



ELSEVIER

Evolution and Human Behavior 27 (2006) 401–416

---

---

Evolution  
and Human  
Behavior

---

---

## A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis

Elizabeth Hampson\*, Sari M. van Anders, Lucy I. Mullin

*Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2*

Initial receipt 13 June 2005; final revision received 8 May 2006

---

### Abstract

A set of computerized tasks was used to investigate sex differences in the speed and accuracy of emotion recognition in 62 men and women of reproductive age. Evolutionary theories have posited that female superiority in the perception of emotion might arise from women's near-universal responsibility for child-rearing. Two variants of the child-rearing hypothesis predict either across-the-board female superiority in the discrimination of emotional expressions ("attachment promotion" hypothesis) or a female superiority that is restricted to expressions of negative emotion ("fitness threat" hypothesis). Therefore, we sought to evaluate whether the expression of the sex difference is influenced by the valence of the emotional signal (Positive or Negative). The results showed that women were faster than men at recognizing both positive and negative emotions from facial cues, supporting the attachment promotion hypothesis. Support for the fitness threat hypothesis also was found, in that the sex difference was accentuated for negative emotions. There was no evidence that the female superiority was learned through previous childcare experience or that it was derived from a sex difference in simple perceptual speed. The results suggest that evolved mechanisms, not domain-general learning, underlie the sex difference in recognition of facial emotions.

© 2006 Elsevier Inc. All rights reserved.

*Keywords:* Sex difference; Gender difference; Facial expression; Emotion; Valence; Child-rearing

---

---

\* Corresponding author. Department of Psychology, University of Western Ontario, London, ON, Canada N6A 5C2. Tel.: +1 519 661 2111x84675; fax: +1 519 661 3213.

*E-mail address:* ehampson@uwo.ca (E. Hampson).

## 1. Introduction

The ability to decode facial expressions of emotion is fundamental to human social interaction. Elements of facial decoding, including the immediate preverbal detection of a facial signal, are believed to represent evolved mechanisms that enable the receiver to predict another individual's emotional state and anticipate future actions (Ekman, 1997; Izard, 1994; cf. Russell, Bachorowski, & Fernández-Dols, 2003). Ekman and others (Ekman, 1994; Ekman and Friesen, 1971; Izard, 1994) have argued that a limited set of facial expressions is innate and universally recognized as signals for happiness, sadness, anger, fear, disgust, and surprise. While the verbal labels and cultural rules governing the expression of these emotions may vary, the expressions themselves have a universal signal value. Thus, both the production of specific facial expressions and their interpretation by a receiver are thought to be innate.

It is often claimed that women are superior to men at recognizing facial expressions of emotion (see below). Explanations for the sex difference range from sexual inequalities in power and social status (e.g., see Hall, 1984; Henley, 1977; Weitz, 1974) to evolutionary perspectives based on women's near-universal responsibility for child-rearing (e.g., Babchuk, Hames, & Thompson, 1985). The primary caretaker hypothesis proposed by Babchuk et al. (1985) contends that females, as a result of their evolutionary role as primary caretakers, will display evolved adaptations that enhance the probability of survival of their offspring. In humans, these adaptations are hypothesized to include the fast and accurate decoding of facial affect, an important means of communication especially in preverbal infants.

The child-rearing hypothesis is more complex than it first appears. It gives rise to two different predictions. According to one interpretation of the theory, the "attachment promotion" hypothesis, women should display across-the-board superiority, relative to men, in decoding all facial expressions of emotion because mothers who are highly responsive to infants' cries, smiles, and other nonverbal signals are likely to produce securely attached infants (Ainsworth, 1979; Hall, Lamb, & Perlmutter, 1986), and secure infants display optimal long-term health and immune function and social outcomes (Goldberg, 2000). A second interpretation of the theory, the "fitness threat" hypothesis, assigns a special status to negative emotions. It predicts a female superiority that is limited to expressions of negative emotion including fear, disgust, sadness, and anger.<sup>1</sup> Because negative emotions signal a potential threat to infant survival (e.g., threats to safety, loss, pain, or the ingestion of a toxin) that calls for action on the caretaker's part—whereas positive expressions carry no such imperative—it is specifically facility in the recognition of negative expressions that may have been selected in the primary caretaker and in which a female superiority may therefore be found. By tying the sex difference to parental roles, the fitness threat hypothesis offers an alternative to theories based on

---

<sup>1</sup>According to Babchuk et al. (1985), anger calls for a response from the mother because it signifies frustration on the part of the infant, a form of distress that may signal a survival issue.

individual survival, which predict either no sex difference in the ability to discriminate threat or a female advantage limited to single emotions (e.g., anger, where a sex difference would be adaptive in allowing physically weaker females to preemptively avoid acts of physical aggression, usually initiated by males; Goodall, 1986; Konner, 1982). Although both sexes have a stake in infant survival, the ability to swiftly and accurately identify potential threat is a basic adaptation to the role of primary caretaker and would be maximized in the sex having the largest investment in each offspring. Finding a female advantage that is selective to negative emotions would constitute support for the fitness threat hypothesis.

