UNIVERSITY OF CALIFORNIA

Los Angeles

Properties of Facial Signaling in Captive Chimpanzees (*Pan troglodytes*)

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of

Philosophy in Anthropology

by

Brittany Nicole Florkiewicz

2022

ABSTRACT OF THE DISSERTATION

Properties of Facial Signaling in Captive Chimpanzees (*Pan troglodytes*)

by

Brittany Nicole Florkiewicz

Doctor of Philosophy in Anthropology

University of California, Los Angeles, 2022

Professor Brooke Scelza, Chair

Since the 4th century B.C. there has been great interest in facial signaling in both human and non-human animals (Fridlund 1994). Early naturalists believed that the face was linked to the soul, and facial signals were the product of emotional experience (Bell 1806; Descartes 1649). As a result, most facial signaling studies to date have focused on placing facial muscle movement into discrete categories of emotion (Ekman 1970). Facial signals are assumed to be spontaneously produced and inflexible due to their strong ties to emotion; these signals are contrasted to other forms of non-vocal communication such as manual gestures, which are intentionally and flexibly produced (Pollick and de Waal 2007). In recent years, the idea that facial signals can be categorized based on physical form or corresponding emotion has been contested (Fridlund 1994; Waller et al. 2017). Some studies have even found evidence for goal association in great ape facial signals types such as the 'play face', which is considered to be an important gesture property (Cartmill and Byrne 2010). But, it is unclear as to whether great ape

facial signals are capable of taking on all important gesture properties, and if so, how they compare to manual gestures.

The goal of this dissertation was to examine the physical form and social function of facial signals in our closest living relatives, the chimpanzee (*Pan troglodytes*). The history of facial signaling inspired two different research questions in this project:

1. Can chimpanzee facial signals take on gesture properties, and if so, how do they compare to manual gestures?
2. Is there variation in the physical form of chimpanzee facial signals, and how does this compare to other ape species?

To address these two questions, I collected data using a combination of two previously described sampling methods: the focal individual sampling method and the opportunistic sampling method. A comparison of these two methods can be found in Chapter 1. The results of this dissertation suggest that chimpanzee facial signals can be used as gestures (Chapter 2), and that there is extensive variation in the physical form of chimpanzee facial signals when compared to other ape species (such as gibbons; Chapter 3). These findings have important implications for the evolution of sociality, gestural communication, and human language, which is discussed at the end of each chapter.

The dissertation of Brittany Nicole Florkiewicz is approved.

Harold Clark Barrett

Matthew Campbell

Joseph Howard Manson

Brooke Scelza, Committee Chair

University of California, Los Angeles

2022

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: A Comparison of Focal and Opportunistic Sampling Methods When Studying Chimpanzee Facial and Gestural Communication

No figures are presented in this chapter.

CHAPTER 2: Chimpanzee Facial Gestures and the Implications for the Evolution of Language

FIGURE 1. Great ape manual gestures are often defined as flexibly and intentionally produced bodily movements during bouts of communication to achieve a goal. It is possible that chimpanzee facial expressions are also capable of being used as gestures. In this photo, an adult male (right) beckons to an adult female (left) using a lower lip relaxer face and a reach gesture. This suggests that both signals are being used to achieve a goal. This goal is eventually met when the female approaches the male for affiliative contact. By comparing facial expressions with the described gesture repertoires of chimpanzees, we sought evidence of whether facial expressions like this might also be gestural.

CHAPTER 3: Primate Socio-Ecology Shapes the Evolution of Distinctive Facial Repertoires

FIGURE 1. A visual representation of how to code facial signals using FACS. Single muscle movements (which are referred to as Action Units or AUs) are used in combination leading to facial movement combinations (AU combinations).

FIGURE 2. A stacked histogram chart showing the proportion of observations (x-axis) for each AU type (y-axis) across chimpanzees and gibbons.

FIGURE 3. Complexity of AU combinations in chimpanzees (blue) and gibbons (red). Complexity is defined as the number of unique AUs (x-axis) an AU combination is composed of.

LIST OF TABLES

CHAPTER 1: A Comparison of Focal and Opportunistic Sampling Methods When Studying Chimpanzee Facial and Gestural Communication

TABLE 1. A list of all facial signal types that were considered in this study.

TABLE 2. Number and proportion of facial signal types observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport).

TABLE 3. Number and proportion of facial signals per individual observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport).

TABLE 4. Number and proportion of gesture signal types observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport).

TABLE 5. Number and proportion of gestures per individual observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport).

CHAPTER 2: Chimpanzee Facial Gestures and the Implications for the Evolution of Language

TABLE 1. A list of the four key gesture properties (communicative, intentional, flexible, and goal associated), their corresponding variables, and how they were measured. For binary variables, if the criteria outlined in the operational definition was met, the variable was coded as being present (i.e, 1); if not, the variable was coded as being absent (i.e., 0). A list of relevant citations (i.e., publications where each property is identified and described) is provided in the last column.

TABLE 2. A list of all facial expression types and manual gesture types that were considered. Categories for facial expressions are based on a facial expression ethogram initially developed by Parr et al. (2007). Additional categories were also created based on the unique signaling properties of this group (which were identified in a 2016 pilot study), such as Lipsmacking Face, Lower Lip Relaxer Face, and Raspberry Face. Categories for manual gestures are based on a gesture ethogram initially developed by Hobaiter (2011) and Byrne et al. (2017). We removed manual gesture types which were seldom observed in the current study (N < 10). We also modified manual gesture types based on the results of inter-observer reliability.

TABLE 3. Context tie index (CTI) scores for all facial expression types and manual gesture types considered in this study. CTI scores were assigned using the methods outlined in Pollick & De Waal (2007). We selected the most frequently observed behavioral context (and its corresponding proportion) as each signal type's CTI score. Scores closer to 1.00 mean that the signal type has less contextual flexibility, whereas scores closer to 0.00 mean that the signal type has greater contextual flexibility. Signal types highlighted in gray were used to compare our results to those outlined in Pollick & De Waal (2007). Neither the overall comparison between

facial expressions and manual gestures nor the Pollick & De Waal (2007) subset were significantly different.

TABLE 4. Binomial tests. The number and proportion of facial expressions and manual gestures that exhibited each gesture variable. We used binomial test results to determine whether gesture variables in facial expressions and manual gestures occurred significantly above or below chance level.

TABLE 5. Comparing GLMMs to null models. Comparing gesture variables GLMMs to their null counterparts using AIC scores.

TABLE 6. GLMM results. The results of the binomial GLMMs examining gesture variables as the outcome variable with signal modality and signaler ID as explanatory variables (with signaler ID set as a random explanatory variable). Odds ratios (OR) are presented for each model (FE/GE). FE stands for facial expression, and GE stands for manual gesture.

CHAPTER 3: Primate Socio-Ecology Shapes the Evolution of Distinctive Facial Repertoires

TABLE 1. A list of the 65 AU combinations that we observed in our current study. We have presented AU combinations, along with the number of observations for each ape type, in ascending order of combination complexity. We also present data on the number of observations per AU combination for chimpanzees and gibbons.

TABLE 2. A brief description of our two hypotheses, their corresponding predictions, whether these predictions were supported, and how we analyzed the evidence.

ACKNOWLEDGEMENTS

Throughout my Ph.D. I have received tremendous support from many people. First, I would like to thank the individual members of my dissertation committee. I am thankful to Clark Barrett and Joe Manson for their continuous insight, feedback, and support during my 6 years here at UCLA. I am grateful to my committee chair Brooke Scelza, who helped me overcome numerous obstacles during my research. I am especially indebted to Matthew Campbell, who served as my external committee member, co-chair, research collaborator, and role model. This dissertation would not have been possible without their tremendous support.

I am thankful to Bridget Waller, Linda Oña, and Leonardo Oña, who helped me acquire the skills and tools needed to conduct research on primate facial signaling. Special thanks to Linda and Leonardo for serving as my research collaborators; I look forward to continuing our work together in the future! I am grateful to Alan Fridlund, Jamie Whitehouse, and Jérôme Micheletta, whose work inspired the theoretical foundations of this dissertation. Special thanks to Dan Fessler, Greg Bryant, and other members of the eXperimental Biological Anthropology workgroup for giving extensive feedback on my research ideas that led to the development of this dissertation project. The knowledge, support, and friendship of graduate students and postdoctoral students in the Department of Anthropology have also been critical in my success and well-being. I am especially indebted to Adam Sparks, who has been both a mentor and role model to me during my time here at UCLA.

Thank you to Cathleen Cox for allowing me to conduct my research at the Los Angeles Zoo. I am grateful to the keepers, researchers, education specialists, and photographers at the Los Angeles Zoo for their advice and insight on chimpanzee behavior. I am especially thankful to Lori Conley, who provided me with extensive behavioral and historical information on the

chimpanzees. This information inspired new and interesting research questions, some of which have been included in this dissertation project. I am so grateful to my undergraduate research assistants and collaborators, which include Alexandra Molina, Cheyanne Krause, Lauren Scott, Misque Boswell, Rosa Munoz, and Sarah Yadegari. Special thanks to Alexandra, Cheyanne, and Sarah for their help with data collection and inter-observer reliability. It has been an honor and a privilege to work with these young aspiring scientists. I wish them nothing but happiness and success in their future research endeavors.

Finally, I would like to thank my family for their unwavering support and patience throughout this process. I would like to thank my parents Kelly and Richard, my brother Bryan, and their dogs Kasey, Skittles, and Kirby. I am incredibly grateful for my fiancé Jonathan Ryan, our dog Coco, our turtles Nausicaa, Totoro, and Ponyo, and our tortoise Eggbert. My entire academic career would not have been possible without their continuous love and support.

VITA

Education

2010 - 2014	B.A. (Anthropology), University of North Carolina Wilmington
2012	UNCEP Study Abroad Program, University of Technology: Sydney
2014	Field School Student, Primates of South Africa Program (SEPELA)
2014 - 2016	M.A. (Anthropology), Southern Illinois University Carbondale Thesis: Facial Behavior and Pair Bonds in Hylobatids

Publications (Selected)

2021	Florkiewicz B, Campbell MW. 2021. Chimpanzee Facial Gestures and Implications for the Evolution of Language. PeerJ 9:e12237.
2021	Florkiewicz B, Campbell MW. 2021. A Comparison of Focal and Opportunistic Sampling Methods in Chimpanzee Communication Research. Folia Primatologica 92(3):164-174.
2019	Molina A, Florkiewicz B, Cartmill EA. Exploring Sources of Variation in Inter-observer Reliability Scoring of Facial Expressions Using the ChimpFACS. FASEB Journal 33(1):872.
2018	Florkiewicz BN, Skollar G, Reichard, UH. 2018. Facial Expressions and Pair-Bonds in Hylobatids. American Journal of Physical Anthropology 167(1):108-123.

Conference Presentations (Selected)

2021	California Workshop on Evolutionary Social Sciences Title: A Comparison of Focal and Opportunistic Sampling Methods in Chimpanzee Communication Research
2018	International Primatological Society Title: Infant-Adult Similarity in Facial and Gestural Signals of Captive Chimpanzees
2017	European Federation for Primatology Title: Studying Multi-Component Facial Behavior in Primates
2016	American Association of Physical Anthropologists Title: Facial Behavior and Pair Bonds in Hylobatids

Honors and Awards (Selected)

2019	Dorothy Jewell Graduate Travel Award Awarded by: UCLA Anthropology Department

2019   Graduate Student Research Mentorship Fellowship
        Awarded by: UCLA Graduate Division

2017   Summer Research Equipment Award
        Awarded by: UCLA Center for Language, Interaction, and Culture

2016   Graduate Dean's Scholar Award
        Awarded by: UCLA Anthropology Department

Teaching (Selected)
2017 - 2021   University of California Los Angeles, Los Angeles, CA
        Role: Anthropology Teaching Assistant/Associate/Fellow

2014 - 2016   Southern Illinois University Carbondale, Carbondale, IL
        Role: Anthropology Teaching Assistant

Invited Talks (Selected)
03/2021   UCLA Center for Behavior, Evolution & Culture
        Lecture Title: At Face Value: The Role of Chimpanzee Facial
        Expressivity in the Evolution of Gestural Communication and Social
        Bonding

10/2018   University of California Los Angeles
        Lecture Title: Primate Communication
        Course: Animal Human Interactions

01/2016   University of North Carolina: Wilmington
        Lecture Title: Primate Cognition
        Course: Intro to Primate Behavior and Ecology

Service (Selected)
2020 - 2021   UCLA Biological Anthropology Diversity, Equity, and Inclusion
        Working Group
        Role: Graduate Student Representative

2017 - 2019   UCLA Anthropology Admissions Committee
        Role: Graduate Student Representative

2017   UCLA Center for Language, Interaction, and Culture
        Role: CLIC GSA Annual Conference Committee Member

2015 - 2016   SIUC Anthropology Graduate Student Association
        Role: President

CHAPTER 1

A Comparison of Focal and Opportunistic Sampling Methods When Studying Chimpanzee Facial and Gestural Communication

BRITTANY N. FLORKIEWICZ[1], MATTHEW W. CAMPBELL[2]

[1] *Department of Anthropology and Center for Behavior, Evolution, & Culture, University of California, Los Angeles, Los Angeles, California, U.S.A.*

[2]*Department of Psychology, California State University, Channel Islands, Camarillo, California, U.S.A.*

A Comparison of Focal and Opportunistic Sampling Methods when Studying Chimpanzee Facial and Gestural Communication

Brittany N. Florkiewicz[a]   Matthew W. Campbell[b]

[a]Department of Anthropology and Center for Behavior, Evolution, and Culture, University of California, Los Angeles, CA, USA; [b]Department of Psychology, California State University, Camarillo, CA, USA

Abstract

Researchers frequently use focal individual sampling to study primate communication. Recent studies of primate gestures have shown that opportunistic sampling offers benefits not found in focal individual sampling, such as the collection of larger sample sizes. What is not known is whether the opportunistic method is biased towards certain signal types or signalers. Our goal was to assess the validity of the opportunistic method by comparing focal individual sampling to opportunistic sampling of facial and gestural communication in a group of captive chimpanzees (*Pan troglodytes*). We compared: (1) the number of observed facial and gestural signals per signal type and (2) the number of observed facial and gestural signals produced by each signaler. Both methods identified facial signals, gesture signals, and gesture signalers at similar relative rates, but the opportunistic sampling method yielded a more even distribution of signalers and signal types than the focal individual sampling method. In addition, the opportunistic sampling method resulted in larger sample sizes for both facial and gestural communication. However, the opportunistic method did not allow us to calculate the signals per time for each individual, which is easily done with the focal individual method. These results suggest that the opportunistic sampling method is (1) comparable to the focal individual sampling method in multiple important measures, (2) associated with additional sampling benefits, and (3) limited in measuring some variables. Thus, we recommend that future studies use a mixed-methods approach, as focal individual and opportunistic sampling have distinct strengths that complement each other's limitations.

© 2021 S. Karger AG, Basel

Introduction

Whereas Altmann [1974] outlined seven types of sampling methods in animal behavior research, primate facial and gestural studies have used almost exclusively just one: focal individual sampling [Liebal et al., 2004; Candiotti et al., 2012; Scheider et al., 2014; Hobaiter et al., 2017]. The focal individual sampling method offers three main benefits in the study of visual and vocal communication.

Correspondence to:
Brittany N. Florkiewicz, bnf6001@gmail.com

First, focal individual sampling helps to reduce the likelihood of individual/subgroup bias by having all members of the group sampled for equal amounts of time. Primates exhibit variation in signaling: some primates interact with others frequently using communicative signals, whereas others seldom use communicative signals during social interactions. Variation in signaling behavior can be attributed to numerous factors, such as age, activity budgets, social rank, and social bond strength [Rose, 2000; Bergman and Sheehan, 2013; Florkiewicz et al., 2018]. Using the focal individual sampling method ensures that researchers do not become biased towards individuals who are more expressive, present, active, or visible than others. Second, focal individual sampling helps to reduce the likelihood of bias towards a given signal type or social context. Some communicative signals are subtle, including only minor muscle movements and/or soft vocalizations. In contrast, other signals are overt, including a large range of muscle movement and/or loud vocalizations.

Focal individual sampling is used to record all signals that are produced by a given individual/subgroup for a set amount of time. This leads to the third benefit of focal individual sampling, which is that it can help reduce the probability of conditional sampling bias [Rose, 2000]. Conditional sampling bias occurs when activities that are typically present in difficult viewing conditions are underrepresented in the final data set [Rose, 2000]. Conditional sampling bias is a common problem in communication studies when conducted with wild populations of primates who occupy densely forested habitats [Rose, 2000; Hobaiter and Byrne, 2011]. Required recording of each individual may increase the probability of viewing a subtle signal in more difficult viewing conditions.

While the focal individual sampling method is associated with multiple benefits, there is one major drawback: adequate sample sizes. In primate communication research, an adequate number of samples per individual and per signal type are required to make meaningful inferences about signaling properties. Adequate sample sizes are especially important for research on primate gestural communication. Great apes appear to be capable of producing over 60 distinct types of gestures [Byrne et al., 2017]. Some of these gesture types are frequently observed across individuals, whereas others are relatively rare. Gesture researchers typically strive for larger sample sizes to help capture rare behaviors and to account for issues pertaining to sample size. Six studies between the years of 2007 and 2018 averaged 14.01 gestures recorded per hour [Liebal et al., 2006; Pollick and de Waal, 2007; Genty et al., 2009; Cartmill and Byrne, 2010; Hobaiter and

Byrne, 2011; Graham et al., 2018]. This average is higher when compared to other types of primate communication studies, such as those examining primate facial signaling. Four studies between 2007 and 2015 averaged just 0.90 facial signals recorded per hour [Pollick and de Waal, 2007; Parr et al., 2010; Waller et al., 2015].

To achieve larger sample sizes, some researchers use alternative sampling techniques. For example: some primate gesture researchers use ad libitum video sampling to capture social interactions which were likely to result in gestural communication [Cartmill and Byrne, 2010; Smith and Delgado, 2015; Roberts and Roberts, 2015]. This ad libitum sampling method will be referred to here as "opportunistic," where opportunities for observing signals (mainly in the form of social interactions) are the focus of continuous video recordings [Pika and Bugnyar, 2011]. In opportunistic sampling, the most active part of the group is followed; individuals who offer the highest likelihood of interacting with others are recorded. Individuals who are more likely to interact with others vary throughout the day, leading to differences in the number of hours recorded for each subject. In contrast, the focal individual sampling method typically results in a fixed number of recorded hours for each subject. The opportunistic sampling method has been used to study the gestural communication of chimpanzees (*Pan troglodytes*) [Roberts and Roberts, 2015], bonobos (*Pan paniscus*) [Smith and Delgado, 2015], orang-utans (*Pongo pygmaeus* and *Pongo abelii*) [Cartmill and Byrne, 2010], mandrills (*Mandrillus sphinx*) [Laidre, 2011], and even nonprimate species (such as ravens, *Corvus corax*) [Pika and Bugnyar, 2011].

The opportunistic sampling method shares some similarities to ad libitum sampling and all occurrence sampling as outlined by Altmann [1974]. In ad libitum sampling, the researcher records as much information as possible on the most salient (i.e., immediately noticeable) behaviors of the social group [Altmann, 1974]. Similar to opportunistic sampling, the amount of time spent during each ad libitum sampling event varies and is at the discretion of the researcher [Altmann, 1974]. However, in opportunistic sampling, the focus is on a specific set of salient behaviors (i.e., facial and gestural signals). In all occurrence sampling, the researcher records all occurrences of a certain type of behavior in a social group [Altmann, 1974], often simultaneously with another method, like focal individual sampling or scan sampling. Similar to all occurrence sampling, opportunistic sampling focuses on certain types of behaviors (mainly those which are communicative). Yet, opportunistic sampling is not limited to

Florkiewicz/Campbell

a single signal type and often includes information on multiple signal types exchanged throughout a social interaction.

One recent gesture study shows that the opportunistic sampling method can be used to achieve larger sample sizes. In Cartmill and Byrne [2010], they identified 31.6 gestures per hour of video footage across all individuals, which is double the number of gestures per hour observed in other orang-utan gesture studies using the focal individual sampling method (such as Liebal et al. [2006], where they observed an average of 13.2 gestures per hour across all individuals). However, it is unclear whether the opportunistic sampling method is comparable to the focal individual sampling method regarding: (1) the number of observed signals per gesture type and (2) the number of gestures produced by each signaler. It is also unclear whether the opportunistic method could be useful for increasing sample sizes for other types of primate communication studies, such as those focused on facial signaling.

Some gesture researchers have started using methods which combine both focal individual sampling and opportunistic sampling in hopes of compensating for the weaknesses of both sampling methods. One example of this is the "focal behavior sampling" method [Hobaiter and Byrne, 2011; Graham et al., 2016, 2018]. In this method, researchers follow a predetermined individual, but they only record when the focal individual comes into close proximity with one or more surrounding individuals [Graham et al., 2018]. The focal behavior sampling method shares strengths and limitations with each of focal individual sampling and opportunistic sampling. As a hybrid method, we did not attempt to include it in the present study. Rather, we chose to focus on the performance of the ends of the spectrum of methods: focal individual sampling and opportunistic sampling.

Our goal was to compare the focal individual sampling method to the opportunistic sampling method with regard to the number and types of communicative signals observed. Our study focused on two research questions associated with the main benefits of the focal individual sampling method (that are applied to both facial and gestural communication). First, is the opportunistic sampling method comparable to the focal individual sampling method regarding the number of observed signals per signal type? And second, is the opportunistic sampling method comparable to the focal individual sampling method regarding the number of observed signals recorded for each signaler? To address both of these questions, we collected a similar amount of video footage us-

ing each sampling method. Then, we compared the number of signals per signaler and per signal type across the two sampling methods. Our results will provide insight and guidance into the benefits and costs of each sampling method when studying primate communication.

Methods

Data Collection

We observed a single group of chimpanzees (*P. troglodytes*) at the Los Angeles Zoo from June 2017 to 2018. The study began with 13 adult individuals, but 1 elderly adult died part way through the data collection. We removed her data from the analysis, yielding a final sample size of 12 individuals. The current study focuses on the behavior of adult chimpanzees, since adulthood is characterized by stability in social behavior [Tkaczynski et al., 2020]. We decided not to include infants and juveniles (i.e., individuals less than 7 years of age at the time of data collection) since individuals can change a lot over 1 year during these times of development. Information regarding the names, birth dates, birthplaces, and age groups of the chimpanzees can be found in supplement Table S1 (see www.karger.com/doi/10.1159/000516315 for all online suppl. material). Data collection for both sampling methods took place Monday through Friday between the hours of 08:00 and 14:00, which we identified in a 2016 pilot study as being peak activity hours (where numerous social interactions take place). The Los Angeles Zoo and Botanical Gardens approved our data collection protocol, and we conducted our study in accordance with the Association for the Study of Animal Behaviour's guidelines for the treatment of animals in behavioral research [2020]. We collected data using a Panasonic Full HD Video Camera Camcorder HC-V770 and a Sennheiser MKE400 external camera-mount shotgun microphone. We formatted and saved videos in AVCHD and MP4 format.

We used two methods to collect data: the opportunistic sampling method and the focal individual sampling method. We used the focal individual sampling method across 51 days from June to August of 2017 and the opportunistic method across 34 days from June to August of 2018. We were initially collecting data for a study of chimpanzee facial signals. At the end of the 2017 data collection, it was clear that focal follows were not generating an adequate sample size of signals. In 2017, we were only able to record 313 gestures. Had we used focal individual sampling in 2018, we could have expected a sample size around 616 gestures across 136 hours of video footage (or 4.53 gestures per hour). When compared to previous studies, our gesture sample size would have been well below average [average of 14.01 gestures per hour across 6 studies: Liebal et al., 2006; Pollick and de Waal, 2007; Genty et al., 2009; Cartmill and Byrne, 2010; Hobaiter and Byrne, 2011; Graham et al., 2018]. Therefore, the following year we switched to opportunistic sampling. The use of two sampling methods provided a unique and unplanned opportunity to compare their data to each other, which is what we are presenting in this study.

