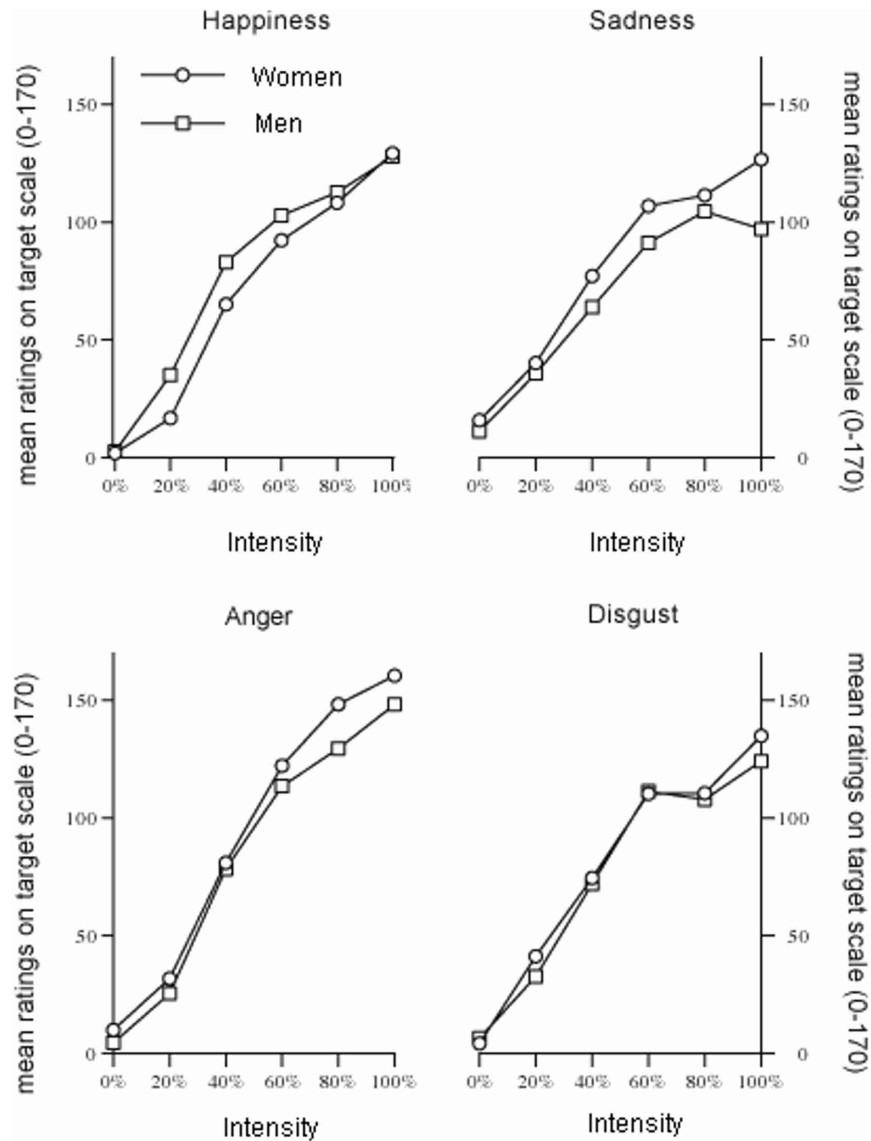
*Figure 2.*    Mean accuracy as a function of emotion and apparent gender of the expressor.

when shown by women. In the present study, all expressions were rated as more intense when shown by an apparent woman, suggesting some level of mediation of perceived dominance and affiliation for lower intensity expressions.

In summary, the results suggest that the presumed gender stereotype effects for ratings of anger, disgust, happiness, and, to some degree, sadness are in fact facial appearance effects. Yet, controlling for facial appearance not only eliminated the gender-stereotypic effects, but actually reversed them for anger and happiness expressions as well as for lower intensity sadness expressions. These findings suggest that women

may actually be perceived as more angry and men as more happy and that this expectation is generally obscured by women's lower levels of perceived dominance and higher levels of affiliation. This explanation would be consistent with the observation that high-dominant women are frequently perceived as unpleasant and aggressive (see, e.g., Hess et al., in press, Study 2).

