Understanding chimpanzee facial expression: insights into the evolution of communication

Lisa A. Parr¹ and Bridget M. Waller²

¹Division of Psychiatry and Behavioral Sciences & Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA and ²Department of Psychology, University of Portsmouth, Portsmouth, UK

To understand the evolution of emotional communication, comparative research on facial expression similarities between humans and related species is essential. Chimpanzees display a complex, flexible facial expression repertoire with many physical and functional similarities to humans. This paper reviews what is known about these facial expression repertoires, discusses the importance of social organization in understanding the meaning of different expressions, and introduces a new coding system, the ChimpFACS, and describes how it can be used to determine homologies between human and chimpanzee facial expressions. Finally, it reviews previous studies on the categorization of facial expressions by chimpanzees using computerized tasks, and discusses the importance of configural processing for this skill in both humans and chimpanzees. Future directions for understanding the evolution of emotional communication will include studies on the social function of facial expressions in ongoing social interactions, the development of facial expression communication and more studies that examine the perception of these important social signals.

Keywords: facial communication; primates; face processing; comparative research

NATURAL ETHOLOGY OF FACIAL EXPRESSION: FORM AND FUNCTION

Many animal species communicate using a variety of highly conspicuous signals, including acoustic, tactile, olfactory and visual displays that have been tuned by natural selection to impact the listener in a reliable way (Smith, 1977; Dawkins and Krebs, 1978). Among primates, the visual and auditory domains have become the two most prominently involved in social communications. There is, for example, a general assumption that facial expressions convey a variety of information about an individual’s motivation, intentions and emotions (van Hooff, 1967; Ekman, 1997; Parr, et al., 2002). As such, facial expressions are critically important for coordinating social interaction, facilitating group cohesion and maintaining individual social relationships (Parr et al., 2002). Despite the importance of facial expressions in the evolution of complex societies, there is little work comparing either the form or function of facial expression across distinct phylogenetic groups. Moreover, apart from a handful of excellent ethograms that describe the communication repertoires for a variety of species, including chimpanzees, bonobos, rhesus monkeys, capuchin monkeys and canids (Hinde and Rowell, 1962; van Hooff, 1962, 1967, 1973; Andrew, 1963; Goodall, 1968; Fox, 1969; Bolwig, 1978; Weigel, 1979; de Waal, 1988; Preuschoft and van Hooff, 1997; Redican, 1982; Parr, et al., 2005), there has been little attempt to standardize the description of their facial and vocal displays in a manner that facilitates comparative and evolutionary studies.

Fridlund (1994) has proposed that facial expressions are best understood as communicative signals and as such, researchers should focus on their functional consequences during social interaction, or how they impact the listener. Moreover, the only way to fully understand why facial expressions have evolved to convey a specific meaning is to compare similar facial expressions between evolutionarily related species, examining any factors that may have influenced their social function, including ecological pressures and social factors like dominance style and social organization (Preuschoft and van Hooff, 1997). Some facial expressions appear to be well represented across diverse taxonomic groups, making them good models for understanding social and emotional function, while others appear to be species-specific. The bared-teeth display, also referred to as the fear grin, or grimace, is one of the most conspicuous and well-studied facial expressions in ethology and has been reported in a variety of mammalian species from canids to primates. Research has shown, however, that the communicative function of this expression can differ quite broadly depending on the species, their type of social organization and social context. In wolves, for example, retraction of the lips horizontally over the teeth results in a ‘submissive grin’ which is used by cubs and subordinates when actively greeting adult conspecifics, or humans (Fox, 1969). Antithetical to this expression is a vertical lip...
retraction which is given by dominant animals during aggressive interactions, very similar facial movements but with vastly different social functions (Darwin, 1872; Fox, 1969).

