Are Effects of Emotion Expression on Trait Impressions Mediated by Babyfaceness? Evidence From Connectionist Modeling

Leslie A. Zebrowitz
Masako Kikuchi
Brandeis University

Jean-Marc Fellous
University of Arizona

Two studies provided evidence that bolsters the Marsh, Adams, and Kleck hypothesis that the morphology of certain emotion expressions reflects an evolved adaptation to mimic babies or mature adults. Study 1 found differences in emotion expressions' resemblance to babies using objective indices *of babyfaceness provided by connectionist models that are impervious to overlapping cultural stereotypes about babies and the emotions. Study 2 not only replicated parallels between impressions of certain emotions and babies versus adults but also showed that objective indices of babyfaceness partially* mediated *impressions of the emotion expressions. babyface effects were independent of strong effects of attractiveness, and babyfaceness did not mediate impressions of happy expressions, to which the evolutionary hypothesis would not apply.*

Keywords: *emotion expressions; babyfaceness; attractiveness; connectionist modeling; face perception*

People's appearance has a profound effect on important social outcomes. To give just one striking example, recent research demonstrated that the more competent looking of two congressional opponents won the election in 70% of the races (Todorov, Mandisodza, & Goren, 2005). Such effects highlight the importance of understanding what facial qualities give rise to particular impressions and why they do so (Zebrowitz & Montepare, 2005). The purpose of this study was to use connectionist modeling to investigate the contribution of babyfaceness to impressions of faces that vary in emotion expression as well as to investigate the contribution of facial attractiveness.

A tendency for impressions of people's stable traits to be influenced by their transient emotional expressions was early identified by Secord (1958) as *temporal extension*: "The perceiver regards a momentary characteristic of the person as if it were an enduring attribute" (p. 313). Consistent with this observation, people with angry expressions are perceived as high in dominance and low in affiliation; people with fear or surprise expressions are perceived as low in dominance and moderate in affiliation; and people with happy expressions are perceived as

Authors' Note: Portions of this research used archived faces from the Intergenerational Studies at the Institute of Human Development, University of California Berkeley. The authors thank Carrie Andreoletti, Deena Grant, Vinayak Sharma, and Mandy Smith for their help computing facial metrics and collecting face ratings, and Alain Mignault, Matt Bronstad, and Yuko Yotsumoto for their help with the connectionist modeling. This work was supported by National Science Foundation (NSF) Grant 97-08619 to the first and third authors and National Institutes of Health (NIH) Grants MH066836 and K02MH72603 to the first author. Correspondence should be addressed to Leslie A. Zebrowitz, Department of Psychology, MS 062, Brandeis University, Waltham, MA 02454: e-mail: zebrowitz@brandeis.edu.

PSPB, Vol. 33 No. 5, May 2007 648-662
DOI: 10.1177/0146167206297399

high in both dominance and affiliation (Knutson, 1996; Marsh, Adams, & Kleck, 2005; Montepare & Dobish, 2003). The most obvious explanation for such effects is temporal extension—the trait impressions associated with emotion expressions correspond to their short-term behavioral consequences. However, this explanation does not address the question of why particular emotional expressions look the way they do. An intriguing answer to that question that also has implications for the associated trait impressions was recently proposed by Marsh, Adams, et al. (2005). Those authors suggested that both the morphology of emotion expressions and the impressions they elicit may derive from the adaptive utility of their mimicking variations in facial maturity. More specifically, they argued that fear and anger expressions evolved to mimic babies' faces and mature faces, respectively. The rationale for this hypothesis was that it is adaptive for those experiencing fear to elicit reactions paralleling those elicited by helpless babies and for those experiencing anger to elicit reactions paralleling those elicited by powerful adults. The hypothesis that impressions of emotion expressions derive from their differential resemblance to babies is a variant of the well-documented babyface overgeneralization effect, whereby evolutionarily adaptive responses to babies are overgeneralized to individuals whose faces merely resemble babies (Montepare & Zebrowitz, 1998; Zebrowitz, 1996, 1997).

Consistent with their reasoning, Marsh, Adams, et al. (2005) found that faces expressing anger were judged as more mature-faced than those expressing fear. Anger and fear expressions were also judged to differ in facial qualities that differentiate faces of babies and adults, with anger expressions judged to have thinner lips, lower brows, more angular features, and smaller eyes. Finally, they also found that anger and fear expressions were judged to differ in traits previously associated with variations in babyfaceness, such as dominance, strength, shrewdness, and warmth.

Whereas Marsh, Adams, et al. (2005) found evidence consistent with their hypotheses that different impressions of anger and fear expressions derive from their differential resemblance to babies' faces, the question arises as to whether the effects are unique to these emotion expressions. We investigated impressions of surprise and happy expressions as well as angry ones to ascertain the specificity of the effect. Surprise is actually a transitory emotion that quickly turns into either fear or happiness, depending upon the nature of the surprise (Posamentier & Abdi, 2003). Perceivers may be more likely to respond to this transitory emotion as they respond to fear rather than to happiness for two reasons. First, there would seem to be more adaptive value to responding appropriately to fear than to happiness because happy people do not need to elicit reactions such as those elicited by helpless babies. Thus, the survival advantage inherent in a resemblance of fear to baby faces may transfer to a resemblance to babies of the precursor surprise faces. Second, consistent with the suggestion that surprise and fear faces both resemble babies, surprise expressions are more morphologically similar to fear than are happy expressions. Whereas fear and surprise expressions are both characterized by widened eyes, happy expressions, particularly those with a Duchenne smile, have narrow eyes due to the contracted orbicularis oculi muscles (Ekman, Davidson, & Friesen, 1990). Indeed, when surprise and fear are mislabeled, they tend to be confused with each other both by human judges (Calder, Burton, & Miller, 2001; de Bonis, 2003; Ekman, 1972; McAndrew, 1986) and by neural networks, whose responses are based solely on the physical configuration of the expressions, not their psychological meaning (Dailey, Cottrell, Padgett, & Adolphs, 2002).

In addition to examining additional emotion expressions, our research extends the findings of Marsh, Adams, et al. (2005) by comparing impressions of each emotion expression to a neutral expression posed by the same person. This enabled us to determine whether each emotion expression differed from neutral in its resemblance to babies, whereas Marsh, Adams, et al. could not answer this question through their single comparison of fear and anger expressions.

Although research has found parallels in impressions of faces varying in emotion and those varying in maturity, Montepare and Dobish (2003) found that trait impressions of faces with particular emotion expressions remained significant when babyface ratings were statistically controlled. Those results suggest that variations in resemblance to babies' faces may not actually mediate impressions of emotion expressions. On the other hand, it is possible that there is *partial* mediation, whereby the effect of emotion expressions on trait impressions would be significantly reduced in magnitude but not eliminated when controlling resemblance to babies. In Study 2, we directly tested the hypothesis that resemblance to babies partially mediates varying impressions of different emotion expressions.

We also examined the attractiveness of emotion expressions and whether this mediates the trait impressions. Attractiveness is of interest because it is enhanced by positive emotion expressions and diminished by negative ones (Hildebrandt, 1983; Karraker & Stern, 1990; Mueser, Grau, Sussman, & Rosen, 1984; O'Doherty, Winston, & Critchley, 2003; Reis, Wilson, Monestere, Bernstein, et al., 1990), and it also has strong influence on trait impressions (Eagly, Ashmore, Makhijani, & Longo, 1991). Thus, the well-documented attractiveness halo effect could also mediate impressions of emotion expressions.

Marsh, Adams, et al. (2005, Study 3) attempted to rule out the possibility that parallels in impressions of emotion expressions and faces varying in maturity derived from knowledge about parallels between fear and anger states and the traits of children and adults rather than from the structural resemblance of the various faces. They did so by creating composite stimuli that were not readily recognized as expressing emotions even though the emotion expression was shown in the eye region. They found that faces with fear or anger eye regions elicited impressions parallel to those showing full-face emotion expressions even though these composite faces were rarely identified as emotion expressions on a free-response measure. In the present studies, we used connectionist modeling to ensure that it is the structural resemblance between emotion expressions and faces varying in maturity that is driving the effects.

