

1 **Higher social tolerance is associated with more complex facial** 2 **behavior in macaques**

3 **Alan V. Rincon^{1*}, Bridget M. Waller², Julie Duboscq³, Alexander Mielke⁴, Claire Pérez¹,**
4 **Peter R. Clark^{1,5}, Jérôme Micheletta^{1*}**

5 ¹Department of Psychology, Centre for Comparative and Evolutionary Psychology, University of
6 Portsmouth, Portsmouth PO1 2DY, UK

7 ²Centre for Interdisciplinary Research on Social Interaction, Department of Psychology, Notting-
8 ham Trent University, Nottingham NG1 4BU, UK

9 ³UMR7206 Eco-Anthropology, CNRS-MNHN-Université Paris Cité, Paris 75016, France

10 ⁴Origins of Mind Group, School of Psychology and Neuroscience, University of St Andrews, St
11 Andrews KY16 9JP, UK

12 ⁵School of Psychology, University of Lincoln, Lincoln LN6 7TS, UK

13 *Corresponding authors

14 Alan V. Rincon: avrincon1@gmail.com

15 Jérôme Micheletta: jerome.micheletta@port.ac.uk

16 **Keywords:** Communicative complexity, social complexity, animal communication, Facial Action
17 Coding System, macaque

18 **Abstract**

19 The social complexity hypothesis for communicative complexity posits that animal societies with
20 more complex social systems require more complex communication systems. We tested the so-
21 cial complexity hypothesis on three macaque species that vary in their degree of social tolerance
22 and complexity. We coded facial behavior in >3000 social interactions across three social contexts
23 (aggressive, submissive, affiliative) in 389 animals, using the Facial Action Coding System for
24 macaques (MaqFACS). We quantified communicative complexity using three measures of uncer-
25 tainty: entropy, specificity, and prediction error. We found that the relative entropy of facial be-
26 havior was higher for the more tolerant crested macaques as compared to the less tolerant Barbary
27 and rhesus macaques across all social contexts, indicating that crested macaques more frequently
28 use a higher diversity of facial behavior. The context specificity of facial behavior was higher
29 in rhesus as compared to Barbary and crested macaques, demonstrating that Barbary and crested
30 macaques used facial behavior more flexibly across different social contexts. Finally, a random
31 forest classifier predicted social context from facial behavior with highest accuracy for rhesus and
32 lowest for crested, indicating there is higher uncertainty and complexity in the facial behavior of
33 crested macaques. Overall, our results support the social complexity hypothesis.

34 **Introduction**

35 Animals must overcome a range of environmental and ecological challenges to survive and re-
36 produce, with group-living species having to overcome additional social challenges to maximize
37 fitness. Communicative signals can be used to navigate a number of different social situations
38 and may need to become more elaborate as social complexity increases. The social complexity
39 hypothesis for communicative complexity encapsulates this idea, proposing that animal societies
40 with more complex social systems require more complex communication systems [1].

41 The social complexity hypothesis has become a topical issue in recent years, with questions regard-
42 ing the definitions, measurement, and selective pressures driving both social and communicative
43 complexity [2,3]. Social complexity as experienced by group members can be affected by the level
44 of differentiation of social relationships, where complexity increases as social relationships become
45 more differentiated [4,5]. In a socially complex society individuals interact frequently with each
46 other in diverse ways and in many different contexts [1]. If the types of interactions that individuals
47 have is constrained, for example, by dominance or kinship, then social complexity decreases [1].
48 Social complexity is also affected by the predictability or consistency of social interactions [5,6].
49 When the behavior of social partners is unpredictable, such as when the dominance hierarchy is un-
50 stable, individuals likely perceive the social environment as more complex [6]. These operational
51 definitions of social complexity are valuable to advance the study of social complexity but are not
52 easy to quantify with a single measure [7].

53 Similarly, communicative complexity is also difficult to quantify. Many studies have used the
54 number of signalling units as a measure of communicative complexity [2]. While a useful mea-
55 sure, it is not always apparent what a signaling unit is. For example, calls are sometimes graded on
56 a continuous scale without a clear separation between different call types [8]. Fewer studies have
57 investigated the complexity of non-vocal communication [1,2], but similar issues exist. One previ-
58 ous study quantified the repertoire of facial behavior in macaques by the number of discrete facial
59 expressions that a species displays and found that it was positively correlated with conciliatory

60 tendency and counter-aggression across species [9]. However, classifying facial expressions into
61 discrete categories (e.g., bared-teeth display) does not capture the full range of expressiveness and
62 meanings that the face can convey. For example, subtle morphological variations in bared-teeth
63 displays are associated with different outcomes of social interactions (e.g., affiliation versus sub-
64 mission) in crested macaques (*Macaca nigra*) [10]. A better approach is to quantify facial behavior
65 at the level of individual facial muscle movements [11], which can be done using the Facial Action
66 Coding System (FACS) [12]. In FACS, visible muscle contractions in the face are called Action
67 Units and allow for a detailed and objective description of facial behavior [11,12]. Indeed, facial
68 mobility, as defined by the number of Action Units that a species has, is positively correlated with
69 group size across non-human primates [13]. However, isolated muscle movements still do not ac-
70 count for the full diversity of facial behavior because facial muscles often contract simultaneously
71 to produce a large variety of distinct facial expressions.

72 One promising avenue to approximate complexity in living organisms is to quantify the uncertainty
73 or predictability of a system [14,15], which are general properties of complex systems [16,17].
74 Shannon's information entropy [18] is a measure of uncertainty that can be applied to animal com-
75 munication. Conceptually, entropy measures the potential amount of information that a communi-
76 cation system holds, rather than what is actually communicated [18,19]. Entropy increases along
77 two dimensions: (i) with increasing diversity of signals, and (ii) as the relative frequency of signal
78 use becomes more balanced. For example, a system with three calls can hold more information
79 than a system with one call and thus would have higher entropy. Likewise, a system with three
80 calls used with equal frequency will have a higher entropy than another system that expresses one
81 call more frequently than the two others. Uncertainty increases with entropy because each com-
82 municative event has the potential to derive from a greater number of units. The relative entropy,
83 or uncertainty, of different systems can be compared by calculating the ratio between the observed
84 and maximum entropy of each system.

85 The predictability and uncertainty of a communication system is also affected by how flexibly sig-
86 nals are used across different social contexts [5]. For instance, if signal A is always used in an

87 aggressive context and signal B is always used in an affiliative context, then it is easy to predict
88 the context from the signal. Conversely, if signals A and B are used in both contexts, then pre-
89 dictability is lower, and complexity is higher. Extremely rare signals do not substantially affect
90 the predictability of a system regardless of whether they have high or low specificity since they are
91 seldom observed in the majority of social interactions. Therefore, predictability is highest when
92 signals are both highly context-specific and occur in that context often. Additionally, predictabil-
93 ity can be measured directly by training a machine learning classifier to predict the social context
94 that a given signal was used in. Differences in prediction error would approximate the relative
95 uncertainty and complexity, with accuracy being lower in more complex systems. However, as
96 complexity lies somewhere between order and randomness [15,19], we should still be able to pre-
97 dict the social contexts better than chance, even in a complex system.

