



The face is central to primate multicomponent signals

Bridget M. Waller¹ · Eithne Kavanagh¹ · Jerome Micheletta² ·
Peter R. Clark² · Jamie Whitehouse¹

Received: 20 April 2021 / Accepted: 29 September 2021
© The Author(s) 2022

Abstract

A wealth of experimental and observational evidence suggests that faces have become increasingly important in the communication system of primates over evolutionary time and that both the static and moveable aspects of faces convey considerable information. Therefore, whenever there is a visual component to any multicomponent signal the face is potentially relevant. However, the role of the face is not always considered in primate multicomponent communication research. We review the literature and make a case for greater focus on the face going forward. We propose that the face can be overlooked for two main reasons: first, due to methodological difficulty. Examination of multicomponent signals in primates is difficult, so scientists tend to examine a limited number of signals in combination. Detailed examination of the subtle and dynamic components of facial signals is particularly hard to achieve in studies of primates. Second, due to a common assumption that the face contains “emotional” content. A priori categorisation of facial behavior as “emotional” ignores the potentially communicative and predictive information present in the face that might contribute to signals. In short, we argue that the face is central to multicomponent signals (and also many multimodal signals) and suggest future directions for investigating this phenomenon.

Keywords Facial communication · Facial expressions · Facial displays · Primates · Multimodal · Multicomponent

Handling Editor: Joanna Setchell

✉ Bridget M. Waller
bridget.waller@ntu.ac.uk

¹ Department of Psychology, Nottingham Trent University, Nottingham, UK

² Centre for Comparative and Evolutionary Psychology, Department of Psychology, University of Portsmouth, Portsmouth, UK

Introduction

Multicomponent and multimodal signals have been proposed as having a potentially important role in communicative complexity, specifically language evolution (Liebal *et al.*, 2014; Partan & Marler, 1999; Rowe, 1999). It has been argued that combining signals in some way (simultaneously or sequentially) increases the generative and/or flexible power of a communication system, which in turn requires more complex cognition for the production and perception of communicative signals. Multimodal refers to communication combining multiple sensory sources, such as auditory and visual (Rowe, 1999), and multicomponent refers to combinations of different signals regardless of the sensory modality (see below for more discussion of definitions). Thus, an evolutionary drive toward communicative complexity may have coincided with both increased multimodality and multicomponent signaling. As such, scientists interested in comparative communication tend to agree that signals should not be studied in isolation and have become “multimodalists” (Fröhlich *et al.*, 2019). However, despite this widespread call to adopt a multimodal or multicomponent approach, in practice any approach beyond single signal types is uncommon (Slocombe *et al.*, 2011). Also, even though there is widespread agreement that signals should be understood as whole entities, research (often necessarily) involves carving up multimodal signals into component parts based on some a priori assumptions. In doing so, this creates a paradox. The multifaceted phenomenon is partitioned to provide a detailed sense of all parts of the whole, but in doing so, it is possible that we lose perspective of the whole. For example, delineating faces from bodies, or calls from faces, might mean that the way that these components interact is overlooked, and/or we make assumptions about what each component does and study them differently.

Many characteristics of animals, in both form and behavior, have the potential to transmit information. Much of this information is available through the visual channel (through bodies, faces). We argue, however, that the physical features of primates are not all equal in terms of communicative salience, and so there are implications of ignoring some characteristics over others. The face, in particular, is an attention grabbing, meaningful, and salient feature of all primates (Mitchell *et al.*, 2014; Nahm *et al.*, 1997), and it is highly unlikely that the presence (or absence) of facial information in all its potential forms is irrelevant when observers are viewing signals in conspecific primates. Any multicomponent signal that includes a visual element, therefore, contains information about the face whether or not this is being recorded by the researcher. Even visual cues on the primate body (for example, sexual swellings in chimpanzees, *Pan troglodytes*) could be interpreted differently if the female is looking at the observer. Consequently, overlooking the face when examining multimodal signals could result in a mis-categorization of the signal and limit our understanding of the communication taking place.

We review the evidence that the face is important in primate communication and make a case for greater focus on the face in multimodal and multicomponent

communication research (regardless of whether facial movements are present). We examine the literature where the face has been included in multimodal and multicomponent communication research to date and discuss the theoretical and methodological reasons why it might have been overlooked. Finally, we propose solutions.

How do primates see the face?

Early primates were characterised by expansion of the visual cortex relative to other brain areas and greater connectivity between visual and motor cortices (Molnár *et al.*, 2014; Sherwood, 2005), which likely coincided with heavier reliance on the visual domain for communication (Dobson & Sherwood, 2011) than other mammals. Likewise, the physical features of the primate face and associated behaviors suggest an evolutionary shift toward greater salience and communicative prominence of the face. Thus, there appears to be specialization in both senders and receivers for facial communication in primates (Fig. 1).

Primate faces are largely hairless. Color vision may have coevolved with this hairlessness in primate faces to facilitate the detection of facial color changes associated with social and sexual displays (Changizi *et al.*, 2006; Santana *et al.*, 2012). For example, some primates have trichromatic color vision system, which allows better detection facial signals, such as the facial redness associated with female intracycle fertility in rhesus macaques (*Macaca mulatta*) (Hiramatsu *et al.*, 2017). Despite the relative hairlessness, there are conspicuous hairy growths in some species that could be associated with facial signaling (Flecha-García, 2010; Sadr *et al.*, 2003; Watt *et al.*, 2007). Human eyebrows, for example, despite having possible protective function against sweat and sun, also seem important in punctuating speech

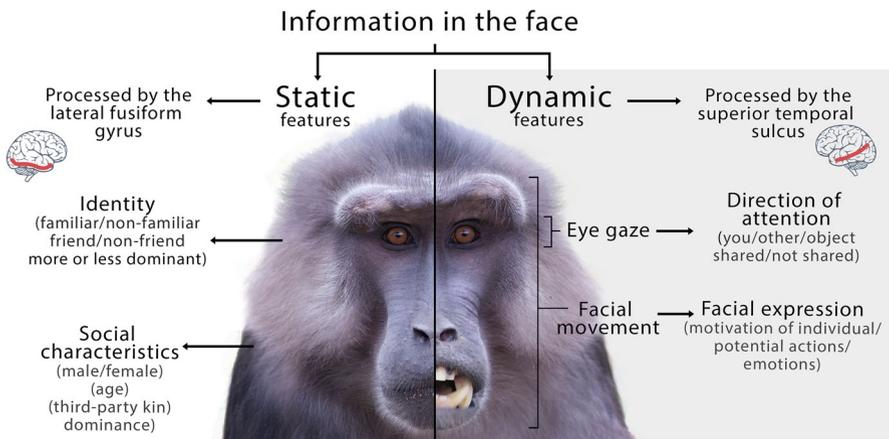


Fig. 1 Information in the face. Example of the static (left) and dynamic (right) information that can be extracted from a conspecific face in primates, which are thought to be separated by distinct neural pathways (Bernstein & Yovel, 2015).