With focal individual sampling, we recorded video in set 30-min intervals while following one member of the troop. We randomly selected individuals to be sampled out of a list of all group members present in the main enclosure that day (group composition could change from day to day), so long as they had not yet been recorded that week. This prevented individual chimpanzees from

4

Table 1. A list of all facial signal types that were considered in this study

Facial signal	Description	Possible context(s)
Bared teeth face	Corners of the lips are drawn backwards, exposing both rows of teeth	Affiliative; agonistic
Play face	Mouth is opened, lips possibly drawn backwards, and bottom row of teeth possibly exposed. Sometimes accompanied with a vocalization	Play
Pant-hoot face	Lips are funneled together, with vocalization being produced	All contexts
Ambiguous face	Communicative movement of the face is observed, but corresponding category is unclear	All contexts
Scream face	Similar to bared teeth face, but jaw is stretched and a loud vocalization is produced.	Agonistic
Pout face	Lips are funneled close together and pushed outwards while the chin is furrowed upward	Food; play
Whimper face	Lips are funneled close together while the corners of the lips are drawn backwards; often accompanied with a low vocalization	Food; play; agonistic
Lower lip relaxer	Bottom lip is relaxed away from the gums and bottom row of teeth	All contexts
Raspberry face	Lips are pressed tightly together (sometimes with the tongue) to create a loud sound	Grooming
Lipsmacking face	Mouth is opened and closed rapidly, producing a low sound	Grooming

Facial signal types are assigned based on: (1) facial muscle movement and (2) standardized behavioral contexts, which were developed by Parr et al. [2007]. Additional categories were also created based on the presence of unique facial signaling types present at the Los Angeles Zoo: these include the lower lip relaxer face, raspberry face, and ambiguous face

being sampled on the same day or the same time each week, or multiple times in one week and not in subsequent weeks. Each day, we selected 3–4 chimpanzees for focal individual sampling so that we recorded all chimpanzees by the end of each week. Videos included the behaviors performed by the focal and neighboring individuals (i.e., who are in close proximity and are included in video footage), as well as periods of inactivity if the focal individual chose to rest. We observed all chimpanzees for an equal amount of time using the focal individual sampling method (4 h each), with each individual sampled at least once during each week of data collection.

With opportunistic sampling, we recorded video during peak activity times while following the largest and most active part of the group. Video recordings varied in length based on the duration of peak activity times. Recording began shortly before the start of a social interaction (i.e., a communicative event) and ended shortly after the end of the social interaction (i.e., when chimpanzees dispersed and/or ceased their communication). We coded all video footage taken using this sampling method for the analysis.

Data Coding
We coded all video in ELAN 5.6-AVFX using a custom annotation template. We identified signals using the definition of Maynard-Smith and Harper [1995], which is an action that is per-

formed by a signaler which attempts to alter the behavior of one or more recipients. In this study, we considered two modes of communication: facial and gestural. We defined facial signals as facial muscle movements used for the purpose of communicating with conspecifics in social interactions. Facial muscle movements used for bodily maintenance (such as blinking, yawning, feeding) or object manipulation (such as chewing, scraping, or tearing at objects) were not considered in the current study since it is difficult to determine whether these movements are used communicatively. We defined gestures as flexibly and intentionally produced bodily movements used for the purpose of communication [Moore, 2016; Byrne et al., 2017; Byrne and Cochet, 2017]. We coded each signal in the interaction for signaler identity and signal modality (i.e., whether it was a facial signal or gesture).

After the signal modality was identified, we coded for signal type. We adapted facial signal types from Parr et al. [2007]. This facial signaling ethogram can be found in Table 1. We adapted gesture types from Hobaiter and Byrne [2011, 2014], which lists over 60 frequently used gestures by apes in captive and wild settings. We added additional gesture types to this ethogram, which we identified to be unique to the Los Angeles Zoo chimpanzees in a 2016 pilot study. These include the following gesture types [which were derived from Byrne et al., 2017]: present body part (general),

5

Table 2. Number and proportion of facial signal types observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport)

Facial signal type	Observations (focal), n	Proportion of observed signals (focal), %	Observations (opport), n	Proportion of observed signals (opport), %
Ambiguous face	18	10.98	60	12.05
Bared teeth face	10	6.10	31	6.22
Lipsmacking face	33	20.12	49	9.84
Lower lip relaxer face	3	1.83	24	4.82
Pant-hoot face	47	28.66	176	35.34
Play face	17	10.37	56	11.24
Pout face	1	0.61	4	0.80
Raspberry face	27	16.46	19	3.82
Scream face	8	4.88	78	15.66
Whimper face	0	0.00	1	0.20
Total	164	100.00	498	100.00

thrust, and unsure/unknown (for bodily movements that are used as communicative gestures but could not be neatly classified into any of the other gesture type categories).

Interobserver Reliability

We conducted interobserver reliability on approximately 10% of observed signals including both facial signals ($n = 149$, or 15.25% of all facial signals) and gestures ($n = 270$, or 12.94% of all gestures). A research assistant coded the subset of facial signals, and the second author coded the subset of gestures. We then compared these results to the results of the first author, who coded both. Facial signal types had 74.83% agreement and a Cohen's κ of 0.604, which was considered to be very good agreement and close to substantial agreement [McHugh, 2012]. Gesture types had 65.56% agreement after modifications had been made to categories, which reduced the number of possible unique gesture types from 62 to 55. We combined categories of gestures that were morphologically similar to one another (such as "hand on," "touch other," and "grab") and had lower agreement when separated. While percentage of agreement was lower for gesture types, the corresponding Cohen's κ-score, which is typically used to evaluate agreement for gesture types [Hobaiter, 2011], was considered to be very good agreement and close to substantial agreement (κ = 0.604; the identical κ-scores for both facial signal types and gesture types is a coincidence).

Data Analysis

We performed statistical tests using base functions in R 3.6.2 [R Core Team, 2017]. Behavioral data, along with the code used to analyze these behavioral data, can be found in the supplement. We used Spearman's rank correlation to determine whether the rank (i.e., relative rate) of signal types/signaler IDs remained consistent between sampling methods for each modality. We chose Spearman's rank correlation (as opposed to a Pearson or Kendall correlation) for 3 reasons: (1) to account for differences in sample size (since the number of signals in 2018 increased by twofold compared to 2017); (2) because the data are categorical (i.e., signal type and signaler ID) as opposed to continuous; and (3) the number of signals in each year was not normally distributed. If a strong correlation value was observed (≥ 0.70), we concluded that the rank of signal

types/signaler IDs was comparable between sampling methods. If a moderate or weak correlation value was observed (<0.70), we concluded that the rank of signal types/signaler IDs was not comparable between sampling methods. We interpreted correlation strength using Akoglu's guide for correlation coefficients [2018].

Next, we used χ^2 tests to determine whether there were significant differences in the proportion of signals for each signal type/signaler ID between sampling methods for each modality. If we observed a significant p value ($p \leq 0.05$), we concluded that the proportion of signals for each signal type/signaler ID was not comparable between sampling methods. If we observed a nonsignificant p value ($p > 0.05$), we concluded that the proportion of signals for each signal type/signaler ID was comparable between sampling methods. We chose χ^2 tests for similar reasons as Spearman's rank correlation (different sample sizes, categorical data, nonnormal distributions), and because the opportunistic sampling method does not contain information on the number of recorded minutes/hours per individual, making it difficult to implement models (such as logistic models) which compare rates of behavior. When sample sizes were small ($n < 5$), we used Fisher's test to verify the results of the χ^2 test, since Fisher's tests are ideal for smaller sample sizes.

Finally, we used statistical F tests [Snedecor and Cochran, 1989] to evaluate the equality of variances between sampling methods for: (1) the number of observed signals produced by each signaler and (2) the number of observed signals per signal type. We performed F tests for both facial and gestural signals. F is calculated by dividing two sample variances; this results in a positive number. We then interpreted the F statistic, coupled with degrees of freedom, using the F table at the desired significance level. We present effect sizes with the results of each F test using Cohen's d.

Results

We collected 142.5 h of video footage (68 h using focal individual sampling and 74.5 h using opportunistic sampling), which is similar to a previous study on chimpan-

6

Table 3. Number and proportion of facial signals per individual observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport)

Signaler	Observations (focal), n	Proportion of observed signals (focal), %	Observations (opport), n	Proportion of observed signals (opport), %
Ben	25	15.24	64	12.85
Glenn	8	4.88	49	9.84
Gracie	14	8.54	19	3.82
Jake	19	11.59	63	12.65
Jean	6	3.66	72	14.46
Jerrard	10	6.10	8	1.61
Julie	10	6.10	31	6.22
Nan	9	5.49	38	7.63
Regina	9	5.49	63	12.65
Shaun	43	26.22	49	9.84
Yoshi	9	5.49	29	5.82
Zoe	2	1.22	13	2.61
Total	164	100.00	498	100.00

zee communication [Hobaiter and Byrne, 2011]. We spent a total of 116 h across 51 days doing focal individual sampling and 101 h across 34 days doing opportunistic sampling. Out of the 51 days spent at the zoo for focal individual sampling, 34 resulted in video recordings that potentially contained signaling information. Out of the 34 days spent at the zoo for opportunistic sampling, 29 resulted in video recordings that potentially contained signaling information.

We identified a total of 1,739 signals (which includes both facial signals and gestures) from the video footage. We observed 477 signals ($n = 164$ facial signals and $n = 313$ gestures) with focal individual sampling and 1,262 signals ($n = 498$ facial signals and $n = 764$ gestures) with opportunistic sampling. The difference in number of observed facial signals and gestures between sampling methods is close to statistical significance ($\chi^2_{1, 1,739} = 3.576$, $p = 0.059$). The opportunistic sampling method resulted in a greater number of signals observed per hour for facial signals, gestures, and both signaling modalities when compared to the focal individual sampling method.

Facial Signals

We observed 164 facial signals using the focal individual sampling method and 498 facial signals using the opportunistic sampling method. The opportunistic sampling method resulted in a greater sample size of observed facial signals (Table 2). There was a moderate correlation in the ranking of facial signal types between sampling methods, which was close to statistical significance

(Spearman's $r = 0.624$; $p = 0.060$). Thus, facial signal types were either common in both methods (e.g., pant hoot face and play face) or rare in both methods (e.g., whimper face and pout face, Table 2). However, the exact proportions of facial signal types differed significantly between sampling methods ($\chi^2_{9, 662} = 54.92$, $p < 0.001$). Due to smaller sample sizes, we verified this result with a Fisher's test ($p < 0.001$). When examining the distribution of observations across facial signal types, it appears that the focal individual sampling method resulted in a more even distribution of signals (SD = 9.37%) compared to the opportunistic sampling method (SD = 10.22%); this difference is statistically significant ($F = 0.091$; $p = 0.001$) with a large effect size ($d = 0.889$; $r = 0.406$). At the same time, the focal individual sampling method yielded a much smaller number of facial signals ($n = 164$) than the opportunistic method ($n = 498$), so it is not clear which method should be viewed as more representative.

We ordered individuals from largest to smallest proportion of facial signals recorded using the focal individual sampling and opportunistic sampling; there was no correlation between the rankings (Table 3, Spearman's $r = 0.120$; $p = 0.7095$). In addition, the opportunistic sampling method was not comparable to the focal individual sampling method with regard to the proportion of signals produced by each facial signaler (Table 3, $\chi^2_{11, 662} = 62.563$, $p < 0.001$). Due to smaller sample sizes, we verified this result with Fisher's test ($p < 0.001$). When examining the distribution of observations across facial signalers between sampling methods, it appears that the focal indi-

Florkiewicz/Campbell

Table 4. Number and proportion of gesture signal types observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport)

Gesture type	Observations (focal), n	Proportion of observations (focal), %	Observations (opport), n	Proportion of observations (opport), %
Arm touch	101	32.27	222	29.06
Arm raise	7	2.24	30	3.93
Arm wave	0	0.00	1	0.13
Beckon	2	0.64	2	0.26
Big loud scratch	0	0.00	4	0.52
Bite	12	3.83	77	10.08
Clap	2	0.64	10	1.31
Directed push	4	1.28	5	0.65
Embrace	31	9.90	44	5.76
Foot present	0	0.00	1	0.13
Hand fling	6	1.92	27	3.53
Hand shake	4	1.28	0	0.00
Hit with object	1	0.32	1	0.13
Jump	0	0.00	7	0.92
Kick	0	0.00	3	0.39
Look	15	4.79	40	5.24
Mouth stroke	1	0.32	0	0.00
Object move	1	0.32	4	0.52
Object shake	0	0.00	2	0.26
Poke	1	0.32	1	0.13
Pounce	0	0.00	1	0.13
Present body part	8	2.56	39	5.10
Present climb on me	6	1.92	4	0.52
Present grooming	1	0.32	3	0.39
Present sexual	4	1.28	14	1.83
Punch object/ground	2	0.64	0	0.00
Punch other	5	1.60	8	1.05
Push	19	6.07	34	4.45
Reach	15	4.79	38	4.97
Roll over	3	0.96	1	0.13
Slap other	3	0.96	22	2.88
Slap other with object	1	0.32	3	0.39
Stomp	0	0.00	1	0.13
Stomp 2 feet	1	0.32	1	0.13
Swing/rock	53	16.93	72	9.42
Tandem walk	0	0.00	3	0.39
Tap/slap/knock object	2	0.64	24	3.14
Throw object	2	0.64	6	0.79
Thrust	0	0.00	3	0.39
Unsure/unknown	0	0.00	2	0.26
Walk (stiff)	0	0.00	4	0.52
Total	313	100.00	764	100.00

vidual sampling method resulted in a less even distribution of signals (SD = 6.72%) compared to the opportunistic sampling method (SD = 4.36%); this difference is statistically significant ($F = 0.257$; $p = 0.034$) with a very large effect size ($d = 1.616$; $r = 0.628$).

Gestures

We observed 313 gestures across 29 gesture types using the focal individual sampling method and 764 gestures across 38 gesture types using the opportunistic sampling method (Table 4). We observed 41 unique gesture

Table 5. Number and proportion of gestures per individual observed using the focal individual sampling method (focal) and the opportunistic sampling method (opport)

Signaler	Observations (focal), n	Proportion of observed signals (focal), %	Observations (opport), n	Proportion of observed signals (opport), %
Ben	36	11.50	84	10.99
Glenn	34	10.86	133	17.41
Gracie	31	9.90	97	12.70
Jake	8	2.56	37	4.84
Jean	9	2.88	73	9.55
Jerrard	32	10.22	46	6.02
Julie	97	30.99	94	12.30
Nan	6	1.92	26	3.40
Regina	13	4.15	44	5.76
Shaun	24	7.67	72	9.42
Yoshi	14	4.47	42	5.50
Zoe	9	2.88	16	2.09
Total	313	100.00	764	100.00

types across both sampling methods. Similar to what was observed with facial signals, the opportunistic method resulted in a greater sample size of gestures. In addition, the opportunistic sampling method resulted in a greater number of unique gesture types observed (Table 4). There was a strong correlation in the ordering of gesture types between sampling methods (Spearman's $r = 0.689$; $p < 0.001$). Thus, gesture types were either common in both methods (e.g., look, swing/rock, and embrace) or rare in both methods (e.g., stomp 2 feet, roll over, and hit with object). However, the exact proportions of gesture types observed across sampling methods were not comparable (Table 5, $\chi^2_{40, 1,077} = 89.406$, $p < 0.001$). Due to smaller sample sizes, we verified this result with Fisher's test ($p < 0.001$). When examining the distribution of observations across gesture types, it appears that the focal individual sampling method resulted in a less even distribution of signals (SD = 5.72%) compared to the opportunistic sampling method (SD = 4.95%). The difference in standard deviations is statistically significant ($F = 0.225$; $p < 0.001$), but the effect size is small ($d = 0.372$; $r = 0.183$).

We ordered individuals from most to least gestures recorded between the two methods, and there was a strong correlation between the rankings (Table 5, Spearman's $r = 0.781$; $p = 0.003$). Thus, individuals either produced a large number of gestures in both methods or a smaller number of gestures in both methods. However, the opportunistic sampling method was not comparable to the focal individual sampling method with regard to the exact proportion of signals produced by each individual (Table 5, $\chi^2_{11, 1,077} = 77.101$, $p < 0.001$). Due to smaller sample sizes,

we verified this result with Fisher's test ($p < 0.001$). When examining the distribution of observations across gesture signalers, it appears that the focal individual sampling method resulted in a less even distribution of signals (SD = 7.99%) compared to the opportunistic sampling method (SD = 4.50%). But this difference in standard deviations is not statistically significant ($F = 0.530$; $p = 0.308$), and the effect size is moderate ($d = 0.625$; $r = 0.298$).

Discussion

Our goal was to compare the focal individual sampling method to the opportunistic sampling method for recording facial and gestural communication of chimpanzees. We found that the opportunistic sampling method is comparable to the focal sampling method when it comes to the relative frequency (i.e., rank) of facial signal types, gesture signal types, and gesture signalers, while collecting more signals per unit time invested. Since the opportunistic method performs similarly to the focal individual method in these key areas, we conclude that the opportunistic method is a valid approach to sampling signal repertoires.

In addition to these similarities, the opportunistic method recorded signal types not recorded in focal individual sampling for both facial signals and gestures, and a more even distribution of signal types/signalers. We found that for facial signal types, gesture signal types, and gesture signalers, the opportunistic sampling method resulted in a smaller standard deviation across signals/signalers when compared to the focal individual sampling method. This

Florkiewicz/Campbell

pattern was statistically significant for facial and gesture signal types. Having a more even distribution of signal types/signalers may be beneficial for studies that wish to examine and compare the properties of communicative signals. For example, studies aimed at examining the proportion of signal types associated with a given communicative property (such as persistence or elaboration) would benefit from a more even distribution of observations.

Additional information would be required to determine why the opportunistic sampling method resulted in a more even distribution of signals/signalers. It is possible that additional recording opportunities, variation in expressivity, or both played a role in the distribution of signals across signal types and signalers using the opportunistic sampling method. Examining the number of hours of video footage collected per individual using the opportunistic sampling method could be used to examine recording opportunities and rates of signaling. While these data exist in theory, logistically it was impractical to code this information. Between the movement of the camera, the movement of the chimpanzees, and the number of individuals in the group ($n = 12$), the chimpanzees in view changed too often to make logging the time on screen for each individual reasonable.

This leads to one of the main benefits of focal individual sampling when compared to opportunistic sampling: the focal individual sampling method can be easily used to calculate baseline rates of facial and gestural signaling by each individual, since each individual is recorded for the same amount of time. Baseline signaling rates are especially helpful for studies that wish to examine the effects of external variables on rates of communication (which could include food, predators, number of social interactions, etc. [Hopkins et al., 2011]). In this study, we sampled each individual using the focal method for a total of 4 h. Having rates of facial and gestural signaling helps to determine whether changes in the number of observed signals/signalers are the result of sampling bias or the result of individual variation in expressivity. This information might explain why the ranking of facial signalers was not consistent across both sampling methods, even though the ranking of facial signal types, gesture signal types, and gesture signalers was comparable.

We suggest three possible explanations for why the focal individual sampling method was not comparable to the opportunistic sampling method when considering the rank of observed facial signals per signaler ID. The first explanation is attributed to differences in activity levels. In opportunistic sampling, we only collected data when there was communicative activity happening within the group,

and even then, we prioritized the largest and most active subgroup of chimpanzees. In focal individual sampling, video footage could contain activity or inactivity. If an individual was resting during its assigned focal follow, then the video contained a lot of inactivity. More bouts of communicative activity could lead to a greater probability of observing a signal that is less commonly used/produced, leading to a skew in observed proportions (Table 2, significant χ^2 test). This leads to the second explanation, which is that focal follows were potentially too long in this given project (30 min). Future studies should examine whether conducting a greater number of shorter focal follows (such as 5 min [Liebal et al., 2004]) changes the distribution of facial signals across signalers. Shorter follows may allow researchers to switch from focals who are engaging in bouts of inactivity to focals who are engaging in social behavior more quickly, resulting in a greater number of signals observed. However, it is also possible that shorter focal follows will result in missing opportunities to record interactions which occur later in longer focal follows.

The final explanation is that changes in signal distributions between sampling methods may be the result of changes in social structure. In late 2017, the oldest and consensus highest-ranking female at the Los Angeles Zoo died at the age of 51 (who was not included in the current study). She maintained her position as a high-ranking female for most of her adult life. After her death, the dominance hierarchy in the female line, and possibly the male line as well, appeared to be in a state of flux for some time. It is possible that the changes in the dominance hierarchy affected communication in at least a couple of potential ways. Individuals may have socialized more or less as their position in the hierarchy changed, which could have affected their opportunity to communicate and thus the number of signals recorded for them. In addition, changes in social rank and role in the group could also have affected the kinds of social situations individuals found themselves in, potentially leading to changes in the types of signals produced, not just the number.

One limitation of our study is time, which is an uncontrolled variable. We used the focal individual sampling method in 2017 and the opportunistic sampling method in 2018. As noted earlier, we initially did not set out to compare sampling methods; rather, we are taking advantage of a decision to change sampling methods to compare their performance in hopes of providing some guidance to the field. In theory, it is possible that the differences between the sampling methods could be a product of the chimpanzees changing their behavior over time, and not differences between the performance of the meth-

ods, but we think that this is unlikely. Firstly, the similarities between the sampling methods (relative frequency of facial signal types, gesture signal types, and gesture signalers) are evidence of stability across the years. If the chimpanzees had undergone a sizable change in behavior, we would expect it to affect these areas as well as the two (signal types and the evenness of the distribution of signal types/signalers) where we saw differences between opportunistic and focal individual sampling.

Secondly, we focused on adult chimpanzees since adulthood is associated with behavioral stability [Tkaczynski et al., 2020]. Adult chimpanzees exhibit long-term consistency in their social behavior (such as consolatory behavior [Webb et al., 2017]) even when controlling for seasonal, demographic, and temporal changes [Tkaczynski et al., 2020]. Adult chimpanzees also exhibit long-term stability in both male and female social bonds [Pusey and Schroepfer-Walker, 2013; Bray and Gilby, 2020]. The results of the current study suggest that chimpanzees exhibit long-term stability in facial and gestural communication as well. While the exact number of observed signals/signalers changed over the 2-year period, the relative ranks of facial signal types, gesture signal types, and gesture signalers remained consistent. Between our results and those of others, similarity across time seems to be the rule for adult chimpanzee social behavior, not drastic changes.

Conclusion

We conclude that the opportunistic sampling method is a viable method for the study of primate communication when the goal is to obtain a large and diverse sample of signaling observations. There was general agreement between opportunistic and focal individual sampling on the order of signal types recorded (Tables 2 and 4). Signals that were commonly seen in one method were commonly seen in the other, and the same goes for uncommon signal types. This consistency between the methods supports the validity of opportunistic sampling, which is helpful because opportunistic sampling has some distinct benefits. Opportunistic sampling resulted in a larger number of signals, including signals per hour, making it a more efficient use of time. Opportunistic sampling also documented some signal types that were not observed in focal individual sampling (Tables 2 and 4). Lastly, opportunistic sampling had a more even distribution of signals/signalers across categories. Nonetheless, focal individual sampling offers the advantage of allowing for rates of signaling to be easily calculated. Thus, we recommend that

researchers employ a mixed-methods design in which they randomize bouts of focal and opportunistic sampling. Combining the two methods would offer the benefits of large signal sample sizes and types from opportunistic sampling with the benefit of rate calculations from focal individual sampling. Hence, the mixed-methods approach would provide the best of both worlds, while controlling for the limitations of each, potentially offering a fuller picture of facial and gestural signaling.

Statement of Ethics

Data collection was approved by the Los Angeles Zoo and Botanical Gardens and was in accordance with the Association for the Study of Animal Behaviour's guidelines for the treatment of animals in behavioral research [2020].

Conflict of Interest Statement

The authors of this publication have no conflicts of interest.

Author Contributions

B.N.F. contributed to the study design, data collection, data analysis, and preparation of the manuscript. M.W.C. contributed to the study design, data analysis, and preparation of the manuscript.

11

References

Akoglu H (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine* 18: 91–93.

Altmann J (1974). Observational study of behavior: sampling methods. *Behaviour* 49: 227–267.

Bergman TJ, Sheehan MJ (2013). Social knowledge and signals in primates. *American Journal of Primatology* 75: 683–694.

Bray J, Gilby IC (2020). Social relationships among adult male chimpanzees (*Pan troglodytes schweinfurthii*): variation in the strength and quality of social bonds. *Behavioral Ecology and Sociobiology* 74: 1–19.

Byrne RW, Cochet H (2017). Where have all the (ape) gestures gone? *Psychonomic Bulletin & Review* 24: 68–71.

Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition* 20: 755–769.

Candiotti A, Zuberbuhler K, Lemasson A (2012). Convergence and divergence in Diana monkey vocalizations. *Animal Behaviour* 8: 382–385.