98 Studying closely related species offers a robust means of testing the social complexity hypothesis
99 due to their homologous communication systems. For this reason, macaques (genus *Macaca*) are
100 excellent taxa to test the social complexity hypothesis. All species have a similar social organiza-
101 tion consisting of multi-male, multi-female groups, but vary in social style in ways that are highly
102 relevant to predictions of the social complexity hypothesis. The social styles of macaques consist
103 of several covarying traits that can be ordered along a social tolerance scale ranging from the least
104 (grade 1) to most tolerant (grade 4) [20,21]. Social interactions for the least tolerant species, such
105 as rhesus (*M. mulatta*) and Japanese (*M. fuscata*) macaques, are generally more constrained by a
106 steep linear dominance hierarchy [22] and kinship [23–25]. Additionally, severe agonistic inter-
107 actions are more frequent [25], instances of counter-aggression and reconciliation after conflicts
108 are rare [22,25], and formal signals of submission are commonly used [26,27]. Combined, these
109 behavioral traits indicate that agonistic interactions of the least tolerant species are more stereo-
110 typed and formalized. Thus, the outcome of such interactions is more certain, whereas the opposite
111 is true for the most tolerant species, such as crested and Tonkean (*M. tonkeana*) macaques. The
112 unpredictability in the outcome of agonistic interactions of tolerant macaques likely results in a
113 social environment that is perceived as more complex by individuals [6], where more subtle means

114 of negotiation during conflicts may be necessary.

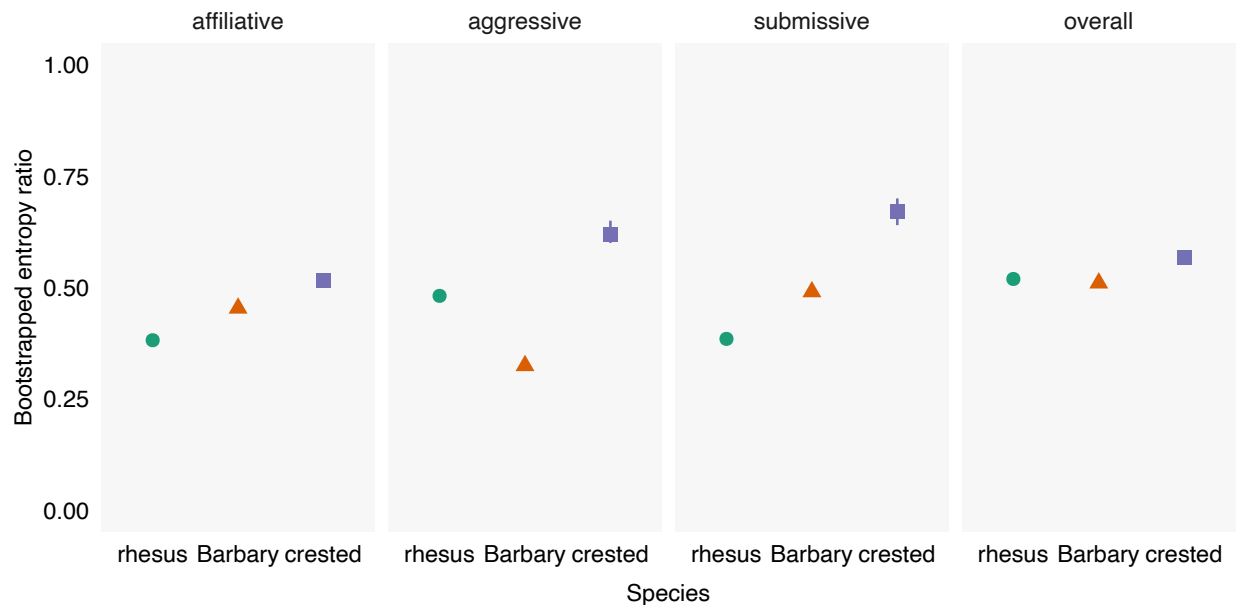
115 In this study we compared the facial behavior of three macaque species that vary in their degree
116 of social tolerance and, therefore, social complexity: rhesus (least tolerant), Barbary (*M. sylvanus*,
117 mid-tolerant), and crested macaques (most tolerant). For macaques (and primates in general), the
118 face is central to communication and is a key tool in allowing individuals to achieve their social
119 goals by communicating motivations, emotions and/or intentions [28,29]. We coded facial behav-
120 ior at the level of individual visible muscle movements using FACS and recorded all observed
121 unique combinations, rather than classifying facial expressions into discrete categories. Based on
122 the social complexity hypothesis [1], we expected that tolerant species would have higher com-
123 municative complexity, given that their social relationships are less constrained by dominance and
124 have higher overall uncertainty in the outcome of agonistic interactions. Specifically, we predicted
125 the following: (1) relative entropy of facial behavior will be lowest in the rhesus and highest in
126 crested macaques, (2) context specificity of facial behavior will be highest in rhesus and lowest
127 in crested macaques, and (3) social context can be predicted from facial behavior most accurately
128 in rhesus and least accurately in crested macaques. For all three metrics, we expected Barbary
129 macaques to lie somewhere in between the rhesus and crested macaques.

130 **Results**

131 **Entropy of facial behavior**

132 To compare the relative uncertainty in the facial behavior of macaques, we defined facial behavior
133 by the unique combination of Action Units (facial muscle movements) that occurred at the same
134 time. We calculated the entropy ratio for each species and social context, defined as the ratio
135 between the observed entropy and the expected entropy if Action Units were used randomly. Values
136 closer to 0 indicate that there is low uncertainty (e.g., when only a few facial movements are used
137 frequently) and values closer to 1 indicate high uncertainty (e.g., when many facial movements are

138 used frequently). To determine whether the entropy ratios for each species differed within social
139 context, we calculated the entropy ratio on 100 bootstrapped samples of the data, resulting in a
140 distribution of possible values. The bootstrapped entropy ratio of facial behavior differed across
141 species and within social contexts (Figure 1). In an affiliative context, the entropy ratio was highest
142 for crested, then Barbary, and lowest for rhesus macaques (crested: mean = 0.52, range = 0.50–
143 0.53; Barbary: mean = 0.45, range = 0.45–0.46; rhesus: mean = 0.38, range = 0.37–0.39). In an
144 aggressive context, the entropy ratio was highest for crested, then rhesus and lowest for Barbary
145 macaques (crested: mean = 0.62, range = 0.60–0.65; Barbary: mean = 0.32, range = 0.32–0.33;
146 rhesus: mean = 0.48, range = 0.47–0.49). In a submissive context, the entropy ratio was highest for
147 crested, then Barbary, and lowest for rhesus macaques. (crested: mean = 0.67, range = 0.64–0.70;
148 Barbary: mean = 0.49, range = 0.48–0.50; rhesus: mean = 0.38, range = 0.37–0.39). Overall, across
149 all contexts, including when the context was unclear, the entropy ratio was highest for crested, and
150 similar for Barbary and rhesus macaques (crested: mean = 0.57, range = 0.56–0.58; Barbary: mean
151 = 0.51, range = 0.51–0.51; rhesus: mean = 0.52, range = 0.51–0.52; Figure 1).