Evidence of a female superiority in identification of facial expressions is mixed. Of 55 studies reviewed by Hall (1978), only 11 (20%) found a significant superiority for females in judging emotions based on visual cues alone (conveyed by the face and/or body). Studies using the Profile of Nonverbal Sensitivity have yielded a median effect size of  $r=.15$  in favor of women when only facial cues were available for decoding (Hall, 1984). A meta-analysis by McClure (2000) found a smaller but statistically significant female advantage among children and adolescents. These effect sizes conceal substantial variability across studies in the size and even the direction of the sex difference. Obtained differences ranged from  $d=1.86$  to  $d=-0.60$  in the 55 studies reviewed by Hall. Inconsistency is to be expected if the female advantage does not encompass all facial expressions of emotion since most studies do not assess the full range. On the other hand, failures to find a sex difference could simply reflect methodological factors. Many studies have used face exposure times in the 10- to 15-s range or up to 1 min. This lengthy time allowance lacks ecological validity since facial expressions are often fleeting and since accuracy of decoding depends on the speed with which an expression can be apprehended.

Female superiority in perceptual speed, the ability to rapidly absorb the details of a visual stimulus, has been recognized since the 1940s (e.g., Harshman, Hampson, & Berenbaum, 1983; Kim & Petrakis, 1998; Tyler, 1965; Wesman, 1949) and generalizes to many types of visual stimuli. Since facial decoding involves, under natural conditions, speeded apprehension of visual detail, it is important to rule out the possibility that any female advantage is based on nothing more than a perceptual speed advantage. In that case, evolutionary explanations based on child-rearing would be inappropriate. Previous work has not included methodological controls to rule out this possibility.

The present study was designed to test whether a female advantage in the discrimination of emotional expressions can be verified among young adults of reproductive age. Differences in accuracy and response times (RTs) were evaluated. Secondly, we wished to investigate whether any advantage applies equally to all emotions regardless of hedonic valence, as predicted by the attachment promotion hypothesis, or is differentially found among the negatively valenced emotions, as predicted by the fitness threat hypothesis.<sup>2</sup>

---

<sup>2</sup>It should be noted that different emotions are not equated in perceptual difficulty (e.g., smiling makes happiness easier to identify than all other emotions). Therefore, the hypothesis does not predict the ease of identification of one emotion relative to another.

## 2. Method

### 2.1. Participants

Sixty-two undergraduate students (31 women, 31 men) participated in the study. The mean age was  $20.77 \pm 3.67$  years. All but four participants were right-handed, according to a standard handedness inventory (Kimura, 1983). All participants had normal visual acuity.

### 2.2. Procedure

The participants were tested individually by a trained examiner. The session began with a demographics questionnaire followed by assessment of visual acuity in each eye using a standard Snellen chart. Participants then completed the control tasks and proceeded to a set of computer-administered face discrimination tasks.

### 2.3. Experimental tasks

Stimuli were presented using a Pentium III 450-MHz computer equipped with 256 MB of RAM. Stimulus presentation and the recording of RTs were controlled by a program created in MATLAB 5.3 (The MathWorks Inc., Natick, MA). Participants were seated 45–60 cm in front of the computer screen. All images were presented centrally and were of fixed size ( $15 \times 10.5$  cm). The faces were gray-scale digitized photographs taken from the Pictures of Facial Affect (Ekman & Friesen, 1976). Permuted faces were used in the Pattern Matching task to provide stimuli that were not recognizable as faces but which preserved the contour information available in the original images. The participant sat with his or her finger on the space bar, which was used to control the initiation and termination of each trial.

Trials were self-paced. Each trial was initiated by a key-press. The stimulus image appeared 1 s later. Timing of the RT began at image onset and was terminated by a key-press to yield the RT (in milliseconds) on each trial. Participants were instructed to press the bar as quickly as possible while still being accurate. RTs were recorded automatically by the computer. A visual mask was displayed for 20 ms at image offset to terminate any further visual processing (Breitneyer, 1984). To verify that participants had made the correct discrimination on each trial, we required them to point to the correct choice on a laminated response card placed on an upright stand immediately adjacent to the monitor and keyboard. Participants' choices were manually recorded by the examiner.

The response card for the Facial Emotion condition is shown in Fig. 1. For each of the four conditions, the response card consisted of six images ( $6.75 \times 4.5$  cm) arranged in two columns with three images each. Before each condition, the participant was shown the response card and asked to familiarize himself or herself with the images. The participant gave a signal when he or she was ready to begin.

The following four conditions were administered. In each condition, equal numbers of male and female faces were used. Order of the trials within each condition was randomized.