Cartmill EA, Byrne RW (2010). Semantics of primate gestures: intentional meanings of orangutan gestures. *Animal Cognition* 13: 793–804.

Florkiewicz B, Skollar G, Reichard UH (2018). Facial expressions and pair bonds in hylobatids. *American Journal of Physical Anthropology* 167: 108–123.

Genty E, Breuer T, Hobaiter C, Byrne RW (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Animal Cognition* 12: 527–546.

Guidelines for the treatment of animals in behavioural research and teaching (2020). *Animal Behaviour* 83: 301–309.

Graham KE, Furuichi T, Byrne RW (2016). The gestural repertoire of the wild bonobo (*Pan paniscus*): a mutually understood communication system. *Animal Cognition* 20: 171–177.

Graham KE, Hobaiter C, Ounsley J, Furuichi T, Byrne RW (2018). Bonobo and chimpanzee gesture overlap extensively in meaning. *PLoS Biology* 16: 1–18.

Hobaiter C (2011). *Gestural Communication in Wild Chimpanzees*. PhD dissertation, University of St. Andrews. DOI: http://hdl.handle.net/10023/2143.

Hobaiter C, Byrne RW (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition* 14: 745–767.

Hobaiter C, Byrne RW (2014). The meanings of chimpanzee gestures. *Current Biology* 24: 1596–1600.

Hobaiter C, Byrne RW, Zuberbühler K (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology* 71(96): 1–13.

Hopkins WD, Taglialatela JP, Leavens DA (2011). Do chimpanzees have voluntary control of their facial expressions and vocalizations? In *Primate Vocalizations and Human Language: Vocalisation, Gestures, Imitation and Deixis in Humans and Non-Humans* (Vilain A, Schwartz JL, Arby C, Vauclair J, eds.), pp 71–88. Philadelphia: John Benjamins Pub. Co.

Laidre ME (2011). Meaningful gesture in monkeys? Investigating whether mandrills create social culture. *PLoS One* 6: e14610.

Liebal K, Call J, Tomasello M (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology* 64: 377–396.

Liebal K, Pika S, Tomasello M (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6: 1–38.

Maynard-Smith J, Harper D (1995). Animal signals: models and terminology. *Journal of Theoretical Biology* 177: 305–311.

McHugh ML (2012). Interrater reliability: the kappa statistic. *Biochemia Medica* 22: 276–282.

Moore R (2016). Meaning and ostension in great ape gestural communication. *Animal Cognition* 19: 223–231.

Parr LA, Waller BM, Vick SJ, Bard KA (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion* 7(1):172–181.

Parr LA, Waller BM, Heintz M (2010). Facial expression categorization by chimpanzees using standardized stimuli. *Emotion* 8: 216–231.

Pika S, Bugnyar T (2011). The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nature Communications* 2: 1–5.

Pollick AS, deWaal FBM (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104: 8184–8189.

Pusey AE, Schroepfer-Walker K (2013). Female competition in chimpanzees. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368: 20130077.

R Core Team (2017). R: a language and environment for statistical computing. https://www.R-project.org/ (accessed May 20, 2020).

Roberts AI, Roberts SGB (2015). Gestural communication and mating tactics in wild chimpanzees. *PLoS One* 10: e0139638.

Rose LM (2000). Behavioral sampling in the field: continuous focal versus focal interval sampling. *Behaviour* 137: 153–180.

Scheider L, Liebal K, Ona L, Burrows A, Waller B (2014). A comparison of the facial expression properties in five hylobatid species. *American Journal of Primatology* 76: 618–628.

Smith LW, Delgado RA (2015). Body language: the interplay between positional behavior and gestural signaling in the genus *Pan* and its implications for language evolution. *American Journal of Physical Anthropology* 157: 592–602.

Snedecor GW, Cochran WG (1989). *Statistical Methods*, 8th ed. Iowa City, Iowa State University Press.

Tkaczynski PJ, Mielke A, Samuni L, Preis A, Wittig RM, Crockford C (2020). Long term repeatability in social behaviour suggests stable social phenotypes in wild chimpanzees. *Royal Society Open Science* 7: 1–17.

Waller BM, Caeiro CC, Davila-Ross M (2015). Orangutans modify facial displays depending on recipient attention. *Peer J* 3: e827.

Webb CE, Romero T, Franks B, de Waal FBM (2017). Long-term consistency in chimpanzee consolation behaviour reflects empathetic personalities. *Nature Communications* 8: 1–8.

12

CHAPTER 2

Chimpanzee Facial Gestures and the Implications for the Evolution of Language

BRITTANY N. FLORKIEWICZ[1], MATTHEW W. CAMPBELL[2]

[1] *Department of Anthropology and Center for Behavior, Evolution, & Culture, University of California, Los Angeles, Los Angeles, California, U.S.A.*

[2]*Department of Psychology, California State University Channel Islands, Camarillo, California, U.S.A.*

Chimpanzee facial gestures and the implications for the evolution of language

Brittany Florkiewicz[1,*] and Matthew Campbell[2,*]

[1] Department of Anthropology, University of California, Los Angeles, Los Angeles, CA, United States of America
[2] Department of Psychology, California State University, Channel Islands, Camarillo, CA, United States of America
[*] These authors contributed equally to this work.

ABSTRACT

Great ape manual gestures are described as communicative, flexible, intentional, and goal-oriented. These gestures are thought to be an evolutionary pre-cursor to human language. Conversely, facial expressions are thought to be inflexible, automatic, and derived from emotion. However, great apes can make a wide range of movements with their faces, and they may possess the control needed to gesture with their faces as well as their hands. We examined whether chimpanzee facial expressions possess the four important gesture properties and how they compare to manual gestures. To do this, we quantified variables that have been previously described through largely qualitative means. Chimpanzee facial expressions met all four gesture criteria and performed remarkably similar to manual gestures. Facial gestures have implications for the evolution of language. If other mammals also show facial gestures, then the gestural origins of language may be much older than the human/great ape lineage.

INTRODUCTION

Great apes are known for their elaborate use of gestures (*Byrne et al., 2017*). Gestures are commonly defined as flexibly and intentionally produced bodily movements used during bouts of communication to achieve a goal (*Byrne et al., 2017*; *Moore, 2016*; *Byrne & Cochet, 2017*). Over 80 different gesture types have been identified across the great apes, which vary in their meaning and usage (*Byrne et al., 2017*). Most gesture types identified in behavioral ethograms describe movement of body or limbs, and we refer to them as 'manual gestures.' For example: over 90% of gesture types listed in the St. Andrews Catalogue of great ape gestures involve the hands, arms, legs, feet, and torso (*Byrne et al., 2017*), with particular focus on the arms and hands ($N = 43$ or 51% of gesture types). One distinguishing feature of great apes is a tendency towards upright posture, which is associated with increased suspensory behavior (such as brachiation, *Andrews, 2020*). Increased reliance on suspensory behavior resulted in anatomical changes associated with greater flexibility and mobility of the hands, wrists, and arms (*Andrews, 2020*), and these anatomical changes may

explain why great apes frequently gesture with these body parts. The transition to bipedal locomotion in the hominid lineage resulted in greater-still freedom of the hands and arms (*Corballis, 1999*). In humans (*Homo sapiens*), manual gestures play an important role in the production, comprehension, and learning of language (*Goldin-Meadow & Alibali, 2014*). However, humans also gesture with the face, with movements of the lips, chins, and eyebrows all documented as deictic gestures (*Enfield, 2002*).

The ability to gesture with the face could be beneficial for species reliant on quadrupedal locomotion. Monkeys, for example, rely almost exclusively on quadrupedalism, whether terrestrial or arboreal, which is associated with decreased flexibility of the hands (*Tamagawa et al., 2020*). Therefore, faces may be a better place to look for gestures in quadrupeds than the arms and hands, since the face is not associated with locomotor constraints. To begin the search for facial gestures in nonhumans, we chose to study chimpanzees (*Pan troglodytes*) for two reasons: first, their well-described manual gestures, and second, their phylogenetic position between humans and monkeys. Firstly, the well-documented manual gestures of chimpanzees (*Byrne et al., 2017*; *Hopkins & Leavens, 1998*; *Leavens & Hopkins, 1998*; *Liebal, Call & Tomasello, 2004*; *Leavens, Russell & Hopkins, 2005*; *Hobaiter, 2011*; *Roberts, Vick & Buchanan-Smith, 2012*; *Roberts, Vick & Buchanan-Smith, 2013*; *McCarthy, Jensvold & Fouts, 2013*; *Hobaiter & Byrne, 2014*; *Graham et al., 2018*; *Heesen et al., 2019*; *Roberts & Roberts, 2019*) provide a benchmark against which to compare the characteristics of facial expressions (Fig. 1). While there is evidence for the use of manual gestures outside of the great apes (*Laidre, 2011*; *Gupta & Sinha, 2019*; *Molesti, Meguerditchian & Bourjade, 2020*), it is limited in comparison to the extensive research available on great ape species. Thus, we would not have the same degree of performance metrics to apply to facial expressions. For these reasons, we feel that the best approach to begin the search for facial gestures is in a species with a documented repertoire of manual gestures.

Secondly, the great apes (*Pan troglodytes, Pan paniscus, Gorilla spp., Pongo spp.*) have unique anatomy that corresponds to their phylogenetic position between humans and the other mammals. The great apes exhibit variation in their locomotor behavior, including quadrupedal knuckle-walking (which is more upright than monkeys who walk on their palms), brachiation, and occasional bipedalism (*Kivell & Schmitt, 2009*). Great apes therefore occupy a transitional stage with a tendency towards more upright posture than monkeys, but less than humans. The result may be greater freedom by great apes to produce a wide variety of manual gestures with their hands (*Hobaiter & Byrne, 2014*; *Graham et al., 2018*; *Genty et al., 2009*), but the continued presence of terrestrial quadrupedalism and brachiation could still favor selection for gesturing with the face. Great apes make frequent use of facial expressions for communication (*Van Hooff, 1967*; *Parr & Waller, 2006*). This has also been observed in platyrrhine monkeys, such as capuchins (*Sapajus apella, Weigel, 1979*; *Visalberghi, Valenzano & Preuschoft, 2006*; *De Marco & Visalberghi, 2007*), but it is currently unclear as to whether the term 'gesture' can also be applied to monkey and great ape facial expressions.

Similar to great apes, research on gestural communication in monkeys focuses predominantly on the hands, arms, legs, and torso (*Laidre, 2011*; *Gupta & Sinha, 2019*; *Molesti, Meguerditchian & Bourjade, 2020*). For example: out of the 67 movements labeled

15

Figure 1 Photograph. Great ape manual gestures are often defined as flexibly and intentionally produced bodily movements during bouts of communication to achieve a goal. It is possible that chimpanzee facial expressions are also capable of being used as gestures. In this photo, an adult male (right) beckons to an adult female (left) using a lower lip relaxer face and a reach gesture. This suggests that both signals are being used to achieve a goal. This goal is eventually met when the female approaches the male for affiliative contact. By comparing facial expressions with the described gesture repertoires of chimpanzees, we sought evidence of whether facial expressions like this might also be gestural. Photo credit: Brittany Florkiewicz.

Full-size 🖼 DOI: 10.7717/peerj.12237/fig-1

as gestures for olive baboons (*Papio anubis*), only nine (or 13.43%) are associated with muscle movement of the face region (such as the eyebrows, eyes, nose, and mouth; *Molesti, Meguerditchian & Bourjade, 2020*). While the literature on great ape gestural communication is more extensive, there is disagreement regarding whether facial expressions can be used as gestures. In some great ape gestural ethograms, such as those used to study orangutan communication, facial expression types such as air bite, bite, duck lips, formal bite, and play face are classified as gestures (*Liebal & Pika, 2006*; *Cartmill & Byrne, 2010*). But it appears that other facial expression types, such as grin, open mouth, relaxed open mouth, pout face, and protruded lips are classified separately as expressions, not gestures (*Liebal & Pika, 2006*). In addition, some facial expression types that were initially classified as gestures (such as the play face, *Cartmill & Byrne, 2010*) were not incorporated into recently constructed gestural ethograms (*Byrne et al., 2017*).

Research on whether facial expressions are capable of being used as gestures would provide important insight into the evolution of human language. While human language consists of many unique properties (such as being able to refer to the past, present, and future, *Fitch, 2010*), it also has shared features with non-human primate communication, such as flexibility and syntax (*Fitch, 2010*; *Hewes et al., 1973*; *Plooj, 1978*; *Arnold & Zuberbühler, 2006*). This has led researchers to suggest that the evolutionary precursor to human language can be found in non-human primate communication (*Fitch, 2010*;

16

Arnold & Zuberbühler, 2006; *Armstrong & Wilcox, 2007*; *Dediu & Levison, 2013*; *Seyfarth & Cheney, 2018*). Some researchers propose that gestural communication of great apes is the most likely candidate, given the similarities between great ape gestures and human language (*Fitch, 2010*; *Corballis, 2002*). According to *Corballis (2002)*, the evolution of human language was a multi-step process: great ape manual gestures led to more symbolic forms of manual communication in the hominid lineage, which then led to the development of spoken language. This multi-step process has often been referred to as the "hand to mouth" theory of language evolution (*Corballis, 2002*).

Those who promote a gestural origin of language sometimes state that other forms of communication, such as vocalizations and facial expressions, cannot be potential precursors due to their strong association with emotion, making them inflexible and spontaneous (rather than flexible and intentional, *Hewes et al., 1973*; *Corballis, 2002*; *Pollick & De Waal, 2007*). The ability to communicate flexibly and intentionally with vocalizations and facial expressions is thought to have appeared later in hominid evolution, closer to the emergence of *Homo sapiens* (*Corballis, 2002*). If facial expressions are capable of being used as gestures, it suggests that both facial gestures and manual gestures were important communication systems in the evolution of human language. This idea has often been referred to as the "multimodal theory" of language evolution, which states that all modes of communication (facial, vocal, and gestural) were important pre-requisites for the establishment of language (*Fröhlich et al., 2019*) and co-evolved closely together (*McNeill, 2012*).

If facial expressions are used as gestures, this would also suggest that the evolutionary precursor to human language could be much older than previously assumed. Those who promote a gestural origin of language argue that the "evolutionary starting point" of human language can be traced to great ape gestural communication (*Tomasello & Call, 2019*). Great ape facial gestures could be consistent with this view, as they could have arisen along with manual gestures within great apes. But, the presence of facial gestures would also raise the possibility that this "starting point" might be traced back further, potentially even to the last common ancestor of mammals. Some facial expression types are found across a wide variety of mammals and are both morphologically and functionally similar across species (*Darwin, 1872*). One example of this includes grins, which are produced by canids (*Canis lupus, Canis domesticus*), cats (*Felis catus*), opossums (*Didelphis virginiana*), tree shrews (family Soricide), lemurs (*Lemur spp., Haplemur spp., Propithecus spp.*), platyrrhine monkeys (*Callitrhix geofroyi, Sapajus spp., Lagothrix spp., Tamarinus illigeri*), haplorrhine monkeys (*Macaca mulatta, Madrillus spinx*), and chimpanzees (*Pan troglodytes*) in response to startling and potentially dangerous stimuli (*Darwin, 1872*; *Andrew, 1963*). If grins and other facial expression types shared among mammals are capable of being produced as gestures in primates, it is possible that gestural communication can also be found in other mammals.

The history of both facial and gestural signaling research inspired our main goal, which was to examine whether chimpanzees use facial expressions as gestures. We hypothesized that if facial expressions can be gestural, then facial expressions should perform similarly to manual gestures on measures of gesture properties and variables (Table 1). To test

17

Table 1 A list of the four key gesture properties (communicative, intentional, flexible, and goal associated), their corresponding variables, and how they were measured. For binary variables, if the criteria outlined in the operational definition was met, the variable was coded as being present (i.e, 1); if not, the variable was coded as being absent (i.e., 0). A list of relevant citations (i.e., publications where each property is identified and described) is provided in the last column.

Gesture properties	Measured variables	Variable type	Statistical test	Operational definition	Relevant citations
Communicative	Mechanical In-effectiveness	Binary	Binomial Test; Composite gesture score; GLMM	The movement(s) are motorically ineffective; they are not used to complete gross motor behaviors, such as walking, climbing, chewing, etc.	*Cartmill (2008)*; *De Marco & Visalberghi (2007)*; *Fröhlich & Hobaiter (2018)*
Communicative	Recipient ID	Binary	Binomial Test; Composite gesture score; GLMM	The signaler directs their behavior towards a conspecific, which suggests that the behavior is 'socially directed'.	*Tomasello & Call (2019)*; *Fröhlich & Hobaiter (2018)*; *Pika & Fröhlich (2019)*
Intentional	Response Waiting [Overall]	Binary	Binomial Test; Composite gesture score; GLMM	One or both forms of response waiting (response waiting while persisting and/or response waiting at the end of the signal) are observed. See below for definitions/criteria. This suggests that the signaler is deliberately communicating with the recipient.	*Byrne et al. (2017)*; *Graham et al. (2018)*; *Roberts & Roberts (2019)*; *Hobaiter & Byrne (2011)*
Intentional	Response Waiting [at End of Signal]	Binary	Binomial Test; Composite gesture score; GLMM	At the end of a signal, the signaler fixates their gaze at the recipient, waiting for a behavioral response. This suggests that the signaler is deliberately communicating with the recipient.	*Byrne et al. (2017)*; *Graham et al. (2018)*; *Roberts & Roberts (2019)*; *Hobaiter & Byrne (2011)*
Intentional	Response Waiting [while Persisting]	Binary	Binomial Test; Composite gesture score; GLMM	As the signaler is persisting with a signal, they fixate their gaze at the recipient, waiting for a behavioral response. This suggests that the signaler is deliberately communicating with the recipient.	New measure introduced in this study, since it is also plausible for a signaler to wait for a response as they persist with the signal.

18

Table 1 (*continued*)

Gesture properties	Measured variables	Variable type	Statistical test	Operational definition	Relevant citations
Intentional	Receiver Attention	Binary	Binomial Test; Composite gesture score; GLMM	The signaler produces the signal while the recipient(s) are looking at and have their body positioned towards the signaler. This suggests that the signaling behavior is sensitive to audience effects and that the signaler is deliberately attempting to communicate with the recipient.	*Byrne et al. (2017); Graham et al. (2018); Tomasello & Call (2019)*
Flexible	Elaboration	Binary	Binomial Test; Composite gesture score; GLMM	After the signal is produced, the signaler modifies the physical form of that signal or switches to a new signal type. Elaboration is often used as a way to 'repair' potential communicative failure that may have occurred during the initial production of the signal.	*Byrne et al. (2017); Leavens, Russell & Hopkins (2005); Roberts, Vick & Buchanan-Smith (2012)*
Flexible	Persistence	Binary	Binomial Test; Composite gesture score; GLMM	The signaler repeats and/or holds the signal that they are producing. Persistence is often used as a way to 'repair' potential communicative failure that may have occurred during the initial production of the signal.	*Byrne et al. (2017); Leavens, Russell & Hopkins (2005); Roberts, Vick & Buchanan-Smith (2012)*
Flexible	Generalized Behavioral Context	Categorical	context tie index score, Mann–Whitney U	Each signal observed was assigned to 1 of 10 possible behavioral contexts which best described the social interaction, which included: affiliation, agonism, arousal (general), feeding, grooming, locomotion, playing, resting, sex, or unsure/unknown.	*Pollick & De Waal (2007)*
Goal Associated	Immediate Interaction Outcome	Binary	Binomial Test; Composite gesture score; GLMM	After the signal is produced, there is an immediate behavioral change in the recipient(s). This behavioral response may/may not satisfy the presumed goal of the signaler.	*Byrne et al. (2017); Hobaiter & Byrne (2014); Cartmill & Byrne (2010)*
Goal Associated	Final Interaction Outcome	Binary	Binomial Test; Composite gesture score; GLMM	After the signaler ceases all communication, there is a clear behavioral response from the recipient(s). This behavioral response may/may not satisfy the presumed goal of the signaler.	*Byrne et al. (2017); Hobaiter & Byrne (2014); Cartmill & Byrne (2010)*
Goal Associated	Presumed Goal	Binary	Binomial Test; Composite gesture score; GLMM	The signaler has a clear and intended behavioral response that they wish to elicit from the recipient(s) which can be hypothesized by the researcher. This behavioral response may be produced immediately following the behavior of the signaler (immediate interaction outcome) or after the signaler ceases all communication (final interaction outcome). If the signaler ceases their communication, it is assumed that this hypothesized goal was achieved (or cannot be achieved).	*Byrne et al. (2017); Hobaiter & Byrne (2014); Cartmill & Byrne (2010); Halina, Liebal & Tomasello (2018)*

this hypothesis, we compared facial expressions to manual gestures using: (1) the average number of gesture variables exhibited for each signaling observation and signal type (which we refer to as composite gesture scores, or CGS); (2) context tie indices (CTI scores), which were initially designed by Pollick and de Waal to test contextual flexibility (*Pollick & De Waal, 2007*); and (3) the prevalence of each gesture property and corresponding variable using binomial tests and generalized linear mixed models.

We experienced three major obstacles in devising our study. First, much of the gesture research is qualitative (*Leavens & Hopkins, 1998*; *Hobaiter & Byrne, 2014*; *Genty et al., 2009*; *Liebal & Pika, 2006*), making comparisons difficult and prone to bias. We therefore devised a novel metric (CGS) to quantify the presence and absence of important gesture properties (and their corresponding variables) for all observed facial expressions and manual gestures. There are four important, widely accepted properties of gestures in the literature: gestures are communicative, intentional, flexible, and goal oriented (*Byrne et al., 2017*; *Moore, 2016*; *Byrne & Cochet, 2017*). Each property was operationalized with two to four measures (Table 1). For example, the gesture property of 'communicative' was operationalized by two measures, mechanical ineffectiveness and direction at a specific receiver (Table 1). For each recorded manual gesture and facial expression we scored presence/absence for each measure. We then took the percentage of all manual gestures and facial expressions that showed each measure and compared them to evaluate performance. Hence, to evaluate being communicative we generated a composite of two measures, thus our term composite gesture score (CGS).

The second obstacle we faced was that there may be variation in the prevalence of gesture properties (and corresponding variables) due to contextual factors. For example, *Leavens, Russell & Hopkins (2005)* found that chimpanzee manual gestures vary in the extent to which they exhibit persistence and elaboration. This variation appears to be attributed to factors such as food quality and availability. Chimpanzees were much more likely to exhibit persistence and elaboration in their gesturing behavior when presented with lower quality food items (such as chow) as opposed to higher quality and highly desired food items (such as bananas) (*Leavens, Russell & Hopkins, 2005*). Because some gesture properties vary in prevalence, it is possible that a given signal type may not exhibit all four gesture properties (and their corresponding variables) simultaneously in each instance of use. To account for this possibility, we compared the variation in gesture properties exhibited by manual gestures to that of facial expressions.

Finally, previous studies have focused on only one or two important gesture properties (such as flexibility or goal-association only, *Leavens, Russell & Hopkins, 2005*; *Hobaiter & Byrne, 2014*) and their corresponding variables. Variables used to measure each gesture property also vary extensively between studies and sometimes overlap with other important gesture properties. For example: studies examining intentionality in gestural communication vary in the number of variables examined, which ranges from one to seven (*Graham et al., 2019*). Some of these variables (such as receiver attention state, persistence, and satisfaction with a goal) are also variables used to measure other important gesture properties (such as flexibility and goal association, *Byrne et al., 2017*; *Leavens, Russell & Hopkins, 2005*; *Roberts, Vick & Buchanan-Smith, 2012*; *Hobaiter & Byrne, 2014*;

20

Cartmill & Byrne, 2010). Therefore, we coded and compared all previously described gesture properties ($N = 4$) and their corresponding variables ($N = 12$) using our novel scoring metric (composite gesture scores), rather than focusing on a select few.

The results of our study will provide insight into the communicative properties associated with facial expressions and their implications for the evolution of human language.

MATERIAL AND METHODS

Subjects and data collection

We studied 18 chimpanzees (*Pan troglodytes*) at the Los Angeles Zoo and Botanical Gardens, which consisted of 13 adults (>seven years old) and five infants (≤seven years old). Out of these 18 chimpanzees, 15 were born and reared at the Los Angeles Zoo. Additional information on the names, birth dates, birth places, genetic relatedness, and husbandry of the chimpanzees can be found in the online supplement (Table S1 and Figs. S1–S2).