Among primates, the function of the bared-teeth also has different meanings depending on the species and their type of social organization. Among macaques species that have despotic social systems characterized by strict, linear dominance hierarchies, i.e. rhesus monkeys, the bared-teeth display appears to be a signal of submission, or rank recognition in that it is only given by subordinates to higher ranking individuals (van Hooff, 1976; de Waal and Luttrell, 1985). This expression has been referred to as a formal signal of dominance in the rhesus monkey because it is highly ritualized in appearance and has long-term predictability in determining dominance relationships despite short-term variation in social contexts (de Waal and Luttrell, 1985). In this study, bared-teeth displays performed by subordinate individuals occurred most often in response to the approach of a dominant monkey, and the most frequent response was for the subordinate to withdraw from any social interaction (de Waal and Luttrell, 1985). However, the meaning of the bared-teeth display is quite different when used by species with more egalitarian social systems, including some macaques, mandrills, Gelada baboons and chimpanzees (van Hooff, 1967; Preuschoft and van Hooff, 1997). In these species, the bared-teeth display is more appeasing functions to increase social attraction and affiliation. It communicates benign intent in that the signaler wishes no harm, and that there is no risk of aggression (van Hooff, 1967; van Hooff, 1976; Waller and Dunbar, 2005). It can also occur during affiliative contexts, such as grooming, sexual solicitation and reconciliations, and thus functions to increase affiliative tendencies and reduce proximity between individuals (van Hooff, 1973; Preuschoft and van Hooff, 1997; Parr et al., 2005; Waller and Dunbar, 2005).

Numerous authors have gone on to propose that the bared-teeth display is homologous with the human smile, meaning that they are the result of common evolutionary descent (van Hooff, 1972; Preuschoft and van Hooff, 1997; Waller and Dunbar, 2005). This conclusion is based in part on the physical similarity in the appearance of the bared-teeth display and the human smile, which are both characterized by retraction of the lip corners exposing, in most cases, the upper and lower teeth. However, the homology is also based on similarity in the social function of these expressions, indicating appeasement, reassurance, increasing social bonding, and thus its important role in facilitating social cohesion among primates (Preuschoft and van Hooff, 1997). These examples suggest that while similar appearing expressions can often have different meanings, or serve different social functions, depending on the species and their type of social organization, they share a common evolutionary root. Therefore, in order to understand the evolution of human emotional expressions, like the smile, it is extremely informative to take a comparative approach and evaluate whether the form and function of these expressions are uniquely human, present in many related species, or perhaps only shared by very closely related species, like Hominoids (Fridlund, 1994). The picture is less clear for other expressions, and only a few other direct comparisons have been made, i.e. homology between human laughter and the nonhuman primate ‘play face’ (Preuschoft and van Hooff, 1967; van Hooff, 1972). This is undoubtedly due to the fact that researchers, to date, have lacked a common language, or standardized system that could be used to identify potential homologues. Ideally, any assessment of homology should be based on something other than physical appearance, so a consideration of the underlying musculature is important (Waller et al., 2006). Among primates, it is highly likely that many expressions may share a common ancestry, and thus a comparative treatment is warranted, even if the expressions look different in terms of their characteristic features.

**DEVELOPMENT OF COMPARABLE CODING SYSTEMS FOR PRIMATE FACIAL EXPRESSIONS**

Recently, researchers have developed an objective, standardized method for measuring facial movement in the chimpanzee that is directly comparable to humans, thus facilitating research on facial expression homologues (Vick et al., in press; www.chimpfacs.com). This system, referred to as ChimpFACS, is based on the well-known human Facial Action Coding System, or FACS, developed by Ekman and colleagues to provide an objective tool for studying the biological basis for human expressions and emotion (Ekman and Friesen, 1978). Both of these systems are unique in that they identify the most minimal units of facial movement according to the function of the underlying musculature. Therefore, both systems enable the objective description of facial appearance changes in the human face and the chimpanzee face that are associated with movements of the underlying facial musculature. Each movement is described using a numeric code, referred to as an Action Unit (AU), and researchers are trained to use these codes reliably through a standardized testing process, making all users of the system statistically reliable with one another (http://face-and-emotion.com). There are several advantages to using such a standardized system for comparative studies. First, because the facial musculature of chimpanzees and humans is highly comparable (Burrows et al., 2006; Waller et al., 2006), the two systems provide a basis for understanding homologous facial expression structure in the two species. Second, by orienting attention towards individual facial movements, researchers are encouraged to view the face objectively and not be biased by the humans’ natural tendency to focus on overall expression configuration (Wallbott and Ricci-Bitti, 1993; Calder et al., 2000). Thus, the face can be described, or coded, objectively using...
minimal units of measurement without interference from its holistic emotional quality. Finally, because the systems are both standardized using a graded testing procedure, researchers from different groups are now able to compare their results directly using a common nomenclature. The value of the system is obvious in the examples given above; when attempting to determine whether expressions are homologous across chimpanzees and humans, like the bared-teeth display and the smile, researchers can begin by using subjective ratings of their similarity, but then utilize the more objective, standardized assessment of their anatomical similarity by comparing individual action units in both species.