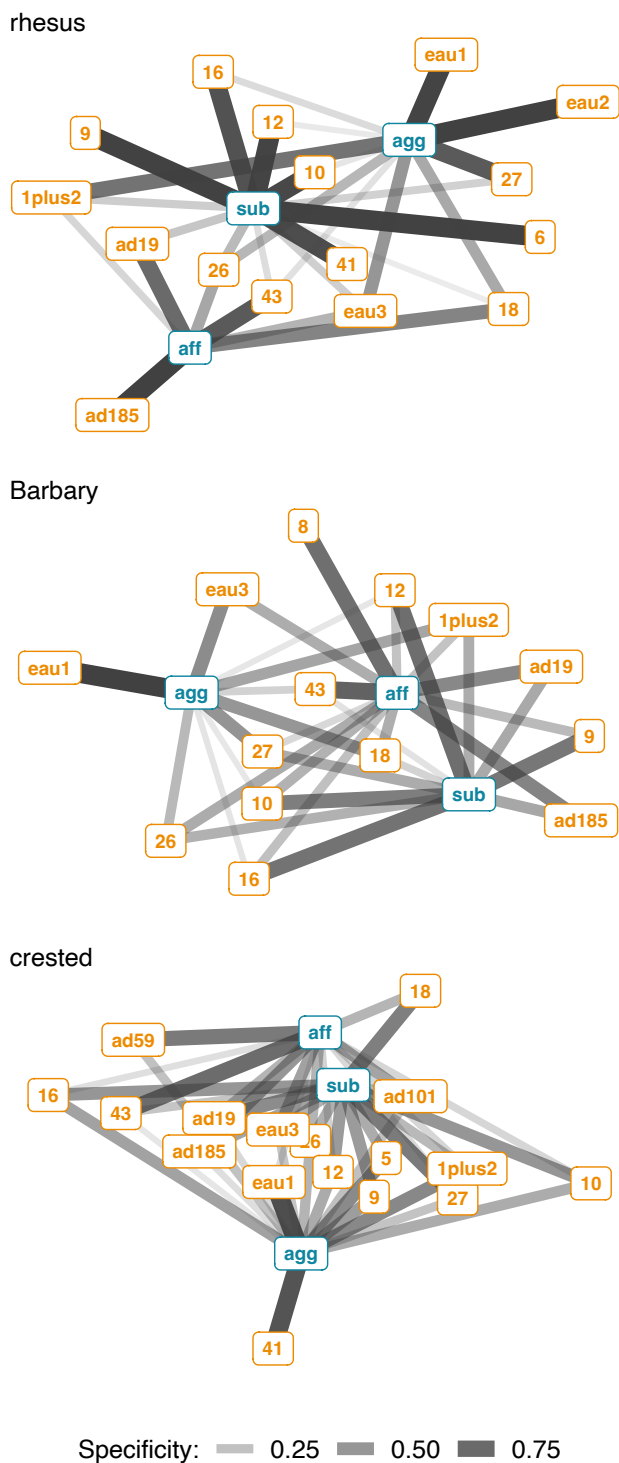
152

153 **Figure 1:** Bootstrapped entropy ratio of facial behavior across social contexts for three species of
154 macaques. The entropy ratio was calculated on 100 bootstrapped samples of the data by dividing
155 the observed entropy by the expected entropy if Action Units were used randomly for each social

156 context. The entropy ratio ranges from 0 to 1, with higher values indicating higher uncertainty.
157 Symbols and whiskers indicate mean and range of bootstrapped values.

158 **Context specificity of facial behavior**

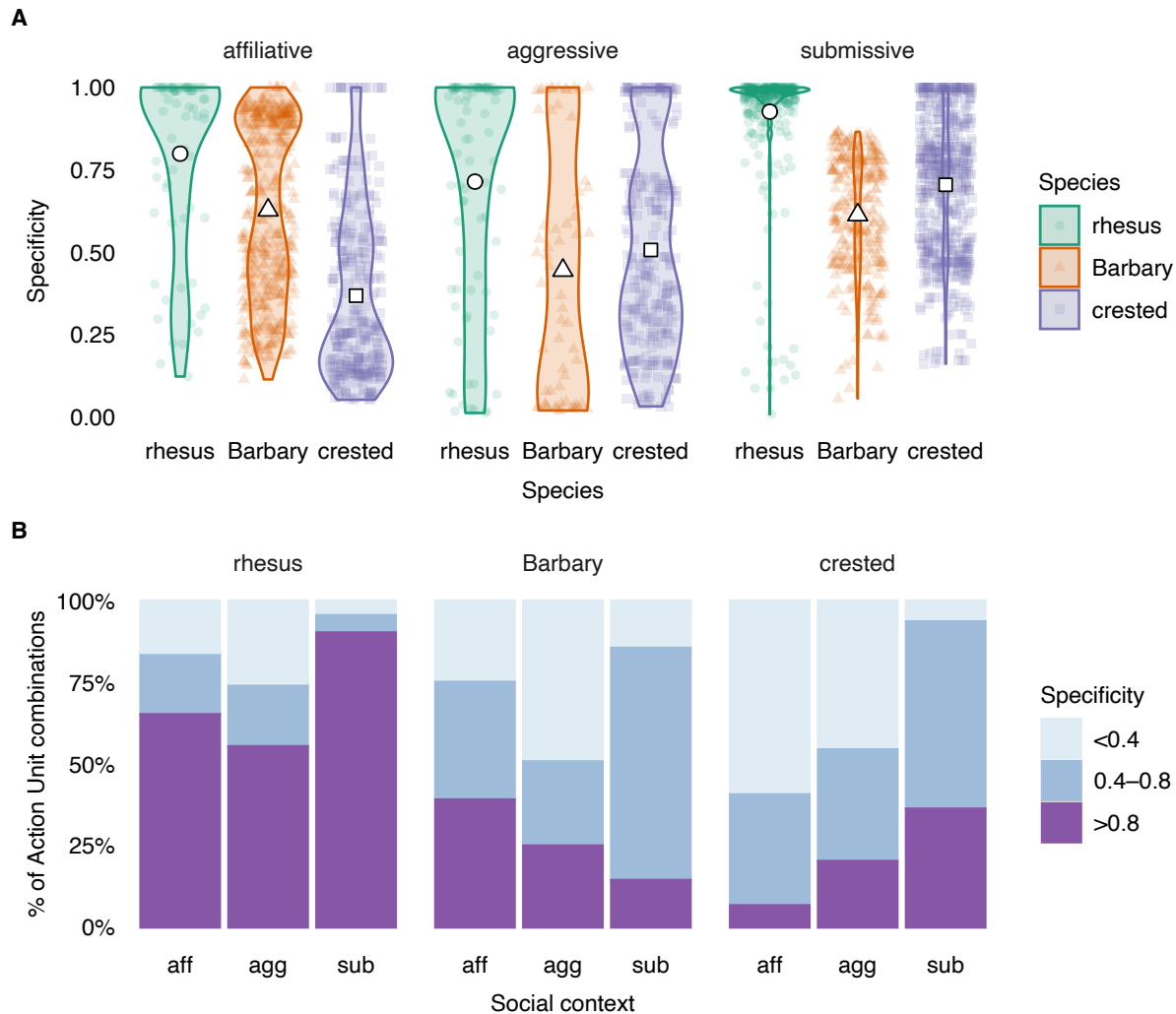
159 We calculated the context specificity for all possible combinations of Action Units. Here we re-
160 port specificity for combinations that were observed in at least 1% of observations per species and
161 social context because extremely rare signals do not affect the predictability of a system substan-
162 tially, regardless of whether they have high or low specificity. Specificity for each Action Unit
163 combination was defined as the number of times it was observed in one context divided by the
164 total number of times it was observed across all contexts. When considering single Action Units,
165 some were observed in only one context, but most were observed at least once in all three contexts
166 for all three species (Figure 2). On average, single Action Units were observed in fewer contexts
167 for rhesus (mean degree = 1.9), compared to Barbary (mean degree = 2.4), and crested macaques
168 (mean degree = 2.6). The specificity of all Action Unit combinations used in an affiliative context
169 was highest for the rhesus macaques, then Barbary, and lowest for crested macaques (rhesus: mean
170 = 0.80, SD = 0.28, n = 69; Barbary: mean = 0.63, SD = 0.26, n = 450; crested: mean = 0.37, SD
171 = 0.26, n = 327; Figure 3a). The specificity of Action Unit combinations used in an aggressive
172 context was highest for rhesus, then crested, and lowest for Barbary macaques (rhesus: mean =
173 0.71, SD = 0.35, n = 83; Barbary: mean = 0.44, SD = 0.38, n = 64; crested: mean = 0.51, SD =
174 0.30, n = 281). The specificity of Action Unit combinations used in a submissive context was also
175 highest for rhesus, then crested, and lowest for Barbary macaques (rhesus: mean = 0.93, SD = 0.18,
176 n = 312; Barbary: mean = 0.61, SD = 0.18, n = 297; crested: mean = 0.70, SD = 0.21, n = 595).
177 The majority (>50%) of Action Unit combinations used by rhesus macaques had high specificity
178 (>0.8) in all three social contexts, whereas only a minority (<50%) of Action Unit combinations
179 used by Barbary and crested macaques had high specificity (Figure 3b).