Although Study 1 seems to strongly support the notion that sex-of-expressor effects, such as those observed by Hess et al. (1997), are mediated by facial dominance and affiliation, it is also possible that the use of drawings as stimuli created an artifact. There-

fore, Study 2 was conducted using real faces that were digitally manipulated in much the same way as the drawings. For this, actual expressions were framed by either the hairline of a woman or the hairline of a man. Study 2 consisted of three experiments that are referred to as Experiments 2a, 2b, and 2c in what follows.

### Studies 2a–2c

Studies 2a–2c all focused on the same question—that is, whether emotional facial expressions in actual gender-manipulated faces are rated differently as a function of the apparent sex of the expressor. However, the use of these stimuli poses additional methodological problems. First, in order to create believable faces, only androgynous faces can be used. Only two male faces from our set could be used for this purpose because skin characteristics and beard shadow made most male faces inappropriate for use with a female hairline. On the one hand, when apparent men and women were presented in within-subject designs, only a limited set of faces could be used without the manipulation becoming evident. On the other hand, a between-subjects design not only allowed the use of a larger set of faces but also introduced biases resulting from anchoring effects when only men or only women were shown. Because all possible methods posed at least some methodological problems, we decided to use a different approach for each of the three studies. Thus, Study 2a used a between-subjects design in which participants saw either only apparent men or only apparent women, whereas Studies 2b and 2c used a within-subjects manipulation. The rating task was also varied. For Studies 2a and 2b, participants were asked to make a categorical choice from a list of emotion labels and to rate the intensity of the expression separately. Study 2c requested participants to provide ratings on an emotion profile (i.e., they rated all faces on several emotion dimensions). The data from these studies were then combined using meta-analytic procedures (Rosenthal & Rosnow, 1991), with the goal to average out effects that were due to the specific combination of the manipulations used and to focus on the convergent effects across studies. These effects were considered to be stable and largely independent of specific manipulations.

### Method

*Participants.* One hundred sixty-five students at Dartmouth College participated for extra course credit. Forty-three additional participants were recruited on campus and paid for their participation. Thirty-four female and 27 male undergraduate volunteers participated in Study 2a. Thirty-six female, 22 male, and 1 gender-unknown undergraduate participated in Study 2b. Forty-eight female and 20 male undergraduates participated in Study 2c.

*Material.* For Study 2a, 12 actual female interior faces were shown either with a female hairline or with a male hairline. An additional 9 male faces were used as filler items. Each face displayed an anger, sadness, fear, disgust, surprise, and happiness expression, respectively. As Studies 2b and 2c used only expressions of anger, fear, sadness, and happiness, only data regarding these expressions are reported. For Study 2b, four different interior faces (2 male and 2 female), each portraying four different emotional expressions (anger, fear, sadness, and happiness), were transposed into four different apparent persons using two male and two female hairlines (yielding 64 total stimuli). For Study 2c, six interior faces that had not been used in Studies 2a or 2b were combined with six different hairlines (3 female and 3 male) to create a total of 36 apparent stimulus persons, each showing four emotions (144 stimuli in total). All interior faces had been rated as *androgynous,* defined as equally masculine and feminine in appearance, in separate pilot tests, conducted for the purpose of selection, and were transposed into apparent men or women using hairstyles that were correctly identified as male or female by 100% of the participants in the pretest. Figure 3 shows an example of an apparent male and an apparent female using this type of manipulation.[2]

*Procedure.* For Study 2a, 29 participants rated the faces with their original female hairlines, and 32 participants rated them with male hairlines in a between-subjects design. Each participant was seated at a computer terminal, and written instructions were presented on the computer screen. The participant's task was to select the label corresponding to the emotion that he or she felt was most clearly conveyed by each face. They then rated the intensity of the expression on a 200-pixels long, bounded rectangle on the screen, the first 30 pixels of which were white and indicated a judgment of 0. The remaining 170 pixels were graded in color from light gray to dark gray,

---

[2] Please note that these are not exemplars of the actual faces used in the study. The expressors had given permission for their faces to be used for research purposes but not for publication. The example shown in Figure 3 used faces drawn from the JACFEE (Matsumoto & Ekman, 1988) and manipulated with permission from the authors.

*Figure 3.* Examples for the gender manipulation of faces used in Study 2. From *Japanese and Caucasian Facial Expressions of Emotion (JACFEE),* by D. Matsumoto and P. Ekman, 1988, San Francisco: San Francisco State University. Copyright 1988 by David Matsumoto and Paul Ekman. Reprinted with permission.

with the darker end of the scale indicating greater intensity of the emotion. Each scale contained an emotion label and was anchored with the phrases "not at all" and "very intensely."