(Flecha-García, 2010), are key features for identity recognition (Sadr *et al.*, 2003) and regulate the visibility of eye gaze direction (Watt *et al.*, 2007). Similarly, it is possible that eyelashes enhance eye stimuli in addition to having protective function. Radial venation patterning in flower petals serve as nectar guides to potential pollinators (Whitney *et al.*, 2013), and so it is worth considering whether attention to the eyes is enhanced by the radial patterning of eyelashes. Eyes are powerful stimuli and provide cues about another's attention (looking at you; looking at other) and whether this is shared with you (we are looking at the same thing; we are looking at each other). Humans are sensitive to even very subtle changes in eyes, such as pupil size (Demos *et al.*, 2008). Pupil size changes can modulate social processes, such as perception of emotion (Harrison *et al.*, 2007) and preference for mutual gaze (Binetti *et al.*, 2016). Thus, any features that draw attention to the eyes are likely to be beneficial in a cooperative system with mutual interests, even if the primary function is something else.

Primates also produce numerous subtle facial movements underpinned by the activation of a complex network of facial muscles. These muscles are well conserved throughout the primate order (Burrows, 2008) and considerably more complex than those found in other mammals. The area of the motor cortex that innervates these facial muscles (facial motor nucleus, VII) has similar structure and organization across mammals, but some features suggest greater control of facial muscles in primates (Sherwood, 2005). For example, there are direct connective projections to facial muscles found only in catarrhine primates and the volume of the facial nuclei of the brainstem is greater in great apes and humans compared with other primates and other mammals (Sherwood, 2005). The extent and manner of voluntary and flexible control also is likely related to the specific properties of the muscle fibers themselves. Slow-twitch muscle fibers are usually involved in activities requiring precise control of weak forces, suggesting that these muscles can be used in a slower and more precise manner. Some facial muscles in the human face have a higher proportion of slow-twitch muscle fibers than both rhesus macaques and chimpanzees (Burrows *et al.*, 2014). Such fine-grained control of facial muscles likely increases as complexity of facial behavior increases.

Primates can use faces to identify key physical and social characteristics from the conspecific primate face (Fig. 1). Often based on the static rather than moveable aspects of the face, primates can distinguish between other individuals in terms of their identity (e.g., many species: Parr, 2011), sex (e.g., rhesus macaques: Paukner *et al.*, 2010), third-party kin relationships (e.g., chimpanzees: Parr & de Waal, 1999), attractiveness (e.g., rhesus macaques: Waitt & Little, 2006), and, potentially, dominance (e.g., bonobos, *Pan paniscus*: Martin *et al.*, 2019). The neural pathways underpinning these perceptual processes are broadly similar across primates (Freiwald *et al.*, 2016), although monkeys seem to have less face expertise than apes on the whole (Parr *et al.*, 2008).

In primates, faces are attention-grabbing phenomena, more so than in other mammals. Like humans, other primates prefer to look at faces compared with areas surrounding the face (common marmosets, *Callithrix jacchus*: Mitchell *et al.*, 2014; rhesus macaques: Nahm *et al.*, 1997). When viewing naturalistic pictures of whole animals, eye-tracking studies showed that both humans and chimpanzees look first

at the face region and look longer at the face region in total compared with other parts of the body (Kano & Tomonaga, 2009). Interestingly, this pattern of increased attention to the face was not seen in domestic dogs (*Canis lupus familiaris*, Correia-Caeiro, Guo, & Mills, 2021a) who attend more to the body than the face in both conspecifics and humans. The chimpanzee stimuli used by Kano and Tomonaga (2009) did not display any prototypical facial expression or notable facial movement, suggesting that primate faces capture attention regardless of whether they contain communicative information. In humans, faces are attended to preferentially from birth (Frank *et al.*, 2014), a mechanism that persists in adulthood (Stein *et al.*, 2011) and can occur below threshold awareness (Lueschow *et al.*, 2004). Therefore, it is likely that similar mechanisms that facilitate bias to faces exist in nonhuman primates.

Given that “neutral” primate faces still command attention from conspecific observers, we argue that a face can never be truly considered “neutral.” First and foremost, if visible, static features are still present and indicative of identity and individual characteristics. In addition, however, the absence of movement in a primate face is still communicating something by the very fact that what *could* be present is not. In a match-to-sample experimental design, crested macaques (*Macaca nigra*) differentially predicted the social outcome of an approach between two unknown individuals depending on the facial expression featured on the face (Waller *et al.*, 2016). The neutral face was more closely associated with a conflict outcome than screams and threat faces, suggesting that the absence of facial expression was perceived as more likely to result in conflict. The interpretation was that the presence of any communicative information (regardless of valence) reduces uncertainty and allows other individuals to respond appropriately, and so a neutral face is conspicuous by the very absence of movement. It also is possible that a neutral face is viewed as a negative signal due to the absence of any affiliative signal. Therefore, as long as the face is visible, it is essentially always of potential importance regardless of whether it displays facial movement. There are circumstances when the face might not be visible, of course, but this should not be confused with a visible face that does not show movement.

