The Discrimination of Faces and Their Emotional Content by Chimpanzees (*Pan troglodytes*)

LISA A. PARR

Division of Psychobiology, Yerkes National Primate Research Center, Atlanta, Georgia 30329, USA

ABSTRACT: The ability to recognize and discriminate conspecific faces and facial expressions has played a critical role in the evolution of social communication. Darwin was one of the first to speculate that human and nonhuman primate facial expressions share similar mechanisms for production and functions in expressing emotion. Since his seminal publication, numerous studies have attempted to unravel the meaning of animal signals, with the most success coming from the field of vocal communication, where researchers have identified the referential and emotional nature of specific vocalizations. Studies specifically addressing nonverbal facial displays, however, have faced numerous methodological challenges, including how to objectively describe facial movements and how to study the perception and production of these signals within a social context. In this paper, I will review my studies on chimpanzee face recognition, their ability to categorize facial expressions, and the extent to which chimpanzee facial expressions may convey information about emotion. Finally, recent studies from my lab have begun to address the role of auditory and visual cues in facial expression categorization. Chimpanzees were given the task of matching expressions according to which sensory modality was more salient, the visual or auditory component. For some expressions the visual modality was preferred, while for others the auditory modality was preferred. These data suggest that different social and ecological pressures may shift attention towards one sensory modality over another, such as during long-distance communication or emotional conflict.

KEYWORDS: face recognition; communication; facial expressions; chimpanzee; social cognition; emotion

Address for correspondence: Lisa A. Parr, Division of Psychobiology, Yerkes National Primate Research Center, 954 Gatewood Rd., Atlanta, GA 30329. Voice: 404-727-3653; fax: 404-727-8470.
parr@rmy.emory.edu

Ann. N.Y. Acad. Sci. 1000: 56–78 (2003). © 2003 New York Academy of Sciences.
doi: 10.1196/annals.1280.005

Faces are highly salient social stimuli for many animal species, including sheep, primates, and birds.[1–4] In humans, faces provide viewers with rapid access to information about another individual's age, sex, individual identity, and emotional state.[5–9] The ability to use the information present in faces and respond to it discriminatively has been critical for the evolution of social communication.[10–12] In primate evolution, for example, there has been an increasing trend towards larger and more complex social groups in which individuals rely less on olfactory than visual cues, such as facial signals, for communication.[11,13] As groups became larger, the ability to acquire social knowledge by recognizing and remembering familiar individuals and their relationships with other group members became highly advantageous.[14–16] Individuals do not simply respond to specific social stimuli in fixed, invariant ways but interact within a fluid social environment that is constantly changing depending on the behavior and motivation of others and their inter-individual relationships.[17] This is particularly important for chimpanzees because of their fission-fusion society, in which individuals travel in small parties that can change composition frequently and periodically come together into a larger group.[18,19] Therefore, individuals must not only be capable of flexibility in their own social interactions, but be able to monitor the relationships of others in order to survive in a constantly changing social environment.

Over the last several decades, research on the recognition of faces and affective signals has been on the rise.[6,21,22] It has even been proposed that a specific area of the brain responds selectively to faces compared to other complex visual stimuli. The fusiform gyrus, or fusiform face area (FFA)—particularly the right fusiform gyrus—is an area of the ventral temporal lobe that responds selectively to the presentation of faces, as opposed to other stimuli.[22] More recently, it has been shown that this area is particularly sensitive to stimuli for which subjects have developed considerable expertise, such as faces.[23] These studies, in particular, have led to a widespread debate over whether face processing is innate, subserved by a domain-specific neural module, as has been suggested of the FFA; or is learned through experience. The studies of Gauthier and colleagues along with recent studies in human infants provide compelling evidence for the role of learning in the development of face recognition.[24–27]

Primate evolution is also marked by an elaboration of the mimetic facial muscles used for the production of facial expressions, resulting in greater variability in the form and number of expressions that are present in more recently evolved species.[14,28] Darwin was one of the first to speculate that the facial expressions of animals and humans may have similar origins and serve similar functions. In his seminal 1872 publication, *Expression of Emotions in Man and Animal*, Darwin described the facial expressions and vocalizations of nonhuman primates in great detail, speculating about their origins as involuntary actions of the nervous system and their associated emotional content.[29]

In this volume, researchers describe the current state of expressive behavior and communication among humans and other animals. Seyfarth and Cheney neatly summarize the debate over whether animal vocalizations convey referential or emotional meaning, concluding that emotional signals can come to convey referential information depending on how reliably the signal is produced in a given context and how specific the signal is to a given referent.[30] But they also raise an important point that is often ignored in the literature on animal communication: expressive displays can be studied in terms of the immediate circumstances that provoke them and in terms of the specific physiological substrates that give them their form, such as the anatomy of the vocal cords and larynx, and the innervation of the facial musculature. These production end mechanisms were those emphasized by Darwin.[29] Expressive displays can also be analyzed in terms of their impact on other social agents, or receivers. A display may tell a receiver something about the motivational state of the sender, something about the immediate environment, or both.

Numerous studies have examined the information content of animal vocalizations, both in the wild and in captivity. Few studies, however, have focused specifically on the facial component of these signals. Ethologists have typically described facial expressions according to individual movements or specific action patterns, staying away from descriptions that imply function or emotionality. De Waal (this volume) describes the process whereby expressive displays become ritualized, divorced from their original function (Darwin's Principle I, serviceable action) to serve a new function. Through this process, displays become very stereotypical and are easily recognized by receivers to maximize their communicative message. Because of these stereotypical movements, researchers have been able to trace the presence of specific facial expressions in related species. This has led to the identification of several facial expressions in macaques and chimpanzees, particularly those that occur during play (e.g., the relaxed open-mouth face) and submission (e.g., the bared-teeth display), that appear to be homologous with the expressions of laughter and smiling in humans (de Waal, this volume).[31,32]

Additionally, de Waal describes how some nonhuman primate expressions, typically those of chimpanzees, have many different elicitors and can be used in many different contexts, unlike the referential specificity of the alarm vocalizations described by Seyfarth and Cheney. Instead, these facial signals may convey information about the motivational state of the sender, which can be similar across different contexts. Thus, the ability of chimpanzees to accurately interpret the referent for a specific gesture—for instance, the begging gesture referred to by de Waal (this volume)—requires a cognitive evaluation of the immediate social context. To accurately decode the meaning of a chimpanzee's facial expression, such as a bared-teeth display, requires a perhaps more challenging cognitive evaluation of the individual's motivation, which is, in turn, dependent on factors such as the immediate social context, the presence and identity of other individuals, the relationship of these individuals to the signaler, and the circumstances leading to the display.

Despite the importance of understanding the evolution of communication and social cognition in species closely related to humans, very few studies have taken a comparative approach to empirically examine face and affect recognition in nonhuman primates. Such an approach would be particularly beneficial because nonhuman primates share a large percentage of their genetic material with humans, they have a long period of postnatal development, have large brains relative to their body size, live in complex social environments, and exhibit advanced cognitive abilities. As stated earlier, the ability to recognize and keep track of individuals and their social relationships is critical for survival and requires early social experience.[33] Experimental studies of the meaning of primate vocalizations using playback studies both in the wild and in more controlled captive settings have focused on both the quality of the sender's signal and how it is perceived by the receiver; however, few studies have focused specifically on primate facial expressions.[15] My research concerns the perception and categorization of primate faces and facial expressions in nonhuman primates. In this paper, I will review several studies on face discrimination and facial expression categorization in chimpanzees and summarize results from a more recent study on the role of auditory and visual modalities in facial expression categorization.