For Study 2b, participants were seated at a table and given a stimulus packet with photographs and a rating booklet. Each participant rated all the expressions of two apparent male and two apparent female expressors for a total of 16 stimuli. To avoid participants seeing the same person once as an apparent male and then again as an apparent female, packets were collated such that no stimulus person appeared twice in a packet with the same expression. This resulted in four different stimulus sets. Participants again first selected the label for the emotion that they felt was most clearly conveyed by each face. They then rated the intensity of the expression using a 7-point scale anchored with 1 = *not at all intense* and 7 = *extremely intense.*

For Study 2c, each participant was seated at a table, randomly assigned one of six stimulus packets, and provided with a rating booklet. Each participant rated the expressions of three apparent male and three apparent female expressors for a total of 24 expressions. No participant saw the same expressor as both a male and a female. The participants rated, on 7-point scales, the degree to which the stimulus person seemed to experience each of seven emotions: anger, fear, sadness, happiness, contempt, disgust, and surprise. Decoding accuracy was assessed by assigning a score of 1 when the highest rated emotion matched the intended emotion expression and a score of 0 when it did not.

### Results and Discussion

For all three studies, we conducted *t* tests to compare the frequency with which each emotion label was chosen for each of the expressions as a function of apparent gender. Table 2 shows the mean intensity ratings and the mean accuracy ratings for expressions of anger, fear, sadness, and happiness.

We combined the results from all three studies using the *z* values for the individual probabilities to derive a *z* statistic (Rosenthal & Rosnow, 1991). Overall, anger displays were more often accurately decoded when displayed by an apparent woman than when displayed by an apparent man ($z = 3.28$, $p < .001$). Across the three studies, anger expressions were rated as more intense when shown by an apparent woman ($z = 4.22$, $p < .001$), whereas happiness expressions were rated as more intense when shown by an apparent man ($z = 2.03$, $p = .021$).[3] In addition, we assessed how often an expression was labeled as showing each emotion, regardless of whether this choice was accurate. Across the three studies, apparent women were more often seen as displaying anger than were apparent men ($z = 4.18$, $p < .001$), whereas apparent men were rated more often as showing happiness ($z = 1.76$, $p = .039$) as well as surprise ($z = 3.09$, $p = .001$).[4]

In summary, the results from Studies 2a–2c provide support for the notion that when facial appearance-equated expressions are manipulated to appear as ei-

---

[3] To combine intensity ratings, mean ratings were calculated across the expressions labeled as *angry, sad, fearful,* or *happy,* respectively, for Studies 2a and 2b. For Study 2c, the mean for the scale on the profile that corresponded to the target emotion was averaged over exemplars. For example, the mean ratings on the Anger scale were calculated for the six anger expressions.

[4] For surprise, only data from Studies 2a and 2c were available.

Table 2
*Mean Intensities and Accuracy Rates as a Function of Apparent Gender for Study 2*

| Emotion expression | Mean intensity | | | | | | Mean accuracy | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | App. women | | App. men | | | | App. women | | App. men | | | |
| | M | SD | M | SD | t(59) | p | M | SD | M | SD | t(59) | p |
| | | | | | | Study 2a | | | | | | |
| Anger | 117.93 | 33.95 | 106.76 | 27.50 | 1.40 | .167 | .69 | .28 | .30 | .31 | 5.22 | .0001 |
| Fear | 146.78 | 27.99 | 126.60 | 35.46 | 2.44 | .018 | .64 | .42 | .45 | .34 | 1.87 | .067 |
| Sadness | 112.05 | 35.44 | 99.89 | 30.80 | 1.42 | .160 | .83 | .27 | .78 | .31 | 0.62 | .539 |
| Happiness | 98.52 | 31.89 | 109.48 | 27.45 | 1.43 | .158 | .98 | .09 | .94 | .21 | 1.10 | .276 |
| | | | | | | Study 2b | | | | | | |
| Anger | 4.64 | 0.94 | 4.35 | 0.98 | 1.90 | .063 | .97 | .13 | .93 | .17 | 1.27 | .209 |
| Fear | 3.84 | 1.21 | 3.90 | 1.15 | 0.26 | .794 | .78 | .28 | .75 | .34 | 0.43 | .672 |
| Sadness | 4.88 | 1.20 | 4.93 | 1.19 | 0.20 | .827 | .83 | .25 | .88 | .24 | 0.96 | .310 |
| Happiness | 4.64 | 0.93 | 4.88 | 1.04 | 1.36 | .179 | .98 | .09 | .98 | .09 | 0.00 | 1.00 |
| | | | | | | Study 2c | | | | | | |
| Anger | 3.47 | 1.45 | 3.02 | 1.41 | 1.70 | .094 | .48 | .34 | .43 | .32 | 0.77 | .484 |
| Fear | 3.18 | 1.30 | 3.41 | 1.30 | 1.33 | .183 | .32 | .86 | .36 | .32 | 0.80 | .428 |
| Sadness | 3.95 | 1.41 | 3.96 | 1.27 | 0.10 | .920 | .71 | .28 | .74 | .23 | 0.63 | .533 |
| Happiness | 5.00 | 1.01 | 5.13 | 1.01 | 0.82 | .416 | .96 | .12 | .96 | .12 | 0.00 | 1.00 |