After the creation of the ChimpFACS, researchers were interested in testing its accuracy to determine whether this bottom-up approach (action unit to expression configuration) could actually validate the existing categories of chimpanzee facial expressions made previously by experts in chimpanzee communicative behavior. This was done using a discriminate functions analysis (DFA) which enabled a statistical comparison of whether the a priori expression categories could be predicted by common patterns of underlying facial movements, or patterns of AUs, coded from each expression using ChimpFACS. Over 250 facial expressions were categorized by Parr using published ethograms of chimpanzee behavior (Parr et al., 2005), and then coded using ChimpFACS by Waller, a certified ChimpFACS expert (Vick et al., in press). Nine expression categories, and 15 action units, accounted for the 250+ examples and this included an ambiguous category that was used for expressions that could not easily be classified. The results of the DFA showed a high percentage of agreement for almost all of the expression categories (Parr et al., in press). The only notable exception was the pout, which was most often identified as a pant-hoot due to a similarity in one AU movement (AU22-lip funneler) which was present in both of these expressions. Therefore, objective coding of naturally occurring chimpanzee facial expressions according to minimal facial movements using ChimpFACS produced functionally meaningful expression categories. Moreover, the analysis enabled the identification of unique combinations of muscle movements (AUs) for all expression categories. Thus, the DFA provided a clear description of prototypical facial expression configurations based on standardized muscle action (Parr et al., in press). These configurations were so predictive that naïve researchers were able to identify expression categories with over 80% accuracy solely by reading the AU configuration (Parr et al., in press). An illustration of each prototypical expression configuration and the percentage of category agreement between AU codes and a priori classifications can be seen in Figure 1.

Fig. 1 An illustration of prototypical chimpanzee facial expressions. These are listed in pairs. The example on the left side of the pair shows the Poser animated expression, while the example on the right shows a naturalistic chimpanzee expression. Under the Poser expression is the prototypical AU configuration as identified by the Discriminant Functions Analysis, and under the naturalistic expression is the percentage agreement between AU configuration and a priori classification for that category.
THE STANDARDIZATION OF CHIMPANZEE FACIAL EXPRESSIONS USING CHIMPFACS

The identification of prototypical expression configurations is extremely important for ongoing research on chimpanzee social cognition as it provides a blueprint for standardizing the images used in computerized tasks, such as expression categorization studies (Parr et al., 1998). Using the results from the DFA, researchers now have a list of prototypical facial configurations that account for each major expression category, and an objective tool in ChimpFACS for identifying these. Even with these tools, however, acquiring photographs of chimpanzee expressions at their peak intensity is extremely difficult because expressions are dynamic, often occur during highly charged social contexts where there is fast movement, and subjects are more often than not faced away from the photographer, making it difficult to capture frontal pictures, or to standardize these in terms of head orientation and posture.