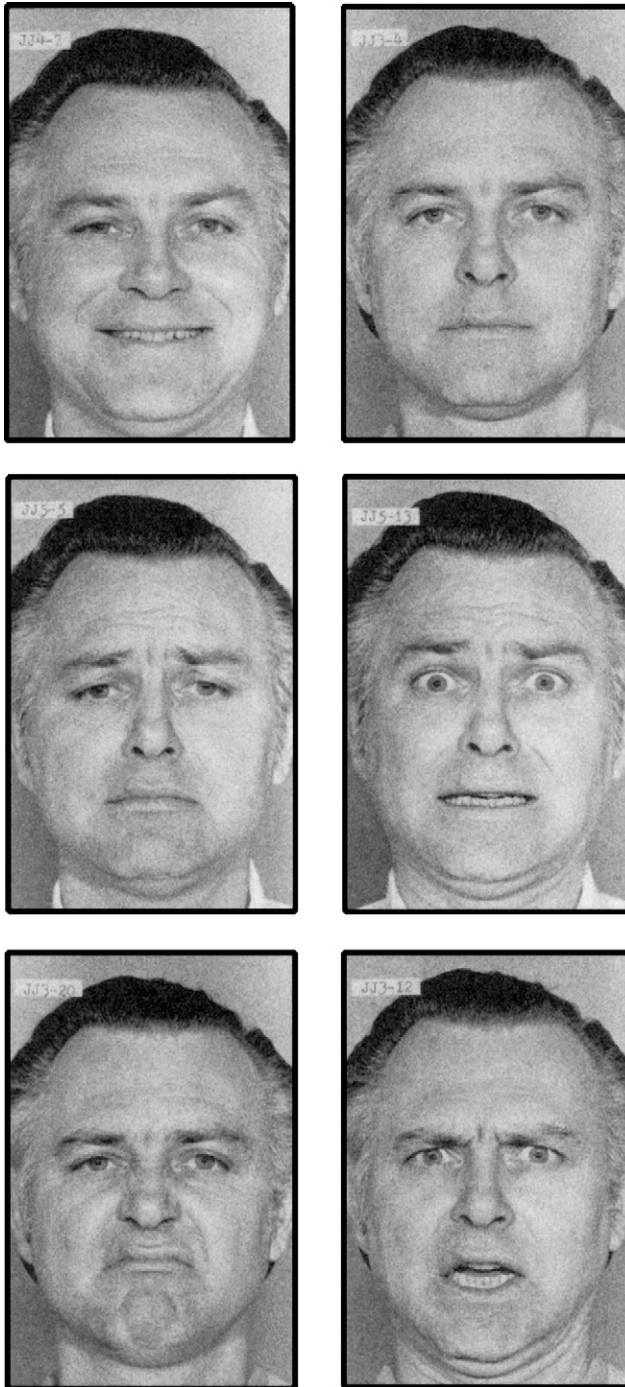


Fig. 1. Example of response card layout (Facial Emotion condition).

The Facial Matching condition was considered a practice condition and was always administered first. The order of the Facial Identity and Facial Emotion conditions was counterbalanced across participants. Before each condition, participants were reminded that their speed and accuracy were being recorded.

### 2.3.1. *Facial matching*

This condition required only a simple perceptual match. Six different faces were presented four times each for a total of 24 trials. All the faces had a neutral expression. On each trial, participants pressed the space bar immediately upon recognizing which face had appeared and then pointed to the identical face on the response card.

### 2.3.2. *Facial identity*

In this condition, participants had to discriminate facial identity. The faces of six different individuals were each presented five times, with a different emotional expression each time, for a total of 30 trials. The purpose of using expressive faces was to ensure that the face stimuli were matched in all respects in this condition and in the Facial Emotion condition. On each trial, the participant pressed the space bar immediately upon recognizing the individual presented, terminating the image and initiating the mask. The participant then pointed to the face on the response card that matched the identity (although not the emotional expression) of the face presented. All faces on the response card had a neutral expression. Thus, a direct perceptual match was not possible. Participants had to decode individual identities in order to make a correct response.

### 2.3.3. *Facial emotion*

In this condition, participants had to discriminate emotional expressions. Each face displayed one of six emotions (happiness, sadness, fear, anger, disgust, or a neutral expression). The neutral faces were deliberately selected to portray contentment or positive satisfaction but without overt markers of positive affect such as smiling or crinkling of the eyes. There were 60 trials in total. Each image appeared only once in the series. On each trial, the face on the computer screen was of a different identity but wore the same facial expression as one of six faces on the response card. The response card showed the six emotions displayed by a single individual who did not appear in any of the 60 trials. On each trial, the participant pressed the space bar immediately upon recognizing the emotion, then pointed to the face on the response card with the matching expression. As in the Facial Identity condition, a direct perceptual match was not possible. A correct response could only be achieved by recognizing the emotional expression, independent of identity.

### 2.3.4. *Pattern matching*

This condition required a direct perceptual match, but nonface stimuli were used. One of the Ekman images was permuted to render the face unrecognizable, and two small black rectangles were superimposed on the image. On each trial, the rectangles appeared at one of six locations on the permuted image. Participants had to point to the identical stimulus on the response card. There were 24 presentations in total.

The dependent variables for all tasks were RT and accuracy. Accuracy was scored as the percentage of correctly identified images. RTs were scored as the mean RT in milliseconds for all correct trials within each condition. Any trial with an RT greater than 3 S.D. from the mean was excluded before calculating the mean RT for each condition. In the Facial Emotion condition, separate accuracy and RT scores were computed for each of the six emotions, as well as two composites that were of theoretical interest. One composite (Positive) represented the mean RT for the conditions in which a positive emotion was presented (happy, neutral). The other composite (Negative) represented the mean RT for all conditions in which a negative emotion was presented (fear, sadness, disgust, and anger).