We used connectionist models to test our hypotheses for two reasons. First, as just noted, such models have an advantage over research using human judge ratings of facial similarity because they are impervious to stereotyped assumptions. If a network trained to differentiate faces of babies from adults subsequently finds anger expressions less physically similar to babies than it finds other expressions, this cannot be attributed to social stereotypes about angry people and babies. Rather, it can be due only to intrinsic similarities in the facial attributes known to the neural network.

The second reason for using connectionist modeling is that the similarity-based generalization effects we tested are a natural property of connectionist models, which are powerful nonlinear statistical modeling tools. Connectionist networks that have been trained to discriminate baby and adult faces will react to other faces according to their similarity to babies versus adults. Network activation to the untrained faces captures the network's overgeneralization of maturity information to those faces. Thus, we operationalized the similarity of a face to a baby as the extent to which it activated a network unit that had been trained to react to babies (baby unit). It should be noted that we are using connectionist modeling as a mathematical technique for generating an objective index of the structural similarity of a face to a particular category of faces. Our aim is *not* to test various alternative cognitive or neural models of face processing (e.g., Valentine, 1995).

In summary, whereas it is not possible to provide a direct empirical test of the Marsh, Adams, et al. (2005) claim that emotion expressions evolved to mimic variations in facial maturity, we report two studies that provide additional evidence bearing on this claim. Study 1 used connectionist modeling to investigate similarities in the facial structure of faces varying in emotion and those varying in maturity, thereby determining whether any similarities have an objective basis rather than merely reflecting social stereotypes that could produce similar judges' ratings such as those reported by Marsh, Adams, et al. Study 2 used connectionist modeling to determine whether the resemblance of emotion expressions to babies mediates impressions of their traits. Both studies also investigated variations in the attractiveness of emotion expressions and extended the emotion set used by Marsh, Adams, et al. to include happy expressions and surprise rather than fear expressions.

STUDY 1: BABYFACENESS AND ATTRACTIVENESS OF EMOTION EXPRESSIONS

In support of their hypothesis, Marsh, Adams, et al. (2005) established similarities between judges' ratings of the structure of faces varying in emotion and those varying in facial maturity. They further found that manipulating emotion expressions in the eye region alone influenced the perceived maturity of the faces and other trait impressions even more strongly than did manipulations using the entire face, which is consistent with other evidence for the salience of eye size in ratings of babyfaceness and associated trait impressions (Montepare & Zebrowitz, 1998). Study 1 employed connectionist modeling as a more objective method for ascertaining the resemblance of whole-face and eye-region emotion expressions to faces that vary in maturity.

We trained a connectionist network to differentiate faces of babies and adults using inputs from the entire face or inputs only from the eye region. We then tested the network on a different set of adult faces in which the same individuals posed different emotion expressions: surprise, anger, happy, and neutral. We used surprise rather than fear expressions for both theoretical and practical reasons. As discussed above, the theoretical justification was to determine whether the Marsh, Adams, et al. (2005) results would extend to surprise expressions, which are often precursors to fear and also show a morphological resemblance that leads these two emotions to be confused with each other more than with other emotions both by human judges and by neural networks (Calder et al., 2001; Dailey et al., 2002; de Bonis, 2003; Ekman, 1972; McAndrew, 1986; Posamentier & Abdi, 2003). The practical justification was that at the time we conducted the study, only surprise expressions were available to us in a sufficiently large number for the connectionist modeling. Eye-region inputs were compared with full-facial inputs because Marsh, Adams, et al. (2005) found stronger effects using eye-region manipulations and also because other research supports the salience of eyes. In particular, there is evidence that eyebrows rank first in the influence of facial features on attention and emotion impressions (Lundqvist & Öhman, 2005; Tipples, Atkinson, & Young, 2002). In addition, research examining perceptions of fear, in particular, revealed that a patient's selective impairment in recognition stemmed from an inability to make normal use of information from the eye region, which was identified as the most important feature for identifying this emotion (Adolphs et al., 2005).

We predicted that surprise expressions would be rated as more babyfaced and would activate the neural network baby units more than would anger, neutral, or happy expressions. We also predicted that anger expressions would be rated as less babyfaced and would activate the baby units less than the three other expressions. The Marsh, Adams, et al. (2005) explanation for the morphology of fear and anger expressions provides no reason to expect happy faces to differ from neutral ones in their resemblance to babies because there is no obvious evolutionary advantage to be derived from happy people eliciting responses such as those directed to helpless babies or powerful adults. However, positive reactions to both babies and happy faces could produce subjective ratings of happy expressions as more baby faced than other expressions. On the other hand, the large eye size of babies in contrast to the small eye size of happy expressions that was noted above could produce lower activation of the baby output units by happy than neutral or surprise expressions, indicating lesser resemblance to babies according to an objective measure.

Based on previous research investigating the attractiveness of positive and negative emotion expressions, we predicted that happy expressions would be rated as more attractive than the other expressions, whereas angry expressions would be rated as less attractive Hildebrandt, 1983; Karraker & Stern, 2003; Mueser et al., 1984; O'Doherty et al., 2003; Reis et al., 1990). Insofar as reactions to surprise are consistent with it being a precursor to fear, we also predicted low attractiveness of surprise expressions like that documented for negative emotions.

Method

Faces

Training/test faces. Facial metrics from digitized black-and-white, neutral-expression portrait photos of 30 White infants (18 males) ranging from 5 to 9 months, and 30 White adults (15 males) with a mean age of 17.4 years were used to train the network to differentiate babies from adults. For more details about these faces, see Zebrowitz, Fellous, Mignault, & Andreolleti (2003).

Generalization faces. Facial metrics from digitized black-and-white portrait photos of 26 men and 26 women, each of whom posed neutral, happy, anger, and surprise expressions, were used to examine network generalization. All photos of the 26 men and 8 of the women were taken from a study by Fellous (1997). Half of these individuals were drama graduate students, and the rest were computer science graduate students and postdoctoral fellows. All were instructed to take a few seconds to produce a happy, angry, neutral, or surprise face. They were further told to imagine an appropriate situation, if needed, and to press the space bar on a keyboard when ready, at which point the computer took the picture instantaneously. Photos of the remaining 18 women were taken by the first author with a camera, expressly for the present study. These women were all graduate students in psychology, and they were simply instructed to pose a happy, angry, neutral, or surprise face. All stimulus persons were in their 20s or 30s, and all were White.

Facial Metrics

Following the procedure reported by Zebrowitz et al. (2003), software was used to mark 64 points on digitized images of each face, from which facial metrics were computed using automatic procedures written in Visual Basic and Excel. After establishing reliability for points marked by two judges on a subset of faces from each category, one judge marked the remaining faces, and those points were used to calculate the final facial metrics. Eighteen nonredundant facial metrics that achieved acceptable reliability were selected as full-facial inputs to the connectionist model. These included facial roundness plus the metrics shown in Figure 1. Seven of these metrics, underlined in Figure 1, were selected as eye-region inputs to the connectionist model.