We selected chimpanzees for the current study because they frequently produce a wide variety of facial signal (*Parr et al., 2007*) and gesture types (*Hobaiter, 2011*; *Hobaiter & Byrne, 2014*). We collected data during the summer months (June to August) from 2017 to 2019, Monday to Friday, between 8:00 and 14:00, which we identified as peak activity hours in a 2016 pilot study. We recorded the chimpanzees with a Panasonic Full HD Video Camera Camcorder HC-V770 with an external shotgun microphone (Sennheiser MKE400) to improve audio quality. We used two methods when recording the chimpanzees: the opportunistic sampling method and the focal sampling method (*Altmann, 1974*; *Florkiewicz & Campbell, 2021*).

Data coding

We defined signals as actions performed by a signaler that attempted to alter the behavior of others (*Maynard-Smith & Harper, 1995*) and recorded the signaler ID. In this study, we defined a facial signal as facial muscle movement used for the purpose of communication. Facial muscle movements used for biological maintenance (such as blinking, feeding, and yawning) and object manipulation (such as chewing or scraping on objects) were not considered in this study. We placed facial expressions into behavioral categories (*i.e.*, facial expression types) based on similarities in key muscle movements (using chimpFACS, *Parr et al., 2007*). We considered a total of nine facial expression types in this study, six of which were derived from Parr et al. *Parr et al. (2007)*. We added three additional facial expression types (lipsmacking face, lower lip relaxer face, and raspberry face) to this ethogram from our 2016 pilot study. We did not consider neutral faces in the current study since it is difficult to evaluate whether they are capable of being used in a communicative manner. We excluded whimper faces due to a small number of observations ($N = 1$). Our final ethogram of facial expressions and manual gestures can be found in Table 2.

We focused primarily on gestures that have already been described in the literature (*Byrne et al., 2017*; *Hobaiter, 2011*). In these studies, gestures are defined as bodily movements produced during bouts of communication (*Hobaiter & Byrne, 2011*). Hereafter, we will refer to these signals as 'manual gestures.' Using these definitions and ethograms, we placed manual gestures into behavioral categories (*i.e.*, gesture types) based on similarity

21

Table 2 **A list of all facial expression types and manual gesture types that were considered.** Categories for facial expressions are based on a facial expression ethogram initially developed by *Parr et al. (2007)*. Additional categories were also created based on the unique signaling properties of this group (which were identified in a 2016 pilot study), such as Lipsmacking Face, Lower Lip Relaxer Face, and Raspberry Face. Categories for manual gestures are based on a gesture ethogram initially developed by *Hobaiter (2011)* and *Byrne et al. (2017)*. We removed manual gesture types which were seldom observed in the current study ($N < 10$). We also modified manual gesture types based on the results of inter-observer reliability.

Facial Expression Ethogram	
Type	**Description**
Ambiguous Face	Communicative movement of the face is observed but does not physically resemble the facial expression types described below.
Bared Teeth Face	Corners of the lips are drawn backwards, exposing both rows of teeth.
Lipsmacking Face	Mouth is opened and closed rapidly, producing a low sound.
Lower Lip Relaxer Face	Bottom lip is relaxed away from the gums and bottom row of teeth.
Pant-Hoot Face	Lips are funneled together, with vocalization being produced.
Play Face	Mouth is opened, lips possibly drawn backwards, and bottom row of teeth possibly exposed. Sometimes accompanied with a vocalization.
Pout Face	Lips are funneled close together and pushed outwards while the chin is furrowed upward.
Raspberry Face	Lips are pressed tightly together (sometimes with the tongue) to create a loud sound.
Scream Face	Similar to bared teeth face, but jaw is stretched, and a loud vocalization is produced.
Manual Gesture Ethogram	
Type	**Description**
Ambiguous Touch	Light contact of the fingertips, fingers, palm, and/or hand onto the recipient's body.
Arm Raise	One or both arms are raised vertically above the shoulder.
Bite	A part of the recipient's body is held between (or against) the lips or teeth of the signaler.
Clap	Hands are brought together to create a loud, audible sound.
Dangle	Signaler hangs from the hands above a recipient while shaking their feet.
Directed Push	The palm is placed on the recipient's body, with force being exerted to try and move the recipient towards a specific location.
Embrace	Both arms are wrapped around a recipient, with physical contact being made and maintained.
Hand Fling	Signaler makes rapid movements of the hands or arms towards the recipient.

22

Table 2 (*continued*)

Jump	Both feet leave the ground simultaneously with the signaler's body being displaced afterwards.
Kick	One or both feet are brought into short contact with the recipient's body.
Look	Signaler moves closely towards the recipient and holds eye contact.
Object Move	An object is displaced by a signaler, with contact being maintained throughout the movement.
Piourette	Signaler twirls around the body's vertical axis.
Present Body Part	A body part is deliberately moved and exposed to a recipient.
Present Climb on Me	The back is deliberately moved and exposed to a recipient. The signaler maintains a quadrupedal stance throughout the movement.
Present Sexual	Genitals are deliberately moved and exposed to a recipient.
Punch Other	Signaler closes their first and makes contact with the recipient's body.
Push	The palm is placed on the recipient's body, with force being exerted to try and move the recipient towards an unspecific location.
Reach	One or both arms are extended towards the recipient, with the palm being oriented upward.
Roll Over	Signaler rolls onto their back, which exposes their stomach to a recipient.
Slap Other	Either the palm of the hand or an object is brought into contact with the recipient's body.
Slap Other with Object	An object is brought into hard contact with the recipient's body.
Stomp	The sole of one or both feet are lifted vertically and brought into quick contact with the ground.
Swing/Rock	Large back and forth movement of the body (and occasionally arms) while being seated, standing quadrupedally, or standing bipedally.
Tandem Walk	Signaler positions their arm over the body of the recipient, with both the signaler and recipient walking forward while remaining side by side.
Tap/Slap/Knock Object	Quick movement(s) of the hand and/or arm are directed towards an object.
Throw Object	Object is thrown into the air and is displaced from the initial starting point.

in movement. Initially, we included 35 manual gesture types in this study as derived *Hobaiter (2011)* and *Byrne et al. (2017)*. The number of gesture types decreased to 27 after conducting inter-observer reliability. Other manual gesture types mentioned in *Hobaiter (2011)* and *Byrne et al. (2017)* were not incorporated into the current study since they were seldom observed (where each type had less than ten observations). Removing manual gestures that are seldom observed helps to promote a more conservative interpretation of our results.

23

We evaluated all facial expressions and manual gestures using the four main properties that are common throughout the gesture literature. Gestures are typically defined as bodily movements which are communicative, intentional, flexible, and associated with a goal (*Byrne et al., 2017*; *Moore, 2016*; *Byrne & Cochet, 2017*). We measured 12 variables to quantify these four gesture properties. Of these, we coded 11 as binary variables, which included: mechanical ineffectiveness, receiver attention, recipient ID, response waiting at the end of the signal, response waiting while persisting with the signal, total response wait time (or response waiting overall), persistence, elaboration, immediate interaction outcomes, final interaction outcomes, and presumed goals (see Table 1 for details). For our 12th variable, we coded the generalized behavioral context that each signal occurred in using Pollick and de Waal's behavioral categories (*Pollick & De Waal, 2007*).

We coded video footage in ELAN 5.6-AVFX (https://archive.mpi.nl/tla/elan) using a custom annotation template, which contained all of the variables mentioned above. A copy of this template can be found in the online supplement (Template S1).

Inter-observer reliability

We conducted inter-observer reliability on 10% of the video clips from 2018, which contained both facial expressions and manual gestures. This video footage contained 149 facial expressions (or 13.67% of all facial expressions) and 270 manual gestures (or 11.46% of all manual gestures). We calculated percentage of agreement for all variables mentioned above. In previous chimpanzee gesture studies, percentages of agreement at or above 70% were classified as good agreement (*Hobaiter, 2011*). We calculated Cohen's Kappa for manual gesture types, which is a common practice in chimpanzee gesture studies (*Hobaiter, 2011*; *Cartmill, 2008*). A Cohen's Kappa of 0.61 or higher is typically considered substantial agreement (*McHugh, 2012*). The average percentage of agreement across all variables was 79.61%. Facial expression types and manual gesture variables had a good level of agreement (with percentages being above 70%). Initially, manual gesture types had a lower agreement (48.52%). This is because manual gesture types that were morphologically similar to one another (such as touch other, hand on, and grab) were difficult to distinguish from one another. This resulted in lower agreement when separated. Therefore, we condensed the 35 manual gesture types into 27 categories (Table 2) based on morphological similarities, and agreement increased to 65.56%. After condensing gesture types, Cohen's Kappa was close to substantial agreement (0.604). Additional details can be found in the online supplement (Table S2).

Data analysis

We exported data from ELAN into R 3.6.2. (*R Core Team, 2020*). For each signal observed, we created a composite gesture score (CGS) by adding the number of gesture variables observed. To create these scores, we used the 11 gesture variables which were coded as being present or absent. The final gesture variable (generalized behavioral contexts) was not included here, since this variable was categorical. The total number of gesture variables which were coded as being present were used for each individual signal's score. For example: a signal with a CGS of 8 means that this signal exhibited eight out of the

24

11 gesture variables considered in this study (see Table 1 for a list of possible gesture variables). We then averaged composite gesture scores across facial expression and manual gesture types. We compared the average CGS for each modality using a Mann–Whitney U test (base functions in R 3.6.2). Next, we used binomial tests to determine whether the proportion of observed binary gesture variables for each modality were the result of chance. Chance occurrence would mean that the observers were as likely to score the variable as absent as they were present, statistically speaking. Thus, a binomial score within the realm of chance would indicate that the gesture variable was not reliably associated with the signal, as on any given observation the variable might be present or might not. By requiring a rate significantly above chance we ensured that we only interpreted gesture variables as representative features of a signal if the variable was consistently scored as present. Chance level was set to 50%, and we evaluated scores at the 95% confidence level.

It is possible for two signal types to have the same average CGS but differ in the types of variables observed. Therefore, we used binomial Generalized Linear Mixed Models (GLMMs) to examine the association between binary gesture variables and each mode of communication. We used GLMMs for two main reasons. First, GLMMs help to account for the pooling fallacy (Waller et al., 2013). The communicative signals of up to 18 individuals are incorporated in this study over the span of three years. This inevitably leads to having more signaling samples than individuals. GLMMs are beneficial for addressing the asymmetrical relationship between number of signals and individuals (Waller et al., 2013). Second, GLMMs help to account for idiosyncratic differences in signal use (Waller et al., 2013). For example, some individuals may use a signal as a gesture while others do not. GLMMs can be used to account for individual differences through the use of random effects (Waller et al., 2013).

We implemented Binomial GLMMs with a logit link function using the "lme4" package in R 3.6.2 (R Core Team, 2020; Bates et al., 2015). We compared each model to a null counterpart model where only signaler ID (not signal type) was included as the outcome variable (as outlined in R Core Team, 2020). This is to determine whether signal modality has a significant influence on each gesture variable. We made these comparisons using a likelihood-ratio test using the ANOVA function in base R (Waller et al., 2013). GLMMs were implemented for the following ten variables: receiver attention, recipient ID, response waiting at the end of the signal, response waiting while persisting with the signal, total response wait time (or response waiting overall), persistence, elaboration, immediate interaction outcomes, final interaction outcomes, and presumed goals. Mechanical ineffectiveness was not considered in GLMMs since almost all facial expressions and manual gestures exhibited this variable. In each model, gesture variables were set as the outcome variable, with signal modality and signaler ID set as explanatory variables. Signaler ID was set as a random variable.

We present odds ratios (OR) comparing the strength of association between facial expressions and gestures with the results of each GLMM. Finally, we created context tie index (CTI) scores for each facial expression and manual gesture type for the generalized behavioral contexts. This metric was initially established by Pollick & De Waal (2007) to examine the strength of association between behavioral contexts and signal types (i.e.,

25

flexibility in the meaning of signals). For each signal type (Table 2), we calculated the proportion of observed behavioral contexts. For example, the 'directed push' gesture was observed in three distinct behavioral contexts: locomotion ($N = 9$ or 75.00%), play ($N = 2$ or 16.67%), and sex ($N = 1$ or 8.33%). We then selected the largest proportion (and corresponding behavioral context) for that signal type's CTI score (*Pollick & De Waal, 2007*). For the 'directed push' gesture, the composite gesture score would be 0.75, since locomotion was the most frequently observed behavioral context. Signal types with larger CTI scores (*i.e.*, close to 100% or 1.00) have a strong association with a single, specific behavioral context, and signal types with low CTI scores (*i.e.*, close to 0% or 0.00) are associated with multiple contexts. Therefore, *Pollick & De Waal (2007)* equated high CTI scores with less flexible usage and low CTI scores with more flexible usage. CTI scores for each signal type can be found in Table 3. We compared CTI scores for facial expressions and manual gestures using a Mann–Whitney U test (using base functions in R 3.6.2).

Datasets (Data S1) and code (R Script S1) used to conduct these analyses can be found in the online supplement.

RESULTS

We observed 3,446 signals across 156.5 h of video footage, of which 1,090 were facial expressions and 2,356 were manual gestures. Both facial expressions and manual gestures varied in the extent to which they exhibited the four gesture properties (and the corresponding 11 gesture variables, Table 4). The median composite gesture score (CGS) for manual gestures (median = 8.48, range = 7.31–9.63) was not significantly different from that of facial expressions (median = 8.13, range = 6.08–9.27, Mann–Whitney $U = 90$, $p = 0.257$, Table 3).

Gestures are defined as being communicative, intentional, flexible, and goal associated (*Byrne et al., 2017*; *Moore, 2016*; *Byrne & Cochet, 2017*). Our primary goal was to evaluate whether facial expressions also possess these traits, and therefore should be considered gestures. To do this, we present the results for each criterion separately, showing how facial expressions performed and how they compared to manual gestures. Proportions and binomial test results can be found in Table 4. GLMM results can be found in Tables 5 and 6. For detailed information on how gesture properties (and their corresponding variables) differed across signal types, see the online supplement (Data S1).

Communicative

We measured communicativeness by whether the movement (facial or other) was mechanically ineffective and directed toward one, clear recipient (Table 1). All facial expressions were mechanically ineffective (100.00%, Binomial Test, $p < 0.05$), and most were produced towards a recipient (75.87%, Binomial Test, $p < 0.05$). We also observed this pattern with manual gestures; 98.94% were mechanically ineffective (Binomial Test, $p < 0.05$) and 95.97% were produced towards one, clear recipient (Binomial Test, $p < 0.05$). Modality was a significant predictor of recipient ID (Likelihood Ratio Test, $X^2(1) = 209.43$, $p < 0.001$), with facial expressions significantly less likely to exhibit this property than manual gestures (GLMM, $p < 0.001$, OR = 0.160).

26

Table 3 Context tie index (CTI) scores for all facial expression types and manual gesture types considered in this study. CTI scores were assigned using the methods outlined in *Pollick & De Waal (2007)*. We selected the most frequently observed behavioral context (and its corresponding proportion) as each signal type's CTI score. Scores closer to 1.00 mean that the signal type has less contextual flexibility, whereas scores closer to 0.00 mean that the signal type has greater contextual flexibility. Signal types highlighted in gray were used to compare our results to those outlined in *Pollick & De Waal (2007)*. Neither the overall comparison between facial expressions and manual gestures nor the *Pollick & De Waal (2007)* subset were significantly different.

Modality	Signal type	CTI score
Facial Expression	Bared Teeth Face	0.25
Facial Expression	Pout Face	0.27
Facial Expression	Lower Lip Relaxer Face	0.31
Manual Gesture	Swing/Rock	0.32
Manual Gesture	Amb. Touch	0.34
Facial Expression	Amb. Face	0.36
Manual Gesture	Jump	0.36
Manual Gesture	Arm Raise	0.38
Manual Gesture	Reach	0.4
Manual Gesture	Bite	0.48
Facial Expression	Pant Hoot Face	0.49
Manual Gesture	Kick	0.5
Manual Gesture	Punch Other	0.5
Manual Gesture	Clap	0.52
Manual Gesture	Hand Fling	0.52
Facial Expression	Scream Face	0.52
Manual Gesture	Tandem Walk	0.55
Manual Gesture	Push	0.55
Manual Gesture	Tap/Slap/Knock Object	0.58
Manual Gesture	Present Body Part	0.61
Manual Gesture	Slap Other	0.65
Manual Gesture	Stomp	0.69
Facial Expression	Raspberry Face	0.72
Manual Gesture	Roll Over	0.74
Manual Gesture	Directed Push	0.75
Manual Gesture	Embrace	0.77
Manual Gesture	Present Sexual	0.79
Manual Gesture	Throw Object	0.84
Manual Gesture	Slap Other with Object	0.88
Manual Gesture	Object Move	0.89
Manual Gesture	Present Climb on Me	0.91
Manual Gesture	Dangle	0.93
Facial Expression	Play Face	0.93
Manual Gesture	Look	0.95
Facial Expression	Lipsmacking Face	0.99
Manual Gesture	Pirouette	1

27

Table 4 Binomial tests. The number and proportion of facial expressions and manual gestures that exhibited each gesture variable. We used binomial test results to determine whether gesture variables in facial expressions and manual gestures occurred significantly above or below chance level.

Gesture variables	Proportion of facial expressions (p <05)?	Binomial test significant for facial expressions (p <05)?	Proportion of manual gestures	Binomial test significant for manual gestures (p <05)?
Mechanical Ineffectiveness	100.00%	$p < 0.001$***	98.94%	$p < 0.001$***
Receiver Attention	49.08%	$p = 0.565$	58.96%	$p < 0.001$***
Recipient ID	75.87%	$p < 0.001$***	95.97%	$p < 0.001$***
Response Waiting (End)[a]	18.35%	$p < 0.001$***	25.21%	$p < 0.001$***
Response Waiting (Persisting)	62.20%	$p < 0.001$***	59.97%	$p < 0.001$***
Response Waiting (Overall)	64.40%	$p < 0.001$***	71.82%	$p < 0.001$***
Persistence	89.45%	$p < 0.001$***	70.76%	$p < 0.001$***
Elaboration	93.94%	$p < 0.001$***	87.14%	$p < 0.001$***
Immediate Interaction Outcome	71.38%	$p < 0.001$***	80.73%	$p < 0.001$***
Final Interaction Outcome	88.17%	$p < 0.001$***	89.43%	$p < 0.001$***
Presumed Goal	70.83%	$p < 0.001$***	91.30%	$p < 0.001$***

Notes.
[a] This was the only gesture variable which occurred significantly below chance level. The remaining gesture variables that were significant occurred above chance level.

Table 5 Comparing GLMMs to null models. Comparing gesture variables GLMMs to their null counterparts using AIC scores.

Gesture variables	Null model	Regular model	P-Value
Receiver Attention	4691.2	4670.8	$p < 0.001$***
Recipient ID	2075.0	1867.6	$p < 0.001$***
Response Waiting (at End of Signal)	3688.2	3677.1	$p < 0.001$***
Response Waiting (while Persisting)	4465.5	4454.0	$p < 0.001$***
Response Waiting (Overall)	4002.4	4001.8	$p = 0.1098$
Persistence	3727.6	3579.2	$p < 0.001$***
Elaboration	2324.6	2277.0	$p < 0.001$***
Immediate Interaction Outcome	3610.0	3581.7	$p < 0.001$***
Final Interaction Outcome	2329.8	2331.8	$p = 0.7949$
Presumed Goal	2612.4	2492.8	$p < 0.001$***

Notes.
*** The p-value indicates whether signal modality is a significant predictor of each gesture variable.

Intentional

We measured intentionality by whether the movement was sensitive to receiver attention, exhibited response waiting at the end of the movement, exhibited response waiting while persisting in the movement, or exhibited response waiting overall (*i.e.*, the presence of any or both forms of response waiting; Table 1). Approximately half of all observed facial expressions were sensitive to receiver attention (49.08%), and this was not significantly different from chance (Binomial Test, $p = 0.565$). Facial expressions exhibited response waiting while persisting (62.20%, Binomial Test, $p < 0.05$) and response waiting overall (64.40%, Binomial Test, $p < 0.05$) significantly above chance. Only a small proportion

28

Table 6 GLMM results. The results of the binomial GLMMs examining gesture variables as the outcome variable with signal modality and signaler ID as explanatory variables (with signaler ID set as a random explanatory variable). Odds ratios (OR) are presented for each model (FE/GE). FE stands for facial expression, and GE stands for manual gesture.

Model	Predictor variable	Estimate	Std. Error	z value	p value	OR (FE/GE)
Receiver Attention	(Intercept)	−0.03989	0.08666	−0.46	$p = 0.645$	0.6936887
	Signal Type GE	0.36573	0.07716	4.74	$p < 0.001$***	
Recipient ID	(Intercept)	1.3308	0.2045	6.508	$p < 0.001$***	0.1599745
	Signal Type GE	1.8327	0.1336	13.717	$p < 0.001$***	
Response Waiting (at End of Signal)	(Intercept)	−1.51649	0.11351	−13.360	$p < 0.001$***	0.7109288
	Signal Type GE	0.34118	0.09549	3.573	$p < 0.001$***	
Response Waiting (while Persisting)	(Intercept)	0.54779	0.13885	3.945	$p < 0.001$***	1.350992
	Signal Type GE	−0.30084	0.08249	−3.647	$p < 0.001$***	
Response Waiting (Overall)	(Intercept)	0.67214	0.16025	4.194	$p < 0.001$***	0.8712955
	Signal Type GE	0.13777	0.08584	1.605	$p = 0.108$	
Persistence	(Intercept)	2.1205	0.1131	18.75	$p < 0.001$***	3.500197
	Signal Type GE	−1.2528	0.1110	−11.29	$p < 0.001$***	
Elaboration	(Intercept)	2.7702	0.1614	17.167	$p < 0.001$***	2.624528
	Signal Type GE	−0.9649	0.1467	−6.576	$p < 0.001$***	
Immediate Interaction Outcome	(Intercept)	0.9864	0.10922	9.031	$p < 0.001$***	0.6081248
	Signal Type GE	0.49738	0.08954	5.555	$p < 0.001$***	
Final Interaction Outcome	(Intercept)	2.1545	0.16075	13.403	$p < 0.001$***	0.9689051
	Signal Type GE	0.03159	0.12076	0.262	$p = 0.794$	
Presumed Goal	(Intercept)	1.0665	0.2272	4.693	$p < 0.001$***	0.3084879
	Signal Type GE	1.1761	0.1070	10.988	$p < 0.001$***	

of facial expressions exhibited response waiting at the end of the signal, and this was significantly below chance level (18.35%, Binomial Test, $p < 0.05$).

Manual gestures exhibited a similar pattern. Approximately half of all manual gestures were sensitive to receiver attention (58.96%), which was significantly above chance level (Binomial Test, $p < 0.05$). Manual gestures exhibited response waiting while persisting (59.97%, Binomial Test, $p < 0.05$) and response waiting overall (71.82%, Binomial Test, $p < 0.05$) significantly above chance. A small proportion of manual gestures exhibited response waiting at the end, and this was significantly below chance level (25.21%, Binomial Test, $p < 0.05$).

Modality was a significant predictor of receiver attention (Likelihood Ratio Test, $X^2(1) = 22.395$, $p < 0.001$), response waiting at the end (Likelihood Ratio Test, $X^2(1) = 13.041$, $p < 0.001$), and response waiting while persisting (Likelihood Ratio Test, $X^2(1) = 13.486$, $p < 0.001$). But modality did not predict response waiting overall (Likelihood Ratio Test, $X^2(1) = 2.557$, $p = 0.110$). Facial expressions were significantly less likely to exhibit sensitivity to receiver attention (GLMM, $p < 0.001$, OR = 0.694) and response waiting at the end of the signal (GLMM, $p < 0.001$, OR = 0.711) when compared to manual gestures; however, facial expressions were significantly more likely to exhibit response waiting while persisting (GLMM, $p < 0.001$, OR = 1.351).

29

Flexible

We measured flexibility by whether the movement was sensitive to persistence and elaboration (Table 1). Facial expressions exhibited persistence (89.45%, Binomial Test, $p < 0.05$) and elaboration (93.94%, Binomial Test, $p < 0.05$) significantly above chance, and manual gestures showed the same pattern (persistence = 70.76%, Binomial Test, $p < 0.05$; elaboration = 87.14%, Binomial Test, $p < 0.05$). Modality was a significant predictor of persistence (Likelihood Ratio Test, $X^2(1) = 150.38, p < 0.001$) and elaboration (Likelihood Ratio Test, $X^2(1) = 50.225, p < 0.001$), with facial expressions being significantly more likely to exhibit both properties when compared to manual gestures (GLMM, p<0.001: persistence OR = 3.500, elaboration OR = 2.625).