In addition to the four conditions described above, a brief Face Labeling task was administered. This task was always given after the Facial Emotion task had already been completed. Participants were asked to verbally provide a name for the emotion depicted on each of the six faces on the Facial Emotion response card. Responses were recorded verbatim. This was done to assess whether the participants accurately recognized which emotions the faces portrayed. Accuracy was scored as the number of correct labels, out of six, provided by the participant. Labels were considered correct if they were synonymous with the name of the emotion depicted (e.g., “mad” was considered equivalent to “angry”). Judgments of the acceptability of the labels were highly reliable across two independent judges ( $r=.93$ ,  $p<.001$ ).

## 2.4. Control tasks

### 2.4.1. Verbal Meaning Test (Thurstone & Thurstone, 1963)

This test was administered as an index of general level of ability to assure that there were no chance differences between the two groups that might affect perceptual scores. Four minutes were allowed to complete 60 items. On each item, the participant had to choose the word from a list of five alternatives that best matched the meaning of a target word. The score was the number of items correct.

### 2.4.2. Identical Pictures Test (Ekstrom, French, Harman, & Dermen, 1976)

This test is a conventional paper-and-pencil measure of perceptual speed and accuracy. On each item, the participant had to choose which of the five alternatives was identical in all respects to a designated target figure. The figures consisted of simple geometric or abstract line drawings. Three minutes were allowed to complete 96 items. The score was the number correct, corrected for guessing.

### 2.4.3. Demographic questionnaire

Participants completed a brief demographic questionnaire inquiring about age, sex, and use of corrective lenses. Questions about two types of experiences that might confer an advantage in the recognition of facial emotion were included. Participants were asked to indicate if they had experience in drama or theater, and if so, to indicate the number of years of involvement. Experience in taking care of young children was self-rated on a 6-point scale ranging from 0

(none) to 5 (*almost everyday*). Separate ratings were obtained for experience in five contexts (e.g., babysitting) and summed to yield a total score (max=25).

### 3. Results

Analysis of the scores on the Verbal Meaning Test revealed that the men ( $M=26.77$ ,  $S.D.=7.54$ ) and women ( $M=25.61$ ,  $S.D.=10.61$ ) were well matched in general level of ability,  $t(60)=0.50$ ,  $p=.621$ . Data from the experimental conditions were evaluated to test whether a female advantage was present in the discrimination of emotional expressions and whether any advantage was differentially seen among the negatively valenced emotions, as predicted by the fitness threat hypothesis.

#### 3.1. Accuracy

Level of accuracy in the practice condition (Facial Matching) was extremely high, demonstrating excellent acquisition of the basic stimulus presentation and response procedure in both sexes. The mean percent correct was 98.25% ( $S.D.=2.59$ ) for men and 98.65% ( $S.D.=2.50$ ) for women. Similar high levels of accuracy were seen in the Facial Identity ( $M=97.20$ ,  $S.D.=3.66$  and  $M=96.34$ ,  $S.D.=3.59$ ) and Pattern Matching conditions ( $M=99.46$ ,  $S.D.=1.78$  and  $M=98.79$ ,  $S.D.=2.45$ ) for men and women, respectively.

Likewise, accuracy of identification in the Facial Emotion condition was very high, with scores of ~90% or above for all emotions except for disgust ( $M=84.67$ ,  $S.D.=14.32$  and  $M=90.00$ ,  $S.D.=7.88$ ) and anger ( $M=50.65$ ,  $S.D.=27.20$  and  $M=60.97$ ,  $S.D.=25.61$ ). Only for the latter two emotions, where accuracy failed to reach ceiling, was there any indication whatsoever of a sex difference:  $t(45)=1.79$ ,  $p=.081$  for disgust and  $t(60)=1.54$ ,  $p=.129$  for anger (two-tailed a priori test). Because the scores for the other emotions were at or near ceiling values, sex differences could not be analyzed meaningfully. Therefore, all further statistical analysis focused on the RT data.

#### 3.2. Response times

Means and standard errors for each of the six emotions are shown in Fig. 2. To evaluate if a sex difference was present in the discrimination of emotional expressions, we entered the RTs for the six emotions and three control conditions into a two-way mixed effects analysis of variance (ANOVA), with sex as the between-subjects factor and condition as a within-subjects factor. Three participants (two females, one male) who had average RTs greater than 3 S.D. from the group mean were omitted from the analysis, resulting in a sample of 59. The results showed a significant main effect of sex,  $F(1, 57)=7.19$ ,  $p=.010$ , and a significant interaction between sex and condition,  $F(5, 272)=5.68$ ,  $p<.001$ . There was also a main effect of condition,  $F(5, 272)=84.08$ ,  $p<.001$ , reflecting the fact that RTs were faster in the control conditions (Facial Matching, Pattern Matching, and Facial Identity) than in the emotion conditions (all  $p$  values  $<.001$ ). The control conditions included the Facial Identity task, in



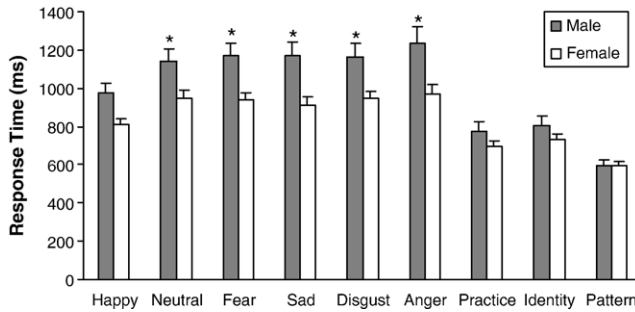


Fig. 2. Mean RTs for men and women in the six emotion conditions and three control conditions. Bars represent standard errors of the means. Asterisks indicate a significant sex difference ( $p < .05$  or less).

which individual identities had to be decoded to make a correct match. Pairwise comparisons showed that of the six emotions, happy faces elicited significantly shorter RTs than all other expressions ( $p$  values  $< .001$ ), while angry faces elicited longer RTs than all others ( $p$  values  $< .025$ ), except fear ( $p = .070$ ) or neutral expressions ( $p = .060$ ).