FACE PROCESSING BY CHIMPANZEES

Previous studies of chimpanzee social cognition have examined subjects' ability to discriminate the faces and facial expressions of unfamiliar conspecifics from photographs. These studies employed a computerized joystick testing paradigm whereby subjects select images on a computer monitor by moving a joystick-controlled cursor. Numerous species have now been trained in this particular paradigm using software developed at the Language Research Center, Georgia State University (Atlanta, GA).[34,35] Images in our studies were presented using a matching-to-sample (MTS) format, whereby a sample image is matched to one of two comparison images. One comparison resembles the sample on some predetermined stimulus dimension, while the other does not match. Because the MTS rule does not vary—subjects are always required to select the comparison image that best matches the sample, it is possible to address additional questions by varying the dimension of stimulus similarity without the need to retrain subjects. One could, for example, match faces according to the identity, sex, or facial expression of the individual presented, which makes the MTS a very powerful and versatile paradigm for studying social cognition in nonverbal organisms.

After first being trained to control the movements of the joystick and cursor and to match abstract shapes presented in a MTS format, chimpanzees were presented with the task of discriminating between two black and white photographs of unfamiliar conspecifics. The correct pair of stimuli was identical photographs, while the nonmatching comparison stimulus was a photograph of another individual. Chimpanzees (N = 6) performed this task above chance on the very first presentation.[4] Although, they performed this task extremely well, the task itself did not specifically address whether the chimpanzees were processing the images as faces per se or whether they were simply viewing the photographs as complex grey-scale patterns and matching them accordingly. So the task was changed to address the ability of chimpanzees to discriminate specific individuals from their facial features. This task presented two different photographs of the same individual as the correct pair, while the nonmatching comparison showed a third photograph of another individual. Correct performance was now dependent on recognizing facial similarities and not similar features of the photographs themselves. With no additional training, subjects performed above chance on this task after two repetitions of each trial. So it appeared that subjects required exposure to each trial at least once before they comprehended that the dimension of matching had changed from the identity of the photograph to the identity of the individual depicted in the photograph.[4] We have performed numerous other studies to assess the specific cognitive strategies used by chimpanzees to discriminate faces. These studies have included testing whether chimpanzees show the face inversion effect, whether they recognize faces when certain facial features have been masked or manipulated,[4] and whether chimpanzee faces convey information about kinship.[36] We have also examined the way in which chimpanzees categorize their facial expressions (see below). Furthermore, we have since replicated most of these initial findings using color presentations of digitized video.[37] In these tasks, the sample stimulus is a short, five-second video clip of an unfamiliar chimpanzee, and the correct comparison image is a photograph of that individual. These and other studies have led to the firm conclusion that conspecific faces represent highly salient and discriminable stimuli for chimpanzees, despite being presented as black-and-white, static photographs; and that these are discriminated in ways that closely resemble human face recognition processes.[38–41]

FACIAL EXPRESSION PROCESSING BY CHIMPANZEES

Initially, we presented chimpanzees with the task of matching two photographs of different individuals making the same facial expression. The nonmatching comparison was a neutral portrait of a third individual. Thus, the dimension of matching in this task changed from the identity of specific individuals to the particular facial expression being made. As in previous face matching tasks, the facial expressions were black and white photographs of individuals who differed in age and sex and were shown with different head

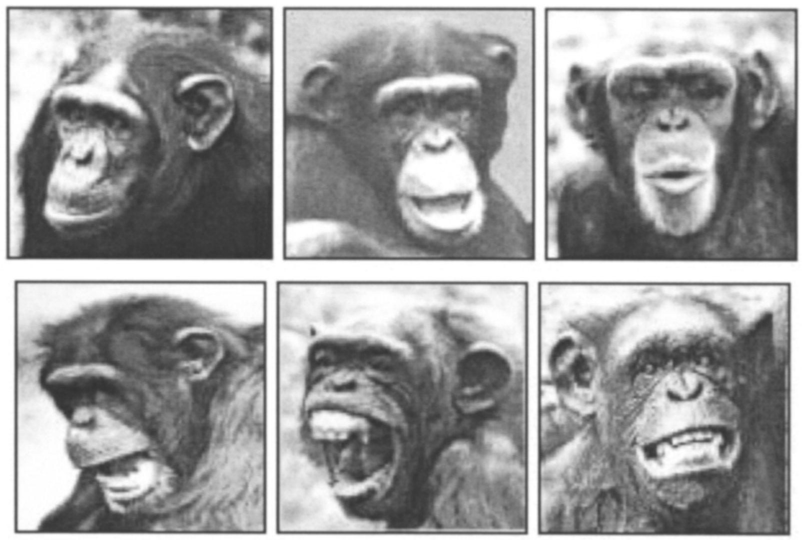

FIGURE 1. Examples of the six chimpanzee facial expressions used in the face matching experiment. **From left to right, top to bottom:** neutral portrait, relaxed-lip face, pant-hoot, play face (relaxed open-mouth face), scream face and bared-teeth display. (Courtesy of Living Links Center, Emory University.)

orientations. An example of each of the six expression types presented in this first experiment can be seen in FIGURE 1. Because individuals differ in the quality and style of their expressions, to select the correct stimulus chimpanzees were required to generalize, within the same trial, their recognition of the type of expression presented in the sample to another individual making the same expression.

FIGURE 2 shows the data from this task graphed across the 5 testing sessions, in which 50 trials were presented per session (i.e., 50 trials represents only 2 repetitions of each of the 25 unique trials presented in the task).[42] Three of the five expressions were discriminated from the neutral portrait on the first testing session, including the bared-teeth display, the play face, and the scream. Thus, the rate of acquisition for facial expression discriminations was comparable to that found for the individual recognition task. Interestingly, the figure shows that the relaxed-lip face was never discriminated above chance from the closed-mouth neutral portrait despite the fact that the relaxed lip is a distinctive feature not present in neutral faces. Subjects did not appear to use this feature to aid them in their discrimination performance and never exceeded 50% performance even after 250 trials, or 10 repetitions.