Network Training

The total set of faces used to train the network was composed of 60 faces (30 babies, 30 adults). Twenty trials were computed. At each trial, 20 babies and 20 adults were randomly selected from the total set to comprise the *training set*. The remaining faces comprised the *test set*. A different network was trained at each trial, using supervised learning and different weights initialized at random

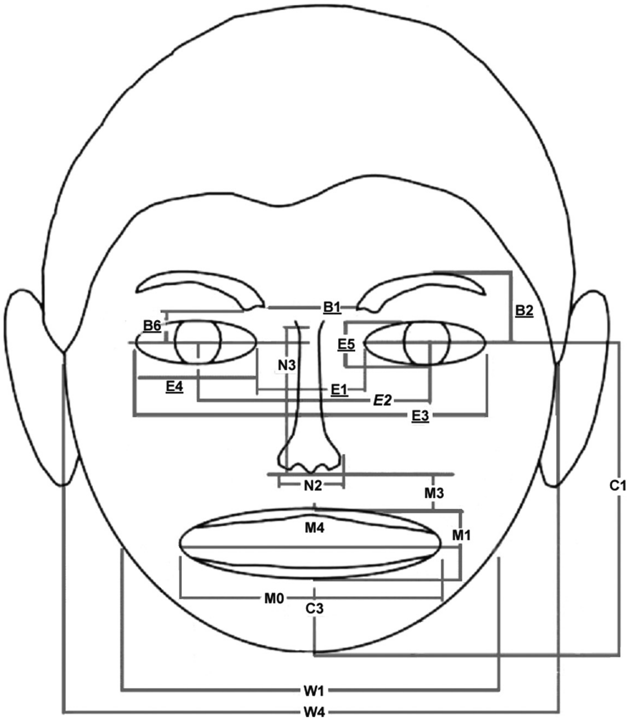

Figure 1 Location of facial metrics used as inputs to the connectionist models trained on full-facial metrics.
NOTE: Underlined distances were used as inputs to models trained on eye-region metrics. All metrics were normed by E2, interpupilary distance.

values. The model was a standard back-propagation neural network with three layers. Two models were trained. The architecture of the *full-facial-metrics model* had 18 input nodes, 4 hidden nodes, and two output units (baby and adult); the *eye-region-metrics model* had 7 input nodes, 4 hidden nodes, and two output units (baby and adult). Each input node projected to any or all of the hidden nodes, and the hidden nodes projected to the output units (baby and adult). The input-weight matrices connecting the layers consisted of numbers between –1 and 1. The output units were rescaled into graded values ranging from 0% to 100% activation. All units were nonlinear and mapped the weighted sum of their inputs to their output using a sigmoidal transfer function. On each trial, after training on a set of 20 baby faces and 20 adult faces, the network was frozen and tested on the test set for that trial to establish successful learning. The training parameters were 3,000 training epochs per trial, a .02 learning rate, and a .2 error goal. It should be noted that the choice of the network architecture and network parameters was not meant to address a specific neural information-processing hypothesis, which was beyond the scope of this work. Rather, the choices were empirical and justified by the performance of the network on the training data set. Because testing was achieved on the basis of faces that were not used for training, the output unit activation for a given face, computed over 20 trials, was considered an "objective" measure of facial quality (i.e., baby or adult). Our emphasis is therefore on the objectivity of the facial quality rather than on a particular way of assessing it.

Network Generalization

After successful training on each trial, the network was presented with input metrics from a new set of 208 generalization faces that included each of the 52 targets posing the four facial expressions (neutral, surprise, angry, happy). This generated two dependent variables for each generalization face: the average activation across 20 trials of the baby unit in the network trained on all facial metrics, and the average activation of the baby unit in the network trained on eye-region metrics.[1]

Emotion and Appearance Ratings

Sixteen judges (eight males) rated the emotions of all faces on 7-point scales, with endpoints labeled *neutral/very angry*, *neutral/very happy*, *neutral/very surprised*. Six additional judges (two males) rated the babyfaceness and attractiveness of all faces on 7-point scales with endpoints labeled *maturefaced/babyfaced*, and *unattractive/attractive*. Faces were presented in a random order.

Results

Reliability of Facial Metrics and Appearance Ratings

High interjudge agreement for the facial metrics of baby and adult faces (cf. training and test set) was previously reported (Zebrowitz et al., 2003). The selected input metrics for the emotion expression faces (cf. generalization set) also showed high agreement both for the normalization interpupilary distance, $r = .98$, and the selected input metrics, mean $r = .70$. The average alpha reliability coefficient was .82 for attractiveness ratings, .86 for babyface ratings, .96 for surprise ratings, and .95 for both anger and happy ratings. Data analyses utilized mean ratings for each face across judges.

Network Training

Training a network to differentiate babies from adults using full-facial metrics as inputs achieved 95.38% correct identification of the 40 training faces and 89% correct identification of the 20 test faces, averaged across 20 trials. Across the 20 trials, activation of the baby unit was significantly higher for babies, M = 85.98, SD = 9.74, than for adults, M = 15.58, SD = 14.15, $F(1, 58) = 504.07$, $p < .0001$. The network was also successfully trained when provided only with inputs from the seven eye-region metrics, achieving 95.63% and 91.25% correct identification of the training and test faces, respectively. Across the 20 trials, activation of the baby unit was significantly higher for babies, M = 86.40, SD = 8.00, than for adults, M = 15.18, SD = 9.37, $F(1, 58) = 1001.79$, $p < .0001$.

Analyses

Face was the unit of analysis. Because the same people posed all four emotions, emotion expression was nested within face, yielding a mixed 2 (face sex) × 4 (emotion expression) ANOVA on all dependent variables with face sex as a between-face variable and posed-emotion expression as a within-face variable. Although we included face sex in the analyses of Studies 1 and 2, in the interest of brevity we do not discuss the effects because we had no theoretical reason to expect sex differences in the resemblance of emotion expressions to babies, and the few effects that were obtained may reflect differences in the method for photographing men and women rather than any meaningful sex differences.[2]

Expression-Manipulation Checks

As shown in the columns of Table 1, the manipulation of emotional expressions was successful. Judges' ratings of surprise were significantly higher for surprise expressions than each of the other emotions; ratings of anger were higher for anger expressions than each of the other emotions, and ratings of happy were higher for happy expressions than each of the other emotions.[3]

Effects of Face Expression on Babyfaceness and Attractiveness

Judges ratings of babyfaceness. Extending the Marsh, Adams, et al. (2005) findings for fear versus anger expressions, judges gave higher babyface ratings to surprise than anger expressions. Moreover, comparing the babyface ratings of these expressions with neutral ones revealed two independent effects: Surprise expressions were rated significantly higher, and anger expressions were rated significantly lower. In addition, happy expressions were rated higher in babyfaceness than either anger or neutral expressions, although they were judged less babyfaced than surprise expressions (see Table 2).

TABLE 1: Mean Emotion Ratings as a Function of Face Expression

	Emotion Rating					
	Surprise		*Angry*		*Happy*	
Face Expression	M	SD	M	SD	M	SD
Neutral	1.54_a	0.50	3.45_a	1.02	2.04_a	0.71
Surprise	$\mathbf{5.47}_b$	1.10	2.82_b	0.78	4.04_b	1.35
Anger	2.37_c	0.62	$\mathbf{4.72}_c$	0.84	2.35_c	0.61
Happy	3.11_d	0.54	1.94_d	0.36	$\mathbf{5.75}_d$	0.60
$F(3, 150)$	309.66****		184.08****		241.25****	

NOTE: Expression means in each column that do not share the same subscript differed from each other at the .05 level determined by LSD test. Due to violations of sphericity, the degrees of freedom were reduced in some analyses. Values in bold correspond to the manipulated expression.
****$p < .001$.

Babyface network activation. As shown in Table 2, whether the network was trained on full-facial metrics or just eye-region metrics, surprise expressions activated the baby unit significantly more than anger expressions, paralleling judges' ratings of babyfaceness. Also consistent with rated babyfaceness, surprise expressions activated the baby unit trained on eye-region metrics more than neutral expressions did, and anger expressions activated the baby unit trained on full-facial metrics less than neutral expressions did. However, in contrast to the findings for rated babyfaceness, surprise expressions did not differ from neutral ones in their activation of the baby unit trained on full-facial metrics, and anger expressions did not differ from neutral ones in their activation of the unit trained on eye-region metrics. Happy expressions did not differ from neutral ones in their activation of the baby unit trained on full-facial metrics, and they activated the baby unit trained on eye-region metrics *less* than neutral expressions did, which was opposite to the effect for judges' babyface ratings.

Attractiveness ratings. As shown in Table 2, happy expressions were significantly more attractive than all other expressions posed by the same people, and anger expressions were significantly less attractive than neutral or happy ones. Although surprise and anger expressions did not differ in attractiveness, surprise also did not differ from a neutral expression.