Tukey tests were used to decompose the significant interaction effect. The sex difference was not significant in any of the control conditions (all  $p$  values  $> .05$ ). There was no sex difference in identification of happy faces ( $p < .10$ ), but women were significantly faster than men at discriminating neutral faces ( $p < .05$ ), as well as faces depicting disgust ( $p < .025$ ), fear ( $p < .025$ ), sadness ( $p < .01$ ), and anger ( $p < .01$ ). Thus, a sex difference in favor of women appeared selectively in the emotion conditions and not in conditions requiring other visual discriminations.

The selectivity of the effect suggested that a female advantage in simple perceptual speed was not the basis for the sex difference. Nevertheless, associations between RTs and performance on a conventional test of perceptual speed, the Identical Pictures Test, were evaluated. As expected, women ( $M = 69.98$ ,  $S.D. = 13.50$ ) tended to show faster performance on the perceptual speed test than men ( $M = 65.79$ ,  $S.D. = 14.17$ ), although this was not significant,  $t(60) = 1.19$ ,  $p = .119$  (one tailed). Scores on the test were correlated significantly with the RTs in five of the six emotion conditions. Therefore, the ANOVA was repeated using the Identical Pictures score as a covariate and, also, in a separate analysis, using the RT in the Pattern Matching condition as a covariate. The Pattern Matching condition was expressly devised to control for perceptual speed while closely matching the emotion conditions in all other stimulus presentation and response characteristics. It provided a direct estimate of visual decoding and RT in the absence of facial stimuli. Therefore, of the two control tasks, it was considered the superior covariate. Correlations between Pattern Matching and the emotion conditions ranged from  $r = .51$  to  $.65$ . With perceptual speed controlled, the main effect of sex was still highly significant,  $F(1, 56) = 6.08$ ,  $p = .017$ , either when Identical Pictures was used as the covariate or when Pattern Matching was used as the covariate,  $F(1, 56) = 14.79$ ,  $p < .001$ . The significant interaction between sex and condition was also preserved,  $F(5, 271) = 5.21$ ,  $p < .001$  and  $F(6, 316) = 4.18$ ,  $p = .001$ , for the two covariates, respectively. Thus, the female superiority in discriminating emotional expressions was retained when perceptual speed was explicitly controlled.

The fitness threat hypothesis predicted that the female superiority would be larger for negative emotions than for positive ones. Inspection of the means in the six emotion conditions revealed that the sex difference was indeed larger for each of the four negative emotions than for either of the two positive emotions. To investigate this formally, we performed a two-way mixed effects ANOVA on the Negative and Positive composite scores that represented the mean RT for each participant averaged across the four negative and two positive conditions, respectively. Pattern Matching was used as a covariate to remove variance associated with nonemotive parts of the task. Sex was the between-subjects factor and valence (Positive or Negative) was a within-subjects factor. The results revealed a significant main effect of sex,  $F(1, 59)=6.88, p=.011$ . In both categories of emotion, women showed consistently shorter RTs than men. Importantly, the interaction between sex and valence was also significant,  $F(1, 59)=4.00, p=.050$ . The sex difference was larger for negative emotions than for positive ones, as predicted by the fitness threat hypothesis (Fig. 3). Although several of the negative emotions were harder to identify than the positive ones, and thus elicited slower RTs (see above), women showed a processing advantage relative to men in decoding the negative emotions. To investigate if the effect was robust, we computed a ratio score using the two composites for each person:  $(\text{Negative RT} - \text{Positive RT}) / \text{Pattern Matching RT}$ , a multiplicative instead of additive adjustment for nonemotive factors. The results were essentially identical. A  $t$  test on the resulting scores showed that, on average, women could identify negative emotions nearly as adeptly as positive ones, with only a 9% change in RT, while men showed a 27% increase in processing time for the negative emotions,  $t(60)=2.01, p=.049$  (two tailed; Fig. 3).

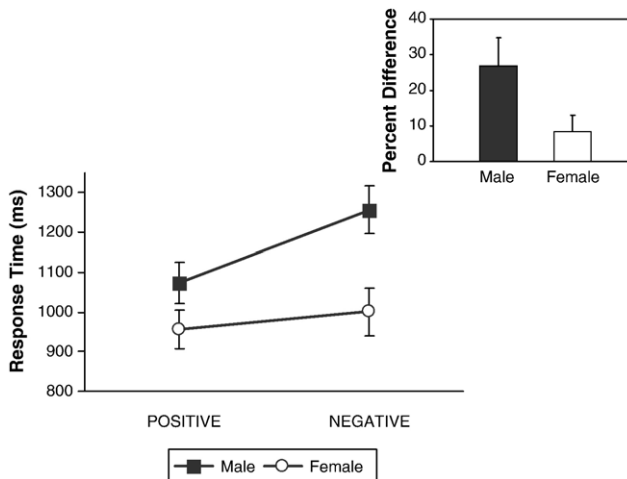


Fig. 3. Analysis of the positive and negative composite scores revealed that the sex difference in RT was larger for negative emotions than for positive emotions. Inset: Men showed a nearly 30% increase in processing time for negative emotions over positive ones, while women showed only a 9% increase. Pattern Matching was used to adjust for individual differences in basal response speed (see text).