The importance of combinationality is, of course, central to the notion of multimodality (Partan & Marler, 1999), in that perception of one feature or behavior is influenced by the presence of another. Despite the long-understood importance of faces in human perception, their salience and meaning are affected by other aspects of the scene. For example, human facial expressions are interpreted differently when paired with different body positions (Aviezer *et al.*, 2008). These prototypical, basic facial expressions, often argued to be categorically discrete and universally recognized (Ekman *et al.*, 1969) thus are sensitive to context and interpretation can be modified at an early perceptual level. There is clear perceptual interaction, therefore, between the face and other aspects of the body during interpretation of signals.

In sum, the evidence that the primate face has evolved as a salient and communicative feature suggests that even when “neutral” in the absence of specific facial movement, there is still meaningful information available to others that has the potential to change the perception of signals. Equally, interpretation of the face can be influenced by the position and movement of the body.

Why are multicomponent signals important?

Multimodality, in essence, refers to communication that provides information from multiple sources, such as auditory information combined with visual (Rowe, 1999). Many scientists consider a signal to be multimodal only when it is received from more than one sensory channel: auditory, visual, tactile, or chemical (CANDOLIN, 2003; Higham & Hebets, 2013). However, others have suggested that any combination of facial, gestural, and vocal signals should be classed similarly (occasionally but less frequently, including olfactory signals: Liebal *et al.*, 2014). The rationale for this divergent usage is that in primates, the way these signal types are cognitively processed and produced can be meaningfully different (Tomasello, 2010; Waller *et al.*, 2013). Thus, a gesture and a facial expression may provide separate concurrent streams of information, despite the fact they are both received through the visual channel. While there may be an argument to use a separate term (e.g., "multicomponent": Micheletta *et al.*, 2013), many of the proposed advantages of multimodal signaling based on multiple sensory channels also apply to this usage, and so studying the combination of these signal types may have important implications for understanding the evolution of language and communication. However, to avoid any confusion, we use multicomponent as a more inclusive term to capture all signals that include one or more visual elements. Regardless of the definition adopted for different types of signals, we argue that the presence or absence of the face in any signal is significant, and so it could be argued that all signals are multicomponent.

The primary value of multimodal (and multicomponent) signaling over single types of signaling is that it can provide increased availability of information to the receiver—either by increasing the amount of information, or by increasing the likelihood that the information is successfully transferred to the receiver (Partan & Marler, 2005). Animals frequently inhabit noisy environments in which signals from one sensory modality are not accessible to the receiver (e.g., visual signals not visible due to dense forestation, or audio signals inaudible due to interfering sounds). The use of multimodal signals may increase the likelihood that information can be received in such environments. For instance, wolf spiders (*Schizocosa retrorsa*) combine visual and seismic signals to increase the likelihood of receiver response in bright and dark conditions (Hebets & Papaj, 2005). This is an example of a redundant function (in the sense that no new information is transferred, but the message is enhanced), but multimodal signals also may serve nonredundant functions (see Partan & Marler, 2005 for a complete framework of multimodal signal functions). The components of multimodal signals may send separate information independently, but they also may combine to send entirely new information than either component could provide separately. Such a function would allow for multimodality to considerably increase complexity in communication systems, but it has not yet been identified in primate communication. Great apes often produce multimodal signals in conditions of good visibility and audibility, which could point toward a nonredundant function, because the likelihood of the message being received is high (Tagliatalata *et al.*, 2015), but this has not yet been empirically demonstrated. Only by examining

the integration of multiple modalities holistically can the full capacity of primate communication be revealed.

Many scientists are interested in multimodality and multicomponent signaling for its value in informing theories of human language evolution, and there is increasing interest in the possibility that language has multimodal origins (Fröhlich *et al.*, 2019; Tagliatalata *et al.*, 2011). One line of inquiry for language evolution scholars is to identify the components and underlying capacities of language and to test their presence in nonhuman primate communication systems (Slocombe *et al.*, 2011). The corresponding argument is that if these capacities can be identified in other primate species, they represent the “building blocks” available to human ancestors before the evolution of language in its current state. Many capacities have been identified in different modalities in different species (e.g., referentiality in vocalizations, Seyfarth & Cheney, 1990; intentionality in gesture, Tomasello, 2010), often provoking claims that the origins of language can be found in the modality displaying this capacity. However, proponents of a multimodal or “amodal” (i.e., human language is not rooted in any specific communication system in the common ancestor) origin of language argue that the modality through which a language-like capacity is expressed is largely irrelevant, because its expression may have diverged from our common ancestor either in the respective species or in humans (Slocombe *et al.*, 2011). Importantly, although human language is primarily expressed through the vocal modality, it can equally be expressed gesturally through sign language (Sandler & Lillo-Martin, 2006), indicating that underlying language capacities are not modality-bound. This points toward flexibility in the expression of different capacities and classes of information across modalities and primate species and emphasises the importance of a multimodal and/or multicomponent approach.

The combination of components into multicomponent signals may in itself be an important precursor to human language and communication by conveying different types of behavioral and affective information at the same time (Waller *et al.*, 2013). The potential for a multicomponent signal to emit additional information than its unimodal components separately, and to do so more efficiently, means that it could have represented an important starting point in human ancestors for the evolution of the rapid complex information transfer associated with human language. More broadly, human communication is inherently multicomponent. Holler and Levinson (2019) proposed that face-to-face interaction is the ecological niche in which language evolved, and language is embedded in multimodal signals between the speaker and the audience. Visual signals are almost always produced alongside spoken conversation, providing important context and facilitate shared understanding between interactants (e.g., eyebrows punctuating speech, Ekman, 1979; iconic gestures disambiguating speech, Holle & Gunter, 2007). Similarly, many nonhuman primate species live in closely bonded, relatively cohesive groups, with frequent visual and audible contact between group members, and their use of multicomponent communication is well-documented (e.g., bonobos: Genty *et al.*, 2014; crested macaques: Micheletta *et al.*, 2013; chimpanzees: Wilke, 2017). It therefore is crucial that the holistic multicomponent interaction between primates is examined and understood in the quest to understand the evolutionary roots of human communication and language.

How has the face been considered in multicomponent signal research so far?