Discussion

The results of Study 1 add to the existing literature in several ways. First, comparing the resemblance to babies of four different emotion expressions allowed us to determine whether there are unique and independent tendencies for particular expressions to differ in their resemblance to babies. By comparing reactions only to

TABLE 2: Babyfaceness and Attractiveness as a Function of Face Expression

	Babyface Rating		*Activation of Baby Output Unit Trainded on Full-Facial Metrics*		*Activation of Baby Output Unit Trained on Eye-Region Metrics*		*Attractiveness Rating*	
Face Expression	M	SD	M	SD	M	SD	M	SD
Neutral	3.60_a	1.32	25.42_a	12.50	18.63_a	9.66	3.38_a	1.17
Surprise	4.07_b	1.23	25.97_a	13.08	24.55_b	11.15	3.23_{ab}	1.07
Anger	3.34_c	1.26	19.02_b	11.42	18.10_a	8.50	3.09_b	1.01
Happy	3.81_d	1.34	25.89_a	12.69	16.20_c	8.85	3.70_c	1.07
F(3, 150)	26.25****		10.14****		13.24****		16.93****	

NOTE: Means in each column that do not share the same subscript are statistically significant at the .05 level, determined by LSD test, except that $p = .08$ for activation of eye-region metrics by angry vs. happy faces. Due to violations of sphericity, the degrees of freedom were reduced in some analyses.
****$p < .001$.

fear versus anger expressions, the Marsh, Adams, et al. (2005) study was unable to make this differentiation. Extending the Marsh, Adams, et al. results for fear versus anger expressions, we found that surprise expressions were rated as more babyfaced than anger expressions posed by the same people. The parallel effects for fear and surprise expressions are consistent with the fact that surprise is often a precursor to fear (Posamentier & Abdi, 2003) and that when surprise and fear are mislabeled, they tend to be confused with each other both by human judges and by neural networks, indicating structural similarities (Calder et al., 2001; Dailey et al., 2002; de Bonis, 2003; Ekman, 1972; McAndrew, 1986). Not only did the rated babyfaceness of surprise and anger expressions differ, but also judges perceived surprise expressions to resemble babies more than neutral ones and anger expressions to resemble babies less, thus documenting two independent effects.

Another significant contribution of Study 1 was to demonstrate that differences in the perceived resemblance of surprise and anger expressions to babies' faces does not merely reflect overlapping cultural stereotypes. Whereas judges' ratings of the babyfaceness of emotion expressions could reflect such stereotypes, the fact that connectionist models responded similarly to babies and certain emotion expressions can be due only to similarities in the facial metrics input to the neural network. Compared with neutral expressions, surprise expressions elicited more activation of a baby unit trained to differentiate babies and adults based on eye-region metrics, and angry expressions elicited less activation of a baby unit trained to differentiate babies and adults based on their full-facial metrics.

Happy, like surprise, expressions were rated as more babyfaced than neutral expressions posed by the same people. However, this difference was not paralleled by the connectionist modeling results, suggesting that the subjective ratings may reflect cultural assumptions about similarities between happy people and babies rather than any true structural resemblance. Indeed, structural resemblance to babies' eyes was *lower* for happy expressions than all others, which is consistent with the eye narrowing that is produced by the contracted orbicularis oculi muscles of a smile (Ekman et al., 1990). The divergence between the subjective ratings and the neural network activations highlights the value of connectionist modeling for ascertaining whether it is actually the physical stimulus characteristics of particular categories of faces that contributes to impressions of their traits.

Another contribution of this study derives from comparing the resemblance between emotions and babies in their full-facial metrics versus their eye-region metrics. That surprise expressions resembled babies in eye-region metrics, but not in full-facial metrics, is consistent with the Marsh, Adams, et al. (2005, Study 3) finding that manipulations of expressions in the eye region alone influenced the perceived maturity of faces more than manipulations using the entire face. The importance of the eye region is also consistent with the fact that large eyes mark surprise and fear expressions as well as babies' faces, the latter due to relatively less growth of the eyes than other facial features from birth to maturity (Enlow, 1990), and with other research that has documented the primacy of the eye region in emotion perception, particularly fear (Adolphs et al., 2005; Lundqvist & Öhman, 2005; Tipples et al., 2002). Whereas resemblance to babies in the eye region characterized surprise expressions, the eye region was less central to identifying anger expressions, which were characterized by low resemblance to babies in full-facial metrics.

As expected, emotion expressions varied in attractiveness as well as babyfaceness, with happy expressions more attractive than all others and anger expressions less attractive than neutral and happy ones. Surprise and anger did not differ significantly in attractiveness, consistent with the hypothesis that surprise would be relatively unattractive due to its similarity to the negative emotion of fear. These findings suggest that trait impressions of emotion expressions may be mediated by an attractiveness halo effect in addition to a babyface overgeneralization effect.

STUDY 2: MEDIATORS OF TRAIT IMPRESSIONS FROM EMOTION EXPRESSION

According to Marsh, Adams, et al. (2005), emotion expressions have evolved to differ in their resemblance to babies because it is adaptive for those experiencing fear to elicit reactions paralleling those elicited by helpless babies and for those experiencing anger to elicit reactions paralleling those elicited by powerful adults. As noted above, there are in fact parallels in the trait impressions of faces varying in emotion expressions and those varying in maturity, with fear expressions and babies' faces eliciting impressions of lower dominance and higher affiliation than anger expressions and adults' faces, respectively (Knutson, 1996; Marsh, Adams, et al., 2005; Montepare & Dobish, 2003; Zebrowitz et al., 2003). Whereas both Marsh, Adams, et al. (2005) and Montepare and Dobish (2003) found such parallels, the former authors did not directly examine mediation, and the latter authors found that trait impressions of emotion expressions remained significant when ratings of babyfaceness were controlled. Although those results suggest that variations in resemblance to babies' faces may not actually mediate impressions of emotion expressions, partial mediation would be consistent both with the Marsh, Adams, et al. hypothesis and the Montepare and Dobish results. In partial mediation, the effect of emotion expressions on trait impressions would remain significant when controlling variations in resemblance to babies, whereas the magnitude of the effect would be significantly reduced.

In Study 2, we directly examined mediation effects using not only rated babyfaceness but also objective indices of resemblance to babies derived from connectionist modeling. The suggestion by Marsh, Adams, et al. (2005) that reactions to fear and anger expressions derive from their resemblance to babies' faces suggests that statistically controlling that resemblance should weaken the impressions of these expressions. We also tested mediation by rated attractiveness because Study 1 found differences in the attractiveness of the emotion expressions. Coupled with the positive association of attractiveness with impressions of affiliation and dominance (e.g., Eagly et al., 1991), this suggests that an attractiveness halo effect could contribute to the trait impressions of emotion expressions.

We predicted that compared with neutral expressions, surprise, like fear, expressions would be perceived as higher in affiliation and lower in dominance; angry expressions would be perceived as lower in affiliation and higher in dominance; and happy expressions would be perceived as higher in both affiliation and dominance.[4] Based on the results of Study 1, we predicted that babyface ratings and resemblance to babies in the eye region would each mediate impressions of surprise as compared with neutral expressions and that babyface ratings and resemblance to babies in full-facial metrics would each mediate impressions of angry expressions. The predictions for impressions of happy expressions were more equivocal. Study 1 results suggested that babyface ratings and resemblance to babies in the eye region would mediate impressions of happy expressions, each in a different direction because happy expressions were rated higher in babyfaceness than neutral ones but activated the baby eye-region unit less than neutral ones. On the other hand, as noted above, the Marsh, Adams, et al. (2005) explanation for the morphology of facial expressions provides no reason to expect babyfaceness to mediate impressions of happy faces because there is no obvious evolutionary advantage to be derived from happy people eliciting responses like those directed to helpless babies or powerful adults. Finally, we predicted that attractiveness ratings would mediate impressions of angry and happy expressions, each of which differed in attractiveness from neutral expressions in Study 1. Because Montepare and Dobish (2003) found that trait impressions of emotion expressions remained significant when babyface and attractive ratings were statistically controlled, we expected to find partial rather than complete mediation by our subjective ratings and objective indices of resemblance to babies.