Facial expressions (median = 0.49, range = 0.25−0.99) exhibited lower average context tie indices (CTI scores) compared to manual gestures (median = 0.61, range = 0.32−1.00, see Table 3). These results suggest that facial expressions exhibit greater contextual flexibility compared to manual gestures. However, these differences were not significantly different (Mann–Whitney $U = 86, p = 0.201$). To compare our results to those published by *Pollick & De Waal (2007)*, we compared a subset of facial expressions ($N = 4$) and manual gestures ($N = 3$) observed in both studies (highlighted in gray in Table 3). CTI scores were very similar across select facial expressions (median = 0.40, range = 0.25−0.52) and manual gestures (median = 0.38, range = 0.34−0.40, Mann–Whitney $U = 6, p = 1.000$), suggesting similarities in contextual flexibility. In contrast, *Pollick & De Waal (2007)* report manual gestures as having significantly greater contextual flexibility when compared to facial expressions.

Goal associated

We measured goal association by whether the movement was sensitive to immediate interaction outcomes, final interaction outcomes, and having a clear presumed goal (Table 1). Facial expressions were associated with an immediate interaction outcome (71.38%, Binomial Test, $p < 0.05$), a final interaction outcome (88.17%, Binomial Test, $p < 0.05$), and a clear presumed goal (70.83%, Binomial Test, $p < 0.05$). We also observed this pattern with manual gestures: immediate interaction outcome (80.73%, Binomial Test, $p < 0.05$), final interaction outcome (89.43%, Binomial Test, $p < 0.05$), and a clear presumed goal (91.30%, Binomial Test, $p < 0.05$). Modality was a significant predictor of immediate interaction outcomes (Likelihood Ratio Test, $X^2(1) = 30.259, p < 0.001$) and presumed goals (Likelihood Ratio Test, $X^2(1) = 121.600, p < 0.001$) but not final interaction outcomes (Likelihood Ratio Test, $X^2(1) = 0.0675, p = 0.7949$). When compared to manual gestures, facial expressions were significantly more likely to be associated with an immediate interaction outcome (GLMM, $p < 0.001$, OR = 0.608) but significantly less likely to be associated with a presumed goal (GLMM, $p < 0.001$, OR = 0.308).

DISCUSSION

Our goal was to examine whether chimpanzee facial expressions meet the criteria for gestures (Table 1) and should be considered as such. In contrast to previous studies, we quantified all previously described gesture properties and corresponding variables, as

30

opposed to a select few or relying on qualitative evidence. Out of the 11 binary gesture variables examined, facial expressions exhibited nine significantly above chance level, and manual gestures exhibited ten (Table 4). The nine gesture variables shown by facial expressions correspond to all four of the main properties of gestural communication, which include: communicativeness, intentionality, flexibility, and goal association.

There was some disagreement between our results for intentionality and those of other studies. Response waiting while persisting and overall were both significantly associated with both manual gestures and facial expressions. However, response waiting at the end of the signal occurred significantly below chance level for both facial expressions and manual gestures. In the gesture literature, response waiting at the end of the signal is often used as a measure of intentionality (*Byrne et al., 2017*; *Graham et al., 2018*; *Roberts & Roberts, 2019*; *Hobaiter & Byrne, 2011*), but previous studies have not quantified this variable to the extent that we did. Thus, it appears that response waiting at the end of a signal is rare and may not be as helpful for the identification of gestures when compared to other forms of response waiting. However, it is important to note that there is extensive disagreement regarding the definition of intentionality (*Tomasello & Call, 2019*) and the validity of variables associated with intentionality (*Graham et al., 2019*). To address these concerns, we recommend examining and comparing multiple gesture properties and corresponding variables simultaneously, as we attempted, since variables associated with other gesture properties (such as flexibility and goal association) can be used to provide additional support for intentionality (*Tomasello & Call, 2019*).

We also examined how the gesture properties observed in facial expressions compared to manual gestures using a new scoring metric (composite gesture scores, or CGS) based on the number of gesture properties (and corresponding variables) exhibited for each signal. We found no significant differences in average CGS across facial expressions and manual gestures. Thus, facial expressions are just as likely to take on as many gesture variables as manual gestures. When directly comparing facial expressions to manual gestures, we found that facial expressions are more strongly predicted by three variables, whereas manual gestures are more strongly predicted by five variables (Table 6). Out of these eight variables, six occur significantly above chance level in both facial expressions and manual gestures. There were no significant differences between facial expressions and manual gestures for the remaining three variables. Overall, this pattern shows that facial expressions performed similarly to manual gestures.

When measuring flexibility using our categorical variable, we also observed no significant differences in context tie indices (CTI scores) across facial expressions and manual gestures (Table 3). This finding suggests that facial expressions are just as likely to demonstrate contextual flexibility as manual gestures. Interestingly, facial expressions exhibited more contextual flexibility on average when compared to manual gestures. Earlier studies have demonstrated the opposite pattern. For example: *Pollick & De Waal (2007)* found that gestures produced by chimpanzees and bonobos were significantly more likely to exhibit contextual flexibility when compared to facial expressions. When making a direct comparison on a subset of gesture types that were included in both studies, we found no significant differences in contextual flexibility between select facial expression and manual

31

gesture types. It is not clear how to reconcile our findings with those of Pollick and de Waal (*Pollick & De Waal, 2007*), but differences in sample sizes, signal types observed, signaling ethograms, or study populations could all have contributed to different results.

Since facial expressions met all four of the key criteria for gestures and performed similarly to manual gestures across these variables, we conclude that chimpanzee facial expressions can be gestural. Our results suggest that a revision to how facial expressions are understood is in order. The term 'facial expression' stems from the idea that the facial muscle movement is attributed to expressions of emotion (*Bell, 1806*; *Descartes,*; *Fridlund, 1994*; *Elliott & Jacobs, 2013*), which are both discrete and universal (*Ekman, 1970*; *Ekman, 1999*). As a result, facial expressions are often perceived as being spontaneous and inflexible (*Ekman, 1993*). Because gestures are defined as intentionally and flexibly produced communicative movements (*Byrne et al., 2017*), they are often described as being a separate mode of communication when compared to facial expressions (*Liebal et al., 2014*). Our results suggest that facial expressions can be used as facial gestures.

The reverse may also be true, that manual gestures can be spontaneous and inflexible. For example: in the current study, manual gestures such as pirouette, look, and dangle had higher CTI scores, which are associated with reduced flexibility in meaning. In contrast, facial expressions such as the bared teeth face, pout face, and lower lip relaxer face had lower CTI scores, suggesting greater flexibility in meaning when compared to these manual gestures. The idea that gestures can be spontaneous and inflexible is also supported by the literature on body language. Researchers of body language argue that both facial expressions and gestures with the hands, arms, and body can be the result of emotion and still exhibit gesture variables (such as being sensitive to audience effects, (*Gelder, 2009*; *Corneau et al., in press*). In some cases, the gesture types described are referred to as bodily expressions and are "recognized as reliably as facial expressions" [66 page 3475]. Thus, the sharp distinction between gestures and expressions, regardless of body part, may need to be revised in favor of a more holistic, unified concept.

The existence of facial gesturing in chimpanzees has important implications for the evolution of human language. As previously mentioned, those who support a gestural origin of language (*Armstrong, 2008*) sometimes argue that other forms of communication, such as facial expressions and vocalizations, cannot be evolutionary precursors due to their strong association with emotion (*Hewes et al., 1973*; *Corballis, 2002*; *Pollick & De Waal, 2007*). As a result, it is believed that the evolutionary starting point of human language can be traced to great ape gestural communication (*Tomasello & Call, 2019*). The results of our study suggest that the evolutionary starting point of human language may be older than the great ape lineage.

Chimpanzee facial expression types and their corresponding muscle movements can be found in other primate (*Van Hooff, 1967*) and non-primate species (*Darwin, 1872*; *Andrew, 1963*). Similar to chimpanzees, pigtailed macaques (*Macaca nemestrina*) produce a facial expression (silent bared-teeth display) that varies in its meaning (*Flack & De Waal, 2007*). This finding suggests that monkey facial expressions may be used in a flexible manner, which is considered to be an important gesture property (*Byrne et al., 2017*). Mice produce a variety of facial expressions (*Langford et al., 2010*; *Dolensek et al., 2020*) and distinct

32

muscle movements, which include bulging of the nose and cheeks (*Langford et al., 2010*). These muscle movements are similar to those observed in chimpanzees, which include the nose wrinkler (AU9) and cheek raiser (AU6, *Parr et al., 2007*), and they have also been observed in humans, macaques, gibbons, orangutans, dogs, and cats (*Waller, Julle-Daniere & Micheletta, 2020*). If chimpanzee facial expressions and their corresponding muscle movements are gestural, it is possible that these other species with shared expressions and muscle movements are also capable of using their face to gesture. Conversely, species outside of the great apes may lack the control or freedom of movement needed to gesture with their faces. Direct tests for facial gestures outside of great apes are needed to resolve these competing hypotheses.

Research on the gesture properties of mammal facial expressions and vocalizations would help to determine whether the ability to gesture with the face is uniquely derived in hominids or more widespread than previously assumed. We may find that rather than the product of recent evolution, the mechanisms needed for language to emerge have been gradually evolving through much, if not all, of the mammalian lineage. In this case, an additional step would need to be added to the "hand to mouth" theory of language evolution (*Corballis, 2002*) to include the emergence of mammalian facial gestures. In this revised theory, mammalian facial (and possibly vocal) gestures provided the basis for great ape manual gestures to evolve. The ability to produce manual gestures then led to more symbolic forms of communication (such as sign language), and subsequently to the development of spoken language (*Corballis, 2002*). It remains to be determined whether facial gestures will be found outside of chimpanzees, but chimpanzees provide a potential important bridge from facial to manual gestures.

Future studies should also address whether vocalizations are capable of taking on important gesture properties. In the current study, we examined the properties of nine types of facial expressions. Some of these facial expressions are associated with vocalizations (such as the lipsmacking face, pant hoot face, raspberry face, and scream face), whereas others are not often associated with vocalizations (such as the ambiguous face, bared teeth face, lower lip relaxer face, and pout face). When separating these eight facial expression types from one another based on modality, interesting patterns emerge. The four facial expression types associated with vocalizations exhibited a higher average context tie index score (average $= 0.68$, median $= 0.62$, range $= 0.52-0.99$) when compared to the four facial expression types which are seldom associated with vocalizations (average $= 0.30$, median $= 0.29$, range $= 0.25-0.36$); these differences appear to be statistically significant (Mann–Whitney $U = 0$, $p = 0.029$). Facial expressions associated with vocalizations had a lower average composite gesture score (average $= 7.01$, median $= 6.91$, range $= 6.08-8.13$) when compared to the four facial expression that are not typically associated with vocalizations (average $= 8.29$, median $= 8.74$, range $= 6.57-9.10$); however, these differences are not statistically significant (Mann–Whitney $U = 14$, $p = 0.114$). These results suggest that facial expressions associated vocalizations exhibit reduced contextual flexibility and fewer gesture variables on average compared to facial expressions without vocalizations. However, it is difficult to make conclusions with these results alone, since there were instances recorded where facial expressions not often associated with vocalizations were produced with one

33

and vice versa. Some facial expression types (such as the play face) also fall between both categories, since they can be produced with a vocalization but it is not required for it to be produced.

Ultimately, the focus on the hands as the origins of gesturing may be a narrow, anthropocentric view. Outside of the great apes, animals are much more dependent upon their hands for locomotion, limiting the ability and opportunity to use them to communicate. But the face has no such restrictions. To take a deeper look into the origins of human flexible and intentional communication, we need to look to the faces and vocalizations of other mammals for gesture properties through a quantitative approach. Our newly devised scoring metric (composite gesture scores) can be applied to a variety of species, including both primate and non-primate mammals. According to *Graham et al. (2019)*, it is currently difficult to evaluate the gesture properties of signals without a more rigorous and quantified approach. In addition, very few studies have examined multiple gesture properties and corresponding variables at once. We hope that our composite gesture scores will be useful for cross-species comparisons and for ensuring replicability in future studies.

CONCLUSION

We conclude that chimpanzees are capable of using their faces to gesture. Current theories about the gestural origins of human language propose that this process began with great ape manual gestures. But, if other mammals are also gesturing with their faces, then the properties associated with human language may have been gradually evolving through much, if not all, of the mammalian lineage. We propose that researchers look to the faces of other mammals for gesture properties to examine this possibility.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

34

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Brittany Florkiewicz conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Matthew Campbell conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This study was approved by the Los Angeles Zoo and followed the Study of Animal Behaviour's guidelines for the treatment of animals in behavioral research. This study also adhered to the American Society of Primatologists Principles for the Ethical Treatment of Non-Human Primates.

Data Availability

The following information was supplied regarding data availability:

The raw data, script, and ELAN templates are available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.12237#supplemental-information.

REFERENCES

Altmann J. 1974. Observational study of behavior: sampling methods. *Behaviour* 49(3/4):227–267 DOI 10.1163/156853974X00534.

Andrew RJ. 1963. The origin and evolution of the calls and facial expressions of the primates. *Behaviour* 20(1):1–109 DOI 10.1163/156853963X00220.

Andrews P. 2020. Last common ancestor of apes and humans: morphology and environment. *Folia Primatologica* 91(2):122–148 DOI 10.1159/000501557.

Armstrong DF. 2008. The gestural theory of language origins. *Sign Language Studies* 8(3):289–314 DOI 10.1353/sls.2008.0005.

35

Armstrong DF, Wilcox SE. 2007. *The gestural origin of language*. Oxford: Oxford University Press.

Arnold K, Zuberbühler K. 2006. Semantic combinations in primate calls. *Nature* 441:303 DOI 10.1038/441303a.

Bates D, Machler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Bell C. 1806. *Essays on the anatomy of expression in painting*. London: Printed for Longman, Hurst, Rees, and Orne.

Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J. 2017. Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition* 20(4):755–769 DOI 10.1007/s10071-017-1096-4.

Byrne RW, Cochet H. 2017. Where have all the (ape) gestures gone? *Psychonomic Bulletin & Review* 24(1):68–71 DOI 10.3758/s13423-016-1071-0.

Cartmill EA. 2008. Gestural communication in orangutans (*Pongo pygmaeus* and *Pongo abelii*): a cognitive approach. Phd dissertation, University of St. Andrews.

Cartmill EA, Byrne RW. 2010. Semantics of primate gestures: intentional meanings of orangutan gestures. *Animal Cognition* 13(6):793–804 DOI 10.1007/s10071-010-0328-7.

Corballis MC. 1999. The gestural origins of language: human language may have evolved from manual gestures, which survive today as a behavioral fossil coupled to speech. *American Scientists* 87(2):138–145 DOI 10.1511/1999.2.138.

Corballis MC. 2002. *From hand to mouth: the origins of language*. Oxford: Princeton University Press.

Corneau CA, Noroozi F, Kaminska D, Sapkinski T, Escalera S, Anbarjafari G. Survey on emotional body gesture recognition. *IEEE Transaction on Affective Computing* in press.

Darwin C. 1872. *The expression of the emotions in man and animals*. Albemarble Street: John Murray.

Dediu D, Levison SC. 2013. On the antiquity of language: the reinterpretation of neandertal linguistic capacities and its consequences. *Frontiers in Psychology* 4(379):1–17.

De Marco A, Visalberghi E. 2007. Facial displays in young tufted capuchin monkeys (*cebus apella*): appearance, meaning, context and target. *Folia Primatologica* 78:118–137 DOI 10.1159/000097061.

Descartes R. *Passions of the soul*. Indianapolis: Hackett Pub. Co, 1649.

Dolensek N, Gehriach DA, Klein AS, Gogolla N. 2020. Facial expressions of emotion states and their neuronal correlates in mice. *Science* 368:89–94 DOI 10.1126/science.aaz9468.

Ekman P. 1970. universal facial expressions of emotion. *California Mental Health Research Digest* 8(4):151–158.

Ekman P. 1993. Facial expressions and emotion. *American Psychologist* 48(4):384–392 DOI 10.1037/0003-066X.48.4.384.

Ekman P. 1999. Basic emotions. In: *Handbook of cognition and emotion*. New York: John Wiley & Sons Ltd.

36

Elliott EA, Jacobs AM. 2013. Facial expressions emotions sign language. *Frontiers in Psychology* **4**(115):1–4.

Enfield NJ. 2002. 'Lip-pointing': a discussion of form and function with reference to data from Laos. *Gesture* **1**(2):185–211 DOI 10.1075/gest.1.2.06enf.

Fitch WT. 2010. *The evolution of language*. Cambridge: Cambridge University Press.

Flack JC, De Waal FBM. 2007. Context modulates signal meaning in primate communication. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5):1581–1586 DOI 10.1073/pnas.0603565104.

Florkiewicz B, Campbell MW. 2021. A Comparison of Focal and Opportunistic Sampling Methods in Chimpanzee Communication Research. *Folia Primatologica* **92**(3):164–174 DOI 10.1159/000516315.

Fridlund AJ. 1994. *Human facial expression: an evolutionary view*. San Diego: Academic Press.

Fröhlich M, Hobaiter C. 2018. The development of gestural communication in great apes. *Behavioral Ecology and Sociobiology* **72**(194):1–14 DOI 10.1007/s00265-017-2413-2.

Fröhlich M, Sievers C, Townsend SW, Gruber T, Van Schaik CP. 2019. Multimodal communication and language origins: integrating gestures and vocalizations. *Biological Reviews* **94**:1809–1829 DOI 10.1111/brv.12535.

Gelder B. 2009. Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**:3475–3484 DOI 10.1098/rstb.2009.0190.

Genty E, Breuer T, Hobaiter C, Byrne RW. 2009. Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality, and possible origins. *Animal Cognition* **12**:527–546 DOI 10.1007/s10071-009-0213-4.

Goldin-Meadow S, Alibali MW. 2014. Gesture's role in speaking, and learning, and creating language. *Annual Review of Psychology* **64**:257–283.

Graham KE, Hobaiter C, Ounsley J, Furuichi T, Byrne RW. 2018. Bonobo and chimpanzee gestures overlap extensively in meaning. *PLOS Biology* **16**(2):e2004825 DOI 10.1371/journal.pbio.2004825.

Graham KE, Wilke C, Lahiff NJ, Slocombe KE. 2019. Scratching beneath the surface: intentionality in great ape signal production. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**:20180403.

Gupta S, Sinha A. 2019. Gestural communication of wild bonnet macaques in the Bandipur National Park, Southern India. *Behaviour Processes* **168**:103956 DOI 10.1016/j.beproc.2019.103956.

Halina M, Liebal K, Tomasello M. 2018. The goal of ape pointing. *PLOS ONE* **13**(4):1–19.

Heesen R, Hobaiter C, Ferrer-i Cancho R, Semple S. 2019. Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society of London* **286**:20182900.

Hewes GW, Andrew RJ, Carini L, Choe H, Gardner RA, Kortlandt A, Krantz GS, McBridge G, Nottebohm F, Pfeiffer J, Rumbaugh DG, Steklis HD, Raliegh MJ, Stopa R, Suzuki A, Washburn SL, Wescott RW. 1973. Primate communication

37

and the gestural origin of language [and comments and reply]. *Current Anthropology* 14(1/2):5–24 DOI 10.1086/201401.

Hobaiter C. 2011. Gestural communication in wild chimpanzees. Phd dissertation, University of St. Andrews.

Hobaiter C, Byrne RW. 2011. The Gestural repertoire of the wild chimpanzee. *Animal Cognition* 4(5):745–767.

Hobaiter C, Byrne RW. 2014. The meanings of chimpanzee gestures. *Current Biology* 24(14):1596–1600 DOI 10.1016/j.cub.2014.05.066.

Hopkins WD, Leavens DA. 1998. Hand use and gestural communication in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 112(1):95–99 DOI 10.1037/0735-7036.112.1.95.

Kivell TL, Schmitt D. 2009. Independent evolution of knuckle-walking in African apes shows that humans did not evolve from a knuckle-walking ancestor. *Proceedings of the National Academy of Sciences of the United States of America* 106(34):14241–14246 DOI 10.1073/pnas.0901280106.

Laidre ME. 2011. Meaningful gesture in monkeys? Investigating whether mandrills create social culture. *PLOS ONE* 6(2):e14610 DOI 10.1371/journal.pone.0014610.

Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, LaCroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, Van den Maagdenberg AMJM, Ferrari MD, Craig KD, Mogil JS. 2010. Coding of facial expressions of pain in the laboratory mouse. *Nature Methods* 7(6):447–452 DOI 10.1038/nmeth.1455.

Leavens DA, Hopkins WD. 1998. Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures. *Developmental Psychology* 34(5):813–822 DOI 10.1037/0012-1649.34.5.813.

Leavens DA, Russell JL, Hopkins WD. 2005. Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Development* 76(1):291–306 DOI 10.1111/j.1467-8624.2005.00845.x.

Liebal K, Call J. 2012. The origins of non-human primates' manual gestures. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1585):118–128 DOI 10.1098/rstb.2011.0044.

Liebal K, Call J, Tomasello M. 2004. Use of gesture sequences in chimpanzees. *American Journal of Primatology* 64:377–396 DOI 10.1002/ajp.20087.

Liebal K, Pika S. 2006. Gestural communication of orangutans (*pongo pygmaeus*). *Gesture* 6(1):1–38 DOI 10.1075/gest.6.1.02lie.

Liebal K, Waller BM, Burrows AM, Slocombe KE. 2014. *Primate communication: a multimodal approach.* New York: Cambridge University Press.

Maynard-Smith J, Harper DGC. 1995. Animal signals: models and terminology. *Journal of Theoretical Biology* 177:305–311 DOI 10.1006/jtbi.1995.0248.

McCarthy MS, Jensvold MLA, Fouts DH. 2013. Use of gesture sequences in captive chimpanzees (*Pan troglodytes*) play. *Animal Cognition* 16:471–481 DOI 10.1007/s10071-012-0587-6.

38

McHugh ML. 2012. Interrater reliability: the kappa statistic. *Biochemical Medicine* 22(3):276–282.

McNeill D. 2012. *How language began: gesture and speech in human evolution.* Cambridge: Cambridge University Press.

Molesti S, Meguerditchian A, Bourjade M. 2020. Gestural communication in olive baboons (*Papio anubis*): repertoire and intentionality. *Animal Cognition* 23:19–40 DOI 10.1007/s10071-019-01312-y.

Moore R. 2016. Meaning and ostension in great ape gestural communication. *Animal Cognition* 19:223–231 DOI 10.1007/s10071-015-0905-x.

Parr LA, Waller BM. 2006. Understanding chimpanzee facial expression: insights into the evolution of communication. *Social Cognitive and Affective Neuroscience* 1(3):221–228 DOI 10.1093/scan/nsl031.

Parr LA, Waller BM, Vick SJ, Bard KA. 2007. Classifying chimpanzee facial expressions using muscle action. *Emotion* 7(1):172–181 DOI 10.1037/1528-3542.7.1.172.

Pika S, Fröhlich M. 2019. Gestural acquisition in great apes: the social negotiation hypothesis. *Animal Cognition* 22(4):551–565 DOI 10.1007/s10071-017-1159-6.

Plooj F. 1978. Some basic traits of language in wild chimpanzee? In: *Action, gesture, and symbol: the emergence of language.* London: Academic Press.

Pollick AS, De Waal FBM. 2007. Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104(19):8184–8189 DOI 10.1073/pnas.0702624104.

R Core Team. 2020. R: a language and environment for statistical computing. *Available at* https://www.R-project.org/ (accessed on 20 May 2020).

Roberts AI, Roberts SGB. 2019. Gestural repertoire size is associated with social proximity measures in wild chimpanzees. *American Journal of Primatology* 81(3):e22954 DOI 10.1002/ajp.22954.

Roberts AI, Vick SJ, Buchanan-Smith HM. 2012. Usage and comprehension of manual gestures in wild chimpanzees. *Animal Behavior* 84(2):459–470 DOI 10.1016/j.anbehav.2012.05.022.

Roberts AI, Vick SJ, Buchanan-Smith HM. 2013. Communicative intentions in wild chimpanzees: persistence and elaboration in gestural signaling. *Animal Cognition* 16:187–196 DOI 10.1007/s10071-012-0563-1.

Seyfarth RM, Cheney DL. 2018. *The social origins of language.* Oxford: Princeton University Press.