To overcome the difficulties in obtaining high quality naturalistic photographs of chimpanzee expressions, researchers have now turned to custom three-dimensional (3D) animation software to create these configurations manually. There are two types of expressions shown in Figure 1, a naturalistic photograph and a cartoon-like chimpanzee face. The cartoon-like face was created using the 3D animation software Poser 6.0 (www.efrontiers.com), which enables detailed, custom animation of facial movements by allowing the user to impose an artificial muscle structure onto the face and then animating this ‘muscle’ to affect facial appearance in a highly controlled and naturalistic manner. To illustrate, the Poser expressions in the figure were made using anatomically accurate muscle structure and realistic AU movements and are posed in the prototypical facial configurations identified by the DFA (Parr et al., in press). Because chimpanzee researchers cannot position chimpanzee faces using verbal instruction, as is the primary method for creating human expression stimuli for use in emotion perception studies, the Poser chimpanzee allows researchers control over head orientation, gaze direction, expression configuration and intensity. The application of these techniques will considerably advance comparative cognition research as it eliminates the aforementioned difficulties in creating controlled stimulus libraries where each example expression is required to be photographed at peak intensity, in full-frontal posture from freely behaving chimpanzees. Moreover, in combination with the ChimpFACS, researchers can use these models to pose individual action units, realistic combinations of action units, such as those seen in Figure 1, and even artificial combinations of facial movements to specifically examine which elements of expressions are most meaningful for the chimpanzee’s emotional perception.

COMPARING HOMOLOGOUS FACIAL MOVEMENTS IN CHIMPANZEES AND HUMANS

With the development of a standardized system for measuring facial movement in chimpanzees (ChimpFACS) that is anatomically and organizationally comparable to humans (FACS), the stage is set for a more comprehensive investigation into the evolution of facial expressions and facial emotion (Parr et al., in press; Vick et al., in press; Waller et al., 2006). There are several ways to attempt such an investigation and it will likely be quite some time before there are enough data to make any definitive conclusions. At this very preliminary state, one can reasonably describe similarities in the facial movements between prototypical chimpanzee expressions and human emotional expressions (Parr et al., in press). In their study, these authors most closely matched the AUs’ configurations of the chimpanzee bared-teeth display and human smile, chimpanzee screams to human screams, the chimpanzee bulging lip face to human anger, chimpanzee laughter to human laughter, but could not find a good human equivalent for the chimpanzee pant-hoot. When humans have been asked to judge the emotional quality of chimpanzee expressions using basic emotions labels, the following comparisons have emerged: chimpanzee bared-teeth display = happiness, chimpanzee play face = happiness, chimpanzee pant-hoot = happy, chimpanzee scream = anger (Fernandez-Carriba et al., unpublished data). Thus, in these very preliminary comparisons, similarities emerged.

An alternative, and arguably a more intuitive, approach is to first match human and chimpanzee expressions based on the muscular components, and then make inferences about the emotional quality of human faces that are structurally homologous to chimpanzee expressions. Figure 2 shows such a comparison where the human images are taken directly from the FACS manual (Ekman et al., 2002) and show facial movements that are structurally homologous to the prototypical chimpanzee expressions. The identical AUs shared by the two expression examples are highlighted in bold italics. In some cases there is an extra movement in either the human or chimpanzee example. Not surprisingly, one of the most striking comparisons here is between the chimpanzee bared-teeth display and its human equivalent. As discussed earlier, previous researchers have suggested a homology between the chimpanzee bared-teeth display and the human smile (van Hooff, 1972; Preuschoft and van Hooff, 1995; Waller and Dunbar, 2005; Parr et al., in press). However, the human bared-teeth expression in Figure 2 strongly resembles a grimace, or a forced smile, but it does not give the countenance of happiness. Contrast this to the human play face, which gives a strong impression of happiness and supports others who have suggested that this expression is homologous to laughter in humans (van Hooff, 1972; Preuschoft and van Hooff, 1995; Waller and Dunbar, 2005; Parr et al., in press). One explanation may be that genuine human smiles, those associated with emotional happiness...
or enjoyment, include an AU6-cheek raiser, which is missing in the human bared-teeth expression (Ekman and Friesen, 1982; Ekman et al., 1990; Frank et al., 1993). Interestingly, when people have been asked to rate human faces that contain either individual movements (single AUs) or combinations of movements (multiple AUs), researchers report that the configuration AU10 + 12 + 16 + 25 (the same movements as in the prototypical chimpanzee bared-teeth face) is most often described as showing fear (Wallbott and Ricci-Bitti, 1993). While it might seem unlikely that the addition of a single AU (the AU6) could change the entire emotional interpretation of an expression, the data strongly support this conclusion. Because humans process facial expressions configurally, and do not selectively attend to individual features, the configuration AU10 + 12 + 16 + 25 is interpreted differently than AU6 + 10 + 12 + 25 (Wallbott and Ricci-Bitti, 1993). If the production of AU6 does indeed correlate with genuine happiness, then it is arguably very important to distinguish between these two expressions, and thus configural processing has a clear and crucial function. Interestingly, the human scream face in Figure 2 contains an AU6, and it does not appear to have a strong negative emotional countenance. These findings provide a strong argument for the need to more fully understand how chimpanzees process their own facial expressions, what the role of configural processing is in expression categorization and whether chimpanzees are as sensitive to individual movements as humans. Therefore, in order to more fully understand the evolution of communication and the social function of facial expressions, a comparative perspective on facial expression categorization is needed.