Method

The facial metrics, connectionist model, network training, and network generalization were identical to Study 1. Trait impressions of the baby and adult training faces were taken from Zebrowitz et al. (2003).

Trait Impressions

Judges. Sixteen judges (8 males) rated one of four sets of faces, earning course credit or payment for their participation.

Face ratings. Judges rated either one of three sets of emotion expression faces (happy and angry, happy and surprise, angry and surprise) or a set of neutral-expression faces. Two subsets of each of the three emotion face sets were created so that judges would not make trait ratings for two versions of the same face; judges did rate one version of all faces. Faces within each set were presented in one of two random orders, blocked by sex, with the order of the sexes counterbalanced across judges. Each face was rated on eight 7-point trait scales with half the judges receiving one order and half receiving the reverse order (dominant or submissive, sociable or unsociable, naïve or shrewd, unhealthy or healthy, intelligent or unintelligent, physically weak or physically strong, cold or warm).[5] Judges' ratings of babyfaceness and attractiveness were taken from Study 1.

Results

Trait-Rating Reliability

High reliability for ratings of the baby and adult training faces was reported by Zebrowitz et al. (2003). For the emotion generalization faces, the average alpha reliability coefficient across ratings was .85 for both the angry/happy and the angry/surprise face sets; .81 for the happy/surprise face set; and .80 for the neutral face set. Data analyses utilized mean ratings for each face across judges.

Trait composites. Two trait composites were created, paralleling dimensions previously found to differentiate impressions of facial expressions. The dominance composite included ratings of dominant, strong, and shrewd: alpha = .95 for the baby and adult training faces; alpha = .88 for the emotion expression generalization faces. The affiliation composite included ratings of warm and sociable: alpha = .95 for the baby and adult training faces; alpha = .94 for the emotion expression generalization faces.

Impressions of Babies Versus Adults and Emotion Expressions

The Marsh, Adams, et al. (2005) hypothesis provides a potential explanation for impressions of emotion expressions only for traits that are judged to differentiate babies from adults. We performed 2 (face sex) × 2 (face age) ANOVAs on dominance and affiliation composites created from individual trait ratings of the baby and adult training faces previously reported by Zebrowitz et al. (2003). As shown in Table 3, babies were judged to be lower in dominance and higher in affiliation than adults. Table 3 also shows the means and standard deviations for impressions of the emotion expression faces. These impressions will be described in the context of the hierarchical linear modeling (HLM) analyses reported below.

TABLE 3: Mean Trait Ratings of Training Faces (Babies and Adults) and Emotion Generalization Faces (Neutral, Surprise, Angry, and Happy)

	Dominance		*Affiliation*	
	M	SD	M	SD
Training set				
Baby	2.83	0.52	4.50	0.85
Adult	4.78	0.62	3.74	1.19
$F(1, 58)$	175.46****		8.19***	
Generalization set				
Neutral	4.07	0.64	3.78	0.90
Surprise	3.54	0.52	4.38	0.73
Angry	4.60	0.63	3.02	0.62
Happy	3.87	0.63	5.09	0.64

***$p < .01$. ****$p < .001$.

Predicting Impressions of Emotion Expressions From Babyfaceness and Attractiveness

Analyses. Face was the unit of analysis. Because emotion expressions were nested within face, we tested our hypotheses with multilevel modeling (e.g., Raudenbush & Bryk, 2002) using HLM 6.02 (Scientific Software International, http://www.ssicentral.com) that takes into account the dependency between the emotion expressions. All variables were standardized prior to analyses so that the γ statistic could be interpreted as a standardized regression coefficient.

In Model 1, we entered one emotion dummy variable (surprise vs. neutral, angry vs. neutral, *or* happy vs. neutral, all within-face Level 1 variables), face sex (a between-face Level 2 variable), and the Face Sex × Emotion interaction effect. This yielded three HLM models, one for each emotion, predicting each of the two trait composites, for a total of six versions of Model 1 (see Table 4). Although we included face sex as a control variable, as in Study 1, we do not discuss the effects both in the interest of brevity and also because they may reflect differences in the method for photographing men and women rather than any meaningful sex differences (see Note 2).

In Model 2, we added the attractiveness ratings plus one of three indices of babyfaceness, all of which were within-face Level 1 variables: Model 2a, rated babyfaceness; Model 2b, activation of the baby unit trained on full-facial metrics; or Model 2c, activation of the

TABLE 4: Standardized Coefficients (and standard errors) of Multilevel Models Predicting Impressions of Emotion Expression Faces

	(Controlling Sex, Sex × Expression)		*Babyface Rating*		*Baby Full-Facial Metrics*		*Baby Eye-Region Metrics*	
	Model 1		*Model 2a*		*Model 2b*		*Model 2c*	
Predictor	γ	(SE)	γ	(SE)	γ	(SE)	γ	(SE)
Surprise								
Perceived dominance								
Surprise vs. neutral expression	–.41****	(.06)	–.28****	(.05)	–.41****	(.06)	–.36****	(.06)
Babyfaceness			**–.65******	(**.07**)	–.11	(.09)	**–.18****	(**.08**)
Attractiveness			**.24******	(**.05**)	.05	(.10)	.08	(.10)
Perceived affiliation								
Surprise vs. neutral expression	.35****	(.08)	.36****	(.07)	.38****	(.07)	.33****	(.07)
Babyfaceness			.10	(.10)	.08	(.07)	**.18***	(**.09**)
Attractiveness			*.41*****	(*.09*)	.44****	(.07)	.42****	(.07)
Anger								
Perceived dominance								
Anger vs. neutral expression	.39****	(.05)	.36****	(.05)	.35****	(.07)	.40****	(.06)
Babyfaceness			**–.61******	(**.08**)	**–.20****	(**.10**)	–.22**	(.09)
Attractiveness			*.24*****	(*.06*)	.11	(.10)	*.17**	(*.09*)
Perceived affiliation								
Anger vs. neutral expression	–.44****	(.06)	–.36****	(.06)	–.34****	(.06)	–.38****	(.06)
Babyfaceness			.17*	(.10)	**.14***	(**.08**)	.14	(.09)
Attractiveness			**.46******	(**.10**)	**.51******	(**.08**)	**.47******	(**.08**)
Happy								
Perceived dominance								
Happy vs. neutral expression	–.16**	(.06)	–.14**	(.06)	–.17***	(.06)	–.21***	(.06)
Babyfaceness			–.68***	(.08)	–.13	(.11)	–.21	(.13)
Attractiveness			*.28*****	(*.05*)	.14	(.10)	*.19***	(*.09*)
Perceived affiliation								
Happy vs. neutral expression	.65****	(.05)	.58****	(.05)	.59****	(.05)	.61****	(.05)
Babyfaceness			.11	(.07)	.15***	(.05)	*.14***	(*.06*)
Attractiveness			**.39******	(**.08**)	**.42******	(**.07**)	**.39******	(**.07**)

NOTE: Expression dummy variables were coded 0 for neutral expressions and 1 for emotion expressions (surprise, angry, happy). All the variables were standardized. Models 2a-2c report results of three separate hierarchical linear modeling analyses that tested mediation of impressions by attractiveness together with either rated babyfaceness, activation of the baby unit trained on full-facial metrics, or activation of the baby unit trained on eye-region metrics. Effects showing partial mediation are in bold. Effects showing partial suppression are in italics.
$^{*}p < .10$. $^{**}p < .05$. $^{***}p < .01$. $^{****}p < .001$.

baby unit trained on eye-region metrics. This yielded three alternative models predicting the two trait composites for each of the three emotions, for a total of 18 versions of Model 2. We included attractiveness together with an indicator of babyfaceness in order to ascertain whether the predicted effects for babyfaceness and attractiveness were independent because attractiveness and babyface ratings of surprise expressions were positively correlated, $r(50) = .27$, $p = .05$.