*Note.* App. = Apparent.

ther male or female, apparent women are rated as showing more intense anger and apparent men as showing more intense happiness. Furthermore, apparent women were more often labeled as *angry* and less often as *happy* than were apparent men when showing the same expressions. That is, the gender-stereotypic effect was reversed for facial appearance-equated expressions even when real faces were used. However, for sadness displays, the findings were less clearly patterned and overall did not suggest a clear facial appearance effect. Thus, the results from Study 2 largely replicated the findings from Study 1, using real faces instead of drawings. However, one should note that in both studies, static faces were used. This limitation is necessary because of the persisting difficulty of manipulating the appearance of moving faces. However, although the addition of vocal information has been found to attenuate certain facial appearance effects, dynamic movement alone does not seem to have the same impact (Zebrowitz-McArthur & Montepare, 1989), suggesting that this restriction in ecological validity has only minor consequences for the generalizability of the results.

## General Discussion

The goal of this series of studies was to assess the influence of gender and perceived dominance on social judgments involving emotions. Specifically, we wanted to assess whether beliefs about men's and women's emotionality are related to perceived dominance and affiliation and whether such perceptions influence how emotion displays shown by men and women are judged.

The results can be summarized as revealing that in the absence of a social context, the gender-stereotypic effect was reversed for anger and happiness expressions. That is, Studies 1 and 2 support the notion that when faces are equated for facial appearance in various ways, the commonly found bias to rate women's smiles as happier and men's anger and disgust display as angrier and more disgusted disappears or even reverses. Specifically, women's anger displays were rated as angrier than men's, whereas men's smiles were rated as happier than women's. As facial appearance-equated expressions control for both dominance and affiliation, it is likely that these two effects can be attributed to the influence of both variables. Women are generally considered to be more likely to smile and less likely to show anger as well as to be more affiliative and less dominant than men. We submit that when perceived dominance and affiliation are controlled for in facial appearance-equated faces, men are perceived as more happy and women as more angry.

This finding may be due to a contrast effect. Given that women are generally expected to smile and to show less anger as well as to be more affiliative and

less dominant, whereas the converse is the case for men, the present stimuli may have violated the observers' expectations. Specifically, it is possible that naturally occurring gender differences in facial appearance might interact with the appearance changes produced by anger and happiness expressions. That is, anger expressions emphasize some of the features that make a face appear dominant (e.g., the mouth region often appears more square, and frowning reduces the distance between eyebrows and eyes). Conversely, smiling enhances the appearance of roundness of the faces associated with affiliative and baby faces. Because of the construction of the present stimuli, the expressive cues for anger and happiness were not "compensated for" by gender typical appearance (as both drawings and real faces were chosen to be androgynous). In some ways, one could say that by depriving the inside of the face of clear gender cues, we actually amplified the expressive cues to anger in women and happiness in men, which are normally "obscured" by the gender-typical facial appearance.

The results reported here are obviously relevant to the shifting standards model (e.g., Biernat & Manis, 1994). This model proposes that social judgments are made according to standards or expectations, which may shift for members of different groups as a function of group stereotypes. Biernat (1995) noted, as an example relevant to the present line of inquiry, that men are generally expected to be more aggressive than are women. Hence, a given aggressive behavior is perceived as more aggressive when shown by a woman than when shown by a man.