180

181 **Figure 2:** Bipartite network of single Action Units (orange) and social context (blue) for three
182 species of macaques. Edges are shown for Action Units that occurred in at least 1% of obser-

183 vations per context. Edge thickness and transparency are weighted by specificity, which ranges
 184 from 0 (indicating an Action Unit is never observed in a context) to 1 (indicating an Action Unit
 185 is only observed in one context). Context abbreviations: agg = aggressive, aff = affiliative, sub =
 186 submissive.



187

188 **Figure 3:** Specificity of Action Unit combinations that were used in at least 1% of observations per
 189 species per social context. Specificity ranges from 0 (indicating an Action Unit is never observed
 190 in a context) to 1 (indicating an Action Unit is only observed in one context). **(A)** Distribution
 191 of Action Unit combination specificity. Width of violin plots indicate the relative density of the
 192 data. Colored symbols indicate unique Action Unit combinations. White symbols indicate mean
 193 specificity. **(B)** Proportion of Action Unit combinations used with high (>0.8), moderate (0.4–

194 0.8) or low (<0.4) specificity. Context abbreviations: agg = aggressive, aff = affiliative, sub =
195 submissive.

196 **Predicting social context from facial behavior**

197 A random forest classifier was able to predict social context (affiliative, aggressive or submissive)
198 from facial behavior with a better accuracy than expected by chance alone for all three species
199 of macaques. The classifier was most accurate for rhesus ($\kappa = 0.92$), then Barbary ($\kappa =$
200 0.68), and least accurate for crested macaques ($\kappa = 0.49$). The confusion matrices for model
201 predictions are shown in table S1.

202 **Discussion**

203 We investigated the hypothesis that complex societies require more complex communication sys-
204 tems [1] by comparing the complexity of facial behavior of three species of macaques that vary in
205 their degree of social tolerance and complexity. We defined facial behavior by the unique combina-
206 tions of muscle movements visible in the face. Doing so allows for a much more precise description
207 of facial behavior and captures subtle differences that are lost if facial expressions are classified as
208 discrete categories. We quantified communicative complexity using three measures of uncertainty
209 and predictability: entropy, context specificity, and prediction error. Collectively, our results sug-
210 gest that the complexity of facial behavior is higher in species with a more tolerant—and therefore
211 more complex—social style; complexity was highest for crested, followed by Barbary, and low-
212 est in rhesus macaques. In light of what we know about the differences between macaque social
213 systems, our results support the predictions of the social complexity hypothesis for communicative
214 complexity.

215 The entropy ratio of facial behavior was highest in crested compared to Barbary and rhesus
216 macaques, both overall and within each social context (affiliative, aggressive, submissive). This

217 result suggests that crested macaques use a higher diversity of facial signals within each social
218 context more frequently, resulting in the higher relative uncertainty in their use of facial behavior.
219 Information theory defines information as the reduction in uncertainty once an outcome is learned
220 [18]. By this definition, our data suggest that the facial behavior of crested macaques has the
221 *potential* to communicate more information, compared to Barbary and rhesus macaques, although
222 this would need to be explicitly tested in future studies. Our findings are in line with predictions
223 of the social complexity hypothesis [1] given the differences in social styles between tolerant and
224 intolerant macaques. In tolerant macaque societies, social interactions are less constrained by
225 dominance [22] such that rates of counter aggression and reconciliation post-conflict are higher
226 [25,30]. Thus, there is a greater variability in the kind of interactions that individuals have,
227 potentially requiring the use of more diverse facial behavior to achieve social goals, particularly
228 during conflicts. Similarly, strongly bonded chimpanzee (*Pan troglodytes*) dyads exhibit a larger
229 repertoire of gestural communication than non-bonded dyads, presumably due to the former
230 having more varied types of social interactions [31].

231 The overall entropy ratio of rhesus and Barbary macaques was similar, suggesting that they have
232 similar communicative capacity using facial behavior. However, the entropy ratio differed when
233 compared within social contexts; while relative entropy was higher for Barbary macaques in af-
234 filiative and submissive contexts, it was higher for rhesus macaques in aggressive contexts. One
235 possible explanation may be due to the use of stereotyped signals of submission and dominance in
236 each species. For example, subordinate rhesus macaques regularly exhibit stereotyped signals of
237 submission (silent-bared-teeth), whereas dominant Barbary macaques regularly exhibit stereotyped
238 threats (round-open-mouth) [26,27]. Frequent use of a stereotyped signal within a context reduces
239 the overall diversity of signals, resulting in a lower entropy ratio for submission and aggression in
240 rhesus and Barbary macaques, respectively. It has been suggested that in societies with high power
241 asymmetries between individuals, such as in rhesus macaques, spontaneous signals of submission
242 serve to prevent conflicts from escalating as well as increasing the tolerance of dominant individuals
243 toward subordinates [27]. In societies with more moderate power asymmetries, such as in Barbary

244 macaques, subordinates may be less motivated to spontaneously submit and thus dominants may
245 need to assert their dominance with formalized threats more frequently [27].

246 While the entropy ratio captures the uncertainty of facial behavior used within a social context,
247 context specificity captures the uncertainty generated when the same facial behavior is used flexi-
248 bly across different social contexts. Overall, the context specificity of facial behavior was higher
249 for the intolerant rhesus macaques as compared to the more tolerant Barbary and crested macaques
250 across all three social contexts. This pattern occurred for both the mean specificity values and the
251 proportion of Action Unit combinations used that had high (>0.8) specificity. Similarly, a previous
252 study demonstrated that vocal calls of tolerant macaques are less context specific than in intoler-
253 ant macaques [32]. There was not a clear difference in specificity between Barbary and crested
254 macaques; specificity was higher for Barbary macaques in affiliative contexts, similar for both
255 species in aggressive contexts, and higher for crested macaques in submissive contexts. These dif-
256 ferences in context specificity of communicative signals across macaque species may be related
257 to differences in power asymmetry in their respective societies, particularly as it relates to the risk
258 of injury. For macaques, bites are far more likely to injure opponents than other types of contact
259 aggression (e.g., grab, slap) and thus provide the best proxy for risk of injury [21]. The percentage
260 of conflicts involving bites is much higher in the less tolerant rhesus macaque, compared to the
261 more tolerant Barbary and crested macaques who have similar low rates of aggression involving
262 bites [25,33]. Risky situations may promote the evolution of more conspicuous, stereotypical sig-
263 nals to reduce ambiguity [34]. Indeed, intolerant macaques such as the rhesus more commonly use
264 formal signals of submission [26,27]. In our study, rhesus macaques used facial behavior with high
265 specificity across all contexts but particularly in submissive contexts. If the same facial behavior
266 (or signal in general) is used in multiple social contexts, its meaning may be uncertain and must
267 be deduced from additional contextual cues [35]. When facial behavior is highly context specific,
268 there is less uncertainty about the meaning of the signal and/or intention of the signaler. In a society
269 where the risk of injury from aggression is high, it may be adaptive for individuals to use signals that
270 are highly context specific or ritualized to reduce uncertainty about its meaning. By contrast, the

271 lower risk of injury in Barbary and crested macaques may allow room for more nuanced exchanges
272 of information during conflicts as well as higher rates of reconciliation post conflict [25,30].

273 In all three species of macaques, at least some facial muscle movements had low specificity and
274 were therefore used across multiple social contexts that likely differed in valence. This finding is
275 in line with the idea that communicative signals in primates are better interpreted as the signaler
276 announcing its intentions and likely future behavior [36,37], and not necessarily as an expression
277 of emotional state [28,29,36,38].

278 We found that a random forest classifier was least accurate at predicting social context from facial
279 behavior for crested, followed by Barbary, and then rhesus macaques. The behavior of complex
280 systems is generally harder to predict than simpler ones [16,17]. Thus, the relatively poorer perfor-
281 mance of the classifier in crested macaques suggests that they have the most complex facial behav-
282 ior. Nevertheless, the classifier was able to predict social context from facial behavior with better
283 accuracy than expected by chance alone for all three species of macaque, including the crested.
284 This result confirms the assumption that facial behavior in macaques is not used randomly and
285 most likely has some communicative or predictive value [39]. Completely random systems are not
286 considered complex [19], but the communications systems of living organisms are unlikely to be
287 observed as random. Therefore, measuring uncertainty becomes a good proxy for complexity [14].

288 In addition to social complexity, it is possible that other factors are related to the complexity of
289 facial behavior. For example, primates with a larger body size have greater facial mobility [13,40],
290 which could allow for greater complexity of facial behavior. However, differences in mean body
291 mass across the three macaques species of this study are small (rhesus: 6.5 kg; Barbary: 11.5
292 kg; crested: 7.4 kg) [41] with substantial overlap in body weight across adult individuals of the
293 different species [42], and so it is unlikely to explain the differences in the complexity of facial
294 behavior that we report in this study. The degree of terrestriality could also influence the evolution
295 of facial signals due to more limited visibility in the canopy. However, differences in facial mobility
296 across terrestrial and non-terrestrial primates are not significant once body size is controlled for

297 [13]. Furthermore, all three species included in this study have comparable levels of terrestriality,
298 spending the majority (52-72%) of the time on the ground [43–45]. Spatial spread and predation
299 pressure could potentially also influence the use of facial signals. For example, when group spread
300 is higher, reliance on facial signals could be lower, or when predation pressure is higher, reliance
301 on facial signals could be higher. There are currently no reliable data on predation pressure and
302 spatial spread of the three species in their natural habitat but it could be a good avenue for future
303 studies.

304 Our results on the complexity of facial behavior in macaques is mirrored by previous studies show-
305 ing that the complexity of vocal calls is similarly higher in tolerant compared to intolerant macaques
306 [32,46]. Although not all macaque facial expressions have a vocal component, vocalizations are
307 fundamentally multisensory with both auditory and visual components, where different facial mus-
308 cle contractions are partly responsible for different-sounding vocalizations [47]. Indeed, some areas
309 of the brain in primates integrate visual and auditory information resulting in behavioral benefits
310 [48]. For example, macaques detect vocalizations in a noisy environment faster when mouth move-
311 ments are also visible, where faster reaction times are associated with a reduced latency in auditory
312 cortical spiking activity [49]. Combined, these findings suggest that the evolution in the complex-
313 ity of vocal and facial signals in macaques may be linked and the same may be true of primates in
314 general. For instance, humans not only have the most complex calls (language) and gestures, but
315 most likely use the most complex facial behavior as well, given that their general facial mobility
316 is highest among primates (most Action Units) [12,50]. In lemurs (Lemuriformes), the repertoire
317 size of vocal, visual, and olfactory signals positively correlate with group size and each other, sug-
318 gesting that complexity in all three communicative modalities coevolved with social complexity
319 [51]. While the complexity of different communication modalities is likely interlinked and cor-
320 related with each other, future studies would ideally integrate signals from all modalities into a
321 single communicative repertoire for each species. While collecting and analyzing data on multiple
322 modalities of communication has historically been a challenge, such endeavors would be an im-
323 portant next step in the study of animal communication [52]. By breaking down signaling units to

324 their smallest components, as we have done for facial behavior in this study, we may be able to
325 define a “signal” by temporal co-activation of visual, auditory, and perhaps even olfactory cues,
326 which would provide the most comprehensive picture of animal communication.

327 **Methods**

328 **Study subjects and data collection**

329 Behavioral data and video recordings were collected on one adult male and 31 adult female rhesus
330 macaques (*M. mulatta*), on 18 adult male and 28 adult female Barbary macaques (*M. sylvanus*),
331 and 17 adult male and 21 adult female crested macaques (*M. nigra*). See supplementary text for
332 further details.

333 For all study groups and subjects, focal animal observations [53] lasting 15-30 minutes were con-
334 ducted throughout the day in a pseudo-randomized order such that the number of days and time of
335 day that each individual was observed was balanced. Videos of social interactions were recorded
336 with a recording camera (Panasonic HDC-SD700, Bracknell, UK) during focal animal observations
337 as well as *ad libitum*. Social behavior, including grooming, body contact, and agonistic interactions
338 were recorded using a handheld smartphone or tablet with purpose-built software (rhesus: Animal
339 Behavior Pro [54]; Barbary: CyberTracker (<http://cybertracker.org>), crested: Microsoft Excel).

340 **Facial behavior and social context coding**

341 Facial behavior was coded at the level of observable individual muscle movements using the Fa-
342 cial Action Coding System (FACS) [12], adapted for each species of macaque (MaqFACS): rhesus
343 [55], Barbary [56], crested [10]. In FACS, individual observable muscle contractions are coded as
344 unique Action Units (AU; e.g., upper lip raiser AU10). Some common facial movements where
345 the underlying muscle is unknown are coded as Action Descriptors (AD; e.g., jaw thrust AD29). In

346 MaqFACS, the lip-pucker AU18 has two subtle variations normally denoted as AU18i and AU18ii
347 [55,56]. However, it was often difficult to reliably distinguish between these two subtle variations
348 when coding videos, and so the lip-pucker was simply coded as AU18. We added a new Action De-
349 scriptor 185 (AD185) called jaw-oscillation, to denote the stereotyped movement of the jaw up and
350 down. When combined with existing Action Units of lip movements, the jaw-oscillation AD185
351 allows for a more detailed and accurate coding of some facial behaviors that would otherwise be
352 labeled as lipsmack (AD181), teeth-chatter, or jaw-wobble [10,55]. A complete list of Action Units
353 and Action Descriptors coded in this study is given in table S2.

354 We coded facial behavior of adult individuals but included their interactions with any other group
355 member regardless of age or sex. Each social interaction was labeled with a context; aggressive,
356 submissive, affiliative, or unclear. We did not consider interactions in a sexual context because data
357 for the rhesus macaques were only collected during the non-mating season. Social context was la-
358 beled from the point of view of the signaler based on their general behavior and body language (but
359 not the facial behavior itself), during or immediately following the facial behavior. An aggressive
360 context was considered when the signaler lunged or leaned forward with the body or head, charged,
361 chased, or physically hit the interaction partner. A submissive context was considered when the
362 signaler leaned back with the body or head, moved away, or fled from the interaction partner. An
363 affiliative context was considered when the signaler approached another individual without aggres-
364 sion (as defined previously) and remained in proximity, in relaxed body contact, or groomed either
365 during or immediately after the facial behavior. In cases where the behavior of the signaler did not
366 match our context definitions, or displayed behaviors belonging to multiple contexts, we labeled
367 the social context as unclear. Social context was determined from the video itself and/or from the
368 matching focal behavioral data, if available. Videos were FACS coded frame-by-frame using the
369 software BORIS [57] by AVR, CP and PRC, who are certified FACS and MaqFACS coders. Table
370 1 shows the number of social interactions per species and context from which FACS codes were
371 made.

372 **Table 1:** Total number of social interactions per species and social context that were MaqFACS
373 coded.

Species	Context	N interactions	N subjects
rhesus	affiliative	193	29
	aggressive	413	32
	submissive	318	31
	unclear	121	30
Barbary	affiliative	683	43
	aggressive	585	44
	submissive	529	34
	unclear	603	45
crested	affiliative	241	35
	aggressive	62	23
	submissive	25	18
	unclear	107	25

374 **Statistical analyses**

375 Prior to analyses, MaqFACS data were formatted as a binary matrix with Action Units and Action
376 Descriptors (hereafter simply Action Units) in the columns. Each row denoted an observation time
377 block of 500ms, where if an Action Unit was active during this time block, it was coded 1 and coded
378 0 if not. Thus, each row contained information on the combination of facial muscle movements
379 that were co-activated within a 500ms time window. All statistical analyses were conducted in R
380 (version 4.2.1) [58].

381 The observed entropy for each social context was calculated using Shannon's information entropy
382 formula [18]:

$$H = - \sum_i^n p_i \log p_i$$

383 where n is the number of unique Action Unit combinations and p is the probability of observ-
384 ing each Action Unit combination in each social context. The expected maximum entropy was
385 calculated by randomizing the data matrix while keeping the number of active Action Units per
386 observation (row) the same. This process was repeated 100 times and the mean of the randomized
387 entropy values was used as the expected entropy. Therefore, the expected entropy indicated the
388 entropy of the system if facial muscle contractions occurred at random, while keeping the com-
389 bination size of co-active muscle movements within the range observed in the data. The entropy
390 ratio was calculated by dividing the observed entropy by the expected (maximum) entropy. To
391 determine whether the entropy ratios for each species differed within social context, the entropy
392 ratio was calculated on 100 bootstrapped samples of the data, resulting in a distribution of possible
393 entropy ratios. If the distribution of bootstrapped entropy ratios did not overlap, the differences
394 between entropy ratios were considered to be meaningful.

395 We calculated the specificity with which Action Unit combinations are associated with a social con-
396 text within each species using the function “specificity” from the R package “NetFACS” (version
397 0.5.0) [59]. Due to an imbalanced number of observations across social contexts, contexts with
398 fewer observations were randomly upsampled prior to the specificity calculation. During the up-
399 sampling procedure all observations of the minority contexts were kept, and new observations were
400 randomly sampled to match the number of observations in the majority context. This procedure
401 corrects for any bias in the specificity results from an imbalanced dataset (see fig. S1). Specificity
402 is the conditional probability of a social context given that an Action Unit combination is observed,
403 and ranges from 0 (when an Action Unit combination is never observed in a context) to 1 (when
404 an Action Unit is only observed in one context). Low specificity values indicate that Action Units
405 were used flexibly across multiple contexts whereas high values indicate that Action Units were
406 used primarily in a single context. Specificity was calculated for all Action Unit combination sizes

407 ranging from 1 to 11 (the maximum observed combination size) co-active Action Units. When
408 reporting context specificity results, we excluded Action Unit combinations that occurred in less
409 than 1% of observations within a social context because extremely rare signals do not impact the
410 predictability of a communication system regardless of whether specificity is low or high. There-
411 fore, excluding rare Action Unit combinations removes noise from the specificity results. We report
412 the mean specificity of Action Unit combinations per social context and the proportion of Action
413 Unit combinations that have high, moderate, or low specificity. For single Action Units we plotted
414 bipartite networks that show how Action Units are connected to social context weighted by their
415 specificity.

416 To predict social context from the combination of Action Units we fit a random forest classifier us-
417 ing the “tidymodels” R package (version 1.0.0) [60] using the function “ran_forest” with the engine
418 set to “ranger” [61], 500 trees, 4 predictor columns randomly sampled at each split, and 10 as the
419 minimum number of data points in a node required for splitting further. The data were randomly
420 split into a training set (70%) and a test set (30%), while keeping the proportion of observations
421 per social context the same in the training and test sets. Due to an imbalanced number of obser-
422 vations across social contexts, contexts with fewer observations were over-sampled in the training
423 set using the SMOTE algorithm [62] to improve the classifier predictions. To assess the classifier
424 performance, we report the kappa statistic, which denotes the observed accuracy corrected for the
425 expected accuracy [63]. Kappa is 0 when the classifier performs at chance level and 1 when it
426 shows perfect classification. Kappa values between 0 and 1 indicate how much better the classi-
427 fier performed than chance (e.g., kappa of 0.5 indicates the classifier was 50% better than chance).
428 Kappa is a more reliable estimate of model performance than accuracy alone when the relative
429 sample size for each context is imbalanced, as was the case with our data.

430 **Acknowledgments**

431 We thank the German Primate Center (DPZ) for permission to collect data on the rhesus macaques,
432 Uwe Schönmann for logistical support, and Julia Ostner for being our host at the DPZ. We thank
433 Matt Lowatt and Ellen Merz for permission to collect data on the Barbary macaques at Trentham
434 Monkey Forest. We thank the Indonesian State Ministry of Research and Technology (RISTEK),
435 the Directorate General of Forest Protection and Nature Conservation (PHKA) and the Department
436 for the Conservation of Natural Resources (BKSDA), North Sulawesi, for permission to access
437 groups of crested macaques in the Tangkoko-Batuangus Nature Reserve. We thank Christof Neu-
438 mann for statistical advice. This work was funded by the Leverhulme Trust (RPG2018-334).

439 **Ethics**

440 This work adhered to the Guidelines for the treatment of animals in behavioral research and teach-
441 ing [64] and was approved by the Animal Welfare and Ethical Review Body of the University of
442 Portsmouth (AWERB, approval number: 919B). The AWERB uses UK Home Office guidelines
443 on the Animals (Scientific Procedures) Act 1986 when assessing proposals and adheres to the reg-
444 ulations of the European Directive 2010/63/EU. The German Primate Center also complies with
445 the European Directive 2010/63/EU, as well as with the provisions of the German Animal Welfare
446 Act.

447 **Data availability**

448 The data and R code used for all statistical analysis is available on GitHub, [https://github.com/](https://github.com/avrincon/macaque-facial-complexity)
449 [avrincon/macaque-facial-complexity](https://github.com/avrincon/macaque-facial-complexity).

450 **References**

- 451 1. Freeberg TM, Dunbar RIM, Ord TJ. 2012 Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B* **367**, 1785–1801.
452 (doi:10.1098/rstb.2011.0213)
- 453 2. Peckre L, Kappeler PM, Fichtel C. 2019 Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behav. Ecol. Sociobiol.* **73**, 11.
454 (doi:10.1007/s00265-018-2605-4)
- 455 3. Raviv L, Peckre LR, Boeckx C. 2022 What is simple is actually quite complex: A critical note on terminology in the domain of language and communication. *J. Comp. Psychol.*
456 **136**, 215–220.
- 457 4. Bergman TJ, Beehner JC. 2015 Measuring social complexity. *Anim. Behav.* **103**, 203–209.
458 (doi:10.1016/j.anbehav.2015.02.018)
- 459 5. Aureli F, Schaffner CM, Schino G. 2022 Variation in communicative complexity in relation to social structure and organization in non-human primates. *Phil. Trans. R. Soc. B* **377**,
460 20210306. (doi:10.1098/rstb.2021.0306)
- 461 6. Aureli F, Schino G. 2019 Social complexity from within: How individuals experience the structure and organization of their groups. *Behav. Ecol. Sociobiol.* **73**, 6.
462 (doi:10.1007/s00265-018-2604-5)
- 463 7. Kappeler PM. 2019 A framework for studying social complexity. *Behav. Ecol. Sociobiol.*
464 **73**, 13. (doi:10.1007/s00265-018-2601-8)
- 465 8. Keenan S, Lemasson A, Zuberbühler K. 2013 Graded or discrete? A quantitative analysis of Campbell’s monkey alarm calls. *Anim. Behav.* **85**, 109–118.
466 (doi:10.1016/j.anbehav.2012.10.014)
- 467 9. Dobson SD. 2012 Coevolution of facial expression and social tolerance in macaques. *Am. J. Primatol.* **74**, 229–235. (doi:10.1002/ajp.21991)
468

- 469 10. Clark PR, Waller BM, Burrows AM, 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*). *Am. J. Phys. Anthropol.* **173**, 411–422.
470 (doi:10.1002/ajpa.24129)
- 471 11. Waller BM, Julle-Daniere E, Micheletta J. 2020 Measuring the evolution of facial
'expression' using multi-species FACS. *Neurosci. Biobehav. Rev.* **113**, 1–11.
472 (doi:10.1016/j.neubiorev.2020.02.031)
- 473 12. Ekman P, Friesen WV, Hager JC. 2002 *Facial Action Coding System*. Salt Lake City: Re-
474 search Nexus. See <https://doi.apa.org/doiLanding?doi=10.1037%2Ft27734-000>.
- 475 13. Dobson SD. 2009 Socioecological correlates of facial mobility in nonhuman anthropoids.
476 *Am. J. Phys. Anthropol.* **139**, 413–420. (doi:10.1002/ajpa.21007)
- 477 14. Rebout N, Lone J-C, De Marco A, Cozzolino R, Lemasson A, Thierry B. 2021 Mea-
suring complexity in organisms and organizations. *R. Soc. Open. Sci.* **8**, 200895.
478 (doi:10.1098/rsos.200895)
- 479 15. Sambrook T, Whiten A. 1997 On the nature of complexity in cognitive and behavioural
480 science. *Theory Psychol.* **7**, 191–213. (doi:10.1177/0959354397072004)
- 481 16. McDaniel RR, Driebe DJ, editors. 2005 *Uncertainty and surprise in complex systems*.
482 Springer Berlin, Heidelberg. See <https://link.springer.com/book/10.1007/b13122>.
- 483 17. Schuster P. 2016 How complexity originates: Examples from history reveal additional roots
484 to complexity. *Complexity* **21**, 7–12. (doi:10.1002/cplx.21841)
- 485 18. Shannon CE. 1948 A mathematical theory of communication. *The Bell System Technical*
486 *Journal* **27**, 379–423. (doi:10.1002/j.1538-7305.1948.tb01338.x)
- 487 19. Adami C. 2002 What is complexity? *BioEssays* **24**, 1085–1094. (doi:10.1002/bies.10192)
- 488
- 489 20. Thierry B. 2007 Unity in diversity: Lessons from macaque societies. *Evol. Anthropol.* **16**,
490 224–238. (doi:10.1002/evan.20147)

- 491 21. Thierry B. 2021 Where do we stand with the covariation framework in primate societies?
492 *Am. J. Biol. Anthropol.* **128**, 5–25. (doi:10.1002/ajpa.24441)
- 493 22. Balasubramaniam KN *et al.* 2012 Hierarchical steepness, counter-aggression, and macaque
494 social style scale. *Am. J. Primatol.* **74**, 915–925. (doi:10.1002/ajp.22044)
- 495 23. Sueur C, Petit O, De Marco A, Jacobs AT, Watanabe K, Thierry B. 2011 A com-
496 parative network analysis of social style in macaques. *Anim. Behav.* **82**, 845–852.
(doi:10.1016/j.anbehav.2011.07.020)
- 497 24. Berman C, Thierry B. 2010 Variation in kin bias: Species differences and time constraints
498 in macaques. *Behaviour* **147**, 1863–1887. (doi:10.1163/000579510X539691)
- 499 25. Duboscq J, Micheletta J, Agil M, Hodges K, Thierry B, Engelhardt A. 2013 Social tolerance
500 in wild female crested macaques (*Macaca nigra*) in Tangkoko-Batuangus nature reserve,
Sulawesi, Indonesia. *Am. J. Primatol.* **75**, 361–375. (doi:10.1002/ajp.22114)
- 501 26. de Waal FBM, Luttrell LM. 1985 The formal hierarchy of rhesus macaques: An investiga-
502 tion of the bared-teeth display. *Am. J. Primatol.* **9**, 73–85. (doi:10.1002/ajp.1350090202)
- 503 27. Preuschoft S, Schaik C. 2000 Dominance and communication: Conflict management in
504 various social settings. In *Natural Conflict Resolution* (eds F Aureli, FBM De Waal),
- 505 28. Waller BM, Whitehouse J, Micheletta J. 2017 Rethinking primate facial expres-
506 sion: A predictive framework. *Neurosci. Biobehav. Rev.* **82**, 13–21.
(doi:10.1016/j.neubiorev.2016.09.005)
- 507 29. Fridlund AJ. 1994 *Human facial expression: An evolutionary view*. San Diego, CA: Aca-
508 demic Press.
- 509 30. Thierry B, Aureli F, Nunn CL, Petit O, Abegg C, de Waal FBM. 2008 A comparative study
510 of conflict resolution in macaques: Insights into the nature of trait covariation. *Anim. Behav.*
75, 847–860. (doi:10.1016/j.anbehav.2007.07.006)

- 511 31. Amici F, Liebal K. 2022 The social dynamics of complex gestural communication in great
and lesser apes (*Pan troglodytes*, *Pongo abelii*, *Symphalangus syndactylus*). *Phil. Trans.*
512 *R. Soc. B* **377**, 20210299. (doi:10.1098/rstb.2021.0299)
- 513 32. Rebut N *et al.* 2022 Tolerant and intolerant macaques differ in the context specificity of
their calls and how they ‘comment’ on the interactions of others. *Behav. Ecol. Sociobiol.*
514 **76**, 67. (doi:10.1007/s00265-022-03177-7)
- 515 33. Tyrrell M, Berman CM, Duboscq J, Agil M, Sutrisno T, Engelhardt A. 2020 Avoidant so-
cial style among wild crested macaque males (*Macaca nigra*) in Tangkoko nature reserve,
516 Sulawesi, Indonesia. *Behaviour* **157**, 451–491. (doi:10.1163/1568539X-bja10009)
- 517 34. Clark PR, Waller BM, Agil M, Micheletta J. 2022 Crested macaque facial movements are
more intense and stereotyped in potentially risky social interactions. *Phil. Trans. R. Soc.*
518 *B* **377**, 20210307. (doi:10.1098/rstb.2021.0307)
- 519 35. Seyfarth RM, Cheney DL. 2017 The origin of meaning in animal signals. *Anim. Behav.*
520 **124**, 339–346. (doi:10.1016/j.anbehav.2016.05.020)
- 521 36. Cheney DL, Seyfarth RM. 2018 Flexible usage and social function in primate vocalizations.
522 *Proc. Natl. Acad. Sci.* **115**, 1974–1979. (doi:10.1073/pnas.1717572115)
- 523 37. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication.
524 *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)
- 525 38. Barrett LF, Adolphs R, Marsella S, Martinez AM, Pollak SD. 2019 Emotional expressions
reconsidered: Challenges to inferring emotion from human facial movements. *Psychol.*
526 *Sci. Public Interest* **20**, 1–68. (doi:10.1177/1529100619832930)
- 527 39. Waller BM, Whitehouse J, Micheletta J. 2016 Macaques can predict social outcomes from
528 facial expressions. *Anim. Cogn.* **19**, 1031–1036. (doi:10.1007/s10071-016-0992-3)
- 529 40. Santana SE, Dobson SD, Diogo R. 2014 Plain faces are more expressive: Comparative
study of facial colour, mobility and musculature in primates. *Biol. Lett.* **10**, 20140275.
530 (doi:10.1098/rsbl.2014.0275)

- 531 41. Jones KE *et al.* 2009 PanTHERIA: A species-level database of life history, ecology, and ge-
ography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648. (doi:10.1890/08-
532 1494.1)
- 533 42. Smith RJ, Jungers WL. 1997 Body mass in comparative primatology. *J. Hum. Evol.* **32**,
534 523–559. (doi:10.1006/jhev.1996.0122)
- 535 43. Khatiwada S, Paudel PK, Chalise MK, Ogawa H. 2020 Comparative ecological and be-
havioral study of *Macaca assamensis* and *M. mulatta* in Shivapuri Nagarjun National Park,
536 Nepal. *Primates* **61**, 603–621. (doi:10.1007/s10329-020-00810-9)
- 537 44. O’Brien TG, Kinnaird MF. 1997 Behavior, diet, and movements of the Sulawesi
crested black macaque (*Macaca nigra*). *Int. J. Primatol.* **18**, 321–351.
538 (doi:10.1023/A:1026330332061)
- 539 45. Alami AE, Chait A. 2014 Comparison of the terrestriality of Barbary macaques (*Macaca*
sylvanus) between tourist and wild sites in the region of Ouzoud, Morocco. *Mammalia* **78**,
540 539–542. (doi:10.1515/mammalia-2013-0061)
- 541 46. Rebut N *et al.* 2020 Tolerant and intolerant macaques show different levels of struc-
tural complexity in their vocal communication. *Proc. R. Soc. B.* **287**, 20200439.
542 (doi:10.1098/rspb.2020.0439)
- 543 47. Ghazanfar AA, Takahashi DY. 2014 The evolution of speech: Vision, rhythm, cooperation.
544 *Trends Cogn. Sci.* **18**, 543–553. (doi:10.1016/j.tics.2014.06.004)
- 545 48. Ghazanfar AA, Eliades SJ. 2014 The neurobiology of primate vocal communication. *Curr.*
546 *Opin. Neurobiol.* **28**, 128–135. (doi:10.1016/j.conb.2014.06.015)
- 547 49. Chandrasekaran C, Lemus L, Ghazanfar AA. 2013 Dynamic faces speed up the onset of
auditory cortical spiking responses during vocal detection. *Proc. Natl. Acad. Sci.* **110**,
548 E4668–E4677. (doi:10.1073/pnas.1312518110)

- 549 50. Dobson SD. 2009 Allometry of facial mobility in anthropoid primates: Implications
for the evolution of facial expression. *Am. J. Phys. Anthropol.* **138**, 70–81.
550 (doi:10.1002/ajpa.20902)
- 551 51. Fichtel C, Kappeler PM. 2022 Coevolution of social and communicative complexity in
552 lemurs. *Phil. Trans. R. Soc. B* **377**, 20210297. (doi:10.1098/rstb.2021.0297)
- 553 52. Liebal K, Slocombe KE, Waller BM. 2022 The language void 10 years on: Multimodal
primate communication research is still uncommon. *Ethol. Ecol. Evol.* **34**, 274–287.
554 (doi:10.1080/03949370.2021.2015453)
- 555 53. Altmann J. 1974 Observational study of behavior: Sampling methods. *Behaviour* **49**, 227–
556 266. (doi:10.1163/156853974X00534)
- 557 54. Newton-Fisher NE. 2020 *Animal behaviour pro*. Mobile App, version 1.5. See [https://
558 itunes.apple.com/gb/app/animal-behaviour-pro/id579588319?mt=8](https://itunes.apple.com/gb/app/animal-behaviour-pro/id579588319?mt=8).
- 559 55. Parr LA, Waller BM, Burrows AM, Gothard KM, Vick SJ. 2010 Brief communication:
MaqFACS: A muscle-based facial movement coding system for the rhesus macaque. *Am.*
560 *J. Phys. Anthropol.* **143**, 625–630. (doi:10.1002/ajpa.21401)
- 561 56. Julle-Danière É, Micheletta J, Whitehouse J, Joly M, Gass C, Burrows AM, Waller
BM. 2015 MaqFACS (Macaque Facial Action Coding System) can be used to doc-
ument facial movements in Barbary macaques (*Macaca sylvanus*). *PeerJ* **3**, e1248.
562 (doi:10.7717/peerj.1248)
- 563 57. Friard O, Gamba M. 2016 BORIS: A free, versatile open-source event-logging software
for video/audio coding and live observations. *Methods Ecol. Evol.* **7**, 1325–1330.
564 (doi:10.1111/2041-210X.12584)
- 565 58. R Core Team. 2022 *R: A language and environment for statistical computing*. Vienna,
566 Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.

- 567 59. Mielke A, Waller BM, Pérez C, Rincon AV, Duboscq J, Micheletta J. 2022 NetFACS: Using
network science to understand facial communication systems. *Behav. Res. Methods* **54**,
568 1912–1927. (doi:10.3758/s13428-021-01692-5)
- 569 60. Kuhn M, Wickham H. 2020 *Tidymodels: A collection of packages for modeling and ma-*
570 *chine learning using tidyverse principles*. See <https://www.tidymodels.org>.
- 571 61. Wright MN, Ziegler A. 2017 Ranger: A fast implementation of random forests for high
572 dimensional data in C++ and R. *J. Stat. Softw.* **77**, 1–17. (doi:10.18637/jss.v077.i01)
- 573 62. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. 2002 SMOTE: Synthetic minority
574 over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357. (doi:10.1613/jair.953)
- 575 63. Cohen J. 1960 A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**,
576 37–46. (doi:10.1177/001316446002000104)
- 577 64. ASAB Ethical Committee, ABS Animal Care Committee. 2022 Guidelines for the
treatment of animals in behavioural research and teaching. *Anim. Behav.* **183**.
578 (doi:10.1016/S0003-3472(21)00389-4)