### 3.3. Other control tasks

Men and women were equally accurate at generating verbal labels on the Facial Labeling task,  $t(60)=1.34$ ,  $p=.187$ . The mean for men was 4.68 correct (S.D.=1.05) and the mean for women was 5.03 correct (S.D.=1.05), out of a possible value of 6. Both sexes were able to capture the emotions portrayed with a high level of verbal accuracy.

Experiences with children or the theater were analyzed using  $t$  tests. The purpose of the analyses was to discover whether any experiential differences existed between the two sexes, which might confer an advantage in the recognition of emotional expressions. Sixteen of the 62 participants (26%) did report drama or theater experience and 75% of these were women. However, Pearson correlations revealed that theater experience did not correlate significantly with RT on any of the emotion tasks ( $-.12 < r < .20$ ). Scores on the childcare variable ranged from 0 to 20 (out of 25). Women reported more childcare experience ( $M=8.42$ , S.D.=4.15) than men ( $M=5.45$ , S.D.=3.13),  $t(60)=3.18$ ,  $p=.002$ . However, there was no evidence that greater experience with children was associated with better recognition times, either for the six emotions individually or for the two valence composites (men:  $r=-.15$  to  $-.02$ ; women:  $r=-.03$  to  $.10$ , n.s.).

## 4. Discussion

We evaluated sex differences in the speed and accuracy of facial decoding in a group of healthy men and women of reproductive age. Six emotions considered to have universal signal value (Ekman & Friesen, 1976) and three control conditions that did not require the discrimination of emotional signals served as stimuli. Consistent with the attachment promotion hypothesis, we found faster identification by women than men when emotions but not other types of visual stimuli were presented. We also found that hedonic valence was a moderator of the sex difference. The female advantage was most prominent in the rapid discrimination of negatively valenced emotions.

Accuracy of performance was close to ceiling on nearly all the experimental tasks. Consequently, our data analysis focused on the RT scores as the primary dependent variable. The high level of accuracy is not surprising considering the tasks were self-paced and shows that participants cooperated well with the instructions and sustained satisfactory motivation. The lack of sex differences is consistent with many previous studies that have failed to detect any major sex difference in accuracy when unlimited exposure times are allowed or with exposures as long as 10–15 s (e.g., Duhaney & McKelvie, 1993; Kirouac & Doré, 1985; Wagner, MacDonald, & Manstead, 1986; but see also Brunori, Ladavas, & Ricci-Bitti, 1979; Zuckerman, Hall, DeFrank, & Rosenthal, 1976). In the present data, scores approached ceiling for most emotions after less than 1.5 s of viewing time. It was only where the scores departed substantially from ceiling (angry or disgusted faces) that any trace of a sex difference was apparent. Lower accuracy for these emotions is in accord with previous work showing that anger is less reliably recognized than other emotions (Ekman & Friesen, 1976; McAndrew, 1986; Rotter & Rotter, 1988).

Analysis of the RT data showed that identification was faster in the control conditions than when an emotion had to be decoded. This included faster performance in the Facial Identity condition where emotional faces served as the stimuli, but the task was to decode the identity, not the emotion, of the face. RTs in this condition were shorter than in any of the emotion conditions. Although the faces were matched for color of the hair and eyes and for the side where the hair was parted, it is conceivable that the participants matched on the basis of some distinctive feature instead of the broad configural properties of the face and that this explains the relatively short RTs. More importantly, we did not find a sex difference in the Facial Identity condition. Studies of episodic memory have found that women outperform men in facial recognition tasks when recognition of identity is tested after a period of minutes to hours (Herlitz & Yonker, 2002; Lewin & Herlitz, 2002). These studies, however, measured retrieval from memory stores, not the immediate perceptual decoding of identity information as sampled in the present study.

Although none of the control conditions showed a sex difference in RT (Facial Matching, Facial Identity, and Pattern Matching), a significant sex difference was observed in the emotion conditions. Five of the six emotions elicited a significant female advantage. Importantly, the female superiority was not seen for visual processing in general or even for processing of faces but was seen only when discrimination of emotion from facial cues was required. An important outcome of the present work was the demonstration that the female advantage was not due to a sex difference in perceptual speed. In fact, the advantage was retained and even strengthened with perceptual speed controlled. A female advantage in perceptual speed has been reported frequently but not invariably in many studies since the 1940s (Harshman, et al., 1983; Schaie & Willis, 1993) and occurs on many types of tasks that involve rapid apprehension of visual detail. The moderately large correlations found in the present study between the measures of perceptual speed and the RT scores suggest that perceptual speed was an important constituent of the RT. However, the selectivity of the sex difference and the failure of statistical controls to eliminate it indicated that the sex difference in emotion processing did not reside in the perceptual speed component. The fact that a female advantage in discriminating facial emotion was found to exist apart from any female advantage in perceptual speed supports the view of facial decoding as a unique evolved capacity.