To explore the role of distinctive features in facial expression categorization, we presented a second task that combined each of the five expression

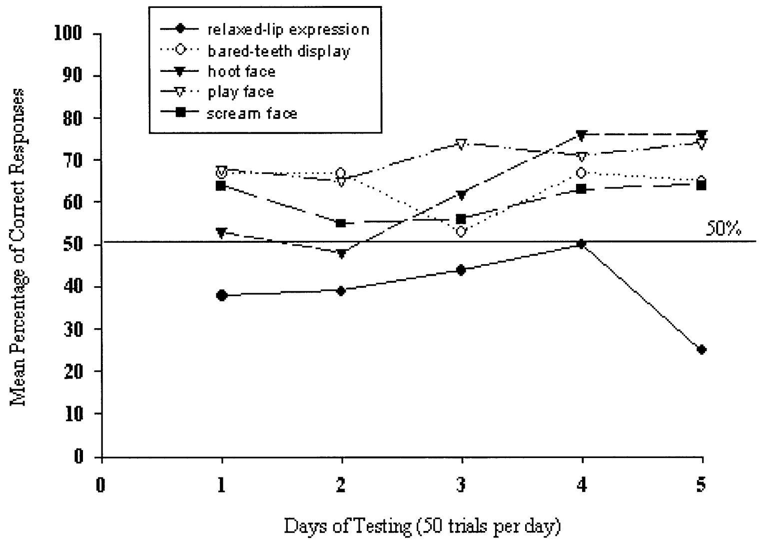

FIGURE 2. Performance of subjects over the first five expression discrimination testing sessions. Three of five expressions were discriminated from the neutral portrait on the first testing session, but the relaxed-lip face was never discriminated above chance levels from the plain neutral portrait. The *50% line* indicates chance levels.

types (we did not use the plain neutral face in this task) with every other expression, totaling 20 different expression dyads. We then selected specific facial features, such as teeth visibility, mouth shape, and eye shape, and rated whether the two expressions presented in each trial shared features in common (e.g., 3 or 4 features shared), or whether the features were distinct (e.g., <2 features shared). Ten trials of each type, similar and distinct, were identified. FIGURE 3 shows an example of each trial type. We then compared whether facial expression discriminations differed according to the number of features the two expressions shared. If chimpanzees were attending to specific salient features when discriminating facial expressions (they did not appear to be doing this in the previous task), they would be expected to perform better on the distinct trials than the similar trials. This was, in fact, the overall finding: subjects performed significantly better discriminating pairs of expressions that looked distinct than they did pairs of expressions that looked similar.[42] However, the pattern was not consistent for all of the 20 dyadic combinations of expression types. Some expressions, like the scream, were discriminated well regardless of whether they were paired with an expression that looked similar or distinct. Thus, overall it appeared as though chimpanzees were relying on something other than, or in addition to, distinctive visual features to discriminate facial expressions; or that visual features are differ-

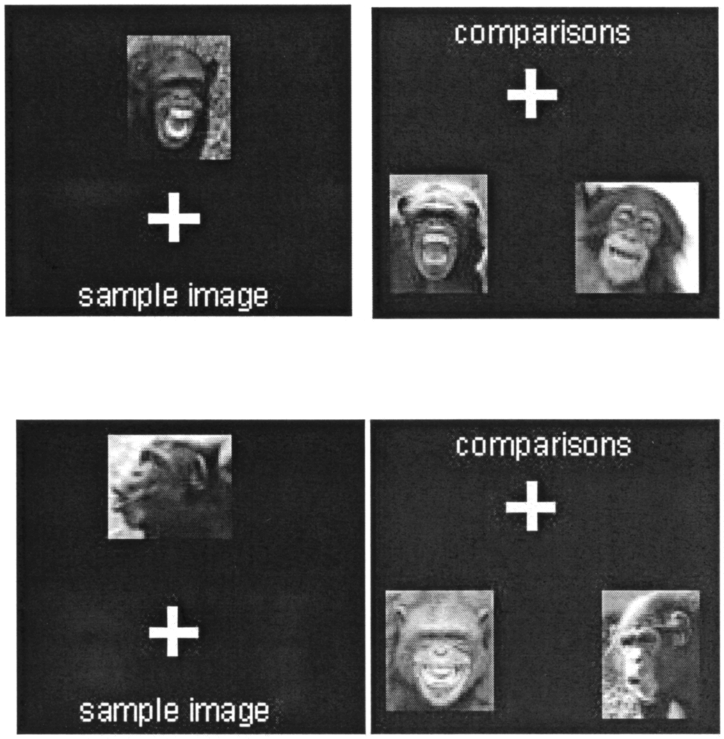

FIGURE 3. An example of two trial types used in chimpanzee facial expression categorization tasks. The **top images** show a trial that combines a scream face with a play face (**bottom right**). These two expressions have features in common and look similar. The **bottom images** show a trial that combines a pant-hoot with a bared-teeth display (**bottom left**). These two expressions do not share many features and look distinct.

entially salient depending on the expression type. This led us to speculate about the role of multimodal features, such as dynamic movement, vocalizations, and specific visual features in contributing to the identity and salience of chimpanzee facial expressions. These typically dynamic, affective expressions would best be presented using short video displays to capture the range of features present and would possibly help to identify the feature or features that are critical for facial expression categorization in this species.

A follow-up study presented chimpanzees with short digitized video scenes of conspecific facial expressions, approximately 5 seconds in length. These contained the vocalizations made by the subject during that display and, where possible, did not contain any conflicting signals, such as different expressions from other nearby group members. Using videos of facial expressions as the sample stimuli, we repeated the dyadic version of the facial ex-

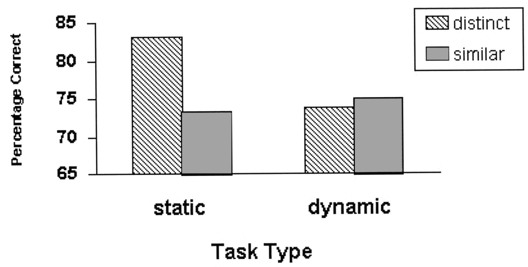

FIGURE 4. A comparison of performance on similar and distinct expression matching trials using both static and dynamic (video) sample images. Subjects performed significantly better on distinct than on similar trials in the static version of the task, but this advantage was not found when the sample images were changed to video displays.

pression categorization task described above. The sample showed a video of an unfamiliar individual's facial expression, while the correct comparison was a black-and-white photograph of a different individual making the same facial expression as shown in the sample video. The nonmatching comparison showed a different expression made by a third individual. After receiving approximately 200 trials on this task (i.e., 10 repetitions of each of the 20 novel trials), the subject's mean performance on similar versus distinct trials was statistically compared. Contrary to our previous finding, that subjects were significantly better at discriminating pairs of expressions that looked distinct than those that looked similar, no such effect was found when the sample stimulus was a video of a facial expression. FIGURE 4 shows the performance on these two trial types.[38] It seemed that presenting the sample as a video eliminated any advantage conferred from the presence of distinctive features, but without improving overall performance.

THE ROLE OF AUDITORY AND VISUAL FEATURES

A more recent direction in my laboratory has been to expand these findings by examining the salience of the auditory and visual modalities in discriminations involving facial expressions. This was done by questioning whether chimpanzees would show a preference towards one modality over the other when discriminating specific facial expressions in a multimodal version of the MTS task. Four expressions were chosen because they are typically accompanied by distinctive vocalizations.[43] These included the following vocal and visual elements: screaming and scream faces, pant-hooting and pant-hoot faces, laughter and play faces, and low-intensity screams and bared-teeth displays.

Stimuli in this experiment were combined to form one of four categories; congruent, multimodal, cross-modal, and intramodal trials. Congruent trials paired a video of a facial expression and its accompanying vocalization with a still photograph of that facial expression as the correct comparison. The congruent audiovisual combinations included those visual-auditory pairings listed above. The vocalizations used in these congruent trials were the original vocalizations that accompanied that video scene, so the audio and video tracks in the sample presentation were temporally synchronized.