We followed the procedure outlined by Baron and Kenny (1986) to determine whether the variables entered in Model 2 mediated trait impressions of emotion expressions, which is an appropriate procedure for our design (Kenny, Korchmaros, & Bolger, 2003). We used the Sobel test even though it is likely to underestimate the significance of mediation effects for moderate sample sizes such as ours in contrast to the alternative bootstrap method (Shrout & Bolger, 2002), because the latter has not been integrated into HLM.

Impressions of surprise expressions. In Model 1, surprise expressions were perceived as lower in dominance and as higher in affiliation than neutral expressions, as predicted. Although surprise expressions continued to be perceived as less dominant than neutral expressions when controlling potential mediators in Model 2, Sobel tests revealed significant partial mediation by surprise expressions' higher babyface ratings, $z = 4.90$, p < .001, and by their higher activation of the baby unit trained on eye-region metrics, z = 1.95, $p = .05$, as predicted. Attractiveness also qualified as a mediator, but only in Model 2a, which included babyface ratings. The Sobel test revealed that the lower perceived dominance of surprise expressions was partially mediated by their lower attractiveness in this model, $z = 1.98$, $p = .05$, an unpredicted effect. As expected, activation of the baby unit trained on full-facial metrics did not meet the criteria for mediation of the lower perceived dominance of surprise expressions.

Although surprise expressions also continued to be perceived as more affiliative than neutral expressions in Model 2, a Sobel test revealed marginally significant partial mediation by the greater resemblance of surprise expressions to babies in the eye region, $z = 1.67$, $p < .10$, as predicted. Neither babyface ratings nor activation of the baby unit trained on full-facial metrics met the criteria for mediation of the higher perceived affiliation of surprise expressions. Attractiveness had positive effects on perceived affiliation in all three models. Moreover, the lower attractiveness of surprise expressions in the model including babyface ratings qualified it as a suppressor of the tendency to perceive surprise expressions as higher in affiliation than neutral ones. The Sobel test revealed that this partial suppression was significant, $z = 2.01$, $p = .05$, an unpredicted effect.

Impressions of anger expressions. In Model 1, anger expressions were perceived as higher in dominance and lower in affiliation than neutral expressions, as predicted. Although anger expressions continued to be perceived as more dominant than neutral expressions when controlling potential mediators in Model 2, Sobel tests revealed significant partial mediation by anger expressions' lower babyface ratings, $z = 2.37$, $p = .02$, and by their lower activation of the baby unit trained on full-facial metrics, $z = 1.93$, $p = .05$, as predicted. Activation of the baby unit trained on eye-region metrics did not meet the criteria for mediation of perceived dominance because it did not differ for anger and neutral expressions, as expected. Consistent with predictions, the lower attractiveness of anger expressions qualified as a partial suppressor of the tendency to perceive them as higher in dominance in the models that included babyface ratings, $z = 2.48$, $p = .01$, and activation of the baby unit trained on eye-region metrics, although this effect was only marginally significant, $z = 1.64$, $p = .10$.

Anger expressions also continued to be perceived as less affiliative than neutral expressions when controlling potential mediators in Model 2, but a Sobel test revealed marginally significant partial mediation by activation of the baby unit trained on full-facial metrics, $z = 1.68$, $p = .09$, as predicted. As in the case of perceived dominance, activation of the baby unit trained on eye-region metrics did not meet the criteria for mediation. Although babyface ratings did meet the criteria for mediation, the Sobel test was not significant, $z = 1.43$, $p = .15$. Finally, as predicted, there was significant partial mediation by anger expressions' lower attractiveness ratings in the models that included babyface ratings, $z = 2.55$, $p = .01$; activation of the baby unit trained on full-facial metrics, $z = 3.15$, $p = .002$; and activation of the baby unit trained on eye-region metrics, $z = 3.37$, $p = .001$.

Impressions of happy expressions. Consistent with previous research, happy expressions were perceived as higher in affiliation than neutral expressions in Model 1. They were also perceived as significantly lower in dominance, which ran counter to previous research showing higher dominance attributed to happy expressions. Happy expressions continued to be perceived as less dominant than neutral expressions when controlling potential mediators in Model 2. As predicted, none of the indices of babyfaceness qualified as mediators or suppressors. However, as expected, the higher attractiveness of happy expressions qualified as a significant partial suppressor of the tendency to perceive them as less dominant in the model including babyface ratings, $z = 2.74$, $p = .01$, and as a marginally significant partial suppressor of the tendency to perceive them as less dominant in the model including activation of the baby unit trained on eye-region metrics, $z = 1.76$, $p = .08$.[6]

Although happy expressions also continued to be perceived as more affiliative than neutral expressions when controlling potential mediators in Model 2, Sobel tests revealed that happy expressions' higher attractiveness was a significant partial mediator of their higher perceived affiliation in the models including babyface ratings or activation of the baby unit trained on eye-region metrics, both zs $= 2.71$, $p = .01$, and in the model including activation of the baby unit trained on full-facial metrics, $z = 3.06$, $p = 002$, as predicted. In addition, happy expressions' lower babyfaceness on the index of eye-region metrics was a marginally significant partial suppressor of their higher perceived affiliation in that model, $z = 1.76$, $p = .08$. The other indices of babyfaceness did not qualify as mediators.

Summary. Compared with neutral expressions, surprise created impressions of lower dominance and higher affiliation, anger created impressions of higher dominance and lower affiliation, and happy created impressions of lower dominance and higher affiliation. As predicted, impressions of surprise and anger expressions were mediated by babyfaceness, and this was confirmed with objective indices from the connectionist model—high babyfaced eye-region metrics for surprise and low babyfaced overall facial metrics for anger. These mediation effects were significant for impressions of dominance and marginally significant for impressions of affiliation. Impressions of happy faces were not mediated by babyfaceness. On the contrary, the lower babyfaceness of happy faces in the eye region partially suppressed the impressions of higher affiliation. Attractiveness also contributed to trait impressions, with the higher attractiveness of happy expressions partially mediating impressions of their higher affiliation and partially suppressing impressions of their lower dominance, consistent with predictions. Similarly, the lower attractiveness of surprise and anger faces partially mediated impressions of surprise expressions' lower dominance and anger expressions' lower affiliation while partially suppressing impressions of surprise expressions' higher affiliation and anger expressions higher dominance, effects that were predicted for anger but not for surprise.

Discussion

Consistent with the Marsh, Adams, et al. (2005) argument that reactions to emotion expressions derive from variations in their facial maturity, impressions of lower dominance and higher affiliation in surprise than neutral expressions mimicked impressions of babies, whereas impressions of higher dominance and lower affiliation in anger than neutral expressions paralleled impressions of mature-faced adults. In addition to demonstrating these parallels, Study 2 replicated Montepare and Dobish's (2003) finding that impressions of the affiliation and dominance of emotion expressions held true independent of their resemblance to a baby's face, as evidenced in significant emotion effects when indices of babyfaceness were controlled. However, Study 2 further showed that impressions of emotion expressions were partially mediated by their resemblance to babies rather than simply paralleling impressions of babies versus adults, as shown by Marsh, Adams, et al. (2005). The particular indices of babyfaceness that mediated the impressions of the emotion expressions were those that had differentiated the expressions in Study 1.