We found no evidence that the female advantage was linked to the amount of previous childcare experience. This is contrary to explanations based on lifetime interaction with children and experience in decoding their expressions. Likewise, Babchuk et al. (1985) found that experience with children, defined as currently having one or more children under age 5 versus never having had child caretaking responsibilities, had no effect on the sex difference in emotion recognition. While there are limitations to the use of self-report, the lack of any association with childcare experience in either our study or Babchuk's suggests that the sex difference is unlikely to be due to learning factors associated with care of young children.

Two competing variants of the primary caretaker hypothesis proposed by Babchuk et al. (1985) were tested in the present study: the attachment promotion hypothesis and the fitness threat hypothesis. Aspects of the present data were consistent with each interpretation, but neither was wholly supported by our data. We found a robust main effect of sex that

transcended the valence of the emotions portrayed. This effect was evident in the analysis of the emotions individually and in the analysis of the positive and negative composites. Females exhibited significantly shorter RTs than males in identifying both positive and negative emotions. This supports the attachment promotion hypothesis, which suggests that women evolved innately greater facility in identifying emotions than men due to their importance in long-term parental bonding processes and due to the female role as primary caretaker (Babchuk et al., 1985). On the other hand, we also found evidence that the hedonic valence of a facial expression was important, as suggested by the fitness threat hypothesis. When the emotions were examined individually, women showed faster RTs than men to each of the four negative emotions. The sex difference was attenuated for happy and, to a lesser extent, neutral faces. (Given the neutral faces were less overtly positive than the happy ones, a milder attenuation effect is not surprising.) The importance of valence was confirmed in the analysis of the positive and negative composites, where a significant interaction between sex and valence was found. The female superiority in RT was larger for negative than positive emotions, as predicted. The interaction may signify that negative emotions have indeed been subject to differential selection pressures.

It should be emphasized, though, that in its pure form, the fitness threat hypothesis would predict a female superiority only in the discrimination of negative emotions. This is not what we observed. This suggests that the fitness threat hypothesis must be considered in addition to, not instead of, a more generalized female processing superiority. A further complication in the present data is that we found longer RTs for negative emotions than positive ones in both sexes, suggesting that negative emotions are generally more difficult to discriminate. This is consistent with many other studies (Ekman & Friesen, 1976; McAndrew, 1986; Mufson & Nowicki, 1991; Ogawa & Suzuki, 1999; Rotter & Rotter, 1988). Perhaps the smile in a happy face obviates the need to attend to finer features of the emotional display. This meant that the RTs for negative emotions were not faster than those for positive emotions, despite their signaling importance in fitness threat. We did find, however, that women identified negative emotions at a much faster latency than did men and showed a proportionately larger advantage for negative than positive emotions. Given the layout of the data, it could be argued that our data show that men are selectively disadvantaged in discriminating negative emotions, not that women have any special facility. Inhibition of emotion processing in males in response to emotional stimuli, especially negative stimuli, was hypothesized by Burton and Levy (1989). This possibility will need to be evaluated by future empirical work, but there is no obvious evolutionary explanation for selective slowness in men. In sum, the interaction with valence suggests that the fitness threat hypothesis deserves further investigation and that the capacity to rapidly discriminate and respond to negative emotions may enhance reproductive success in women.

The possibility that the sex difference does not apply equally to all emotions is plausible from a neurobiological perspective. Perception of facial identity is separate from perception of facial emotion and can be differentially impaired by brain lesions (Tranel, Damasio, & Damasio, 1988). Contemporary imaging studies support a fractionation in the cerebral representation of emotion, with activation of both common and specific sites in response to different emotions (Phan, Wager, Taylor, & Liberzon, 2002). In patients with right

hemisphere lesions, Adolphs, Damasio, Tranel, and Damasio (1996) found that some patients were impaired at recognizing negative emotions, especially fear and sadness, while showing a preserved ability to recognize happy expressions. The amygdala, a structure implicated in fear, contains abundant alpha-type estrogen receptors (Österlund, Keller, & Hurd, 2000), suggesting that it is one estrogen-responsive site in the human brain. The possibility of hormonal modulation might imply that the size of the sex difference in emotion perception is dependent on the hormonal state of the individuals tested.

One caveat concerns the fact that only adult faces were used in the present study. Although purely speculative, it is possible that women learn to pay more attention to facial expressions for cultural reasons that are unrelated to the time spent with children (e.g., more time spent with others in close proximity) or for evolutionary reasons (e.g., avoiding male violence). It is unlikely that male violence is involved in the sex differences we observed since it predicts a female advantage only for discriminating anger, which is not what we found. Nevertheless, a question of considerable interest is whether the female advantage is enhanced under testing conditions involving infants, such as the infant face stimuli used previously by Babchuk et al. (1985). This would provide further support for the child-rearing hypothesis as opposed to explanations that rely on experience with adult faces and learned or evolutionary explanations relating to affiliation or power structure.

The present study adds to the literature on facial affect recognition by establishing that there is a larger female superiority for the decoding of negative than positive emotions. This conceivably explains some of the inconsistencies in past literature, as a female superiority has not been found universally. The present results favor the possibility of an evolved computational mechanism for emotion recognition and demonstrate domain specificity in facial processing. The female advantage may be an evolved adaptation related to the care of preverbal offspring.

## Acknowledgments

We thank R. Harshman for writing the MATLAB programs for the face presentation and reaction time recording. We also thank K. Humphrey for advice on visual masking. S. van Anders is a graduate student at Simon Fraser University. L. Mullin is now at McGill. This work was supported by a grant to E. Hampson from the Natural Sciences and Engineering Research Council of Canada.

## References

- Adolphs, R., Damasio, H., Tranel, D., & Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, *16*, 7678–7687.
- Ainsworth, M. D. S. (1979). Attachment as related to mother–infant interaction. *Advances in the Study of Behavior*, *9*, 2–52.
- Babchuk, W. A., Hames, R. B., & Thompson, R. A. (1985). Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis. *Ethology and Sociobiology*, *6*, 89–101.

- Breitneyer, B. G. (1984). *Visual masking: An integrative approach*. Oxford: Clarendon Press.
- Brunori, P., Ladavas, E., & Ricci-Bitti, P. E. (1979). Differential aspects in the recognition of facial expression of emotions. *Italian Journal of Psychology*, 6, 265–272.
- Burton, L. A., & Levy, J. (1989). Sex differences in the lateralized processing of facial emotion. *Brain and Cognition*, 11, 210–228.
- Duhaney, A., & McKelvie, S. J. (1993). Gender differences in accuracy of identification and rated intensity of facial expressions. *Perceptual and Motor Skills*, 76, 716–718.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, 115, 268–287.
- Ekman, P. (1997). Should we call it expression or communication? *Innovation*, 10, 333–344.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124–129.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. San Francisco, CA: Human Interaction Laboratory, University of California Medical Center.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Goldberg, S. (2000). *Attachment and development*. London, UK: Hodder Arnold.
- Goodall, J. (1986). *The chimpanzees of Gombe*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Hall, J. A., Lamb, E., & Perlmutter, M. (1986). *Child psychology today*. (2nd ed.). New York: Random House.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85, 845–857.
- Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore: Johns Hopkins University Press.
- Harshman, R. A., Hampson, E., & Berenbaum, S. A. (1983). Individual differences in cognitive abilities and brain organization: Part I. Sex and handedness differences in ability. *Canadian Journal of Psychology*, 37, 144–192.
- Henley, N. M. (1977). *Body politics: Power, sex, and nonverbal communication*. Englewood Cliffs, NJ: Prentice-Hall.
- Herlitz, A., & Yonker, J. (2002). Sex differences in episodic memory: The impact of intelligence. *Journal of Clinical and Experimental Neuropsychology*, 24, 107–114.
- Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115, 288–299.
- Kim, H. S., & Petrakis, E. (1998). Visuoperceptual speed of karate practitioners at three levels of skill. *Perceptual and Motor Skills*, 87, 96–98.
- Kimura, D. (1983). Speech representation in an unbiased sample of left-handers. *Human Neurobiology*, 2, 147–154.
- Kirouac, G., & Doré, F. Y. (1985). Accuracy of the judgment of facial expression of emotions as a function of sex and level of education. *Journal of Nonverbal Behavior*, 9, 3–7.
- Konner, M. (1982). *The tangled wing: Biological constraints on the human spirit*. New York, Holt: Rinehart and Winston.
- Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition—Women's faces make the difference. *Brain and Cognition*, 50, 121–128.
- McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology*, 17, 211–224.
- McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, 126, 424–453.
- Mufson, L., & Nowicki, S. (1991). Factors affecting the accuracy of facial affect recognition. *Journal of Social Psychology*, 131, 815–822.
- Ogawa, T., & Suzuki, N. (1999). Response differentiation to facial expression of emotion as increasing exposure duration. *Perceptual and Motor Skills*, 89, 557–563.
- Österlund, M. K., Keller, E., & Hurd, Y. L. (2000). The human forebrain has discrete estrogen receptor  $\alpha$  messenger RNA expression: High levels in the amygdaloid complex. *Neuroscience*, 95, 333–342.

- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*, 331–348.
- Rotter, N. G., & Rotter, G. S. (1988). Sex differences in the encoding and decoding of negative facial emotions. *Journal of Nonverbal Behavior*, *12*, 139–148.
- Russell, J. A., Bachorowski, J., & Fernández-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology*, *54*, 329–349.
- Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging*, *8*, 44–55.
- Thurstone, L. L., & Thurstone, T. G. (1963). *Primary mental abilities*. Chicago: Science Research Associates.
- Tranel, D., Damasio, A. R., & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, *38*, 690–696.
- Tyler, L. E. (1965). *The psychology of human differences*. (3rd ed.). New York: Appleton-Century-Crofts.
- Wagner, H. L., MacDonald, C. J., & Manstead, A. S. R. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology*, *50*, 737–743.
- Weitz, S. (1974). *Nonverbal communication: Readings with commentary*. New York: Oxford University Press.
- Wesman, A. G. (1949). Separation of sex groups in test reporting. *Journal of Educational Psychology*, *40*, 223–229.
- Zuckerman, M., Hall, J. A., DeFrank, R. S., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. *Journal of Personality and Social Psychology*, *34*, 966–977.