Cross-modal trials paired an audio recording of the vocalization that accompanied each of the facial expressions listed above with a still photograph of that same expression (i.e., laughter paired with a photograph of a play face). While the audio clip played, the computer monitor remained black. The cross-modal trials presented a different example of the vocalization from each category than that used in the congruent trails, so there was no repetition of the vocalizations subjects heard during the congruent trials.

Intramodal trials paired a video of each facial expression, without an accompanying audio track, with a still photograph of that same expression. The intramodal trials showed the same sample videos as in the other categories, but with the audio removed. The nonmatching comparison stimuli in all categories listed above were neutral portraits. Correct responses were to select the photograph depicting the expression represented in the sample video, regardless of the modality in which it was presented. Thus, the dimension of matching in these studies, as in other expression discrimination studies, was the expression portrayed in the sample stimulus. These three categories were considered control trials because a correct response was present in each trial.

In the multimodal trials, the conditions were slightly different than for the control categories described above. For multimodal trials, the sample facial expression videos were edited so that the audio track contained a vocalization of each of the other three expression types. Hence these trials combined incongruent audio and visual tracks (i.e., a scream face paired with laughter vocal, a pant-hoot paired with screaming, etc.). These trials paired novel examples of vocalizations from each category with the original visual component of the sample videos, so there was no overlap of any of the individual vocalizations presented in this experiment. FIGURE 5 shows an illustration of a congruent trial described above and an incongruent multimodal trial.

The comparison stimuli for the multimodal trials consisted of a still photograph depicting a facial expression that matched each sensory modality presented in the sample. Thus, one comparison matched the facial expression represented visually in the sample, while the other comparison matched the facial expression represented by the vocalization. Every combination of expressions was presented, 3 possible combinations for each of the 4 expression types (i.e., scream with laughter, scream with low-intensity scream, and scream with pant-hoot) totaling 12 individual multimodal trials. Subjects were nondifferentially reinforced for the multimodal trials, meaning they

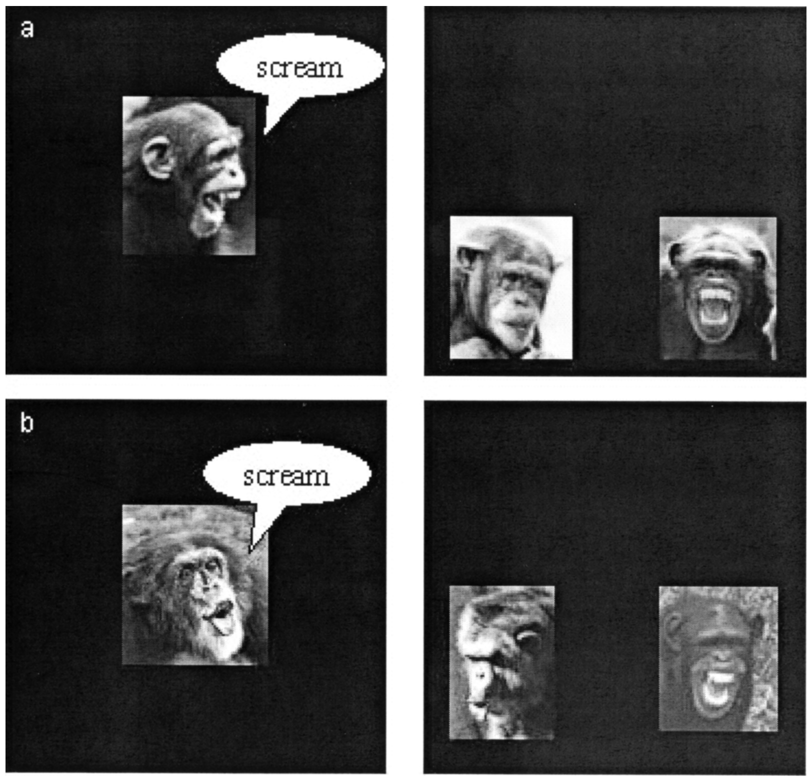

FIGURE 5. An illustration of **(a)** a congruent and **(b)** multimodal matching trial. The congruent trial shows a scream video and vocalization as the sample paired with a neutral portrait on the **left** and a scream face on the **right** (correct response). The multimodal matching trial **(b)** illustrates a pant-hoot video and scream vocalization as the sample, paired with a hoot face on the **left** (visual modality) and a scream face on the **right** (auditory modality).

were reinforced for any response they made, selecting the comparison stimulus that represented either the visual or the auditory modality presented in the sample. This enabled subjects to choose freely the modality of the sample video they preferred to match, the auditory or the visual. The modality they chose was believed to represent the most salient modality for discriminating that particular facial expression.

We hypothesized that subjects would perform best on the congruent trials, since these contained the most consistent information as to the type of expression being presented; and that they would perform well on the intramodal trials, since we knew that subjects were already competent at discriminating static photographs of facial expressions that contained no vocal information.

We hypothesized that subjects would perform above chance on the cross-modal trials; but since they had no previous experience with auditory discriminations, their performance was not expected to be as good as these as on the congruent and intramodal trials. Overall, we expected that subjects would show a preference for one modality over the other for the expressions presented in the multimodal trials. The pant-hoot, for example, is a long-distance signal; so we anticipated that the auditory modality would be more salient than the visual modality, resulting in a vocal bias for discriminations involving pant-hoots. Similarly, we anticipated that the auditory modality of laughter would be more salient than the visual signal of the play face because this latter modality is often concealed from view when animals are engaged in play wrestling. The bared-teeth display is a highly ritualized visual signal in many species.[11,44] Researchers have speculated that the silent bared-teeth expression is homologous to the human smile.[32] However, the bared-teeth display can be accompanied by a low-intensity scream or a tonal squeal, or it can be silent. Because of this, we anticipated that subjects would show a visual preference for this expression and select the bared-teeth display whenever it was shown in the sample video.

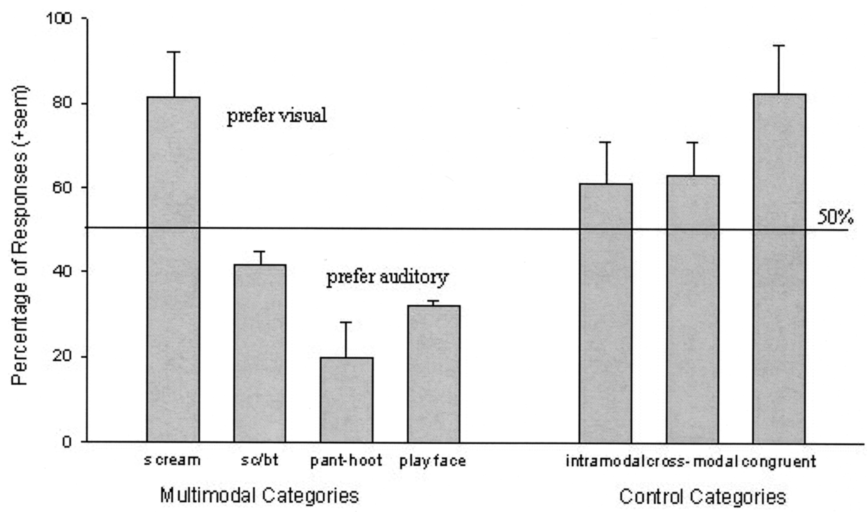

FIGURE 6. The mean performance on matching multimodal and control trials for all subjects over the five testing sessions. The *error bars* indicate the standard errors of the means. Control trials are plotted according to the percentage of correct responses—that is, selecting the comparison facial expression that matched that portrayed in the sample, regardless of sensory modality.