In 2011, a review of the primate communication literature revealed striking differences in the methodological and theoretical focus of published research depending on the species class under study (Slocombe et al., 2011). Notably, studies investigating multicomponent communication represented only 28 of 553 studies examined (5%); 22 examined facial expression alongside at least one of gestural or vocal communication. Since then, multicomponent research in primatology has gained popularity in principle but still accounts for the minority of publications in primate communication and often ignores facial signals (Liebal et al., 2014).

Experimental research focussing on cross-modal integration is a notable exception. A number of studies investigating the evolutionary precursors of humans' ability to perceive and integrate spoken utterances with facial movements found evidence for this multisensory integration across a number of nonhuman primate species. Using a preferential-looking paradigm, Ghazanfar et al. (2005) presented rhesus macaques with video clips of conspecifics producing two different vocalisations, paired with the sound of one of these vocalizations. The macaques spent more time looking at the matching visual stimulus than the nonmatching stimulus, indicating that they recognized the correspondence between auditory and visual information for these two vocalizations. Following this seminal work, further "preferential looking" studies were conducted in a number of nonhuman primate species, with similar results (e.g., grey-cheeked mangabeys (*Lophocebus albigena*), Bovet & Deputte, 2009; capuchin monkeys (*Cebus apella*), Evans et al., 2005; chimpanzees, Izumi & Kojima, 2004). Parr (2004) obtained similar findings using a matching-to-sample procedure in chimpanzees.

A few observational studies have attempted to document facial signals with a multicomponent perspective. In captivity, 50% of chimpanzees' vocalizations were accompanied by signals from one or more other modalities within 2 s (Tagliatalata et al., 2015). Although the authors did not provide data on the relative frequencies of different signal modalities, facial signals did occur in a number of these multicomponent bouts. In a more detailed account, Wilke and colleagues (2017) attempted to quantify the multicomponent repertoire of wild chimpanzees. They considered eight different facial signals, of which six temporally overlapped with signals from different modalities. While the frequency of production of these multicomponent facial signals was low, they constituted a nonnegligible part of the multicomponent repertoire. The most complete quantitative account of multicomponent communication in a nonhuman primate to date is Sarah Partan's (2002) study of rhesus macaques. Components of facial signals (i.e., mouth, ear, eyebrow, and head position, as well as gaze direction) were considered alongside body postures and vocalization, and the composition of the multicomponent signals was analyzed using Principal Component Analysis and Multiple Correspondence Analysis. With this method, she identified clusters corresponding to threatening, submissive, and affiliative behavior with an unprecedented level of detail and demonstrated the multicomponent nature of rhesus macaques' communication.

Although these overviews of species' multicomponent repertoires are useful, the efforts required to collect enough data to be able to examine the function of multicomponent signals often are restrictive and the analyses are problematic. Instead, several researchers have chosen to focus on a single facial signal and its combination with signals from other modalities. For example, Fig. 2 illustrates how different face positions and movements can influence the form (and potential interpretation) of gestures and vocalization combinations in chimpanzees. This approach has the benefit of providing a richer understanding of the function of multicomponent signals (although at the cost of ignoring a large part of the communicative repertoire). Chimpanzees can produce facial expressions that are visually identical but appear to have different functions when used singly and when used in combination with a vocalization (Davila-Ross *et al.*, 2015) and also use different combinations of gesture and facial expression that elicit different responses from conspecifics (Oña *et al.*, 2019). In Japanese macaques (*Macaca fuscata*), females exhibit variations in facial color parameters alongside changes in vocal behaviors, which might influence male behavior and reduce the probability of mating with females who are already pregnant (Rigaill *et al.*, 2015). Crested macaques frequently produce lipsmacks, a facial signal that can be accompanied by a close-range vocalization: the soft grunt. In this species (Micheletta *et al.*, 2013), as well as in rhesus macaques (Partan, 1998), multimodal lipsmacks were more likely to elicit positive responses in the form of affiliative body contacts, demonstrating the importance of a multimodal approach to better characterize the complexity of communication in nonhuman primates.

Why has the face been overlooked in primate multicomponent communication research?

We argue that there are two main reasons why the face is hard to integrate in primate multicomponent communication literature. The first is a purely methodological point. It is difficult to record primate signals in general, but we argue that it

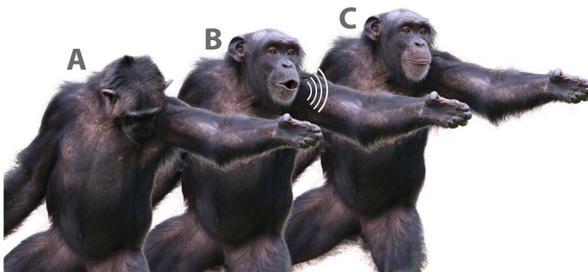


Fig. 2 The importance of the face in multicomponent signals. A chimpanzee gesture paired with different head positions and movements of the face, illustrating the potential for different resulting meanings. **A.** Arm stretch gesture with face down/not visible; **B.** Arm stretch gesture with AU22 (lip funneler) + AU25 (lips parted) + AU26 (jaw drop) and vocalization; **C.** Arm stretch gesture with “neutral” face. Note: these chimpanzees are composite images featuring multiple individuals (i.e., head, body, and face taken from different individuals were combined to illustrate a single individual). Picture credit: Matt Henderson.

is particularly difficult to capture the detail of facial expressions, and the methods that are available (e.g., Facial Action Coding Systems, Waller *et al.*, 2020) can be onerous. Second, there is a long history of approaching facial expressions as emotional entities, which has perhaps side-lined them from mainstream discussion of the evolution of communication. Investigation of the evolutionary roots of language has been a particular focus within the field (Liebal *et al.*, 2014), and because facial expressions are rarely considered as precursors to language, this also might have deflected attention. To clarify, although primate facial communication research is a fairly active field independently, we argue that it is not well integrated into other modalities of primate communication research (as evidenced by a comprehensive literature review (Slocombe *et al.*, 2011) and a recent follow-up to this review (Liebal *et al.*, 2014).