Tamagawa T, Lundh T, Shigetoshi K, Nitta N, Ushiro N, Inubushi T, Shiino A, Karlsson A, Inoue T, Mera Y, Hino K, Komori M, Morikawa S, Sawajiri S, Naka S, Honma S, Kimura T, Uchimura Y, Imai S, Egi N, Otani H, Udagawa J. 2020. Correlation between musculoskeletal structure of the hand and primate locomotion: morphometric and mechanical analysis in prehension using the cross- and triple-ratios. *PLOS ONE* 15(5):e0232397 DOI 10.1371/journal.pone.0232397.

Tomasello M, Call J. 2019. Thirty years of great ape gestures. *Animal Cognition* 22:461–469 DOI 10.1007/s10071-018-1167-1.

39

Van Hooff JARAM. 1967. The facial displays of the catarrhine monkeys and apes. In: *Primate ethology.* London: Weidenfeld & Nicolson.

Visalberghi E, Valenzano DR, Preuschoft S. 2006. Facial displays in *Cebus apella. International Journal of Primatology* 27(6):1689–1707 DOI 10.1007/s10764-006-9084-6.

Waller BM, Julle-Daniere E, Micheletta J. 2020. Measuring the evolution of facial 'expression' using multi-species FACS. *Neuroscience and Biobehavioral Reviews* 113:1–11 DOI 10.1016/j.neubiorev.2020.02.031.

Waller BM, Warmelink L, Liebal K, Micheletta J, Slocombe KE. 2013. Psuedoreplication: a widespread problem in primate communication research. *Animal Behavior* 86(2):483–488 DOI 10.1016/j.anbehav.2013.05.038.

Weigel RM. 1979. The facial expressions of the brown capuchin monkey (*Cebus apella*). *Behaviour* 68(3/4):250–276 DOI 10.1163/156853979X00331.

40

CHAPTER 3

Primate Socio-Ecology Shapes the Evolution of Distinctive Facial Repertoires

BRITTANY N. FLORKIEWICZ[1], LINDA OÑA[2], LEONARDO OÑA[3], MATTHEW W. CAMPBELL[4]

[1] *Department of Anthropology and Center for Behavior, Evolution, & Culture, University of California, Los Angeles, Los Angeles, California, U.S.A.*

[2] *Max Planck Research Group Naturalistic Social Cognition, Max Planck Institute for Human Development, Berlin, Germany*

[3] *Department of Ecology, Universität Osnabrück, Osnabrück, Germany*

[4] *Department of Psychology, California State University Channel Islands, Camarillo, California, U.S.A.*

ABSTRACT

Primates are capable of producing a wide variety of facial muscle movements during bouts of communication, which is likely due to similarities in underlying facial musculature. While shared facial musculature results in shared facial muscle movement, it is currently unclear if shared facial muscle movement leads to morphologically similar facial signals. One hypothesis is that morphological constraints shape the evolution of facial repertoires: species that have higher facial mobility will produce large and diverse repertoires. An alternative hypothesis is that social complexity shapes the evolution of facial repertoires: as social complexity increases, so does communicative complexity. We tested these two hypotheses (the *facial mobility hypothesis* and the *socio-ecological complexity hypothesis*) by comparing chimpanzees (*Pan troglodytes*) and gibbons (family Hylobatidae), two distantly related apes who vary in both their facial mobility and social organization. From the facial mobility hypothesis, we would predict gibbons to have a larger, more diverse facial repertoire compared to chimpanzees, whereas according to the socio-ecological complexity hypothesis we would predict chimpanzees to have a larger, more diverse facial signaling repertoire than gibbons. We compared the morphology and complexity of facial repertoires for both apes using Facial Action Coding Systems designed for chimpanzees and gibbons. We found that the chimpanzee facial signaling repertoire was larger and more complex than that of gibbons, which is consistent with the socio-ecological complexity hypothesis. Chimpanzees produced a greater number of morphologically distinct Action Unit (AU) combinations than gibbons, even when focusing on AUs which are present in both apes. On average, chimpanzees also produced AU combinations consisting of a greater number of morphologically distinct AUs. In contrast, gibbons produced AU combinations consisting of

fewer AUs on average. Differences in facial repertoires may be attributed to a variety of socio-

ecological factors (such as group size) differing between chimpanzees and gibbons.

INTRODUCTION

Compared to other animals, primates appear to demonstrate high facial mobility: they can produce a wide variety of facial muscle movements during bouts of communication that other animals cannot (Darwin, 1872; Van Hooff, 1967). But why do primates exhibit such high facial mobility? Different primate species exhibit similar muscle movements due to their shared underlying facial musculature (Liebal et al., 2013). Facial signals might help facilitate social activities in primates (Roberts & Roberts, 2019; Silk, 2002). However, while shared facial musculature results in the presence of shared facial muscle movements, it is unclear if shared facial muscle movements result in morphologically similar facial signals. While some stereotyped facial signals are produced across a wide variety of species (such as the play face; Preuschoft, 2000; Preuschoft & van Hooff, 1997), others appear to be restricted to fewer primate species (such as the raspberry face; Hopkins et al., 2007).

One hypothesis for why variation exists in stereotyped facial signals exists is that greater facial mobility may lead to a large and more diverse repertoire of facial signals (Wathan et al., 2015). For example: gibbons (family Hylobatidae) have a large number of documented facial muscle movements (N=20; Waller et al., 2012) and are capable of producing up to 80 morphologically distinct facial signals (Florkiewicz et al., 2018; Scheider et al., 2014). In contrast, chimpanzees (*Pan troglodytes*) have fewer documented facial muscle movements (N=15; Vick et al., 2007), so they may produce fewer facial signals than gibbons. Variation in facial mobility may also be associated with differences in the connective projects of facial muscles and variation in properties associated with these muscle fibers (Sherwood, 2005). For example: when compared to humans (*Homo sapiens*), chimpanzees show a smaller proportion of slow-twitch facial muscles, which are necessary for fine-grained control of the face (Sherwood,

2005). In contrast, gibbons have a proportion of slow-twitch facial muscle which falls in between the human and chimpanzee range (Burrows et al., 2016).

However, the 'social complexity hypothesis' (Freeberg et al., 2012) suggests that complex social systems (such as those with numerous individuals) result in greater communicative complexity. Being able to produce a wide variety of morphologically distinct facial signals may be useful for navigating a greater number of possible social interactions (Parr et al., 2005; Waller et al., 2017), and the restricted use of stereotyped facial signals by certain primate species may be attributed to differences in their social behavior and organization (Clutton-Brock, 2009). It is possible that primates living in larger social groups utilize not only a large number of morphologically distinct facial signals but also produce more complex facial signals. Many primates use facial signals as reliable predictors of future behavior, such as crested macaques (*Macaca nigra*; Waller et al., 2016), capuchin monkeys (*Cebus apella*; Morimoto & Fujita, 2012), Bornean orangutans (*Pongo pygmaeus*), gorillas (*Gorilla gorilla*), bonobos *(Pan paniscus)*, and chimpanzees (*Pan troglodytes*; Buttelmann et al., 2009). Individuals may be able to increase the accuracy of their predictions or predict novel social behaviors by increasing the size and complexity of their communicative repertoires. This fine-tuned predictive framework could then be used to manage multiple social relationships simultaneously and navigate complex social bond management activities (Roberts & Roberts, 2019; Silk, 2002; Whiten, 1997).

Our goal was to test two hypotheses regarding the relationship between facial mobility, facial signaling, and sociality by studying the behavior of two distantly related apes who vary in their social behavior and organization: chimpanzees (*Pan troglodytes*) and gibbons (family Hylobatidae).

45

(1) The <u>facial mobility hypothesis</u> posits that signal repertoire size is based on the underlying mobility of the face. Gibbons are capable of producing more facial muscle movements during bouts of communication than chimpanzees (Vick et al., 2007; Waller et al., 2012), which implies that gibbons have higher facial mobility. According to the facial mobility hypothesis, we predict that gibbons will have a larger repertoire size than chimpanzees because gibbons are capable of producing more facial movements (i.e., facial action units) and have a greater proportion of slow-twitch facial muscles (<u>prediction 1A</u>). We also predict that gibbon facial signals will be more complex (i.e., consist of more facial muscle movements) than chimpanzees (<u>prediction 1B</u>).

(2) The <u>socio-ecological complexity hypothesis</u> states that species living in larger social groups should have larger signal repertoire size than species living in smaller social groups. Chimpanzees live in much larger groups (anywhere from 15 to 150 individuals; (Campbell et al., 2010) than gibbons (typically 2 to 6 individuals; Ryne, 1996). Chimpanzees often engage in group-level social activities (such as cooperative hunting; Boesch & Boesch, 1989) and learn complex behavioral patterns from other members of their social group (such as tool use; (Lonsdorf, 2006), many of which have not been documented in gibbons. According to the socio-ecological complexity hypothesis, we predict that chimpanzees will have a greater number of morphologically distinct facial signals documented than gibbons (<u>prediction 2A</u>), and that chimpanzees will produce facial signals that are more complex (i.e., consist of more facial muscle movements) than gibbons (<u>prediction 2B</u>).

To examine variation in facial mobility and facial signaling repertoires, we made use of Facial Action Coding Systems (or FACS) specifically designed for chimpanzees (Vick et al., 2007) and gibbons (Waller et al., 2012). FACS are considered to be one of the most systematic

46

and standardized methods of studying facial signals because they limit observation bias by placing equal emphasis on all facial muscle movements (Parr et al., 2008). While multiple studies have examined facial mobility in primates using FACS (Florkiewicz et al., 2018; Parr et al., 2007; Scheider et al., 2014; Scheider et al., 2016), they tend to focus on a single species or set of closely related species. In addition, studies comparing facial mobility across a wide variety of species tend to focus on the presence or absence of individual facial muscle movements in a species repertoire (Dobson, 2009) rather than how these facial muscle movements are combined to create a facial signal.

To date, only 3 gibbon studies have reported on the extent of variation in facial muscle movement combinations produced during bouts of communication using FACS (Florkiewicz et al., 2018; Scheider et al., 2014; Scheider et al., 2016). We make use of data from 2 of these studies (Florkiewicz et al., 2018; Scheider et al., 2014), along with newly collected chimpanzee data, to examine whether there are differences in the production of individual facial muscle movements and combinations between gibbons and chimpanzees. Our study provides insight into the evolution of facial communication by disentangling the relationship between repertoire size, repertoire complexity, musculature, and sociality among apes.

METHODS

Chimpanzee Data Collection

We collected data at the Los Angeles Zoo from 2017 to 2019 with a single troop of chimpanzees (*Pan troglodytes*). We observed a total of 18 chimpanzees throughout the study. We used two different sampling methods to gather our data: the focal individual sampling method, which took place in 2017 (Altmann, 1974) and the opportunistic sampling method, which took place from 2018 to 2019 (Florkiewicz & Campbell, 2021). For the focal individual sampling method, we

recorded each member of the troop in 30-minute intervals, with each individual being sampled once per week. The order in which we recorded individuals (along with the time that they were recorded) was randomized. The focal individual sampling method resulted in 72 hours of video footage (or 4 hours per individual). For the opportunistic sampling method, we followed the most active portion of the troop and recording all social interactions that occurred. We began recordings just before the start of a social interaction and ended them when chimpanzees dispersed and/or ceased their communication. Our video recordings varied in length based on the duration of social interaction(s). The opportunistic sampling method resulted in 84.5 hours of video footage. We collected our data Monday through Friday from 8:00 to 14:00, which we identified as peak activity hours during a pilot study in 2016. We combined data across the two methods, since our previous comparison of the sampling methods showed that they performed similarly (Florkiewicz & Campbell, 2021). When combined, these sampling methods resulted in 156.5hr of footage for chimpanzees.

Gibbon Data Collection

We collected data across 5 different institutions. Data reported in Scheider et al. (2014) took place at the Twycross Zoo (N=6 individuals), the Zurich Zoo (N=4), the Mullhouse Zoo (N=4), and the Rheine Zoo (N=2). The 16 individuals represented five different species: *Symphalangus syndactylus* (N=6), *Hylobates pileatus* (N=4), *Hylobates lar* (N=2), *Nomascus gabriellae* (N=2), and *Nomascus siki* (N=2). Data reported in Florkiewicz et al. (2018) took place at the Gibbon Conservation Center comprising 20 individuals across 4 species: *Hylobates moloch* (N=6), *Hylobates pileatus* (N=2), *Hoolock leuconedys* (N=8), and *Nomascus leucogenys* (N=2). In total, our analysis includes 36 individuals from 8 species and all 4 genera of Hylobatids. Both studies used the focal pair sampling method to continuously video record the subjects (Altmann, 1974).

We collected a total of 227hr of footage between the two studies; 21hr in Scheider et al. 2014, and 206hr in Florkiewicz et al. 2018.

Information regarding the names, age groups, birthplaces and sexes of the chimpanzees and gibbons incorporated into this study can be found in the appendix (Table S1).

Data Coding

In the current study, we defined a facial signal as facial muscle movement that was performed by a signaler with the attempt to alter the behavior(s) of potential recipient(s), which was based off of the broader definition of a communicative signal (Smith & Harper, 1995). Each facial signal can consist of one or more facial muscle movements, defined as the contraction and/or relaxation of muscles present on the face (i.e., above the neck region). Each facial signal we identified was coded using Facial Action Coding Systems (or FACS). We referred to each facial muscle movement produced in the signal as an action unit (or AU), and we assigned each facial signal a numerical combination consisting of multiple AUs (or AU combination; Ekman & Rosenberg, 2005, see Figure 1).
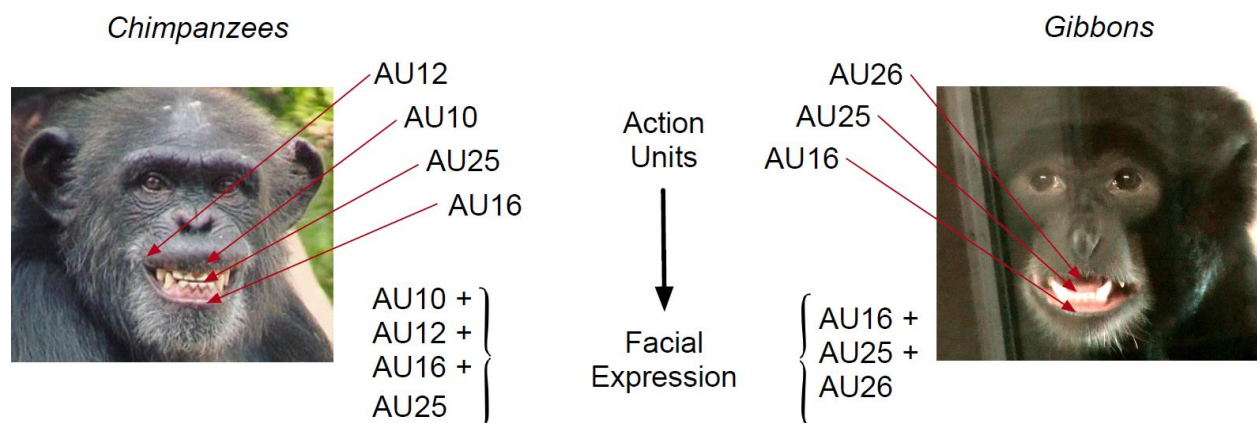

FIGURE 1. A visual representation of how to code facial signals using FACS. Single muscle movements (which are referred to as Action Units or AUs) are used in combination leading to facial movement combinations (AU combinations).

We coded all chimpanzee facial signals using the ChimpFACS (Vick et al., 2007), and we coded gibbon facial signals using the GibbonFACS (Waller et al., 2012). Additional information about the AUs documented in the ChimpFACS and GibbonFACS can be found in the appendix (Table S2).

Inter-Observer Reliability (IOR)

We assessed agreement using Wexler's Ratio, which is a common practice in studies that use FACS (Parr et al., 2007). The equation for Wexler's ratio is:

$$\frac{2 \times (\# \, of \, AUs \, Agreed \, by \, Both \, Coders)}{(\# of \, AUs \, coded \, by \, R1) + (\# \, of \, AUs \, coded \, by \, R2)}$$

with R1 denoting researcher 1 and R2 denoting researcher 2 (Parr et al., 2007). We calculated Wexler's ratio for each facial signal with the average ratio used to assess agreement. We used a minimum Wexler's ratio of 0.70 for 'good' agreement, which is a common threshold to pass human and some non-human FACS certification tests (Lewinski et al., 2014). We performed IOR on a subset of facial signaling data in each study. Both Scheider et al. (2014) and Florkiewicz et al. (2018) extracted 10% of observed facial signals from their datasets, and calculated Wexler's ratio for each facial signal. Scheider et al. (2014) reported an average Wexler's ratio of 0.83, and Florkiewicz et al. (2018) reported an average Wexler's ratio of 0.73. Chimpanzee AU combinations had an average Wexler's ratio of 0.75. As all of the ratios are above the 0.70 threshold, we considered them all to be in good agreement.

Data Analysis

We performed statistical tests in R 3.6.2 (R Core Team 2017). To test the facial mobility hypothesis and socio-ecological complexity hypothesis, we made comparisons between chimpanzees and gibbons regarding the morphology and complexity of their facial signals. The facial mobility hypothesis predicts that gibbons will have: (1A) a greater number of

morphologically distinct facial signals; and (1B) more complex facial signals compared to chimpanzees. In contrast, the socio-ecological complexity hypothesis predicts that chimpanzees will have: (2A) a greater number of morphologically distinct facial signals; and (2B) more complex facial signals compared to gibbons.

In our current study, we draw comparisons between greater and lesser apes (*Pan troglodytes* and family Hylobatidae) rather than species for two main reasons. Most gibbon species exhibit similarities in their social behavior and organization (Campbell et al., 2010; Gittins, 1980; Ryne, 1994; Southwick, 1985), and previous studies have identified consistency in facial signaling behavior across gibbon species (Florkiewicz et al., 2018; Scheider et al., 2014). In addition, previous studies have found no relationship between facial signaling behavior (in the form of rates, repertoire composition, and diversity of repertoires) and socio-ecological factors across gibbons, such as group size and extent of monogamy (Scheider et al., 2014). Similarity in facial signaling behavior across gibbons may be attributed to similarities in their overall socio-ecology.

We faced two main obstacles for analyzing our data. First, there were differences in the number of hours recorded for each ape and species across the 3 studies. As a whole, we had more hours for gibbons (N=227hr) than chimpanzees (N=156.5hr). Some species (such as *Nomascus gabriellae*) were only sampled in one study, whereas others were sampled in two (Table S1). Second, there could be idiosyncratic variation in facial signaling behavior. Our current study includes data from 54 individuals: 36 gibbons and 18 chimpanzees. It is possible that some AUs or AU combinations are restricted to certain individuals. It is also possible that other factors, such as age and sex, contribute to differences in facial signaling behavior. If unaccounted for,

these differences could impact our analyses pertaining to the diversity and complexity of facial signals in chimpanzees and gibbons.

To account for idiosyncratic differences and differences in recording time between apes, we made use of: (1) corrected repertoire sizes; and (2) generalized linear mixed models. Corrected facial repertoire sizes (R_{CM}) were first introduced by Scheider et al. (2014) as a method to account for differences in recording time between species. To calculate R_{CM} in our current study, we divided the number of AU combinations observed for each ape type (gibbons and chimpanzees) by their recording time (in minutes). We then compared our R_{CM} scores using a two-sample test for equality. By using corrected repertoire sizes, we were able to make direct comparisons between chimpanzees and gibbons to test predictions 1A and 2A. To account for both idiosyncratic differences and differences in recording time, we also made use of generalized linear mixed models (or GLMMs). GLMMs help to account for the pooling fallacy and idiosyncratic differences in signaling behavior through random effects (Waller et al., 2013). By using an offset term, we can also account for differences in recording time across species.

We ran three different sets of models to examine differences in: (M1) the use of individual AU's; (M2) the use of AU combinations; and (M3) the complexity of facial signals (i.e., length of AU combinations). We used model sets M1 and M2 to examine differences in facial signaling morphology (i.e., test predictions 1A and 2A), whereas we used models in set M3 to examine differences in facial signaling complexity (i.e., test predictions 1B and 2B). For all of our models, we set signaler ID as a random variable to help account for idiosyncratic differences in facial signaling behavior. We also used an offset term to account for differences in recording time (in minutes) across the 9 species.

Facial Signaling Morphology Models (M1 and M2)

To make comparisons between chimpanzees and gibbons facial signaling morphology, we analyzed facial signals at two different levels: (1) the AU; and (2) the AU combination. This is because each facial signal (i.e., AU combination) comprises multiple facial muscle movements (i.e., AUs). Chimpanzees and gibbons may differ in the production and use of AUs, which could impact the composition of their AU combination. To account for differences in the production of AUs and AU combinations, we ran two sets of models.

In the first set of models (M1 series), we set the number of observations per AU as the outcome variable. For model set M1, we tested five different models, which contain a combination of four different explanatory variables. In the first model (M1.1), we set ape type as an explanatory variable to determine whether there are significant differences in the production of AUs and AU lengths between gibbons and chimpanzees. In the second model (M1.2), we examined the interaction between categories of AUs and AU lengths with ape type to better understand why gibbons and chimpanzees are similar/different from one another. In the remaining three models, we included age (M1.3), sex (M1.4), and a combination of both (M1.5) as explanatory variables in addition to our interaction term. For our M1 series, we ran Negative-Binomial GLMMs (NB-GLMMS) since our outcomes were discrete counts of each AU observed and the variance was greater than the mean.

In the second set of models (M2 series), we set the number of observations per AU combination as the outcome variable. We followed a similar procedure as the M1 series for our M2 model set, which examines differences in gibbons and chimpanzees in the production of AU combinations. However, we did not include an interaction between ape type and AU combination type for two main reasons: (1) there were a large number of AU combinations identified across apes (N=65); and (2) the data were zero-inflated, which led to convergence

issues when including interaction terms. For model M2.1, we included ape type as an explanatory variable. In the remaining four models, we included age (M2.2), sex (M2.3), and a combination of both (M2.4) as explanatory variables in addition to ape type. In our AU combination dataset, approximately 84.90% of all data points had a value of 0 assigned for the outcome variable, resulting in a zero-inflated dataset resembling a negative-binomial distribution. To accommodate this, we made use of Zero-Inflated Negative-Binomial GLMMs (ZINB-GLMM).

Facial Signaling Complexity Models (M3)

We defined complexity as the number of AUs used to produce a given AU combination (i.e., more AUs = more complexity). After coding all facial signals observed with the chimpFACS and the gibbonFACS, we counted the number of AUs used to produce each facial signal (or AU combination). We used this information to construct our categorical outcome variable "AU Length." Initially, we had 7 categories associated with AU Length (with AU combinations ranging in length from 1 to 7 AUs). However, we had a small number of facial signaling observations which consisted of 7 AUs (N=3), so we removed this category from our analyses. For model M3.1, we included ape type as an explanatory variable. In the remaining four models, we included age (M3.2), sex (M3.3), and a combination of both (M3.4), in addition to ape type. For our M3 series, we ran Ordinal GLMMs (O-GLMMS), since our outcome variable (AU Length) consisted of naturally ordered (based off of the number of AU's present within a given AU Combination, with each step above the previous consisting of +1 AUs).

Model Selection Procedure

We compared and selected our best fitting models for each model series using Akaike's Information Criterion (or AIC) scores. To compare our models, we calculated the difference

(ΔAIC) between each model *i* (AIC*i*) and the model with the lowest AIC score *l* (AIC*l*; Preininger et al. 2013). We considered models with an ΔAIC value of ≤2 to be good fitting models, and we considered the model with the lowest ΔAIC to be the best fitting model (Preininger et al., 2013). AIC and ΔAIC scores for all models can be found in the appendix (Table S3).

Our best fitting models were M1.3 (ApeType*AU + Age), M2.1 (ApeType), and M3.1 (ApeType). For our M1 series, models M1.3 and M1.5 had matching scores for both AIC and ΔAIC. Therefore, we went with the simpler model (M1.3: ApeType*AU + Age) since adding the variable sex in M1.5 did not significantly improve the fit of the model (which would result in a lower AIC and ΔAIC than M1.3). We will discuss these models in our results section, and their outputs can be found in the appendix (Table S3). For our best fitting models, explanatory variables were only considered to be significant if $p<0.05$. For M1.3, we separated the data into smaller subsets (based on AU category) to calculate significant effects (Preininger et al., 2013) for our interaction term (ApeType*AU). This allowed us to determine whether there were significant effects for each AU between apes. We also performed pairwise comparisons to examine differences between the four age groupings in model M1.3. For model 3.1, we conducted two-sample tests for equality of proportions to determine the source of significance between thresholds.