PREVIOUS STUDIES ON FACIAL EXPRESSION CATEGORIZATION BY CHIMPANZEES

Studies of emotional communication have typically taken one of two approaches, either they focus on the perception of the display, how it is processed and an examination of salient features or configuration necessary for accurate interpretation, or they focus on the social function of the expression, how it affects the behavior of a receiver within a social environment. Ideally, it is the combination of these two approaches that will provide the most complete understanding of emotional communication in ongoing social interactions. Therefore, two critical questions for understanding the evolution of emotional communication are; first, whether chimpanzees and humans use similar perceptual cues to discriminate among facial expression categories, and second, what is the functional outcome of these expressions for ongoing social interactions. In addressing the first approach, human studies have shown strong configural preferences for both face identity recognition and facial expression categorization (Calder et al., 2000; Etcoff and Magee, 1992; Wallbott and Ricci-Bitti, 1993). Configural cues refer to the relative size, shape and spatial arrangement of features within the face (Maurer et al., 2002). While numerous studies have demonstrated a strong configural bias for face identity processing in the chimpanzee (Parr et al., 2000; Parr et al., 2006; Parr and Heintz, in press), the perceptual cues important for facial expression categorization remain unclear (Parr et al., 1998).

Parr and colleagues have performed several studies to address this first approach: how do chimpanzees discriminate facial expressions, and what are the relevant features for this perceptual process (Parr et al., 1998;
The goal of these studies was first to determine whether chimpanzees could visually match different examples of expression types and second, to try and understand how these categorizations were achieved. Were subjects, for example, extracting specific salient features, like the presence of teeth, mouth position, etc., or were they relying on the overall expression configuration, like humans? These studies have been performed using a computerized joystick-testing paradigm and matching-to-sample (MTS) format. Subjects have had years of expertise with this testing situation and eagerly participate in daily testing sessions (Parr et al., 2000, Parr and de Waal, 1999). The basic format for MTS is that subjects are first presented with a sample stimulus, a conspecific facial expression, for example, and a cursor on the computer screen over a black background. They have been trained that when they can control the movements of the cursor on the monitor by manipulating the joystick. The sample stimulus is the image to match and they first must orient towards it by contacting it with the joystick-controlled cursor. After this, the sample stimulus clears the screen and they are presented with two alternative stimuli, one matches the sample on predetermined stimulus dimension, i.e. expression type (target image), while the other does not match (foil).

In a previous expression matching task, Parr and colleagues (1998) presented chimpanzees with a sample photograph showing an unfamiliar conspecific making one of the six facial expressions, bared-teeth, pant-hoot, play face, relaxed-lip face, scream and neutral. The target image was another individual making the same facial expression, so the matching pair of photographs was unique, and the foil showed a third individual making a neutral expression. Successful matching, therefore, required discrimination based on expression type, not the identity of the individual making the expression. On the first testing session, all six chimpanzees performed significantly above chance (50%) on the majority of the expression categories (bared-teeth = 67%, play face = 68%, scream = 64%), demonstrating that not only are expressions highly salient, but that expression type can be used spontaneously as a means for discrimination (Parr et al., 1998).