Methodological reasons

As discussed earlier, the primate face has the capacity to produce numerous dynamic appearance changes—both subtle and extreme. The specific social impact of these facial movements on other individuals is not always easy to determine (and in many cases there might be no impact at all), but to determine this impact, the facial movements need to be recorded accurately. There are practical considerations that might make this hard to achieve. Given the subtlety of some facial movement, human observers might underestimate the occurrence of facial movements due to distance and thus potentially miss their contribution to social interactions, particularly in wild populations. Likewise, body orientation can affect our ability to collect data on facial movements but has less of an impact on the visibility of gestures. For example, Hobaiter *et al.* (2017) aimed to collect data on vocalization, gesture, and facial movement in chimpanzees but were unable to collect sufficient information on faces due to the above difficulties. It also is difficult to determine for humans to assess eye contact between individuals (although see Bard *et al.*, 2005; Bethell *et al.*, 2007), which could be an important factor in whether facial movements are likely relevant in a social interaction. Another problem is that because the production of vocalizations always involves some movement of the jaw and/or face, it is difficult to identify meaningful facial movements that augment an audible signal.

There also are perceptual reasons why accurate recording of facial expression is difficult. Unlike vocalizations and gestures, there are clear counterparts to human facial expressions in other primate species. Some facial expressions appear so similar that scientists have investigated similarity of form and function for decades (e.g., human smile and bared-teeth display, Van Hooff, 1972). These similarities likely indicate some shared history, but this also can lead to problematic subjective interpretation and false judgements of physical properties (Waller *et al.*, 2007). Historically, scientists have tended toward a holistic approach to primate facial movements and identified configurations of facial movements as the unit for analysis (e.g., bared-teeth display, threat face). Such a priori classification can be problematic for many reasons: relevant detail can be missed; meaningful variants can be lumped together (Clark *et al.*, 2020); and the labels often are loaded with emotional and/or

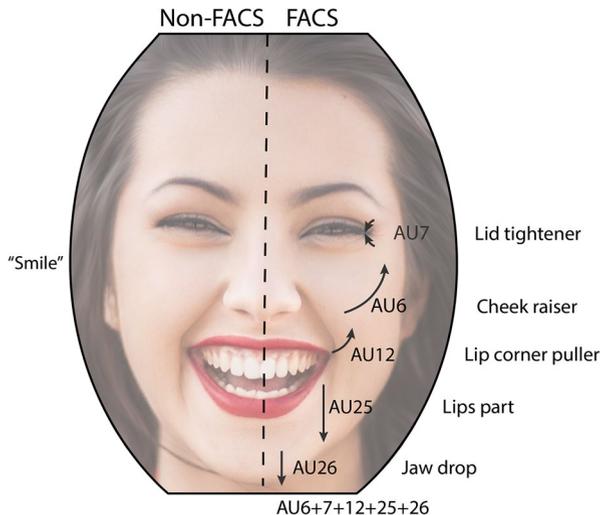
functional assumptions (see Waller & Micheletta, 2013 for a more extensive discussion of this issue).

To solve these problems in human facial expression research, FACS (Facial Action Coding System; Ekman *et al.*, 2002; Ekman & Friesen, 1978) was developed (Fig. 3). FACS divides the face into the minimal units of facial movements (Action Units [AUs]), which are based on specific muscles movements. The goal was to make identification of facial movement anatomically based and as detailed and objective as possible. FACS has now been modified for use with multiple primate species (chimpanzees, macaques, orangutans, hylobatids; see Waller *et al.*, 2020 for a review), which should help to minimize the problems stated above. The systems require training and can be time consuming, however, and reliability is not always easy to achieve (Molina *et al.*, 2019). FACS also might not capture the specific details that scientists might be interested in, and for most understudied primate species, there is still no FACS developed (although attempts to integrate FACS systems developed for different, closely related species have been successful, Correia-Caeiro, Holmes, & Miyabe-Nishiwaki, 2021b; Julle-Danière *et al.*, 2015). It can be difficult to use the data resulting from FACS and create meaningful units for analysis, although there have been some good efforts to create indices of expressivity and diversity from FACS data (Scheider *et al.*, 2014)

Theoretical reasons

In addition to the methodological difficulties stated above, there may be theoretical barriers to the integration of facial movements into the study of multimodal and multicomponent primate communication. Human and nonhuman facial displays have long been studied within an emotional framework that we argue has limited the scope of investigation (Waller & Micheletta, 2013). Since Darwin first made comparisons between the “expressive” behavior of humans and other animals (Darwin,

Fig. 3 Facial Action Coding System (FACS) in action. In FACS, individual muscle movements of the face (or, Action Units) are coded independently and objectively, each of which are assigned a number (e.g., AU7). This provides a more objective and detailed description of the dynamic movements of the face compared with a more subjective and holistic descriptions, such as “smile” or “frown.” Photo credit to Michael Dam on [Unsplash.com](https://unsplash.com).



1872), the scientific approach has been dominated by an assumption that facial displays are involuntary, innate, and reflexive expressions of emotion strongly associated with physiological and subjective processes. As such, despite almost always occurring within social interaction, their role as communicative agents has received only minor attention. Instead, the focus has been on what facial behavior can reveal about the emotional state of the sender. Signal value has been considered mainly in terms of whether and how observers can discern the feeling state of the sender, rather than what they can do with that information. Communicative content therefore is, *a priori*, understood to relay the emotion of the sender, with the potential adaptive consequences of this communication rarely considered.

This assumed close link between felt emotion, facial behavior, and perceived emotion, however, has been challenged in humans (Barrett *et al.*, 2019), and so it likely that a similar complex and messy relationship between emotion and facial movement exists in nonhumans. There is an argument for some correlation between internal state and behavior in primates (Kret *et al.*, 2020), but whether this relationship exists for facial expression is debated. The alternative approach (that we support) is rooted in behavioral ecology, is arguably more adaptationist, and posits that facial displays signal the likely behavior of the sender and can be conceptualised as predictive cues (Fridlund, 1994, 2017; Waller, 2017). Releasing facial displays from a narrow definition of emotion opens up new avenues for investigation, including their likely role in social communication, be it multimodal, multicomponent, or unimodal. Whether the communicative signals are intentionally produced and underpinned by complex cognition is irrelevant to ascertain whether information is transmitted and whether there are behavioral consequences. However, acknowledging that facial signals *can* be communicative makes questions of intentionality, flexibility, referentiality, and complexity more pertinent.