Finally, screams may initially appear to be an easy category to predict because they have such loud vocalizations that are easily identified. Unlike macaque screams, however, there is little evidence that chimpanzee screams convey referential information about the nature of the conflict or that they function to recruit allies during agonistic encounters.[45] Chimpanzee screams seem to be primarily affective in nature. Because the auditory component of the scream is so salient, however, we predicted that subjects would show an auditory bias for discriminations involving screams.

FIGURE 6 illustrates the performance for matching each expression type according to its visual cues, thus selecting the expression that matched the sample based on a preference for the visual modality. This shows a clear preference for the auditory modality for discriminations involving pant-hoots and play faces. Counter to our predictions, however, subjects showed a visual preference for discriminating screams, despite the fact that scream vocalizations are very loud and distinctive. Subjects showed no significant preference for the auditory or visual modality when discriminating the bared-teeth displays. These data may best be interpreted with reference to the differences between graded and discrete signals. Researchers have proposed, for example, that vocalizations given over relatively short distances, such as screams, should be graded and show more intraindividual variability than signals given over long distances.[46,47] Long-distance signals are believed to be discrete and contain little intraindividual variability compared to the more graded close-range signals. Thus, applying these hypotheses to the multimodal matching data, the auditory modality of the pant-hoot should be highly salient, while the auditory component of the scream should be highly variable, perhaps biasing discrimination preferences towards the visual modality.

MATCHING-TO-MEANING

Finally, do chimpanzee facial expressions convey emotional information, as has been demonstrated for human facial expressions?[6,48] We designed a task, named *matching to meaning*, in which subjects were presented with short emotional video scenes.[49] Subjects were then required to select one of two facial expressions that communicated a similar emotional valence, either positive or negative, as that presented in the sample video. The positive facial expression was the play face, and negative facial expressions included scream faces and bared-teeth displays. These were paired with other facial expressions, such as pant-hoot, relaxed-lip, and whimper as the nonmatching comparisons. In contrast to previous tasks—in which subjects were required to match stimuli based on their physical similarity, such as matching facial expressions—this task required subjects to match stimuli according to their emotional similarities. Five experimental scene categories and two control categories were presented. The negative experimental video categories included (1) chase = scenes of veterinarians preparing to anesthetize chimpanzees for routine medical procedures; (2) inject = scenes of chimpanzees being injected with darts and needles during these procedures; (3) dart = pictures of darts and needles; and (4) kd (knock down) = anesthetized chimpanzees. The (5) pos = (positive) experimental video category showed scenes of highly preferred items, such as bottles of grape juice and favorite toys.

The first control category (con1) showed previously learned video discriminations. These were positive and negative social scenes, such as aggression and play; they were paired with the facial expression that occurred most often in those social contexts—screams and play faces, respectively. The subjects were trained so that they performed >85% correctly prior to the matching-to-meaning experiment; this served to control for subjects' overall motivation to perform the task and conformation to the MTS rule. If, for example, subjects performed well on con1 trials but poorly on the experimental trials, then it could be concluded that they understood the goals of the task but did not see the emotional similarity between the sample and correct comparison. If, however, subjects performed poorly on both con1 and experimental trials, it may be that they were distracted, or unmotivated, during that testing session. The second control category (con2) showed scenes of chimpanzees sleeping. These were paired with facial expression that would not typically occur during sleeping (i.e., basically any expression except a neutral one). One expression was arbitrarily designated as the correct response, and subjects were rewarded if they chose that expression. Therefore, con2 controlled for performance based on subjects' history of reinforcement, because there was nothing similar between the sample and correct comparison. If subjects performed well on con2 trials in addition to the experimental trials, we could not rule out the explanation that their performance on experimental trials was not due to their history of being reinforced for selecting that expression. If, however, subjects performed well on experimental trials but poorly on the con2 trials, we could conclude that their performance was due to something other than learning based on history of reinforcement—for example, that they inferred something about the emotional similarity between the sample and correct comparison.

The performance on the MTM task can be seen in FIGURE 7.[49] Subjects performed above chance on the experimental categories after only two exposures to each trial, similar to their performance on both the individual recognition and facial expression discrimination tasks described earlier. Several categories were discriminated above chance on the very first trial. Performance on con1 trials was best, as expected, because the subjects of these trials had been previously trained to a high level of performance. The performance on con2 trials, however, never reached above chance performance after two testing sessions, nor was performance on the second testing session of con2 better than the first testing session for any experimental category. Therefore, this experiment provides evidence that the affective valence

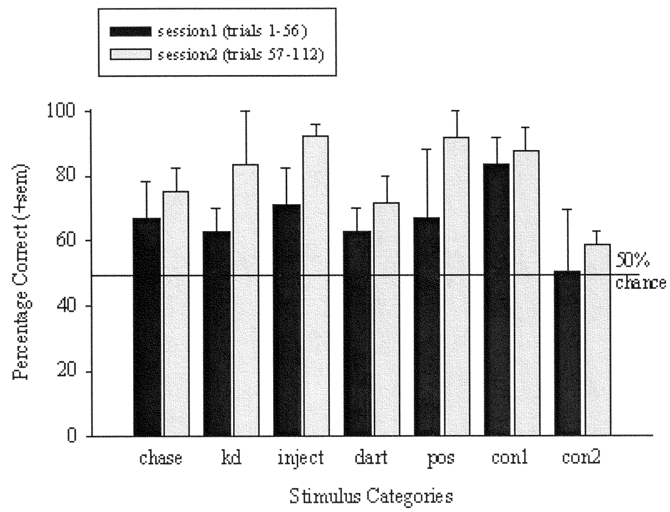

FIGURE 7. The mean performance on matching-to-meaning trials for subjects over two testing sessions. The *error bars* indicate standard errors of the means. kd = knock down; pos = positive; con = control.

of salient images may be a natural dimension by which chimpanzees can discriminate, or categorize, social stimuli. The data do not, however, go so far as to support a conscious understanding of the relationship between emotion and facial expressions. The most reasonable explanation for performance may be that because the subjects had all had experience with the situations presented in the videos, they responded to them with a form of emotional contagion, a similar state of affect to that conveyed by the images in the video.[50] This passive acquisition of emotion then biased their responses towards selecting facial expressions that shared a similar emotional content to that presented in the sample; this is similar to what has been reported in humans in tasks involving subliminal or mere exposure effects, when emotional information is acquired but not consciously perceived.[51,52] The data cannot elucidate whether subjects consciously processed the emotional information communicated by their facial expressions or whether they processed basic emotions, like anger, fear, sadness, disgust, happiness, and surprise.[6] Because of the control data and the overall design of the experiment, however, we feel confident that subjects were not using elaborate perceptual strategies to perform the task.

RECOGNITION OF FACIAL EXPRESSIONS AND THEIR EMOTIONAL CONTENT IN HUMANS

Studies of facial expression recognition in humans have had to overcome several important issues related to experimental procedures, levels of explanation, and cognitive and affective interpretation. In discussing the findings from chimpanzee facial expression studies, it would be useful to review these basic issues. First, researchers have struggled to determine the most complete and appropriate methodology for classifying human facial expressions and how these expressions are recognized by observers. What perceptual features or combination of features are important for facial expressions to be accurately identified? Additionally, what is the importance of additional contextual information that accompanies the perception of facial expressions, such as the identity of the individual making the expression, the external environment, the attitude, current mood, and personal history of the perceiver?[53]

Several lines of research have contributed significantly to overcoming most of these issues, although debate remains. This research includes methodological advancements in facial expression categorization, cognitive studies in humans, and developmental research on human infants. In order to identify facial expressions objectively and minimize the impact of contextual and subjective biases in their interpretation, Paul Ekman and colleagues developed a facial action coding system (FACS) that identifies facial expressions according to the movements of individual facial muscles.[8,48,54] The FACS has significantly advanced our understanding and identification of facial emotions, in addition to helping standardize experimental procedures in which facial expressions are used as stimuli or interpreted behaviorally.

Cognitive studies have also aided in understanding not only the relationship between facial expressions and emotion, but also how face perception may differ from the way in which facial expressions are processed. These studies have shown that facial expressions are perceived categorically, not by individual features or a configuration of physical properties.[55] Thus, while specific features and spatial orientations contribute importantly to the accuracy of facial expression recognition, this ability is not dependent on constructed bottom-up processing.[56–58] Recent studies have blended facial expressions and confirmed that facial expressions are indeed categorically perceived; but they also show important features of how categories are constituted.[49,50] Brain imaging studies have identified the contribution of different neural systems for face processing than for processing facial expressions, suggesting that facial expression recognition is not simply a subset of face processing.[61–66]

Studies of human infants have confirmed that facial expression recognition occurs in developmental stages over the first 2 years of life.[67] Infants from a very early age are able to extract certain perceptual features to discriminate some expressions over others, but it is not until after 7 months of age that infants are able to respond categorically to specific expressions.[67] Understanding the emotional significance of facial expressions may take even longer. Thus, facial expression discrimination involves more than the recognition of distinctive perceptual features. It follows a specific developmental trajectory that involves categorical processing and the recruitment of neural systems that are distinct from face processing.[66]

The recognition of facial expressions is inherently different from the recognition of perceptual patterns. Expressions are not simply perceived as mere movements, but their recognition can occur at the level of the perception of meaning.[68] Thus, integrally tied to this perceptual event is the message that the expression conveys and how individuals recognize the meaning of expressions. This is further complicated by the fact that facial expressions can convey meaning on multiple levels, including semantics, intentionality, communicative interaction, and emotion.[68] The techniques of psychophysiology, measuring subtle somatic changes in response to psychological stimuli, have contributed significantly to our understanding of how individuals respond emotionally to facial expressions. The perception of facial expressions, for example, produces a low-level motor mimicry in the perceiver that can be measured using electromyographic recordings.[69,70] These subtle movements correlate to the self-perception of emotion, suggesting an integral link between facial action and emotional experience.[71–73]

In summary, human facial expression research has identified procedures for objectively categorizing specific expressions based on muscle movements; elaborated on the neural mechanisms involved in face and facial expression processing; identified developmental stages for the discrimination, categorization, and understanding of affective facial signals; and demonstrated the integral link between facial movements and associated emotional experience. Each of these levels of analysis and explanation would be facilitated by the addition of comparative data on facial expression processing in nonhuman primate species, particularly the chimpanzees, our closest living relative. It is only through comparative efforts that we will be able to identify and understand what aspects of facial expression processing and emotional communication are indeed unique to humans and which represent older adaptations.

DISCRIMINATION OF FACIAL EXPRESSIONS AND THEIR EMOTIONAL CONTENT IN CHIMPANZEES

The data reviewed in the first half of this paper suggests that chimpanzees process facial information in ways that are similar to those of human face processing. Chimpanzees, for example, spontaneously discriminate their facial expressions when presented in an MTS task. Chimpanzees, like humans, do not automatically respond to the presentation of facial expressions with expressions of their own. This is in contrast to previous reports in monkeys, in which the presentation of photographs showing facial expressions produced emotional responses in the viewer similar to those that would occur if the presented expressions were made by a live animal.[74,75] In other words, presenting facial expressions triggered monkeys to produce facial expressions also, as though they were responding to a conspecific. The discrimination performance of chimpanzees does not appear to be dependent on the recognition of specific distinctive facial features, such as the presence of teeth in the expression, the mouth position, the eye shape, or overall head and body orientation.[42] Only a weak relationship was found for discriminations in which distinctive facial features were present, compared to discriminations in which expressions shared features in common and looked similar.[42] When facial expressions were presented as dynamic stimuli using video, subjects showed no preference for trials in which distinctive features were present. Therefore, the addition of movement, vocalizations, and context significantly changed the manner by which chimpanzees discriminated some facial expressions. Like humans, chimpanzees seemed to process conspecific facial expressions categorically. It should be noted that no studies to date have reported the successful discrimination of facial expressions in a monkey species, perhaps because they do not process facial stimuli in categorical ways and because reliance on individual facial features alone is too invariant to produce accurate results.[76]

More recent studies from my lab involving multimodal matching have furthered previous findings on facial expression processing in chimpanzees by examining the role of both auditory and visual modalities in the spontaneous categorization of conspecific facial expressions. This study demonstrated a consistent preference for one sensory modality over the other for three of the four expressions presented: for scream and scream vocalizations, visual preference; for play face and laughter, and pant-hoot and hooting, auditory preference. No significant preference was shown for either the auditory or visual modality when discriminating the bared-teeth display and low-intensity scream vocalizations. The performance of subjects was consistent across the five testing sessions, suggesting that subjects did not learn or develop these intermodal preferences over the course of repeated trials. Interestingly, the expression type that required the most training to discriminate above chance in the initial static expression matching task described above was the pant-hoot, which we found to be more salient in the vocal modality, perhaps explaining why initial performance was low for this expression type. Because their preferences were not the same for every expression type, each sensory modality appears to convey a specific salience for that expression type. Subjects did not, in general, prefer visual rather than auditory information, as has been suggested in humans.[77] This suggests that the salience for one sensory modality over another may be related to the social and ecological function of the expression types, and the salience of graded versus discrete signals.[46,47]

Finally, previous studies have demonstrated the ability of chimpanzees to match facial expressions and short emotional video scenes according to their shared emotional valence.[49] A negative video, such as a veterinary procedure, was spontaneously matched to a facial expression, such as a bared-teeth display, that represents a similar negative emotional valance to that of the video. This ability was not related to reinforcement history and seemed to be facilitated by the presence of affective information in the videos, like vocalizations. While further studies are needed to confirm or deny the presence of basic facial emotions in chimpanzees, this task provides preliminary support for the spontaneous use of emotion as a basis for discriminating naturalistic stimuli, and for the association of emotion and facial expressions in chimpanzees.[49]

Future studies are important to further understand the cognitive and perceptual processes involved in facial expression categorization in chimpanzees. Additionally, studies should continue to examine the relationship between facial expression and emotion in chimpanzees. Having a means to communicate and infer emotional information from one another would be highly adaptive, both in terms of close proximity through the use of visual signals, but also across long distance through the use of vocal signals. Studies have shown that not only can vocalizations vary according to their referential meaning, but subtle differences in acoustic parameters can convey important aspects of an individual's affective state.[44,78–80] The combination of detailed acoustic analyses and contextual categorization with controlled playback studies are necessary if we are to understand fully the complexity and potential social function of affective communication in chimpanzees.

ACKNOWLEDGMENTS

This investigation was supported by Grant RR-00165 from the NIH/NCRR to the Yerkes Regional Primate Research Center, and NSF Grant IBN-987675 to Emory University. Special thanks to Paul Ekman, Frans de Waal, and the New York Academy of Sciences; and to the Living Links Center for the use of photographic material. The Yerkes Primate Center is fully accredited by the American Association for Accreditation of Laboratory Animal Care.

REFERENCES

1. BROWN, S.D. & R.J. DOOLING. 1992. Perception of conspecific faces by budgerigars (Melopsittacus undulates): I. Natural faces. J. Comp. Psychol. **106:** 203–216.
2. BROWN, S.D. & R.J. DOOLING. 1992. Perception of conspecific faces by budgerigars (Melopsittacus undulates): II. Synthetic models. J. Comp. Psychol. **107:** 48–60.

3. KENDRICK, K. M., K. ATKINS, M.R. HINTON, *et al.* 1995. Facial and vocal discrimination in sheep. Anim. Behav. **49:** 1665–1676.
4. PARR, L.A., J.T. WINSLOW, W.D. HOPKINS & F.B.M. DE WAAL. 2000. Recognizing facial cues: individual recognition in chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). J. Comp. Psychol. **114:** 47–60.
5. BURT, D.M. & D.I. PERRETT. 1995. Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information. Proc. R. Soc. **259:** 137–143.
6. EKMAN, P. 1992. Facial expressions of emotion: new findings, new quesitons. Psychol. Sci. **3:** 34–38.
7. EKMAN, P. & H. OSTER. 1979. Facial expressions of emotion. Ann. Rev. Psychol. **30:** 527–554.
8. IZARD, C.E. 1971. The Face of Emotion. Appleton-Century-Crofts. New York.
9. TRANEL, D., A.R. DAMASIO & H. DAMASIO. 1988. Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. Neurology **38:** 690–696.
10. ANDREW, R.J. 1963. The origin and evolution of the calls and facial expressions of the primates. Behaviour **20:** 1–109.
11. ANDREW, R.J. 1963. Evolution of facial expression. Science **142:** 1034–1041.
12. BROTHERS, L. 1990. The neural basis of primate social communication. Motiv. Emotion **14:** 81–91.
13. MARLER, P. 1965. Communication in monkeys and apes. *In* Primate Behavior. I. DeVore, Ed.: 544–584. Holt, Rinehart & Winston. New York.
14. ANDERSON, J.R. 1994. Valeur ethologique des visages et des mimiques chez les primates non humains. Psychol. Franc. **39:** 345–355.
15. CHENEY, D.L. & R.M. SEYFARTH. 1990. How Monkeys See the World. University of Chicago Press. Chicago.
16. HINDE, R.A. 1976. Interactions, relationships, and social structure. Man **11:** 1–17.
17. KUMMER, H. 1971. Primate Societies. Aldine. Chicago.
18. GOODALL, J. 1986. The Chimpanzees of Gombe. Cambridge University Press. London.
19. NISHIDA, T. 1979. The social structure of chimpanzees of the Mahale Mountains. *In* The Great Apes. D.A. Hamburg & E.R. McCowen, Eds.: 73–121. Benjamin-Cummings. Menlo Park, CA.
20. FRIDLUND, A.J. 1994. Human Facial Expression. Academic Press. New York.
21. RUSSELL, J.A. & J.M. FERNANDEZ-DOLS. 1997. The Psychology of Facial Expression. Cambridge University Press. Cambridge.
22. KANWISHER, N., J. MCDERMOTT & M.M. CHUN. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. **17:** 4302–4311.
23. GAUTHIER, I., M.J. TARR, J. MOYLAN, *et al.* 2000. The fusiform "face area" is part of a network that processes faces and the individual level. J. Cogn. Neurosci. **12:** 495–504.
24. GAUTHIER, I. & C.A. NELSON. 2001. The development of face expertise. Curr. Opin. Neurobiol. **11:** 219–224.
25. LE GRAND, R., C.J. MONDLOCH, D. MAURER & H.P. BRENT. 2001. Early visual experience and face processing. Nature **410:** 890.
26. PASCALIS, O., M. DE HAAN & C.A. NELSON. 2002. Is face processing species-specific during the first year of life? Science **296:** 1321–1323.

27. TARR, M.J. & I. GAUTHIER. 2000. FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. Nat. Neurosci. **3:** 764–769.
28. HUBER, E. 1931. The Evolution of Facial Musculature and Facial Expression. Johns Hopkins Press. Baltimore, MD.
29. DARWIN, C. 1872. Expression of the Emotions in Man and Animals, 1955 edit. Philosophical Library. New York.
30. SEYFARTH, R.M. & CHENEY, D.L. 2003. Signalers and receivers in animal communication. Ann. Rev. Psychol. **54:** 145–173.
31. PREUSCHOFT, S. 1992. "Laughter" and "smile" in Barbary macaques (Macaca sylvanus). Ethology **91:** 220–236.
32. PREUSCHOFT, S. & J.A.R.A.M. VAN HOOFF. 1995. Homologizing primate facial displays: a critical review of methods. Folia Primatol. **65:** 121–137.
33. HARLOW, H.F. & C.E. MEARS. 1983. Emotional sequences and consequences. *In* Emotion: Theory, Research, and Experience, Vol. 2. R. Plutchik & H. Kellerman, Eds.: 171–197. Academic Press. New York.
34. RICHARDSON, W.K., D.A. WASHBURN, W.D. HOPKINS, *et al.* 1990. The NASA/LRC computerized test system. Behav. Res. Methods, Instrum. Computers **22:** 127–131.
35. WASHBURN, D.A. & D.M. RUMBAUGH. 1992. Testing primates with joystick-based automated apparatus: lessons from the Language Research Center's Computerized Test System. Behav. Res. Methods Instrum. Computers **24:** 157–164.
36. PARR, L.A. & F.B.M. DE WAAL. 1999. Visual kin recognition in chimpanzees. Nature **399:** 647–648.
37. PARR, L.A. & D. MAESTRIPIERI. Nonvocal communication in nonhuman primates. *In* Primate Psychology: The Mind and Behavior of Human and Nonhuman Primates. D. Maestripieri, Ed. Chicago University Press. Chicago. In press.
38. BAUER, H.R. & M. PHILIP. 1983. Facial and vocal individual recognition in the common chimpanzee. Psychol. Record **33:** 161–170.
39. BOYSEN, S.T. & G.G. BERNTSON. 1986. Cardiac correlates of individual recognition in the chimpanzee (Pan troglodytes). J. Comp. Psychol. **100:** 321–324.
40. BOYSEN, S.T. & G.G. BERNTSON. 1989. Conspecific recognition in the chimpanzee (Pan troglodytes): cardiac responses to significant others. J. Comp. Psychol. **103:** 215–220.
41. TOMONAGA, M., S. ITAKURA & T. MATSUZAWA. 1993. Superiority of conspecific faces and reduced inversion effect in face perception by a chimpanzee. Folia Primatol. **61:** 110–114.
42. PARR, L.A., W.D. HOPKINS & F.B.M. DE WAAL. 1998. The perception of facial expressions in chimpanzees (*Pan troglodytes*). Evol. Comm. **2:** 1–23.
43. VAN HOOFF, J.A.R.A.M. 1973. A structural analysis of the social behavior of a semi-captive group of chimpanzees. *In* Social Communication and Movement. M. von Cranach & I. Vine, Eds.: 75–162. Academic Press. London.
44. DE WAAL, F.B.M. & L.M. LUTTRELL. 1985. The formal hierarchy of rhesus macaques: an investigation of the bared-teeth display. Am. J. Primatol. **9:** 73–85.
45. GOUZOULES, S., H. GOUZOULES & P. MARLER. 1984. Rhesus monkey (Macaca mulatta) screams: representational signalling in the recruitment of agonistic aid. Anim. Behav. **32:** 182–193.
46. MARLER, P. 1976. Social organization, communication, and graded signals: the chimpanzee and the gorilla. *In*. Growing Points in Ethology. P.P.G. Bateson & R.A. Hinde, Eds.: 239–279. Cambridge University Press. London.

47. MORTON, E.S. 1982. Grading, discreteness, redundancy, and motivation-structural rules. *In* Acoustic Communication in Birds. D.E. Kroodsma, E.H. Miller & H. Ouellet, Eds.: 183–212. Academic Press. New York.
48. EKMAN, P. & W.V. FRIESEN. 1975. Unmasking the Face. Prentice Hall. Englewood Cliffs, NJ.
49. PARR, L.A. 2001. Cognitive and physiological markers of emotional awareness in chimpanzees, *Pan troglodytes*. Anim. Cogn. **4:** 223–229.
50. HATFIELD, E., J.T. CACIOPPO & R.L. RAPSON. 1994. Emotional Contagion. Cambridge University Press. Paris.
51. ESTEVES, F., F. PARRA, U. DIMBERG & A. OHMAN. 1994. Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. Psychophysiology **31:** 375–385.
52. MORRIS, J.S., A. OHMAN & R.J. DOLAN. 1998. Conscious and unconscious emotional learning in the human amygdala. Nature **393:** 467–470.
53. EKMAN, P. 1997. Should we call it expression or communication? Innovations Soc. Sci. Res. **10:** 333–344.
54. EKMAN, P., W.V. FRIESEN & P. ELLSWORTH. 1972. Emotion in the Human Face: Guidelines for Research and an Integration of Findings. Pergamon Press. New York.
55. ETCOFF, N.L. & J.J. MAGEE. 1992. Categorical perception of facial expressions. Cognition **44:** 227–240.
56. ARONOFF, J. & A.M. BARCLAY. 1988. The recognition of threatening facial stimuli. J. Pers. Soc. Psychol. **54:** 647–655.
57. FRANK, M.G., P. EKMAN & W.V. FRIESEN. 1993. Behavioral markers and recognizability of the smile of enjoyment. J. Pers. Soc. Psychol. **64:** 83–93.
58. KAPPAS, A., U. HESS, C.L. BARR & R.E. KLECK. 1994. Angle of regard: the effect of vertical viewing angle on the perception of facial expressions. J. Nonverbal Behav. **18:** 263–280.
59. CALDER, A.J., A.M. BURTON, P. MILLER, *et al*. 2001. A principal component analysis of facial expressions. Vision Res. **41:** 1179–1208.
60. YOUNG, A.W., D. ROWLAND, A.J. CALDER, *et al*. 1997. Facial expression megamix: tests of dimensional and category accounts of emotion recognition. Cognition **63:** 271–313.
61. HAMANN, S.B., L. STEFANACCI, L.R. SQUIRE, *et al.* 1996. Recognizing facial emotion. Nature **379:** 497.
62. KANWISHER, N., D. STANLEY & A. HARRIS. 1999. The fusiform face area is selective for faces not animals. NeuroReport **10:** 183–187.
63. KESLER-WEST, M. L., A.H. ANDERSEN, C.D. SMITH, *et al.* 2001. Neural substrates of facial emotion processing using fMRI. Cogn. Brain Res. **11:** 213–226.
64. NAKAMURA, K., R. KAWASHIMA, K. ITO, *et al.* 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. J. Neurophysiol. **82:** 1610–1614.
65. SERGENT, J. & J. SIGNORET. 1992. Functional and anatomical decomposition of face processing: evidence from prosopagnosia and PET study of normal subjects. Phil. Trans. R. Soc. Lond. B **335:** 55–62.
66. HAXBY, J.V., E.A. HOFFMAN & M.I. GOBBINI. 2002. Human neural systems for face recognition and social communication. Biol. Psychiat. **51:** 59–67.
67. NELSON, C.A. 1987. The recognition of facial expressions in the first two years of life: mechanisms and development. Child Dev. **58:** 889–909.

68. FRIJDA, N.H. 1986. Facial expression processing. *In* Aspects of Face Processing. E.H.D. Ellis, M.A. Jeeves, F. Newcombe & A. Young, Eds.: 319–325. Martinus Nijhoff. Boston.
69. CACIOPPO, J.T., L.K. BUSH & L.G. TASSINARY. 1992. Microexpressive facial actions as a function of affective stimuli: replication and extension. Pers. Soc. Psychol. Bull. **18:** 515–526.
70. DIMBERG, U. 1982. Facial reactions to facial expressions. Psychophysiology **19:** 643–647.
71. DIMBERG, U. 1987. Facial reactions, autonomic activity and experienced emotion: a three component model of emotional conditioning. Biol. Psychol. **24:** 105–122.
72. HESS, U., A. KAPPAS, G.J. MCHUGO, *et al.* 1992. The facilitative effect of facial expression on the self-generation of emotion. Int. J. Psychophysiol. **12:** 251–265.
73. LANZETTA, J.T. & G.J. MCHUGO. 1989. Facial expressions and psychophysiological correlates of emotion. *In* Emotions and the Dual Brain. G. Caltagirone & C. Caltagirone, Eds.: 92–118. Springer-Verlag. New York.
74. SACKETT, G.P. 1965. Response of rhesus monkeys to social stimulation presented by means of colored slides. Percep. Mot. Skills **20:** 1027–1028.
75. SACKETT, G.P. 1966. Monkeys reared in isolation with pictures as visual input: evidence for an innate releasing mechanism. Science **154:** 1468–1473.
76. KANAZAWA, S. 1996. Recognition of facial expressions in a Japanese monkey. Primates **37:** 25–38.
77. BURNS, K.L. & E.G. BEIER. 1973. Significance of vocal and visual channels in the decoding of emotional meaning. J. Commun. **23:** 118–130.
78. BANSE, R. & K.R. SCHERER. 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. **70:** 614–636.
79. SEYFARTH, R.M., D.L. CHENEY & P. MARLER. 1980. Vervet monkey alarm calls: semantic communication in a free-ranging primate. Anim. Behav. **28:** 1070–1094.
80. SEYFARTH, R.M., D.L. CHENEY & P. MARLER. 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science **210:** 801–803.