That mediation was shown by resemblance to babies' face as assessed by the connectionist models adds credence to the Marsh, Adams, et al. (2005) hypothesis by providing evidence that it is indeed the structural similarities between emotion expressions and faces varying in facial maturity that account for the effects rather than similar cultural stereotypes about the traits of babies and those with particular emotion expressions. It is noteworthy that impressions of happy expressions' lower dominance were not mediated by any of the indices of babyfaceness even though those impressions paralleled impressions of surprise (and fear) expressions. Indeed, although judges rated happy expressions as more babyfaced in Study 1, not only did this perception fail to mediate the trait impressions in Study 2, but also happy expressions' objectively *lower* resemblance to babies in the eye region tended to suppress impressions of them as higher in affiliation than neutral expressions. The uniqueness of babyfaceness's positive contribution to impressions of surprise and fear expressions strengthens the argument that they evolved to mimic babies because it is adaptive for those experiencing these emotions to elicit reactions paralleling those elicited by helpless infants. This argument would be weakened had babyfaceness also mediated impressions of happy expressions, as there is no apparent adaptive value in the evolution of happy expressions to mimic babies, because happy people do not need to elicit care and protection.

That the mediating effects of babyfaceness on impressions of surprise and anger expressions held true over and above strong effects of attractiveness strengthens support for the Marsh, Adams, et al. (2005) hypothesis regarding the effects of babyfaceness. The finding that attractiveness made a positive contribution to the perceived affiliation and dominance of emotion expressions is consistent with the well-documented halo effect (e.g., Eagly et al., 1991). Although its contribution to impressions of surprise expressions was not predicted, it is consistent with the tendency for surprise expressions to be rated as somewhat less attractive than neutral ones in Study 1. This trend attained significance in Study 2 with babyfaceness controlled, thereby qualifying attractiveness as a mediator.

The contributions of attractiveness to trait impressions complemented some babyface effects and opposed others. Like surprise's higher babyfaceness as compared with neutral expressions, its lower attractiveness mediated impressions of lower dominance. Also, like anger's lower babyfaceness, its lower attractiveness mediated impressions of lower affiliation. On the other hand, surprise's lower attractiveness than neutral expressions suppressed impressions of higher affiliation, opposite to the purported adaptive effect of its higher babyfaceness, and anger's lower attractiveness suppressed impressions of higher dominance, opposite to the purported adaptive effect of its lower babyfaceness. One might speculate about the adaptive value of these opposing attractiveness effects. First, it may be functional for individuals experiencing surprise (or fear) to elicit a guarded approach response that a combination of high babyfaceness and low attractiveness might produce. Consistent with this reasoning, fear faces do facilitate approach responses more than angry faces (Marsh, Ambady, & Kleck, 2005), and probably less than happy ones, although that has not been investigated. Second, although the lower attractiveness of anger faces suppressed impressions of their higher dominance, the tendency for low attractiveness to mediate impressions of their lower affiliation was even stronger. Thus, the net effect of anger's low attractiveness may be functional for warding off unwanted advances. This argument is strengthened by evidence that both angry faces and anomalous faces facilitate avoidance as compared with approach responses (Marsh, Ambady, et al., 2005), coupled with the fact that anomalous and unattractive faces show a structural resemblance as ascertained by connectionist models (Zebrowitz et al., 2003).

GENERAL DISCUSSION

Although an evolutionary account of the morphology of emotion expressions can be dated to Darwin's (1872/1965) treatise on *The Expression of Emotions in Man and Animals*, many previous accounts have focused on the nonsocial adaptive functions of expression morphology. The Marsh, Adams, et al. (2005) account is unique in considering adaptive social functions served by the morphology of fear and anger. The present research provides several findings to bolster Marsh, Adams, et al.'s innovative hypothesis that the morphology of anger and fear expressions evolved to vary in facial maturity because it is adaptive for those experiencing anger to elicit reactions paralleling those elicited by powerful adults and for those experiencing fear to elicit reactions like those elicited by helpless babies.

First, connectionist models demonstrated objective differences in the babyfaceness of emotion expressions versus neutral ones, with lower babyfaceness in anger and higher babyfaceness in surprise, an expression that should be comparable to fear because it is a precursor with similar morphology. These results clearly demonstrate that differences in the facial maturity of emotion expressions reported by Marsh, Adams, et al. (2005) do not merely reflect judges' overlapping cultural stereotypes of babies and these emotions. Rather, they are tied to structural similarities in the facial attributes of emotion expressions and babies, which is all that is known to the connectionist model. Second, objective indicators of babyfaceness mediated trait impressions of anger and surprise expressions, for which it was argued that age-related impressions would be adaptive. These results demonstrate that the results reported by Marsh, Adams, et al. do not merely reflect unrelated parallels in impressions of certain emotions and babies versus adults. Third, the fact that babyfaceness did not mediate impressions of happy expressions also supports the Marsh, Adams, et al. hypothesis because their evolutionary argument would be undercut if babyfaceness also mediated impressions of happy people, for whom there is no obvious adaptive value to mimicking helpless babies or powerful adults. Finally, the mediating effects of babyfaceness on trait impressions were independent of strong effects of attractiveness.

The mediating effects of attractiveness did not unequivocally support a general claim that the morphology of emotion expressions evolved to serve some adaptive social function. Although the effects of attractiveness sometimes complemented those of babyfaceness, attractiveness also contributed to impressions opposite to the purported adaptive impressions elicited by variations in facial maturity. Although one can construct a functional account of these opposing effects, variations in the attractiveness of emotion expressions may not have an adaptive social value for the people expressing the emotions.

The finding that both attractiveness and babyfaceness mediated impressions of some emotion expressions is consistent with other evidence for an attractiveness halo effect and a babyface overgeneralization effect. Research documenting the attractiveness halo effect has shown that nonemotional faces of all ages, races, and either sex that are judged more attractive are also perceived as more affiliative and dominant although the latter varies with perceiver culture (Langlois & Stephan, 1977; Wheeler & Kim, 1977; Zebrowitz, Montepare, & Lee, 1993). The present results show that these well-documented impressions of neutral faces that vary in attractiveness hold equally true for emotional faces that vary in attractiveness. Research documenting the babyface overgeneralization effect has shown that nonemotional faces of all ages, races, and either sex that show more structural resemblance to babies are perceived as more affiliative and less dominant, paralleling impressions of actual babies (Montepare & Zebrowitz, 1998; Zebrowitz, 1997; Zebrowitz et al., 2003). The present results show that the same is true for emotional faces. However, the Marsh, Adams, et al. (2005) suggestion that variations in the babyfaceness of emotion expressions is an evolved adaptation provides a significant twist on the babyface overgeneralization hypothesis, which makes no such claim regarding adults who vary in structural babyfaceness.

Some limitations to the current research should be noted. One is that we did not use fear faces, which would provide the most direct test of the Marsh, Adams, et al. (2005) hypothesis. Although, as discussed above, there is good reason to expect similar results for surprise and fear faces, it is possible that the support we provided for the Marsh, Adams, et al. hypothesis using surprise faces would have been even stronger had we used fear faces. Research using samples of male and female faces with expressions photographed under the same conditions is also needed to investigate whether the effects hold equally true for faces of both sexes. Another limitation concerns the quality of the emotion expressions, which were not validated by any objective standard, such as the Facial Action Coding System (FACS; Cohn & Ekman, 2006), and were not "pure" expressions of a single emotion. For example, the neutral expressions were rated relatively high in anger. However, this particular limitation actually works against our hypotheses that anger and neutral faces would be differentiated by their babyfaceness in Study 1 and by their perceived traits in Study 2, both of which were confirmed. Perhaps more problematic than the impurity of the posed expressions is the question of whether the effects obtained using posed facial expressions will generalize to dynamic, spontaneous expressions, which would seem necessary to support the argument that the morphology of emotion expressions evolved to mimic babies or adults due to the adaptive value of that resemblance. Support for that argument also requires that the present findings generalize to faces and perceivers of other races because the evolutionary basis of the hypothesis would predict universality. Evidence for cross-cultural agreement in the identification of emotion expressions (Ekman, 1972), babyfaceness (Zebrowitz et al., 1993), and attractiveness (Dion, 2002) suggests that the present findings would replicate across faces and judges of different races. Nevertheless, empirical documentation of such generalizability is desirable, particularly because there are some cultural variations in emotion recognition in addition to the overall agreement (Elfenbein, & Ambady, 2002).

That babyfaceness and attractiveness only partially differentiated the emotion expressions and only partially mediated the trait impressions indicates that other factors contribute to the morphology and impressions of emotion expressions. One other influence on trait impressions is undoubtedly behavioral expectations derived from previous interactions with people expressing different emotions, what Secord (1958) called "temporal extension." Other sources of influence on morphology are the nonsocial ones that affect the person expressing the emotion rather than the perceiver. For example, research has demonstrated that facial actions influence emotional experience, as predicted by Zajonc, Murphy, and Inglehart's (1989) theory of emotional efference. It may be that the morphology of certain facial expressions evolved to vary in the physiological consequences for the expresser in addition to the social reactions elicited from the perceiver.

NOTES

1. We did not analyze activation of the adult unit because it is a reciprocal of the baby unit activation.

2. For the interested reader, we note the following qualifications to the emotion main effects by sex of face: In Study 1, the manipulation check ratings of anger and surprise were stronger for female faces, although highly significant for both sexes; the tendency for surprise expressions to elicit more activation of the eye-region baby unit than neutral expressions was stronger for female than male faces, with the latter marginally significant; the tendency for anger expressions to be less attractive than neutral ones was significant only for female faces; and the tendency for happy expressions to be more attractive than neutral ones was stronger for female than male faces, with the latter marginally significant. In Study 2, female faces were rated significantly lower in dominance than male faces in the angry and happy models and significantly higher in affiliation in the surprise and happy models. Although there was a significant Face Sex × Emotion interaction for perceived affiliation in the anger model, the lower perceived affiliation of anger than neutral expressions was highly significant for both sexes. More details about these results are available from the first author.

3. Additional findings of interest were that surprise expressions were rated more happy than angry, and happy expressions were rated more surprised than angry, which is consistent with the fact that surprise can be a precursor to happiness. Also, although neutral faces were rated lowest of all expressions in surprise and happiness, they were rated second only to anger faces in anger, which is consistent with other evidence concerning similar reactions to neutral and anger expressions (e.g., Vrana, 2004).

4. Although Montepare and Dobish (2003) found that surprise expressions elicited impressions of high power in contrast to the lower power perceived in fear expressions, they attributed this result to a confusion between surprise and happy expressions rather than between surprise and fear expressions in their study. Given our Study 1 results, we expected surprise to yield impressions similar to fear expressions in Study 2.

5. In the interest of brevity, we have not reported results for ratings of intelligence and health for which we had no a priori predictions because these traits had not been included in the dominance or affiliation composites in previous research investigating impressions of emotion faces. Compared with neutral expressions, surprise was perceived as less intelligent, anger was perceived as less healthy, and happy was perceived as more intelligent and healthy, but none of these effects were mediated by their resemblance to babies.

6. The gamma coefficient for expression was not strengthened with attractiveness controlled in Model 2a despite its suppressing effect on impressions of happy expressions' low dominance because the simultaneous control of babyface ratings exerted an influence in the opposite direction, even though babyface ratings failed to qualify as a mediator. The gamma coefficient for expression with only babyfaceness added to the model, $\gamma = -.11$, $p < .10$, was strengthened to $\gamma = -.14$, $p < .05$ when attractiveness was also added to the model, revealing the suppression effect.

REFERENCES

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Calder, A. J., Burton, A. M., & Miller, P. (2001). A principal component analysis of facial expressions. *Vision Research*, *41*, 1179-1208.

Cohn, J. F., & Ekman, P. (2006). Measuring facial action. In J. Harrigan, R. Rosenthal, & K. R. Scherer, (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 9-64). New York: Oxford University Press.

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience*, *14*, 1158-1173.

Darwin, C. E. (1965). *The expression of the emotions in man and animals* (Rev. ed.). Chicago: University of Chicago Press. (Original work published 1872) de Bonis, M. (2003). Causes and reasons in failures to perceive fearful faces. In M. Katsikitis (Ed.), *Human face: Measurement and meaning* (pp. 149-167). Dordrecht, the Netherlands: Kluwer Academic.

Dion, K. K. (2002). Cultural perspectives on facial attractiveness. In C. Rhodes & L. A. Zebrowitz (Eds.), *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Westport, CT: Ablex.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109-128.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation* (pp. 207-283). Lincoln: University of Nebraska Press.

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

Elfenbein, H., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*, 203-235.

Enlow, D. H. (1990). *Facial growth* (3rd ed.). Philadelphia: Harcourt, Brace Jovanovich.

Fellous, J. M. (1997). Gender discrimination and prediction on the basis of facial metric information. *Vision Research*, *37*, 1961-1973.

Hildebrandt, Katherine A. (1983). Effect of facial expression variations on ratings of infants' physical attractiveness. *Developmental Psychology, 19,* 414-417.

Karraker, K. H., & Stern, M. (1990). Infant physical attractiveness and facial expression: Effects on adult perceptions. *Basic and Applied Social Psychology*, *11*, 371-385.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, *8*, 115-128.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*, 165-182.

Langlois, J. H., & Stephan, C. (1977). The effects of physical attractiveness and ethnicity on children's behavioral attributions and peer preferences. *Child Development, 48,* 1694-1698.

Lundqvist, D., & Öhman, A. (2005). Emotion regulates attention: The relation between facial configurations, facial emotion, and visual attention. *Visual Cognition*, *12*, 51-84.

Marsh, A. A., Adams, R. B., Jr., & Kleck, R. E. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin*, *31*, 73-86.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*, 119-124.

McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology*, *17*, 211-224.

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions, *Journal of Nonverbal Behavior*, *27*, 237-254.

Montepare, J. M., & Zebrowitz, L. A. (1998). "Person perception comes of age": The salience and significance of age in social judgments. In M. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 30, pp. 93-163). San Diego, CA: Academic Press.

Mueser, K. T., Grau, B. W., Sussman, M. S., & Rosen, A. J. (1984). You're only as pretty as you feel: Facial expression as a determinant of physical attractiveness. *Journal of Personality and Social Psychology, 46,* 469-478.

O'Doherty, J., Winston, J., & Critchley, H. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. The cognitive neuroscience of social behavior [Special issue]. *Neuropsychologia*, *41,* 147-155.

Posamentier, M. T., & Abdi, H. (2003). Processing faces and facial expressions. *Neuropsychology Review*, *13*, 113-143.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Reis, H. T., Wilson, I. M., Monestere, C., Bernstein, S., Clark, K., & Seidl, E. (1990). What is smiling is beautiful and good. *European Journal of Social Psychology*, *20*, 259-267.]

Secord, P. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and Interpersonal Behavior* (pp. 300-315). Stanford, CA: Stanford University Press.

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow: A salient social signal. *Emotion*, *2*, 288-296.

Todorov, A., Mandisodza, A. N., & Goren, A. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*, 1623-1626.

Valentine, T. (Ed.). (1995). *Cognitive and computational aspects of face recognition*. London: Routledge.

Vrana, S. R. (2004). Reactions to facial expressions: Effects of social context and speech anxiety on responses to neutral, anger, and joy expressions. *Biological Psychology*, *66*, 63-78.

Wheeler, L., & Kim, Y. (1997). What is beautiful is culturally good: The physical attractiveness stereotype has different content in collectivistic cultures. *Personality and Social Psychology Bulletin, 23,* 795-800.

Zajonc, R. B., Murphy, S. T., & Inglehart, M. (1989). Feeling and facial efference: Implications of the vascular theory of emotions. *Psychological Review*, *96*, 395-416.

Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In N. MacRae, M. Hewstone, & C. Stangor (Eds.), *Foundations of stereotypes and stereotyping* (pp. 79-120). New York: Guilford.

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview.

Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*, 194-215.

Zebrowitz, L. A., & Montepare, J. M. (2005). Appearance DOES matter. *Science*, *308*, 1565-1566.

Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don't all look alike: Differentiating same versus other race individuals. *Journal of Personality and Social Psychology*, *65*, 85-101.

Received August 1, 2005
Revision accepted October 20, 2006