From our perspective, this effect would need to be qualified by the facial appearance of the specific individuals engaged in the aggressive behavior. That facial appearance can create such shifting standards is suggested by the observation that more baby-faced or more attractive delinquents experience more adverse effects on their family relationships, whereas more baby-faced or more attractive nondelinquents experience more favorable effects, a finding that the authors attribute to the counterexpectedness of delinquent behavior in baby-faced individuals, making this behavior therefore appear worse (Zebrowitz & Lee, 1999). Specifically, the present results would suggest that were a given aggressive behavior paired with a woman's or a man's face showing characteristic levels of dominance and affiliation cues present in the two genders, the aggressive behavior would be seen as less aggressive in the case of women because typical female facial cues would attenuate perceptions of the intensity of the anger associated with the aggressive

behavior. In the absence of facial cue information, the perceptions of the behavior may well be driven primarily by the different stereotypes for men and women.

Although the experiment necessary to support this line of reasoning has not been done, the results of the present studies can be seen as consistent with the shifting standards model (Biernat & Manis, 1994). Specifically, when observers judge "normal" male and female faces, as was the case in Hess et al. (1997), then women's anger expressions are rated as less angry and men's smiles as less happy. We submit that this is because the actual signal perceived by the observer has been attenuated for women's anger and men's happiness. Thus, even though the morphological features of the expressions (in the case of Hess et al., 1997, the carefully controlled pictures of the JACFEE; Matsumoto & Ekman, 1988) were perfectly matched for physical intensity, as measured by an objective criterion such as the facial action coding system (Ekman & Friesen, 1978), the perceived intensities were nonetheless not the same.

In contrast, when faces are equated for physical dominance and affiliation cues, the differing standards of judgment for the emotions of women and men result in a specific anger expression being seen as more intense in the case of a woman than a man, whereas for a happiness expression, just the reverse occurs. Put another way, it is only when faces have been equated for morphological cues associated with dominance and affiliation that stereotypes, and their associated judgment standards, come into play as the primary variable determining emotional attributions.

However, it may alternatively be argued that our findings can be explained by differences in perceived masculinity. Specifically, the androgynous faces we have used may have created the impression of relatively more masculine women and relatively more feminine men. Hence, participants may have rated the masculine women as showing more masculine emotions and the feminine men as showing more feminine emotions. Although this remains plausible within the framework of the present study, there is evidence from other research that undermines this argument. First, Adams, Hess, and Kleck (2004), using the same manipulation as in Study 2, but with neutral faces, asked participants how likely they felt the stimulus person was to show a variety of emotions. In this study, apparent women were found to be more likely to show anger—in line with our observation—but also more likely to show the typical feminine emotions of sadness and fear. Conversely, men were rated as more

likely to show happiness. Similarly, in Study 1 reported here, we found no gender differences for disgust, a more typical masculine emotion. Together, these findings are not congruent with the notion that the masculinity–femininity dimension primarily drives our observed effects.

The present research advances the argument that what is generally described as an effect of a gender stereotype, that is, the attribution of anger to men and of happiness to women, is largely a function of facial appearance. We further submit that it is likely that facial cues associated with perceived affiliation and dominance interact with expressive cues to produce perceptions of underlying emotional states that are congruent with stereotypical expectations.

This argument has important implications for the impact of gender-stereotypical expectations on perceptions of men's and women's emotional displays. Specifically, these effects may not be easily counteracted by men's and women's perceived social roles. That is, even if the stereotypical role expectations were to change, women's and men's anger and happiness would still be perceived differently as a function of the facial cues to dominance and affiliation that are strongly confounded with gender. This line of thought is consistent with findings showing that the expression of anger in a leadership situation is interpreted as a sign of competence for men and helps them reassert their dominance status (Lewis, 2000), but not in the case of women. That is, regardless of social status, a woman's angry frown will threaten less convincingly without a matching dominant appearance. By the same token, a man's happy smile will shine less brilliantly without a matching affiliative appearance. And yet, by virtue of sex category membership alone, at equal levels of perceived dominance and affiliation, women appear to require a less dominant appearance to give their frowns that extra bang and men a less affiliative appearance to give their smiles that extra shine.

## References

Adams, R. B., Jr., Hess, U., & Kleck, R. E. (2004). *Emotion in the neutral face: The influence of gender-emotion stereotypes and gender-related facial appearance*. Manuscript submitted for publication.

Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right? Personality correlates of facial babyishness. *Personality and Social Psychology Bulletin, 15,* 266–279.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin, 100,* 3–10.

Biernat, M. (1995). The shifting standards model: Implications of stereotype accuracy for social judgment. In L. Yueh-Ting, L. J. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 82–114). Washington, DC: American Psychological Association.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66,* 5–20.

Brody, L. R., & Hall, J. A. (2000). Gender, emotion, and expression. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (2nd ed., pp. 447–460). New York: Guilford Press.

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology, 46,* 991–1004.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46,* 735–754.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44,* 227–240.

Fischer, A. H. (1993). Sex differences in emotionality: Fact or stereotype? *Feminism & Psychology, 3,* 303–318.

Friedman, H., & Zebrowitz, L. A. (1992). The contribution of typical sex differences in facial maturity to sex role stereotypes. *Personality and Social Psychology Bulletin, 18,* 430–438.

Gaelick, L., Bodenhausen, G. V., & Wyer, R. S. (1985). Emotional communication in close relationships. *Journal of Personality and Social Psychology, 49,* 1246–1265.

Hall, J. A., & Halberstadt, A. G. (1994). "Subordination" and sensitivity to nonverbal cues: A study of married working women. *Sex Roles, 31,* 149–165.

Haugh, S. S., Hoffman, C. D., & Cowan, G. (1980). The eye of the very young beholder: Sex typing of infants by young children. *Child Development, 51,* 598–600.

Henley, N. M. (1977). *Body politics: Power, sex and nonverbal communication*. New York: Prentice Hall.

Henley, N. M. (1995). Body politics revisited: What do we know today? In P. J. Kalbfleisch & M. J. Cody (Eds.),

*Gender, power, and communication in human relationships* (pp. 27–61). Hillsdale, NJ: Erlbaum.

Hess, U., Adams, R. B., Jr., & Kleck, R. E. (in press). Smiling women–frowning men: Dominance and the display of emotions. *Cognition & Emotion.*

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior, 21,* 241–257.

Keating, C. F. (1985). Human dominance signals: The primate in us. In S. L. Ellyson & J. F. Dovidio (Eds.), *Power, dominance, and nonverbal communication* (pp. 89–108). New York: Springer-Verlag.

Keating, C. F., Mazur, A., & Segall, M. H. (1981). A cross-cultural exploration of physiognomic traits of dominance and happiness. *Ethology and Sociobiology, 2,* 41–48.

LaFrance, M., & Hecht, M. A. (1999). Option or obligation to smile: The effects of power and gender on facial expression. In P. Philippot, R. S. Feldman, & E. J. Coats (Eds.), *The social context of nonverbal behavior* (pp. 45–70). Cambridge, England: Cambridge University Press.

LaFrance, M., & Hecht, M. A. (2000). Gender and smiling: A meta-analysis. In A. H. Fischer (Ed.), *Gender and emotion: Social psychological perspectives: Studies in emotion and social interaction. Second series.* (pp. 118–142). Cambridge, England: Cambridge University Press.

LaFrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin, 129,* 305–334.

Lewis, K. M. (2000). When leaders display emotion: How followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior, 21,* 221–234.

Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuf)* [Slides & brochure]. San Francisco: San Francisco State University.

Preuschoft, S., & van Hooff, J. A. R. A. M. (1997). The social function of "smile" and "laughter": Variations across primate species and societies. In U. Segerståle & P. Molnár (Eds.), *Nonverbal communication: Where nature meets culture* (pp. 171–190). Mahwah, NJ: Erlbaum.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Senior, C., Phillips, M. L., Barnes, J., & David, A. S. (1999). An investigation into the perception of dominance from schematic faces: A study using the World-Wide Web. *Behavior Research Methods, Instruments, & Computers, 31,* 341–346.

Shields, S. A. (2000). Thinking about gender, thinking about theory: Gender and emotional experience. In A. H. Fischer (Ed.), *Gender and emotion: Social psychological perspectives* (pp. 3–23). Cambridge, England: Cambridge University Press.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition, 63,* 271–313.

Zebrowitz, L. A., & Lee, S. Y. (1999). Appearance, stereotype-incongruent behavior, and social relationships. *Personality and Social Psychology Bulletin, 25,* 569–584.

Zebrowitz-McArthur, L., & Montepare, J. M. (1989). Contributions of a baby face and a childlike voice to impressions of moving and talking faces. *Journal of Nonverbal Behavior, 13,* 189–203.