To run model sets M1 and M2 in R, we used the packages "r2admb" and "glmmADMB," which run both zero-inflated and non-zero-inflated negative-binomial GLMMs (Fournier et al., 2012). To run model set M3, we used the package "ordinal," which is specifically designed to run ordinal GLMMs (Christensen, 2019). We calculated AIC and ΔAIC scores for all our models using the package 'bbmle' (Bolker, 2021).

RESULTS

After combining datasets, we observed a total of 9,140 facial signaling events (N=1,091 for chimpanzees; N=8,049 for gibbons) across 383.5hr of video footage (N=156.5hr for chimpanzees; N=227hr for gibbons). Initially, we identified a total of 37 distinct AUs and 432 distinct AU combinations. Out of the 37 muscle movements observed, we observed 22 in gibbons and 27 in chimpanzees. Out of the 432 AU combinations observed, we observed 80 in gibbons and 357 in chimpanzees. Our initial results suggest that chimpanzees produce a greater variety of AUs and AU combinations. However, this could be attributed to differences in use, visibility, and/or study design. For example: chimpanzees and gibbons are capable of producing AU1+2 (inner and outer brow raiser), but we only observed this movement in gibbons. It is unclear if this is attributed to differences in use (i.e., that gibbons use AU1+2 for the purpose of communication, whereas chimpanzees do not) or in reduced visibility with our chimpanzee video footage. To account for these possibilities, we decided to take a conservative approach and focus only on AUs which we identified across all 3 studies. We will focus on this filtered dataset for the rest of our results.

Facial Signaling Morphology

In our filtered dataset, we observed a total of 5,521 facial signaling events (N=1,068 for chimpanzees; N=4,453 for gibbons). We identified a total of 12 distinct AUs shared between chimpanzees and gibbons. These 12 distinct AUs were used by apes to produce 65 distinct AU combinations.

AU Combinations

Out of the 65 distinct AU combinations observed, we found 21 in both gibbons and chimpanzees. We found 14 AU combinations exclusive to gibbons and 30 AU combinations

exclusive to chimpanzees. After accounting for differences in recording time between apes, we found that chimpanzees have a significantly larger corrected repertoire size (R_{CM}) than gibbons (χ^2=12.563, p=0.003). A list of all 65 AU combinations we observed can be found in Table 1. Our best fitting model only contained ape type as a fixed explanatory variable (AIC=5665.04; ΔAIC=0.0); the incorporation of age and/or sex as fixed variables reduced the fit of our model (ΔAIC=1.4-6.5). Our ZINB-GLMM revealed that differences in the production of AU combinations between chimpanzees and gibbons are significant (ß=2.173; SE=0.361; z=6.020; p<0.001).

TABLE 1. A list of the 65 AU combinations that we observed in our current study. We have presented AU combinations, along with the number of observations for each ape type, in ascending order of combination complexity. We also present data on the number of observations per AU combination for chimpanzees and gibbons.

AU Combination	Chimpanzees	Gibbons
6	0	2
9	0	9
12	0	3
17	2	13
22	4	0
24	0	48
25	0	16
10+25	0	1
16+25	0	2
17+24	56	5
17+25	1	0
22+25	45	0
25+26	107	2058
25+27	12	63
10+12+25	1	0
10+25+26	0	31
10+25+27	0	34
12+16+25	5	0
12+22+25	2	0
12+25+26	33	9
12+25+27	7	6
16+25+26	18	1333

16+25+27	13	59
16+26+27	1	0
17+22+25	1	0
17+25+26	0	62
19+25+26	47	17
19+25+27	0	8
22+25+26	236	243
22+25+27	38	39
24+25+26	2	0
10+12+16+25	2	0
10+12+25+26	1	2
10+12+25+27	0	7
10+16+25+26	5	16
10+16+25+27	8	100
12+16+25+26	123	5
12+16+25+27	47	6
12+19+25+26	1	0
16+17+25+26	0	1
16+19+25+26	1	0
16+22+25+26	4	0
16+22+25+27	1	2
19+22+25+26	2	0
6+12+16+25	4	0
6+12+25+26	2	0
9+10+25+26	1	0
9+10+25+27	0	2
9+16+25+26	1	0
10+12+16+25+26	88	42
10+12+16+25+27	28	207
12+16+19+25+26	3	0
12+16+19+25+27	2	0
6+10+12+16+25	4	0
6+12+16+25+26	12	0
6+12+16+25+27	2	0
9+10+16+25+26	5	0
9+10+16+25+27	1	1
9+10+19+25+26	3	0
9+12+16+25+26	1	0
6+10+12+16+25+26	50	0
6+10+12+16+25+27	7	0
9+10+12+16+25+26	13	1

9+10+12+16+25+27	12	0
6+9+10+12+16+25+26	3	0

We found that differences in the production of AU combinations for each ape (with production being associated with the number of unique AU combinations and number of observations per AU combination) appear to be associated with differences in: (1) the production of individual AUs; and (2) the production of complex facial expressions.

Action Units (AUs)

Our best fitting NB-GLMM revealed moderate differences regarding the overall production of AUs between chimpanzees and gibbons (ß=0.878; SE=0.516; z=1.700; p=0.089). There were significant differences between chimpanzees and gibbons in the production of the following AUs: AU6 (ß= -3.036; SE=0.788; z=-3.860; p<0.001), AU10 (ß=1.234; SE=0.509; z=2.430; p=0.015), AU16 (ß=2.035; SE=0.518; z=3.93; p<0.001), AU25 (ß=2.863; SE=2.99; z=9.58; p<0.001), AU26 (ß=2.912; SE=0.297; z=9.820; p<0.001), and AU27 (ß=1.508; SE=0.506; z=2.980; p=0.003). Chimpanzees produced significantly more AU6's, whereas gibbons produced significantly more AU10's, AU16's, AU25's, AU26's, and AU27's (Figure 2). There were no significant differences in the production of AU12, AU17, AU19, AU22, AU24, and AU9 (p<0.05). While our best fitting model included age category as a predictor variable (AIC=3786.90; ΔAIC=0.0), there were no significant differences in the production of AUs between infants, subadults, and adults; however, there were significant differences between juveniles and adults (Pair-wise comparison: ß=2.206; SE=0.771; z=2.863; p=0.019).
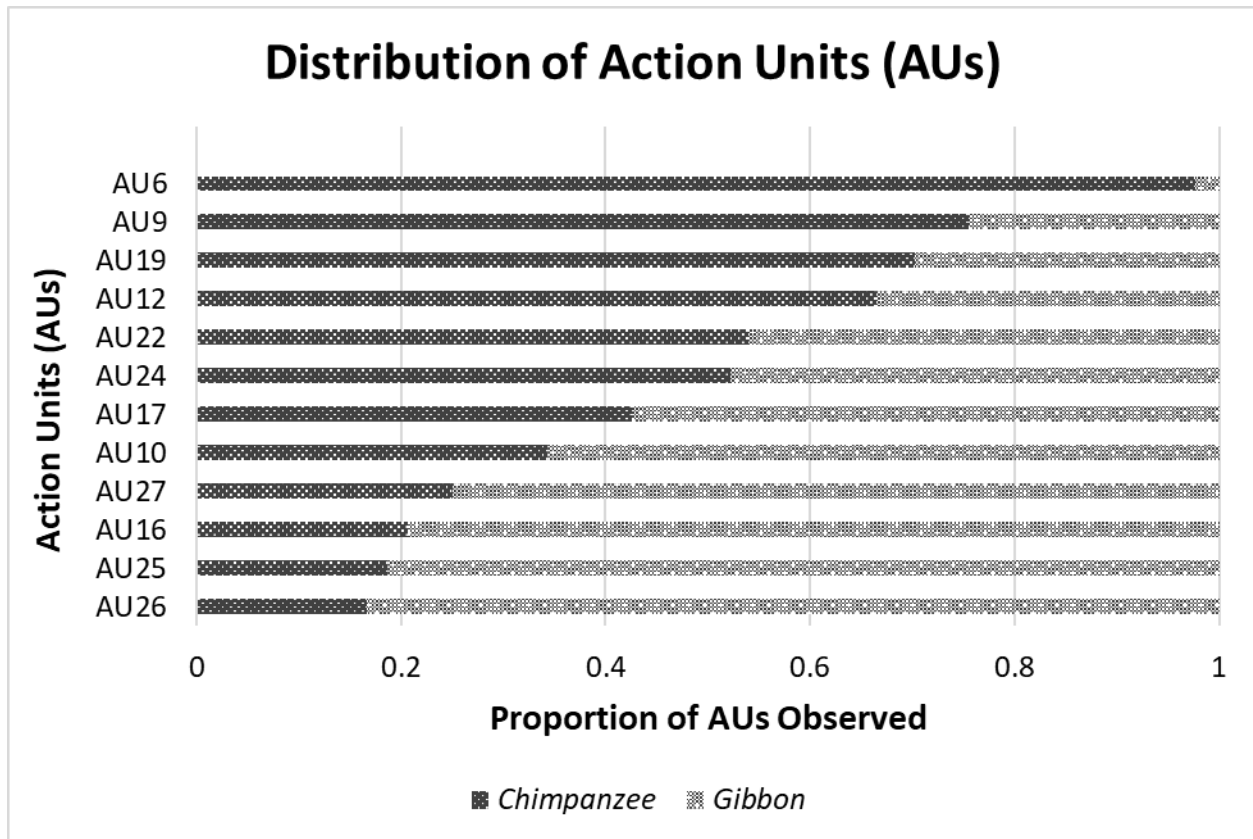
FIGURE 2. A stacked histogram chart showing the proportion of observations (x-axis) for each AU type (y-axis) across chimpanzees and gibbons.

Differences in the number of AUs produced by each ape can partially explain differences in the production of AU combinations for each ape. For example: chimpanzees produced significantly more AU6's compared to gibbons. As a result, chimpanzees produced AU combinations which contained AU6 more frequently (N=84) than gibbons (N=2). Interestingly, out of the 30 AU combinations which were unique to chimpanzees, 8 contained AU6. Only 1 of the 14 unique AU combinations observed in gibbons contained AU6. However, this pattern was not consistent across AUs and apes. For example: gibbons produced significantly more AU10's compared to chimpanzees (β=1.234; SE=0.509; z=2.430; p=0.015). As a result, gibbons produced AU combinations which contained AU10 more frequently (N=444) than chimpanzees (N=232). Out of the 14 unique AU combinations observed in gibbons, 5 contained AU10.

60

However, 10 out of the 30 unique AU combinations observed in chimpanzees contained AU10. It appears that differences in the complexity of chimpanzee and gibbon facial signals can also explain differences in the production of AU combinations produced by each ape.

Facial Signaling Complexity

There was variation in the complexity of AU combinations, with complexity being defined as the number of AUs used to produce an AU combination. Most AU combinations observed (83.23%) consisted of either two (42.57%) or three (40.66%) AUs. It was less common to observe AU combinations consisting of four or more AUs (15.02%). It was also rare to observe an AU combination consisting of a single AU (1.76%). For our O-GLMM, we compared the number of facial signals observed in each level of complexity between chimpanzees and gibbons (Figure 3). Our best fitting model only contained ape type as a fixed explanatory variable (AIC=11461.02; ΔAIC=0.0); the incorporation of age and/or sex as fixed variables reduced the fit of our model (ΔAIC=1.2-2.9).
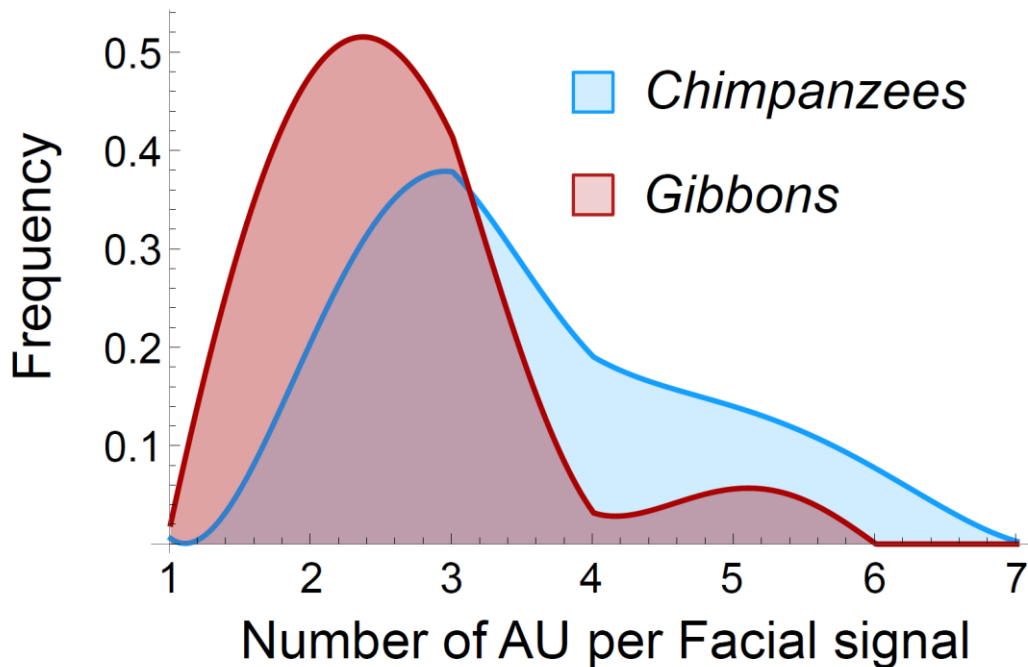
FIGURE 3. Complexity of AU combinations in chimpanzees (blue) and gibbons (red). Complexity is defined as the number of unique AUs (x-axis) an AU combination is composed of.

On average, chimpanzees produced more complex facial signals (mean=3.49, standard deviation=1.21) compared to gibbons (mean=2.63, standard deviation=0.82). The results of our O-GLMM shows that ape type shows a trend for a moderate effect on facial signaling complexity (ß=-0.627; SE=0.378; z=-1.658; p=0.0974). For gibbons, the odds of producing a more complex facial signal (i.e., one which consists of 2 or more AUs) is 46.57% lower than chimpanzees (OR=0.534; CI=0.255-1.121). To better understand why ape type only had a moderate effect on facial signaling complexity, we ran a 2-sample test for equality of proportions for each level of facial signaling complexity. We found that gibbons produce a significantly greater proportion of facial signals with 1 AU (χ^2=10.063, p=0.002), 2 AUs (χ^2=256.27, p<0.001), and 3 AUs (χ^2=3.998, p=0.050) than chimpanzees. In contrast, we found that chimpanzees produce a significantly greater proportion of facial signals with 4 AUs (χ^2=368.74, p<0.001), 5 AUs (χ^2=88.652, p<0.001), and 6 AUs (χ^2=336.74, p<0.001) than gibbons.

Differences in the complexity of facial signals produced by each ape can also explain differences in the production of AU combinations for each ape. Out of the 30 AU combinations which were unique to chimpanzees, 21 contained 4 or more AUs. Only 3 of the 14 unique AU combinations observed in gibbons contained 4 or more AUs. In contrast, out of the 14 unique AU combinations observed in gibbons, 11 contained 1-3 AUs. Only 9 out of the 30 unique AU combinations produced by chimpanzees contained 1-3 AUs.

DISCUSSION

The goal of our current study was to compare how facial muscle movements are produced, combined, and used during bouts of communication between distantly related apes

who vary extensively in their social structure. Specifically, we tested two hypotheses: (1) the facial mobility hypothesis; and (2) the socio-ecological complexity hypothesis (Table 2).

TABLE 2. A brief description of our two hypotheses, their corresponding predictions, whether these predictions were supported, and how we analyzed the evidence.

Hypothesis	Predictions	Support	Evidence
The Facial Mobility Hypothesis	1A. Gibbons will have a larger facial repertoire size than chimpanzees.	No	R_{CM}, ZINB-GLMMs, NB-GLMMs
	1B. Gibbon facial signals will be more complex than those produced by chimpanzees.	No	Ordinal GLMMs, Two-sample test for equality
The Socio-Ecological Complexity Hypothesis	2A. Chimpanzees will have a larger facial repertoire size than gibbons.	Yes	R_{CM}, ZINB-GLMMs, NB-GLMMs
	2B. Chimpanzee facial signals will be more complex than those produced by gibbons.	Yes	Ordinal GLMMs, Two-sample test for equality

We found that chimpanzees produce a greater variety of AUs (i.e., facial muscle movements) and a greater variety of AU combinations (i.e., facial muscle movement combinations, or facial signals) than gibbons, which provides support for prediction 2A and evidence against prediction 1A (Table 2). To verify that these results were not due to differences in coding schemes and/or differences in video quality, we ran a further analysis on only the AUs which were identified and shared across chimpanzees and gibbons. The pattern still remained in this filtered sample: chimpanzees produced a greater variety of AU combinations compared to gibbons despite having fewer recording hours, fewer individuals sampled, and fewer facial signaling observations. Chimpanzees also produced facial signals which were more complex on average (i.e., consisted of a greater number of AUs) than those observed in gibbons; this provides further support for prediction 2B and evidence against prediction 1B (Table 2). Our results suggest that having increased facial mobility does not necessarily lead to large and more complex facial signaling repertoires. Rather, facial signaling repertoires are dependent on how each AU is produced and combined with other AUs by a given ape.

While chimpanzees have larger and more complex facial signaling repertoires than gibbons, both chimpanzees and gibbons still exhibit similarities in their facial signaling morphology. In our current study, we focused on 12 AUs which we observed in both chimpanzees and gibbons. These AUs we observed were used to produce 21 AU combinations observed in both chimpanzees and gibbons. The presence of shared AUs and AU combinations is likely the result of evolutionary continuity. Many of the AUs found in both chimpanzees and gibbons are also found in a wide variety of mammals (Waller et al., 2020) such as orangutans (*Pongo spp.;* Caeiro et al., 2013), macaques (*Macaca mulatta;* Parr et al., 2010), horses (*Equus callabus;* Wathan et al., 2015), dogs (*Canis lupus familiaris;* Waller et al., 2013), and cats (*Felis catus;* Caeiro et al., 2013). This includes AU10, AU12, AU16, AU25, AU26, and AU27 (Waller et al., 2020). These AUs are often used to produce stereotyped facial signals which are found in a wide variety of mammal species (such as the ones mentioned above; Andrew, 1963). One example of this can be seen with 'threat' displays. Threat displays are found in a wide variety of mammal species, and typically involve opening the mouth (AU25, AU26 and 27), drawing the corners of the lips backwards (AU12), and/or exposing both rows of teeth (AU10 and AU16; Andrew 1963).

However, the results of our study show that evolutionary continuity alone cannot explain facial repertoire size and use. While chimpanzees and gibbons share 12 AUs, we discovered differences in how they are produced and used by each ape. Chimpanzees produced significantly more AU6's, whereas gibbons produced significantly more AU10's, AU16's, AU25's, AU26's, and AU27's. Differences in AU production could be attributed to differences in contextual use. For example: AU6 is often associated with the production of bared teeth faces in primates (Parr et al., 2007). While both chimpanzees and gibbons are capable of producing bared teeth faces

(Liebal et al., 2004; Parr et al., 2007), our study found that chimpanzees are significantly more likely to produce this AU. Previous studies have found that chimpanzees frequently use bared teeth faces to signal affiliation (Waller & Dunbar, 2005) and/or submissiveness (Van Hooff, 1967). In contrast, bared teeth faces are mostly produced by adult, male gibbons in the context of sex (Liebal et al., 2004).

Chimpanzees also possess a larger and more complex facial signaling repertoire than gibbons, which provides support for our socio-ecological complexity hypothesis. Differences in facial signaling function could explain differences in the size and composition of facial signaling repertoires in both chimpanzees and gibbons. In chimpanzees, having a large and complex facial signaling repertoire may be important for managing a larger number of social relationships that vary in type, duration, and associated social behaviors (due to fission-fusion dynamics). In gibbons, facial synchrony is important for coordinating activities needed to establish and maintain long-term pair bonds (Florkiewicz et al., 2018). One example of synchronous activity in gibbons includes vocal duetting, where songs strengthen social bonds and advertise territorial boundaries (Geissmann, 1993). Synchronous facial signaling was also observed in one of the two gibbon datasets incorporated into the current project (for more information, see Florkiewicz et al., 2018). However, this does not necessarily suggest that chimpanzees do not exhibit facial synchrony, or that gibbons are unable to maintain multiple relationships with others. Facial synchrony (in the form of rapid and delayed facial mimicry) during bouts of play in chimpanzees is important for modulating play sessions and communicating playful motivations (Palagi et al., 2019). In gibbons, extra-pair copulations suggest that social relationships can be established and maintained outside of the pair bond (Barelli et al., 2013).

Study Limitations and Future Directions

65

Our results support socio-ecology as playing an important role in the evolution of primate facial signaling repertoires, but there are four limitations to our study that we have identified. First, we mainly focused on AUs which were produced by both chimpanzees and gibbons. We took a conservative approach to the study of these individual AUs, since it was unclear if differences between apes (in the production of certain AUs) were attributed to differences in usage or methodological constraints between the three studies. Second, we only considered one of five variables associated with the social complexity hypothesis (Freeberg et al., 2012). In this study, we compared the facial signaling repertoires of two apes who exhibit dramatic differences in their average group size. It is possible that variables other than group size (such as bond quality) play an important role in the evolution of facial signaling repertoires. Third, we did not examine the relationship between ecological variables (such as forest density, food availability, home range size, etc.) and complexity in facial signaling repertoires. To fully test the socio-ecological complexity hypothesis, both social and ecological variables should be considered.

Finally, our study did not focus on primate species outside of superfamily *Hominoidea*. Additional work on other primate species (such as prosimians and monkeys) would be useful for testing whether the relationship between socio-ecological complexity and communicative complexity in facial signaling is widespread both within and across taxonomic groups. For example: there are currently FACS established for multiple macaque species, including rhesus macaques (*Macaca mulatta;* Parr et al., 2010), Barbary macaques (*Macaca sylvanus;* Julle-Danière et al., 2015), crested macaques (*Macaca nigra;* Clark et al., 2020), and Japanese macaques (*Macaca fuscata;* Correia-Caeiro et al., 2021). Because macaques live in large multi-male, multi-female groups, our socio-ecological complexity hypothesis predicts that these species will have large and diverse facial signaling repertoires, similar to that of chimpanzees.

However, it is plausible that limited facial mobility would result in macaques having a smaller and less diverse facial signaling repertoire, similar to that of gibbons; this would provide evidence for our facial mobility hypothesis. A cross-family comparison could test macaques, chimpanzees, and gibbons for the relationship between facial mobility, socio-ecological complexity, and facial signaling diversity/complexity. Macaques provide an opportunity to test these hypotheses within a single genus, as macaque species vary in both socio-ecology (Thierry, 2007) and facial mobility (Correia-Caeiro et al., 2021). Macaques may be a powerful test of whether socio-ecological complexity or facial mobility better explains patterns of facial signaling as the genus, is large, diverse, and yet closely related.

The creation of FACS for species other than humans, chimpanzees, orangutans, gibbons, and macaques would also provide opportunities to test the relationship between socio-ecological complexity and communicative complexity. For example: the creation of FACS for Callitrichids (who are monogamous) and Atelids (who exhibit fission-fusion dynamics) would provide opportunities to test the relationship between social group size, facial repertoire size, and facial signaling complexity.

CONCLUSION

We compared the facial signaling behavior of chimpanzees (*Pan troglodytes*) and gibbons (family Hylobatidae) to test two hypotheses regarding the evolution of facial signaling repertoires: (1) the facial mobility hypothesis; and (2) the socio-ecological complexity hypothesis. We found that chimpanzees exhibit larger and more complex facial signaling repertoires than gibbons, which provides support for the socio-ecological complexity hypothesis and evidence against the facial mobility hypothesis. While there appears to be differences in facial mobility across apes, this does not necessarily limit the number of morphologically distinct

facial signals each ape is capable of producing. Having a larger and more complex facial signaling repertoire may be useful for managing numerous social relationships embedded within multi-level social networks (such as those observed in fission-fusion groups), whereas smaller facial signaling repertoires may be useful for the management of long-term pair bonds (in the form of facial synchrony). Our results suggest that multiple socio-ecological variables act on facial signals to possibly expand or contract repertoire size as befitting the needs of the organisms.

REFERENCES

Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, *49*(3/4), 227-267. http://www.jstor.org/stable/4533591

Andrew, R. J. (1963). The origin and evolution of the calls and facial expressions of the primates. *Behaviour, Leiden*, *20*(1-2), 1-109. https://doi.org/10.1163/156853963X00220

Aureli, F., Schaffner, Colleen M., Boesch, C., Bearder, Simon K., Call, J., Chapman, Colin A., Connor, R., Fiore, Anthony D., Dunbar, Robin I. M., Henzi, S. P., Holekamp, K., Korstjens, Amanda H., Layton, R., Lee, P., Lehmann, J., Manson, Joseph H., Ramos-Fernandez, G., Strier, Karen B., & Schaik, Carel P. v. (2008). Fission-Fusion Dynamics: New Research Frameworks. *Current Anthropology*, *49*(4), 627-654. https://doi.org/10.1086/586708

Barelli, C., Matsudaira, K., Wolf, T., Roos, C., Heistermann, M., Hodges, K., Ishida, T., Malaivijitnond, S., & Reichard, U. H. (2013). Extra-pair paternity confirmed in wild white-handed gibbons. *Am J Primatol*, *75*(12), 1185-1195. https://doi.org/10.1002/ajp.22180

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï National

    Park [https://doi.org/10.1002/ajpa.1330780410]. *American Journal of Physical*

    Anthropology, *78*(4), 547-573. https://doi.org/https://doi.org/10.1002/ajpa.1330780410

Bolker, B. (2021). *Package 'bbmle'*. Retrieved September 9 from

    https://github.com/bbolker/bbmle

Burrows, A. M., Li, L., Waller, B. M., & Micheletta, J. (2016). Social variables exert selective

    pressures in the evolution and form of primate mimetic musculature

    [https://doi.org/10.1111/joa.12440]. *Journal of Anatomy*, *228*(4), 595-607.

    https://doi.org/https://doi.org/10.1111/joa.12440

Buttelmann, D., Call, J., & Tomasello, M. (2009). Do great apes use emotional expressions to

    infer desires? [https://doi.org/10.1111/j.1467-7687.2008.00802.x]. *Developmental*

    Science, *12*(5), 688-698. https://doi.org/https://doi.org/10.1111/j.1467-

    7687.2008.00802.x

Caeiro, C., Waller, B., & Burrows, A. (2013). The Cat Facial Action Coding System manual

    (CatFACS).

Caeiro, C. C., Waller, B. M., Zimmermann, E., Burrows, A. M., & Davila-Ross, M. (2013).

    OrangFACS: A Muscle-Based Facial Movement Coding System for Orangutans (Pongo

    spp.). *International Journal of Primatology*, *34*(1), 115-129.

    https://doi.org/10.1007/s10764-012-9652-x

Camerlink, I., Coulange, E., Farish, M., Baxter, E. M., & Turner, S. P. (2018). Facial expression

    as a potential measure of both intent and emotion. *Scientific Reports*, *8*(1), 17602.

    https://doi.org/10.1038/s41598-018-35905-3

Campbell, C. J., Fuentes, A., MacKinnon, K. C., Bearder, S. K., & Stumpf, R. (2010). *Primates in Perspective* (2 ed.). Oxford University Press.

Christensen, R. H. B. (2019). *Package 'ordinal'*. Retrieved September 14 from https://github.com/runehaubo/ordinal

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) [https://doi.org/10.1002/ajpa.24129]. *American Journal of Physical Anthropology*, *173*(3), 411-422. https://doi.org/https://doi.org/10.1002/ajpa.24129

Clutton-Brock, T. (2009). Structure and function in mammalian societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3229-3242. https://doi.org/10.1098/rstb.2009.0120

Correia-Caeiro, C., Holmes, K., & Miyabe-Nishiwaki, T. (2021). 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* [doi:10.1037/10001-000]. John Murray. https://doi.org/10.1037/10001-000

Dobson, S. D. (2009). Socioecological correlates of facial mobility in nonhuman anthropoids [https://doi.org/10.1002/ajpa.21007]. *American Journal of Physical Anthropology*, *139*(3), 413-420. https://doi.org/https://doi.org/10.1002/ajpa.21007

Ekman, P., & Rosenberg, E. L. (2005). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179644.001.0001

Florkiewicz, B., Skollar, G., & Reichard, U. H. (2018). Facial expressions and pair bonds in hylobatids [https://doi.org/10.1002/ajpa.23608]. *American Journal of Physical Anthropology*, *167*(1), 108-123. https://doi.org/https://doi.org/10.1002/ajpa.23608

Florkiewicz, B. N., & Campbell, M. W. (2021). A Comparison of Focal and Opportunistic Sampling Methods when Studying Chimpanzee Facial and Gestural Communication. *Folia Primatologica*, *92*(3), 164-174. https://doi.org/10.1159/000516315

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, *27*(2), 233-249. https://doi.org/10.1080/10556788.2011.597854

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *367*(1597), 1785-1801. https://doi.org/10.1098/rstb.2011.0213

Geissmann, T. (1993). *Evolution of Communication in Gibbons (Hylobatidae)* [Ph.D. Dissertation, Zürich University]. Switzerland. http://www.gibbons.de/main/theses/pdf_files/1993geissmann.pdf

Geissmann, T., Rosenkranz-Weck, S., J.G.M. Van Der Loo, J., & Orgeldinger, M. (2020). Taxon-Specific Pair Bonding in Gibbons (Hylobatidae). In *Updates on Veterinary Anatomy and Physiology [Working Title]*. IntechOpen.

https://www.intechopen.com/online-first/taxon-specific-pair-bonding-in-gibbons-hylobatidae

Gittins, S. P. (1980). Territorial behavior in the Agile Gibbon. *International Journal of Primatology*, *1*(4), 381-399. https://doi.org/10.1007/BF02692281

Hopkins, W. D., Taglialatela, J., & Leavens, D. A. (2007). Chimpanzees Differentially Produce Novel Vocalizations to Capture the Attention of a Human. *Animal behaviour*, *73*(2), 281-286. https://doi.org/10.1016/j.anbehav.2006.08.004

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-e1248. https://doi.org/10.7717/peerj.1248

Lewinski, P., Uyl, T., & Butler, C. (2014). Automated Facial Coding: Validation of Basic Emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, *7*, 227-236. https://doi.org/10.1037/npe0000028

Liebal, K., Pika, S., & Tomasello, M. (2004). Social communication in siamangs (Symphalangus syndactylus): use of gestures and facial expressions. *Primates*, *45*(1), 41-57. https://doi.org/10.1007/s10329-003-0063-7

Liebal, K., Waller, B., Burrows, A., & Slocombe, K. (2013). *Primate Communication: A Multimodal Approach*. https://doi.org/10.1017/CBO9781139018111

Lonsdorf, E. V. (2006). What is the role of mothers in the acquisition of termite-fishing behaviors in wild chimpanzees (Pan troglodytes schweinfurthii)? *Animal Cognition*, *9*(1), 36-46. https://doi.org/10.1007/s10071-005-0002-7

Matthews, J. K. (2021). Ecological and reproductive drivers of fission-fusion dynamics in

    chimpanzees (Pan troglodytes schweinfurthii) inhabiting a montane forest. *Behavioral*

    Ecology and Sociobiology, *v. 75*(no. 1), pp. 23-23-2021 v.2075 no.2021.

    https://doi.org/10.1007/s00265-020-02964-4

Morimoto, Y., & Fujita, K. (2012). Capuchin monkeys (Cebus apella) use conspecifics'

    emotional expressions to evaluate emotional valence of objects. *Animal Cognition*, *15*(3),

    341-347. https://doi.org/10.1007/s10071-011-0458-6

Palagi, E., Norscia, I., Pressi, S., & Cordoni, G. (2019). *Facial mimicry and play: A comparative*

    study in chimpanzees and gorillas [doi:10.1037/emo0000476].

Parr, L. A., Cohen, M., & Waal, F. d. (2005). Influence of Social Context on the Use of Blended

    and Graded Facial Displays in Chimpanzees. *International Journal of Primatology*,

    26(1), 73-103. https://doi.org/10.1007/s10764-005-0724-z

Parr, L. A., Waller, B. M., Burrows, A. M., Gothard, K. M., & Vick, S. J. (2010). Brief

    communication: MaqFACS: A muscle-based facial movement coding system for the

    rhesus macaque [https://doi.org/10.1002/ajpa.21401]. *American Journal of Physical*

    Anthropology, *143*(4), 625-630. https://doi.org/https://doi.org/10.1002/ajpa.21401

Parr, L. A., Waller, B. M., & Heintz, M. (2008). Facial expression categorization by

    chimpanzees using standardized stimuli. *Emotion (Washington, D.C.)*, *8*(2), 216-231.

    https://doi.org/10.1037/1528-3542.8.2.216

Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial

    expressions using muscle action. *Emotion (Washington, D.C.)*, *7*(1), 172-181.

    https://doi.org/10.1037/1528-3542.7.1.172

Preininger, D., Boeckle, M., Sztatecsny, M., & Hödl, W. (2013). Divergent Receiver Responses to Components of Multimodal Signals in Two Foot-Flagging Frog Species. *PLOS ONE*, *8*(1), e55367. https://doi.org/10.1371/journal.pone.0055367

Preuschoft, S. (2000). Primate Faces and Facial Expressions. *Social Research*, *67*(1), 245-271. http://www.jstor.org/stable/40971384

Preuschoft, S., & van Hooff, J. A. R. A. M. (1997). *The social function of "smile" and "laughter": Variations across primate species and societies*Lawrence Erlbaum Associates, Inc.

Prox, L., & Farine, D. (2020). A framework for conceptualizing dimensions of social organization in mammals. *Ecology and Evolution*, *10*(2), 791-807. https://doi.org/10.1002/ece3.5936

Roberts, A. I., & Roberts, S. G. B. (2019). Persistence in gestural communication predicts sociality in wild chimpanzees. *Animal Cognition*, *22*(5), 605-618. https://doi.org/10.1007/s10071-018-1219-6

Ryne, A. P. (1996). Pair Bonds in Monogamous Apes: A Comparison of the Siamang Hylobates syndactylus and the White-Handed Gibbon Hylobates lar. *Behaviour*, *133*(5/6), 321-356. http://www.jstor.org/stable/4535362

Ryne, P. (1994). Dynamic Pair Bonds in Hylobatids: Implications regarding Monogamous Social Systems. *Behaviour*, *128*(1/2), 65-101. http://www.jstor.org/stable/4535165

Scheider, L., Liebal, K., Oña, L., Burrows, A., & Waller, B. (2014). A comparison of facial expression properties in five hylobatid species [https://doi.org/10.1002/ajp.22255]. *American Journal of Primatology*, *76*(7), 618-628. https://doi.org/https://doi.org/10.1002/ajp.22255

Scheider, L., Waller, B. M., Oña, L., Burrows, A. M., & Liebal, K. (2016). Social Use of Facial

Expressions in Hylobatids. *PLOS ONE*, *11*(3), e0151733.

https://doi.org/10.1371/journal.pone.0151733

Sherwood, C. C. (2005). Comparative anatomy of the facial motor nucleus in mammals, with an

analysis of neuron numbers in primates [https://doi.org/10.1002/ar.a.20259]. *The*

Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary

Biology, *287A*(1), 1067-1079. https://doi.org/https://doi.org/10.1002/ar.a.20259

Silk, J. (2002). Grunts, girneys, and good intentions: The origins of strategic commitment in

nonhuman primates. *Commitment: Evolutionary Perspectives*, 138-157.

Smith, M. J., & Harper, D. G. C. (1995). Animal Signals: Models and Terminology. *Journal of*

Theoretical Biology, *177*(3), 305-311.

https://doi.org/https://doi.org/10.1006/jtbi.1995.0248

Southwick, C. H. (1985). THE LESSER APES: EVOLUTIONARY AND BEHAVIOURAL

BIOLOGY [https://doi.org/10.1111/j.1558-5646.1985.tb05708.x]. *Evolution*, *39*(6),

1399-1400. https://doi.org/https://doi.org/10.1111/j.1558-5646.1985.tb05708.x

Thierry, B. (2007). Unity in diversity: Lessons from macaque societies

[https://doi.org/10.1002/evan.20147]. *Evolutionary Anthropology: Issues, News, and*

Reviews, *16*(6), 224-238. https://doi.org/https://doi.org/10.1002/evan.20147

Van Hooff, J. A. R. A. M. (1967). *The Facial Displays of the Catarrhine Monkeys and Apes*

Vick, S.-J., Waller, B. M., Parr, L. A., Smith Pasqualini, M. C., & Bard, K. A. (2007). A Cross-

species Comparison of Facial Morphology and Movement in Humans and Chimpanzees

Using the Facial Action Coding System (FACS). *Journal of nonverbal behavior*, *31*(1),

1-20. https://doi.org/10.1007/s10919-006-0017-z

Waller, B. M., & Dunbar, R. I. M. (2005). Differential Behavioural Effects of Silent Bared Teeth

Display and Relaxed Open Mouth Display in Chimpanzees (Pan troglodytes). *Ethology*,

111(2), 129-142. https://doi.org/https://doi.org/10.1111/j.1439-0310.2004.01045.x

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/https://doi.org/10.1016/j.neubiorev.2020.02.031

Waller, B. M., Lembeck, M., Kuchenbuch, P., Burrows, A. M., & Liebal, K. (2012).

GibbonFACS: A Muscle-Based Facial Movement Coding System for Hylobatids.

International Journal of Primatology, *33*(4), 809-821. https://doi.org/10.1007/s10764-

012-9611-6

Waller, B. M., Peirce, K., Caeiro, C. C., Scheider, L., Burrows, A. M., McCune, S., & Kaminski,

J. (2013). Paedomorphic Facial Expressions Give Dogs a Selective Advantage. *PLOS

ONE*, *8*(12), e82686. https://doi.org/10.1371/journal.pone.0082686

Waller, B. M., Warmelink, L., Liebal, K., Micheletta, J., & Slocombe, K. E. (2013).

Pseudoreplication: a widespread problem in primate communication research. *Animal

behaviour*, *86*(2), 483-488. https://doi.org/https://doi.org/10.1016/j.anbehav.2013.05.038

Waller, B. M., Whitehouse, J., & Micheletta, J. (2016). Macaques can predict social outcomes

from facial expressions. *Animal Cognition*, *19*(5), 1031-1036.

https://doi.org/10.1007/s10071-016-0992-3

Waller, B. M., Whitehouse, J., & Micheletta, J. (2017). Rethinking primate facial expression: A

predictive framework. *Neuroscience & Biobehavioral Reviews*, *82*, 13-21.

https://doi.org/https://doi.org/10.1016/j.neubiorev.2016.09.005

Wathan, J., Burrows, A. M., Waller, B. M., & McComb, K. (2015). EquiFACS: The Equine

    Facial Action Coding System. *PLOS ONE*, *10*(8), e0131738.

    https://doi.org/10.1371/journal.pone.0131738

Whiten, A. B., Richard W. (1997). *Machiavellian intelligence II: Extensions and evaluations*

    [doi:10.1017/CBO9780511525636]. Cambridge University Press.

    https://doi.org/10.1017/CBO9780511525636

APPENDIX

TABLE S1. A list of all individuals included in the current study. Demographic information, such as species, birth date, age group, sex, and location are provided for each individual. Individuals are grouped into 4 different age categories: infants (0-4 years), juvenile (5-8 years), subadult (9-15 years), and adult (16+ years). Age groups were calculated using birth dates and date of study start.

Study	Species	Name	Birth Date (MM/DD/YY)	Sex	Location	Date of Study Start	Age Group*
Current Study	Pan troglodytes	Ben	7/31/2002	Male	Los Angeles Zoo	6/1/2017	Subadult
Current Study	Pan troglodytes	Glenn	4/21/1994	Male	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Gracie	1/26/1987	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Jake	6/14/1999	Male	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Jean	5/6/1999	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Jerrard	2/20/1990	Male	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Johari	10/27/2014	Male	Los Angeles Zoo	6/1/2017	Infant
Current Study	Pan troglodytes	Julie	3/28/1986	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Kima	4/22/2013	Female	Los Angeles Zoo	6/1/2017	Infant
Current Study	Pan troglodytes	Nan	12/12/1979	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Oliver	8/23/2014	Male	Los Angeles Zoo	6/1/2017	Infant
Current Study	Pan troglodytes	Pandora	3/5/1967	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Regina	10/23/1983	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Shaun	7/8/1988	Male	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Uki	3/8/2013	Female	Los Angeles Zoo	6/1/2017	Infant
Current Study	Pan troglodytes	Yoshiko	7/2/1990	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Zoe	8/9/1999	Female	Los Angeles Zoo	6/1/2017	Adult
Current Study	Pan troglodytes	Zuri	7/23/2012	Female	Los Angeles Zoo	6/1/2017	Infant
Florkiewicz et al. 2018	Hoolock leuconedys	Arthur	1/1/1996	Male	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	Betty	1/1/1999	Female	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	Chan Thar	1/1/2006	Female	Gibbon Conservation Center	6/1/2015	Subadult

Florkiewicz et al. 2018	Hylobates moloch	Chloe	2/24/1990	Female	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hylobates pileatus	Domino	5/19/1995	Male	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	Hmawe Ni	1/1/2004	Female	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Hylobates moloch	Ivan	1/1/1974	Male	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	Khin Maung Win	10/19/2007	Male	Gibbon Conservation Center	6/1/2015	Juvenile
Florkiewicz et al. 2018	Hylobates moloch	Khusus	1/11/1995	Female	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Nomascus leucogenys	ParkerPerak	6/15/2001	Female	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Hylobates moloch	Perak	11/16/2001	Male	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Hoolock leuconedys	Phy Gi	1/1/2003	Female	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Nomascus leucogenys	Pierre	2/13/2004	Male	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Nomascus leucogenys	Ricky	8/24/1985	Female	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hylobates moloch	Shelby	5/18/1983	Male	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hylobates moloch	Simpang	5/23/2000	Female	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Hylobates pileatus	Tuk	6/23/1993	Female	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	U Maung Maung	1/1/2001	Male	Gibbon Conservation Center	6/1/2015	Subadult
Florkiewicz et al. 2018	Nomascus leucogenys	Vok	4/29/1983	Male	Gibbon Conservation Center	6/1/2015	Adult
Florkiewicz et al. 2018	Hoolock leuconedys	Win Bo	1/1/2003	Male	Gibbon Conservation Center	6/1/2015	Subadult
Scheider et al. 2014	Hylobates lar	Bert	5/1/1982	Male	Rheine Zoo	3/1/2009	Adult
Scheider et al. 2014	Nomascus gabriellae	Chloe	1/6/1990	Female	Mulhouse Zoo	3/1/2009	Adult
Scheider et al. 2014	Nomascus gabriellae	Dan	1/1/1991	Male	Mulhouse Zoo	3/1/2009	Adult
Scheider et al. 2014	Symphalangus syndactylus	Daniel	5/26/1996	Male	Twycross Zoo	3/1/2009	Subadult
Scheider et al. 2014	Nomascus siki	Dorian	12/23/1989	Male	Mulhouse Zoo	3/1/2009	Adult
Scheider et al. 2014	Nomascus siki	Fanny	6/13/1993	Female	Mulhouse Zoo	3/1/2009	Subadult
Scheider et al. 2014	Hylobates pileatus	Iaman	1/1/1959	Male	Zurich Zoo	3/1/2009	Adult
Scheider et al. 2014	Hylobates pileatus	Iba	1/1/1974	Female	Zurich Zoo	3/1/2009	Adult
Scheider et al. 2014	Symphalangus syndactylus	Kane	11/2/1990	Male	Twycross Zoo	3/1/2009	Adult

Scheider et al. 2014	Hylobates pileatus	Khmer	11/28/1984	Male	Zurich Zoo	3/1/2009	Adult
Scheider et al. 2014	Hylobates lar	Lissy	1/1/1981	Female	Rheine Zoo	3/1/2009	Adult
Scheider et al. 2014	Symphalangus syndactylus	Sheena	1/30/1991	Female	Twycross Zoo	3/1/2009	Adult
Scheider et al. 2014	Symphalangus syndactylus	Spike	11/25/2000	Male	Twycross Zoo	3/1/2009	Juvenile
Scheider et al. 2014	Symphalangus syndactylus	Tango	3/27/1994	Female	Twycross Zoo	3/1/2009	Subadult
Scheider et al. 2014	Symphalangus syndactylus	Tarragona	11/18/2000	Female	Twycross Zoo	3/1/2009	Juvenile
Scheider et al. 2014	Hylobates pileatus	Willow	5/6/1987	Female	Zurich Zoo	3/1/2009	Adult

TABLE S2. A list of all action units (AUs) that were observed in the current study (before filtering the data and after filtering the data).

AU Code	AU Name	Chimpanzees?	Gibbons?	Before Data Filtering	After Data Filtering
1+2	*Inner + outer brow raiser*	✗	✓	✓	✗
5	*Upper lid raiser*	✗	✓	✓	✗
6	*Cheek raiser*	✓	✓	✓	✓
7	*Lids tight*	✗	✓	✓	✗
8	*Lips towards each other*	✗	✓	✓	✗
9	*Nose wrinkler*	✓	✓	✓	✓
10	*Upper lip raiser*	✓	✓	✓	✓
12	*Lip corner puller*	✓	✓	✓	✓
16	*Lower lip depressor*	✓	✓	✓	✓
17	*Chin raiser*	✓	✓	✓	✓
18	*Lip pucker*	✗	✓	✓	✗
19	*Tongue out*	✓	✓	✓	✓
22	*Lip funneler*	✓	✓	✓	✓
24	*Lip presser*	✓	✓	✓	✓
25	*Lips part*	✓	✓	✓	✓
26	*Jaw drop*	✓	✓	✓	✓
27	*Mouth stretch*	✓	✓	✓	✓
28	*Lip suck*	✗	✓	✓	✗
33	*Cheek blow*	✓	✗	✓	✗
34	*Puff*	✓	✗	✓	✗
37	*Lip wipe*	✗	✓	✓	✗
41	*Glabella lowerer*	✗	✓	✓	✗
43	*Eye closure*	✓	✗	✓	✗
50	*Vocalization*	✓	✗	✓	✗
51	*Head turn left*	✓	✗	✓	✗
52	*Head turn right*	✓	✗	✓	✗
53	*Head up*	✓	✗	✓	✗
54	*Head down*	✓	✗	✓	✗
55	*Head tilt left*	✓	✗	✓	✗
69	*Look towards/fixate gaze*	✓	✗	✓	✗
84	*Head shake (back and forth)*	✓	✗	✓	✗
85	*Head shake (up and down)*	✓	✗	✓	✗
100	*With gesture*	✓	✗	✓	✗

101	*With positional behavior*	✓	✗	✓	✗
160	*Lower lip relaxer*	✓	✗	✓	✗
500	*Throat sac inflation*	✗	✓	✓	✗
EYE	*Eye blink or eye closure*	✗	✓	✓	✗
TOTAL		*27*	*22*	*37*	*12*

TABLE S3. A list of all models that we ran in the present study. Models are organized into sets based on outcome variable, which includes: (M1) AU observations; (M2) AU combination observations; and (M3) AU length observations. All models listed contained a random variable (1|SignalerID) and an offset (which was species recording time in minutes, or offRT). AIC scores and ΔAIC scores are presented with each model. Models with an ΔAIC value of ≤2 are italicized, and the best fitting model is bolded.

Model Code	Model	AIC	ΔAIC
M1.1	ApeType + (1|Signaler ID) + offset(offRT)	4409.56	622.7
M1.2	ApeType*AU + (1|Signaler ID) + offset(offRT)	3789.52	2.7
M1.3	***ApeType*AU + Age + (1|Signaler ID) + offset(offRT)***	***3786.90***	***0.0***
M1.4	ApeType*AU + Sex + (1|Signaler ID) + offset(offRT)	3789.44	2.6
M1.5	*ApeType*AU + Sex + Age + (1|Signaler ID) + offset(offRT)*	*3786.86*	*0.0*
M2.1	***ApeType + (1|Signaler ID) + offset(offRT)***	***5665.04***	***0.0***
M2.2	ApeType + Age + (1|Signaler ID) + offset(offRT)	5671.58	6.5
M2.3	*ApeType + Sex + (1|Signaler ID) + offset(offRT)*	*5666.40*	*1.4*
M2.4	ApeType + Sex + Age + (1|Signaler ID) + offset(offRT)	5667.24	2.2
M3.1	***ApeType + (1|Signaler ID) + offset(offRT)***	***11461.02***	***0.0***
M3.2	*ApeType + Age + (1|Signaler ID) + offset(offRT)*	*11462.54*	*1.5*
M3.3	*ApeType + Sex + (1|Signaler ID) + offset(offRT)*	*11462.21*	*1.2*
M3.4	ApeType + Sex + Age + (1|Signaler ID) + offset(offRT)	11463.91	2.9