Conclusion: Is the face central to multicomponent signals and does this matter?

We have argued that the face is central to primate communication. In sum, the attention-grabbing salience of the primate face, regardless of whether there is facial movement present (or whether the face is displaying an adaptive communicative signal), means that the face is relevant in any behavioral communicative signal and therefore adds an additional communicative component. In a sense, we are arguing that the face *makes* all nonfacial communication multicomponent. Whether the face is visible or not is relevant, of course, but it is still conspicuous via its absence. An absent face means that the sender cannot see the receiver, which is highly relevant information for primates, most of which have been shown to follow the gaze of conspecifics (Rosati & Hare, 2009) and change behavior accordingly.

Fortunately, widening the net to include the face when studying primate communication need not be overly burdensome. As discussed, some of the methods to capture facial behavior can be onerous and time consuming (e.g., FACS), but it also could be sufficient to record simple variables in real time (face visible to receiver: yes/no; face static or moving: yes/no). Gesture researchers do already record much

of this information, but often in the context of the producer and not receiver (Liebal *et al.*, 2014). Alternatively, FACS can be used in a reduced form to record specific AUs, or combinations of AUs, if they are of specific interest. Finally, and most importantly, checking a priori assumptions about what faces do and what information they contain is a necessary step forward to advance our understanding of primate communication. Specifically, we need to move beyond the dominant emotional framework. Empirical data are needed to demonstrate how faces contribute to primate communication (also true for gestures and vocalisations). If we are to make meaningful comparisons between these different behaviors and draw conclusions about differences that justify a priori assumptions, we need to collect comparable data.

In short, the face has grown in importance throughout primate evolution and carries a wealth of information that can be extracted by conspecifics. If we want to understand the form and function of primate multicomponent and multimodal communication, how it evolved, and how human communication is and is not related, we always need to consider the face in primate communication.

Acknowledgments The development of this paper has been partly supported by Horizon 2020 European Research Council Consolidator Grant 864694 (FACEDIFF) to BW. The authors are grateful to the guest editor of this special issue plus two anonymous reviewers for insightful and helpful comments on various iterations of this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–732. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>.
- Bard, K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., & Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (*Pan troglodytes*). *Developmental Psychology*, *41*(4), 616.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. <https://doi.org/10.1177/1529100619832930>.
- Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience and Biobehavioral Reviews*, *55*, 536–546. <https://doi.org/10.1016/j.neubiorev.2015.06.010>.
- Bethell, E. J., Vick, S.-J., & Bard, K. A. (2007). Measurement of eye-gaze in chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, *69*(5), 562–575.
- Binetti, N., Harrison, C., Coutrot, A., Johnston, A., & Mareschal, I. (2016). Pupil dilation as an index of preferred mutual gaze duration. *Royal Society Open Science*, *3*(7), 160086.

- Bovet, D., & Deputte, B. L. (2009). Matching vocalizations to faces of familiar conspecifics in grey-cheeked mangabeys (*Lophocebus albigena*). *Folia Primatologica*, 80(3), 220–232.
- Burrows, A. M. (2008). The facial expression musculature in primates and its evolutionary significance. *Bioessays*, 30(3), 212–225.
- Burrows, A. M., Parr, L. A., Durham, E. L., Matthews, L. C., & Smith, T. D. (2014). Human faces are slower than chimpanzee faces. *PLoS One*, 9(10), e110523.
- Candolin, U. (2003). The use of multiple cues in mate choice. *Biological Reviews*, 78(4), 575–595. Cambridge Core. <https://doi.org/10.1017/S1464793103006158>.
- Changizi, M. A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology Letters*, 2(2), 217–221. <https://doi.org/10.1098/rsbl.2006.0440>
- Clark, P. R., Waller, B. M., Burrows, A. M., Julle-Danière, E., Agil, M., Engelhardt, A., & Micheletta, J. (2020). Morphological variants of silent bared-teeth displays have different social interaction outcomes in crested macaques (*MACACA NIGRA*). *American Journal of Physical Anthropology*, 173(3), 411–422. <https://doi.org/10.1002/ajpa.24129>.
- Correia-Caeiro, C., Guo, K., & Mills, D. (2021a). Bodily emotional expressions are a primary source of information for dogs, but not for humans. *Animal Cognition*, 1–13.
- Correia-Caeiro, C., Holmes, K., & Miyabe-Nishiwaki, T. (2021b). Extending the MaqFACS to measure facial movement in Japanese macaques (*Macaca fuscata*) reveals a wide repertoire potential. *PLoS One*, 16(1), e0245117. <https://doi.org/10.1371/journal.pone.0245117>.
- Darwin, C. (1872). *The expression of the emotions in man and animals by Charles Darwin*. John Murray.
- Davila-Ross, M., Jesus, G., Osborne, J., & Bard, K. A. (2015). Chimpanzees (*Pan troglodytes*) produce the same types of ‘laugh faces’ when they emit laughter and when they are silent. *PLoS ONE*, 10(6):e0127337. <https://doi.org/10.1371/journal.pone.0127337>
- Demos, K. E., Kelley, W. M., Ryan, S. L., Davis, F. C., & Whalen, P. J. (2008). Human amygdala sensitivity to the pupil size of others. *Cerebral Cortex*, 18(12), 2729–2734.
- Dobson, S. D., & Sherwood, C. C. (2011). Correlated evolution of brain regions involved in producing and processing facial expressions in anthropoid primates. *Biology Letters*, 7(1), 86–88. <https://doi.org/10.1098/rsbl.2010.0427>.
- Ekman, P. (1979). About brows: emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human Ethology* (pp. 169–248). Cambridge: Cambridge University Press.
- Ekman, P., & Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial movement. *Palo Alto*, 3(2), 5.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86–88.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System*. Research Nexus.
- Evans, T. A., Howell, S., & Westergaard, G. C. (2005). Auditory-visual cross-modal perception of communicative stimuli in tufted capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 31(4), 399.
- Flecha-García, M. L. (2010). Eyebrow raises in dialogue and their relation to discourse structure, utterance function and pitch accents in English. *Speech Communication*, 13.
- Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, 118, 13–26. <https://doi.org/10.1016/j.jecp.2013.08.012>.
- Freiwald, W., Duchaine, B., & Yovel, G. (2016). Face processing systems: from neurons to real-world social perception. *Annual Review of Neuroscience*, 39(1), 325–346. <https://doi.org/10.1146/annurev-neuro-070815-013934>.
- Fridlund, A. J. (1994). *Human Facial Expression: An Evolutionary View*. Academic Press.
- Fridlund, A. J. (2017). The behavioral ecology view of facial displays, 25 years later. *The Science of Facial Expression*, 77–92.
- Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & Schaik, C. P. (2019). Multimodal communication and language origins: Integrating gestures and vocalizations. *Biological Reviews*, 94(5), 1809–1829. <https://doi.org/10.1111/brv.12535>.
- Genty, E., Clay, Z., & Hobaiter, C. (2014). Multi-modal use of a socially directed call in bonobos. *PLoS One*, 9(1), 12.

- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience*, *25*(20), 5004–5012.
- Harrison, N. A., Wilson, C. E., & Critchley, H. D. (2007). Processing of observed pupil size modulates perception of sadness and predicts empathy. *Emotion*, *7*(4), 724.
- Hebets, E. A., & Papaj, D. R. (2005). Complex signal function: Developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*, *57*(3), 197–214. <https://doi.org/10.1007/s00265-004-0865-7>.
- Higham, J., & Hebets, E. (2013). An introduction to multimodal communication. *Behavioral Ecology and Sociobiology*, *67*. <https://doi.org/10.1007/s00265-013-1590-x>.
- Hiramatsu, C., Melin, A. D., Allen, W. L., Dubuc, C., & Higham, J. P. (2017). Experimental evidence that primate trichromacy is well suited for detecting primate social colour signals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856), 20162458. <https://doi.org/10.1098/rspb.2016.2458>.
- Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, *71*(6), 1–13.
- Holle, H., & Gunter, T. C. (2007). The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience*, *19*(7), 1175–1192. <https://doi.org/10.1162/jocn.2007.19.7.1175>.
- Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. <https://doi.org/10.1016/j.tics.2019.05.006>.
- Izumi, A., & Kojima, S. (2004). Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Animal Cognition*, *7*(3), 179–184.
- Julle-Danière, É., Micheletta, J., Whitehouse, J., Joly, M., Gass, C., Burrows, A. M., & Waller, B. M. (2015). *MaqFACS* (Macaque Facial Action Coding System) can be used to document facial movements in Barbary macaques (*Macaca sylvanus*). *PeerJ*, *3*, e1248. <https://doi.org/10.7717/peerj.1248>.
- Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures: A comparative eye-tracking study. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1664), 1949–1955. <https://doi.org/10.1098/rspb.2008.1811>.
- Kret, M. E., Prochazkova, E., Sterck, E. H. M., & Clay, Z. (2020). Emotional expressions in human and non-human great apes. *Neuroscience & Biobehavioral Reviews*, *115*, 378–395. <https://doi.org/10.1016/j.neubiorev.2020.01.027>.
- Liebal, K., Waller, B. M., Slocombe, K. E., & Burrows, A. M. (2014). *Primate communication: A multimodal approach*. Cambridge University Press.
- Lueschow, A., Sander, T., Boehm, S. G., Nolte, G., Trahms, L., & Curio, G. (2004). Looking for faces: Attention modulates early occipitotemporal object processing. *Psychophysiology*, *41*(3), 350–360. <https://doi.org/10.1111/j.1469-8986.2004.00159.x>.
- Martin, J. S., Staes, N., Weiss, A., Stevens, J. M. G., & Jaeggi, A. V. (2019). Facial width-to-height ratio is associated with agonistic and affiliative dominance in bonobos (*Pan paniscus*). *Biology Letters*, *15*(8), 20190232.
- Micheletta, J., Engelhardt, A., Matthews, L. E. E., Agil, M., & Waller, B. M. (2013). Multicomponent and multimodal lipsmacking in crested macaques (*Macaca nigra*). *American Journal of Primatology*, *75*(7), 763–773.
- Mitchell, J. F., Reynolds, J. H., & Miller, C. T. (2014). Active vision in marmosets: a model system for visual neuroscience. *The Journal of Neuroscience*, *34*(4), 1183. <https://doi.org/10.1523/JNEUROSCI.3899-13.2014>.
- Molina, A., Florkiewicz, B., & Cartmill, E. (2019). Exploring sources of variation in inter-observer reliability scoring of facial expressions using the ChimpFACS. *The FASEB Journal*, *33*(S1), 774–779.
- Molnár, Z., Kaas, J. H., de Carlos, J. A., Hevner, R. F., Lein, E., & Němec, P. (2014). Evolution and development of the mammalian cerebral cortex. *Brain, Behavior and Evolution*, *83*(2), 126–139. <https://doi.org/10.1159/000357753>.
- Nahm, F. K. D., Perret, A., Amaral, D. G., & Albright, T. D. (1997). How Do Monkeys Look at Faces? *Journal of Cognitive Neuroscience*, *9*(5), 611–623. <https://doi.org/10.1162/jocn.1997.9.5.611>.
- Oña, L. S., Sandler, W., & Liebal, K. (2019). A stepping stone to compositionality in chimpanzee communication. *PeerJ*, *7*, e7623. <https://doi.org/10.7717/peerj.7623>.
- Parr, L. A. (2004). Perceptual biases for multimodal cues in chimpanzee (*Pan troglodytes*) affect recognition. *Animal Cognition*, *7*(3), 171–178.

- Parr L. A. (2011). The evolution of face processing in primates. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366(1571), 1764–1777. <https://doi.org/10.1098/rstb.2010.0358>
- Parr, L. A., & de Waal, F. B. M. (1999). Visual kin recognition in chimpanzees. *Nature*, 399(6737), 647–648. <https://doi.org/10.1038/21345>.
- Parr, L. A., Heintz, M., & Pradhan, G. (2008). Rhesus monkeys (*Macaca mulatta*) lack expertise in face processing. *Journal of Comparative Psychology*, 122(4), 390.
- Partan, S. R. (1998). Multimodal communication: The integration of visual and vocal signals by rhesus macaques. *Unpublished Doctoral Dissertation*. University of California, Davis.
- Partan, S. R. (2002). Single and multichannel signal composition: Facial expressions and vocalizations of rhesus macaques (*Macaca mulatta*). *Behaviour*, 139(8), 993–1027.
- Partan, S. R., & Marler, P. (1999). Communication goes multimodal. *Science*, 283(5406), 1272–1273. <https://doi.org/10.1126/science.283.5406.1272>.
- Partan, S. R., & Marler, P. (2005). Issues in the classification of multimodal communication signals. *The American Naturalist*, 166(2), 231–245.
- Paukner, A., Huntsberry, M. E., & Suomi, S. J. (2010). Visual discrimination of male and female faces by infant rhesus macaques. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 52(1), 54–61.
- Rigaill, L., MacIntosh, A. J., Higham, J. P., Winters, S., Shimizu, K., Mouri, K., Furuichi, T., & Garcia, C. (2015). Multimodal advertisement of pregnancy in free-ranging female Japanese macaques (*Macaca fuscata*). *PLoS One*, 10(8), e0135127.
- Rosati, A. G., & Hare, B. (2009). Looking past the model species: Diversity in gaze-following skills across primates. *Current Opinion in Neurobiology*, 19(1), 45–51. <https://doi.org/10.1016/j.conb.2009.03.002>.
- Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour*, 58(5), 921–931. <https://doi.org/10.1006/anbe.1999.1242>.
- Sadr, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception*, 32(3), 285–293. <https://doi.org/10.1068/p5027>.
- Sandler, W., & Lillo-Martin, D. C. (2006). *Sign language and linguistic universals*. Cambridge University Press.
- Santana, S. E., Lynch Alfaro, J., & Alfaro, M. E. (2012). Adaptive evolution of facial colour patterns in neotropical primates. *Proceedings of the Royal Society B: Biological Sciences*, 279(1736), 2204–2211. <https://doi.org/10.1098/rspb.2011.2326>.
- Scheider, L., Liebal, K., Oña, L., Burrows, A., & Waller, B. M. (2014). A comparison of facial expression properties in five hylobatid species: Facial expressions in hylobatids. *American Journal of Primatology*, 76(7), 618–628. <https://doi.org/10.1002/ajp.22255>.
- Seyfarth, R., & Cheney, D. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour*, 40(4), 754–764.
- Sherwood, C. C. (2005). Comparative anatomy of the facial motor nucleus in mammals, with an analysis of neuron numbers in primates. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 287A(1), 1067–1079. <https://doi.org/10.1002/ar.a.20259>.
- Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: The need for multimodality in primate communication research. *Animal Behaviour*, 81(5), 919–924. <https://doi.org/10.1016/j.anbehav.2011.02.002>.
- Stein, T., Peelen, M. V., & Sterzer, P. (2011). Adults' awareness of faces follows newborns' looking preferences. *PLoS One*, 6(12), 1–6. <https://doi.org/10.1371/journal.pone.0029361>.
- Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2011). Chimpanzee Vocal Signaling Points to a Multimodal Origin of Human Language. *PLoS One*, 6(4), e18852. <https://doi.org/10.1371/journal.pone.0018852>.
- Tagliatalata, J. P., Russell, J. L., Pope, S. M., Morton, T., Bogart, S., Reamer, L. A., Schapiro, S. J., & Hopkins, W. D. (2015). Multimodal communication in chimpanzees. *American Journal of Primatology*, 77(11), 1143–1148.
- Tomasello, M. (2010). *Origins of human communication*. MIT press.
- Van Hooff, J. (1972). A comparative approach to the phylogeny of laughter and smiling. *Non-Verbal Communication*, 209–241.
- Waite, C., & Little, A. C. (2006). Preferences for Symmetry in Conspecific Facial Shape Among *Macaca mulatta*. *International Journal of Primatology*, 27(1), 133–145. <https://doi.org/10.1007/s10764-005-9015-y>.

- Waller, B. M., Whitehouse, J., & Micheletta, J. (2017). Rethinking primate facial expression: A predictive framework. *Neuroscience & Biobehavioral Reviews*, *82*, 13–21.
- Waller, B. M., & Micheletta, J. (2013). Facial expression in nonhuman animals. *Emotion Review*, *5*(1), 54–59.
- Waller, B. M., Bard, K. A., Vick, S.-J., & Smith Pasqualini, M. C. (2007). Perceived differences between chimpanzee (*Pan troglodytes*) and human (*Homo sapiens*) facial expressions are related to emotional interpretation. *Journal of Comparative Psychology*, *121*(4), 398.
- Waller, B. M., Liebal, K., Burrows, A. M., Slocombe, K., & E. (2013). How Can a Multimodal Approach to Primate Communication Help Us Understand the Evolution of Communication? *Evolutionary Psychology*, *11*(3), 147470491301100. <https://doi.org/10.1177/147470491301100305>.
- Waller, B. M., Whitehouse, J., & Micheletta, J. (2016). Macaques can predict social outcomes from facial expressions. *Animal Cognition*, *19*(5), 1031–1036.
- Waller, B. M., Julle-Daniere, E., & Micheletta, J. (2020). Measuring the evolution of facial ‘expression’ using multi-species FACS. *Neuroscience & Biobehavioral Reviews*, *113*, 1–11. <https://doi.org/10.1016/j.neubiorev.2020.02.031>.
- Watt, R., Craven, B., & Quinn, S. (2007). A Role for Eyebrows in Regulating the Visibility of Eye Gaze Direction. *Quarterly Journal of Experimental Psychology*, *60*(9), 1169–1177. <https://doi.org/10.1080/17470210701396798>.
- Whitney, H. M., Milne, G., Rands, S. A., Vignolini, S., Martin, C., & Glover, B. J. (2013). The influence of pigmentation patterning on bumblebee foraging from flowers of *Antirrhinum majus*. *Naturwissenschaften*, *100*(3), 249–256. <https://doi.org/10.1007/s00114-013-1020-y>.
- Wilke, C. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, *Pan troglodytes schweinfurthii*. *Animal Behaviour*, *12*.