In a follow-up to this study, Parr and colleagues (1998) investigated how expressions were discriminated, using overall configuration or the extraction of specific features. To do this, each expression was characterized according to the presence or absence of specific salient features using subjective criteria, such as mouth open, eyes open, teeth visible. This was done for five expression types, bared-teeth, pant-hoot, play face, relaxed-lip and scream. In the experimental task, each dyadic combination of expressions was then paired together, totaling 20 different combinations for the five expression categories. For half of the expression pairs (N=10), the target (match) and foil (non-match) expressions shared three or more features in common, while the other 10 pairs were very distinctive, sharing two or fewer features in common. Figure 3 shows examples of a trial where the expression pair shares features in common (Figure 3A) or has little feature overlap (Figure 3B). The performance on expression dyads that were similar versus distinctive was then compared. The hypothesis was that if the chimpanzees were extracting specific salient features, such as teeth visible, when performing expression discriminations, their performance should be better on distinct dyads compared with similar dyads. Moreover, a negative correlation would be expected overall between the number of shared features in each of the 20 expression dyads and performance matching the expressions in those trials, i.e. the more feature overlap, the worse performance would be. The data supported the first prediction: overall, subjects’ performance was significantly better discriminating expression dyads that were distinct compared to similar, \( t(4) = 2.94, P < 0.05 \). There was, however, little support for the second prediction. The overall correlation between shared features and performance was only weakly negative, \( r = -0.11, \text{NS} \) (Parr et al., 1998). In fact, the role of distinctive, non-overlapping facial features in providing an advantage for expression discrimination was only detected for some expression categories and not others. Thus,
no conclusive evidence was found for or against configural facial expression processing in this species.

One serious problem with these earlier studies was the inability to objectively categorize each expression in terms of its relevant features. With the creation of ChimpFACS and the development of the standardized, prototypical Poser expressions described earlier, the question of how chimpanzees process facial expressions, configurally or using feature-based cues, can be re-addressed in a more controlled fashion by manipulating action units individually and in combinations and presenting these to subjects using similar discrimination studies as described earlier. These studies are currently underway at the Yerkes Primate Center and will make an important contribution towards understanding the evolution of emotional communication by addressing similarities and differences in the perceptual discrimination of facial expressions by chimpanzees compared to humans.

In conclusion, a great deal still remains to be learned about the evolution of social signals, like facial expressions. This article has reviewed and expanded on data suggesting homologies between the primate bared-teeth display and the human smile, and the primate play face and human laughter. But to what extent are other facial expressions evolutionarily similar across related species? Emphasis has also been placed on the general lack of knowledge concerning the social function of facial expressions in non-human primates. Only a handful of studies have thus far been attempted, yet the question of how these expressions function in a social context is critical for a more complete understanding of their evolutionary significance and continuity. Finally, this brief review did not address issues of development or social learning, i.e. how do primates learn the social function and/or meaning of these expressions? Is there an innate component, or is learning totally dependent on experience? These developmental questions pertain to more than just understanding the meaning of expressions, but also in knowing how to produce expressions, as in appearance, so as to convey appropriate information to conspecifics, and then knowing when to use them in the appropriate social contexts.

Tools such as the ChimpFACS provide an exciting new methodology for understanding similarities and differences in chimpanzee and human expressions by comparing homologous facial movements. Thus, this new tool will help facilitate comparisons of facial expressions according to their structural similarities and enable researchers to expand on whether facial expressions other than the bared-teeth display and the play face may have structural homologs in humans. Moreover, combining studies on the structural similarity of these expressions with tasks that investigate their perceptual classification will provide insights into which features are important for expression categorization, and how the cognitive processes that underlie the categorization of these signals are indeed similar to what is known about human expression categorization. This latter question may tie in with recent work on human neuroimaging that suggests that neural systems important for face identity recognition and processing emotional content from facial expressions are, in fact, dissociable (Haxby et al., 2000; Kesler-West et al., 2001; Phan et al., 2004). Finally, sorely lacking are studies that examine the social function of facial expressions both in chimpanzees and in humans. Understanding their function is critical for making any inferences about evolutionary continuity, their emotional meaning and the co-evolution of signaler and receiver to maximize efficient communication within social groups.

REFERENCES


References

For reference, here is a list of important references in the document:


