

Wild-captive contrasts in non-vocal communicative repertoires and functional specificity in orang-utans

Marlen Fröhlich^{1*}, Natasha Bartolotta¹, Caroline Fryns¹, Colin Wagner², Laurene Momon², Marvin Jaffrezic², Tatang Mitra Setia³, Caroline Schuppli^{1,4}, Maria van Noordwijk¹, Carel P. van Schaik¹

¹ Department of Anthropology, University of Zurich, Zurich, Switzerland

² DEPE-IPHC – Département Ecologie, Physiologie et Ethologie, University of Strasbourg, Strasbourg, France

³ Fakultas Biologi, Universitas Nasional, Jakarta Selatan, Indonesia

⁴ Leipzig Center for Early Child Development, Leipzig, Germany

*Corresponding author, *email address*: marlen.froehlich@uzh.ch

Abstract

The creation of novel communicative acts is an essential element of human language. Although some research suggests the presence of this ability in great apes, this claim remains controversial. Here, we use orang-utans (*Pongo* spp.) to systematically assess the effect of the wild-captive contrast on the repertoire size of communicative acts. We find that individual communicative repertoires are significantly larger in captive compared to wild settings, irrespective of species, age-sex class or sampling effort. Twenty percent of the orang-utan repertoire in captivity were not observed in the wild. In Sumatran orang-utans, the more sociable species, functional specificity was also higher in captive versus wild settings. We thus conclude that orang-utans, when exposed to a more sociable

and terrestrial lifestyle, have the behavioural plasticity to invent new communicative behaviours that are highly functionally specific. This productive capacity by great apes is a major prerequisite for the evolution of language and seems to be ancestral in the hominid lineage.

Keywords

Close-range communication, gesture, signal repertoire, signal meaning, behavioural plasticity, socio-ecological environment, language evolution, *Pongo abelii*, *Pongo pygmaeus*

Introduction

One of the core features of human language is its productivity, referring to the idea that signallers can create and understand novel utterances with novel meanings¹. While several other building blocks of language, including intentionality, reference and compositional syntax, have increasingly been documented among a wide range of non-human species^{2,3,4}, productivity seems to be very rare in the animal kingdom. Instead, animal communication systems are thought to rely heavily on evolved signals, that is, communicative acts whose form and function evolved under the influence of the effect on the recipient^{5,6}. The current consensus is that all facial, vocal and most gestural expressions of non-human species have evolved through natural selection over long periods of time, and have become innate: the ability to produce them arises spontaneously during ontogeny, whereas its use is often fine-tuned by practice^{7,8,9}.

Recent research on great apes, our closest living relatives, suggests that it may be timely to distinguish between innate animal signals and those that are acquired developmentally¹⁰. Great apes have provided most comparative evidence for the cognitive building blocks and selective pressures

shaping the human communication system^{11, 12, 13}. But recent work has also shown that some of their gestures and sounds are apparently innovated and maintained over time^{14, 15, 16, 17, 18, 19, 20}, and play the same role in the communication process as evolved signals do – we could thus call them invented signals. Because targeted studies to estimate the extent of productivity in great apes are so far lacking, we here examine this question by comparing the same species in the wild and in a novel setting, captivity^{10, 21}. The contrast allows a direct test of how repertoires respond to the changes in the socio-ecological environment. In captivity, individuals face less competition for food, have more spare time, are closer together, are more often visible to each other, and more on the ground than in the wild^{22, 23}. Especially in fission-fusion species such as orang-utans (*Pongo* spp.), interaction rates in contexts such as social play, grooming, conflict situations and mating are boosted in captive settings e.g.^{24, 25, 26}, which may favour the production of innovative communicative acts, and thus cause differences in the communicative repertoires of individuals and groups.

The captive-wild contrast also allows us to examine the extent to which the meaning of communicative acts depends on the context in which they are used. We predict that the learned, “species-atypical” communicative acts have high functional specificity, i.e. a highly context-specific production^{10, 27, 28}, a phenomenon reflecting communicative plasticity. This is because the use of invented signals inevitably implies intentionality: naïve individuals observed and learned them, inferred their meaning from context and reactions, and subsequently used them in the same context with the same function as their original inventor but see^{29, 30} for contrary views on highly conserved signal production. In contrast, evolved signals may be used intentionally or non-intentionally. They may therefore vary in functional specificity, and show more communicative flexibility: the same communicative act serves several different goals or functions, relying on context to provide disambiguation^{2, 31}.

The aim of the present study was to examine repertoires and functional specificity of close-range communicative acts, in both wild and captive populations of orang-utans, the great ape genus which is in our view ideal for this avenue of research. Systematic studies on the gestural repertoire of captive orang-utans have demonstrated that their propensity for elaborate and flexible gesture use parallels that of other great apes^{32, 33, 34}. This suggests that social propensities can be fully expressed in captivity, as individuals do not need to be solitary in order to obtain sufficient food, whereas in the wild they may have fewer interaction opportunities and communication is hampered by arboreality and obscuring vegetation. In addition, mothers are the predominant communication partner of infant orang-utans in the wild^{35, 36}, so that there is a limited need for the production of extensive communicative repertoires. There are no systematic wild-captive comparisons of apes' communicative behaviour to date, but we assume that contrasts must be larger for orang-utans than any other great ape taxon.

We examined non-vocal (i.e. gestural and facial) communicative acts of Bornean and Sumatran orang-utans (*Pongo pygmaeus/abelii*) in two wild populations and five zoos. In a first step, we established the repertoires and functions (presumed goals of communicative acts, with outcomes that apparently satisfied the signaller) of orang-utans' non-vocal communicative acts, building on previous work conducted on chimpanzees and captive orang-utans^{33, 37}, but separately for wild and captive settings. We then tested several predictions about how setting affected individual repertoire sizes and functional specificity of signal types, while controlling for important confounding variables such as age-sex class and sampling effort. First, captivity should result in larger communicative repertoires because of boosted territoriality, sociability and interaction rates. As a result, wild repertoires should be a subset of the captive ones, except for those communicative acts that cannot be expressed in captivity ("wild-only"). Second, we expect that the form of these communicative acts expressed only in captivity should be tightly linked to the more terrestrial lifestyle or the increased

sociability, especially in the less terrestrial Sumatran orang-utans. Third, the “invention” of additional signals in captivity should be accompanied by a wild-captive contrast in functional specificity. We expect to find this contrast especially or even exclusively in Sumatran orang-utans, because sociability and interaction rates are reportedly higher in the Northwest-Sumatran population compared to the Bornean populations^{38,39}. The effect of this setting-species interaction on functional specificity will allow us to derive important conclusions on the plasticity and flexibility underlying communicative interaction in the *Pongo* genus.

Results

Communicative acts across settings

A total of 40 distinct signal types were identified across all settings, out of which 34 were observed in Bornean (captive: N = 27, wild: N = 24) and 39 in Sumatran orang-utans (captive: N = 37, wild: N = 32). Plotting the cumulative number of identified communicative acts over the course of the observation period indicated that study groups have been sufficiently sampled to grasp complete repertoires (Fig. 1, S1, S2, S3), except for two captive groups of Bornean orang-utans (Apenheul and Cologne, see Fig. S1). In Table S1 we provide definitions for all coded behaviours and their relation to previous work on orang-utans’ communicative repertoire. The majority of signal types (N) and cases (n) consisted of manual (N = 19, n = 5106) and bodily signals (N = 18, n = 2212), whereas considerably fewer facial acts (N = 3, n = 110) act were observed (see Tab. S2 for detailed overview of signal presence in relation to settings, species, subjects and age classes). The relatively small repertoire of facial signals may be partly due to our strict criteria of inclusion into the repertoire (see methods).

We first tested the first prediction that captivity should result in enlarged communicative repertoires at the aggregate level owing to boosted sociability and interaction rates. We found that the majority of communicative acts ($N = 27$) was shared across orang-utan species and research settings, but that 9 communicative acts were restricted to captivity (e.g. “roll on back”, “throw object”, “somersault”, see Fig. 2), and 2 to the wild (“loud scratch”, “shake object”, see Fig. 2), thus confirming the prediction. Out of these, seven captivity-specific and one wild-specific acts were observed in the Sumatran species only (e.g. “rub on body”, “head-butt”, “shake object”), whereas one behaviour (“spin”) was observed in captive Borneans only (see Tab. S2 for a detailed overview communicative acts across settings and species). A more conservative way of testing the prediction is by producing a list based on all previous studies. We found that three of the communicative acts we found only in captive settings (“throw object”, “rise up”) or wild settings (“shake object”), were also observed in other species-setting combinations in other studies, which leaves at least seven captivity-only acts and one wild-only act (see Tab. S3, note that Cartmill & Byrne [2010] do not specify which gestures were observed in which orang-utan species). This more conservative test thus also confirms the prediction.

To ensure that these differences between captivity and the wild do not reflect differences in social opportunities (e.g. with regard to the availability same-age play partners), we compiled separate play repertoires for mother-offspring (for which there is no change in partner availability between natural and captive settings) versus same-aged interactions (Tab. S4). A graphical analysis revealed no substantial difference between captive versus wild play repertoires with regard to mother-offspring interactions (25 vs 24 in Sumatrans, 22 vs 18 in Borneans). For peer play interactions, repertoire sizes apparently differed between settings for Sumatrans (26 vs. 19), but not Borneans (12 vs. 11). Differences in repertoire sizes between captive and wild settings are thus not driven by partner availability alone. Given that 13 communicative acts used in same-aged play interactions were

exclusively used in Sumatrans (Tab. S4), we suggest that repertoire size is driven more by social opportunities in Sumatrans than in Borneans.

Individual repertoire sizes across settings

The average Bornean orang-utans had 1.9 more signal types in its repertoire (captive: mean \pm SD = 8.1 ± 3.8 , wild: 10 ± 4.7), while the average Sumatran had 6.8 more signal types (captive: mean \pm SD = 15.3 ± 6.5 , wild: 8.5 ± 7.2), with a larger between-individual variation. With regard to age classes, we found that younger immatures had the largest repertoire on average (mean \pm SD = 14.8 ± 5.9 , N = 42), followed by older immatures (10.1 ± 8 , N = 41), and adults (mean \pm SD = 7.3 ± 3.9 , N = 37).

Using a linear mixed model (LMM), we tested how setting, species and confounding variables such as sex, age class and sampling effort affected the number of communicative acts in individuals recorded during the study (for details see methods). The full model including the key test predictors (i.e. setting and species) fitted the data better than the null models irrespective of the subsets used (LRT all individuals: $\chi^2_3 = 14.058$, $P = 0.003$, $N = 70$; individuals with > 50 interactions: $\chi^2_3 = 15.131$, $P = 0.002$, $N = 44$). As expected, the number of communicative acts were strongly affected by the number of samples contributed to the dataset (see Tab. 1 for output of the model using the restricted dataset, Tab. S5 for the model including all individuals). Irrespective of the effects of sampling effort, however, we found that captive individuals deployed a significantly larger variety of communicative acts than their wild counterparts, again confirming our prediction. We also found that individual repertoires of Sumatran individuals exceeded those of their Bornean counterparts (after removing the non-significant interaction term, Tab. 1 a, Fig. 3 a, b). In addition, younger individuals produced significantly more different communicative acts than adults (Tab. 1 a, Fig. S4). These

results are consistent with our expectations that young individuals (i.e. all immatures, but especially those below the age of 5 years) regularly use a larger communicative set than adults, particularly for the function of soliciting social play and food sharing. This is further supported by descriptive results on presumed goals and outcomes of communicative acts (see ESM, Tab. S6, S7) suggesting that interactions in both affiliative and conflict situations rather than co-locomotion or food-sharing underlie the proliferation of communicative acts in captive settings.

Functional specificity of communicative acts across settings

We systematically tested the second prediction on functional specificity of communicative acts depends on research setting and orang-utan species using a LMM, which also included confounding variables such as outcome and sampling effort. The full model including the key test predictors (i.e. setting, species, dominant outcome) fitted the data better than the null models (LRT: $\chi^2_3 = 22.612$, $P < 0.001$, $N = 114$). There was a significant interaction between research setting and orang-utan species: while specificity scores in captivity did not differ between the species, we found a significantly lower functional specificity in wild Sumatrans compared to their captive counterparts (Fig. 4). Irrespective of this result, specificity was significantly higher for communicative acts predominantly used to solicit play and observed in a larger number of subjects (Tab. 1 b). Our findings thus support the prediction that larger repertoire sizes in captivity should be accompanied by an increase in average functional specificity.

Discussion

Answering the question whether our primate relatives possess the behavioural plasticity, or creative capacity, to complement their species-typical repertoires by inventing novel signals from scratch is highly relevant to theories of language evolution. Yet, to date no study explicitly and systematically

examined communication systems of apes exposed to novel socioecological conditions relative to the wild baseline situation. Here, we adopted a 2 x 2 comparative design, investigating repertoire sizes and functional specificity of communicative acts in zoo-housed and wild groups of two different orang-utan species with different sociability and terrestriality. By examining the captive-wild contrast in these two related species we tested the prediction that captive environments favour the emergence of novel communicative acts, which should also have higher functional specificity. Moreover, comparing species differences related to differential sociability and terrestriality on one hand, and setting on the other, we expected that Sumatran orang-utans, but not Borneans, would show a wild-captive contrast in average functional specificity, offering insights into the degrees of plasticity and flexibility underlying the communicative repertoires of orang-utans.

Consistent with our first prediction, communicative repertoires on both the aggregate and individual level were larger in captivity as compared to the wild, even after controlling for the expected effects of age class and sampling effort (i.e. irrespective of whether all or only highly sampled individuals are included in the analysis). There may be some doubt that a single study can exhaustively sample signal repertoires. We therefore also compared the captive-wild contrast for each species using all available studies. This comparison supported the conclusions based on our study, in that the actual repertoire *composition* found in previous studies in the same setting and species revealed no major differences (Tab. S2, S3). First, our own findings regarding zoo repertoires (i.e. 27 different communicative acts in captive Borneans and 37 in captive Sumatrans, as compared to 24 and 32 communicative acts in the wild, respectively) are broadly consistent with the available systematic studies in single settings. Liebal and colleagues³², studying two captive groups of Sumatran orang-utans, reported a repertoire of 34 signal types (29 gestures, and 5 facial expressions). Like them, we found that the majority of communicative acts were used to solicit social play and food transfers. Cartmill & Byrne³³, examining two zoo groups of Bornean and one group of Sumatran orang-utans,

identified 38 types of gesture and facial expressions that allowed the analysis of “intentional meaning”. Second, the first systematic study on mother-offspring gesture use among wild orang-utans, conducted at the Bornean population of Sabangau Forest, identified 21 gesture types that met the criteria for inclusion into the repertoire³⁵. With 24 different observed communicative acts observed in our study population at Tuanan, it seems like the inclusion of communicative interactions outside the mother-offspring bond does not result in a substantially larger repertoire size. Thus, we can conclude that moving wild orang-utans into captivity leads to a 20 to 25% increase (i.e. 7 to 9 acts “gained”, 1 to 2 acts “lost”) in their repertoire of communicative acts.

The second prediction we made was that the form of these communicative acts expressed only in captivity should be tightly linked to the increased sociability and more terrestrial lifestyle. As expected, differences in repertoire size were particularly pronounced for presumed goals related to seeking body contact (“Play/affiliate”, “Groom”, “Sexual contact”) and social conflict (“Move away” and “Stop action”). Interaction rates with these outcomes are greatly boosted in captivity, where a more differentiated use of bodily communication is both enabled and required¹⁰. Captive facilities are stable, plentiful and predator-free environments that may provide opportunities for, and even require (e.g. due to increasing conflict with limited space) signal inventions and innovations, just like they foster innovations in general^{40, 41}. Our findings thus provide direct evidence that the new environments we have created for great apes boost the invention of new signals, which may spread through social learning. Indeed, captive settings in general have generated extensive and convincing evidence for invented (“species-atypical”) signalling, encompassing novel pant-hoot variants⁴², and “whistling”¹⁵, as well as pointing with hands and fingers^{43, 44}, “raspberries” and “extended grunts” in chimpanzees¹⁶. Although detailed captive-wild comparisons are, to our knowledge, so far lacking for other great apes, wild-captive contrasts are probably larger in orang-

utans than any other great ape taxon. In conclusion, our results support the notion that the new opportunities and needs linked to captivity may lead to a proliferation of signal invention.

We also expected that the form of these communicative acts expressed only in captivity are linked to a more terrestrial lifestyle. Indeed, we found that the communicative acts that are exclusively (or overwhelmingly) produced in captive settings are strongly linked to the more terrestrial nature of their artificial habitat: “somersaults”, “spitting”, “head-stands”, communicative acts that involve either the ground or objects obtained from the ground, would be very difficult to perform by wild orang-utans with their purely (Sumatra) or predominantly Borneo: ⁴⁵ arboreal lifestyle. This setting effect is not attributable to the presence of certain interaction partners alone: by comparing the repertoires for mother-offspring and peer play interactions we demonstrated that differences between interaction dyads with regard to wild-captive contrasts were only strong in Sumatran orang-utans. It thus appears that the new affordances of captive settings, on top of the elevated exposure to certain social contexts, enabled orang-utans to better exploit their (communicative) motion spectrum, resulting in novel communicative movements that may independently and predictably such as spitting as an attention-getter, see ⁴⁶ be invented in several captive colonies and species. This also confirms earlier reports making the case that the complex individual-based fission fusion structure of orang-utans and their sophisticated social-cognitive skills seem to be reflected in a highly variable communicative repertoire ^{32, 47}, illustrating their remarkable behavioural plasticity.

Finally, we predicted that the additional signals in captivity should be accompanied by an increase in functional specificity. In line with our predictions, we found that Sumatrans in the wild exhibited a lower average functional specificity compared to their captive counterparts, while average functional specificity in captivity did not differ between the species. When comparing Sumatrans’

communicative acts used only in captivity (“invented signals”) with those used in both research settings, captivity-only acts appeared to be on average more functionally specific (although the small sample prevented inferential analyses). In other words, wild Sumatrans seem to use their communicative acts more flexibly (i.e. redundantly) across presumed goals than in captivity, which appears to be largely due to captive Sumatrans’ use of invented context-dependent, and therefore functionally specific acts (i.e. those that have “tight meanings” according to ³³) not present in the wild repertoire. In contrast, the relatively low interactions rates (due to few social opportunities) of wild Bornean orang-utans ^{38, 39, 48} seem to be reflected in a lesser need to use signals flexibly across contexts. Together, these results corroborate our expectation that average functional specificity increased with repertoire size in captive Sumatran, though not Bornean orang-utans.

To convey a message that can be understood by a targeted recipient, an intentional agent may use two alternative “strategies”. First, she can use one and the same communicative act for several different functions, relying on context or other information (e.g. possibly age difference or sex relative to recipient) to disambiguate between ambiguous meanings. This “flexible” communicative strategy produces redundancy in the communicative repertoire. Alternatively, she can use a communicative behaviour that is specific to one interaction outcome – and, if not available, invent one for which naïve recipients (over repeated instances of interaction) infer their meaning from context and reactions, which they may subsequently use in exactly the same contexts with exactly the same function as their original inventor. This “plastic” communicative strategy produces productivity in the communicative repertoire. Sumatran orang-utans in the wild seem to rely somewhat more on the flexibility option; their captive counterparts, however, rely more on the plasticity option because repeated interactions with the same partners in captivity allow them to establish novel signal meanings more rapidly. This outcome is expected because novel signals are most likely to be understood and thus maintained when they are highly context-specific.

289

290 We argue that distinguishing species-typical from invented communicative acts matters greatly for
 291 current debates on language origins see also ¹⁰. Language reflects extreme communicative plasticity.
 292 Words that make up language are predominantly invented noises used intentionally, and the
 293 productivity feature of language fundamentally relies on the ability to produce such new noises,
 294 which are used intentionally and rapidly acquire a shared meaning due to their high context-
 295 specificity. Functionally specific signals are often more effective and efficient in complex
 296 interactional exchanges (including language), because they are less dependent on context and thus
 297 less ambiguous. Reliance on context may be less favourable with increasing message complexity
 298 (multiple intertwining messages), because it increases both (i) the risk of misunderstanding, and (ii)
 299 the “decoding work” necessary for the recipient, and thus the risk that the signal is ignored.

300

301 Although our knowledge of the taxonomic distribution of invented signals is still incomplete, they
 302 are so far reported almost exclusively for great apes. Ongoing work supports this pattern, in that
 303 great apes are increasingly documented to make up new vocal e.g. ^{15, 16, 49} and gestural e.g. ^{18, 50}
 304 ‘signals’ in the novel conditions of captivity, whereas reports from other taxa are rare but see e.g. ^{51,}
 305 ^{52, 53, 54}. Such communicative creativity may therefore be most common in, or even limited to, species
 306 with intentional communication. These findings imply that once the conditions were in place that
 307 favoured the open-ended use of invented expressions, our hominin ancestors readily responded to
 308 this opportunity, because they could build on a long evolutionary history of communicative
 309 creativity. This might explain why language evolved in the hominin lineage and not others that found
 310 themselves in similar conditions (e.g. those canids that also rely on interdependent foraging and
 311 cooperative breeding).

312

Materials and Methods

Data collection

Data were collected at two field sites and five captive facilities (zoos). We observed wild orang-utans at the long-term research sites of *Suaq Balimbing* (03°02'N; 97°25'E, Gunung Leuser National Park, South Aceh, Indonesia) and *Tuanan* (02°15'S; 114°44'E, Mawas Reserve, Central Kalimantan, Indonesia), on a population of wild Sumatran (*Pongo abelii*) and Bornean orang-utans (*Pongo pygmaeus wurmbii*), respectively. Both study sites consist mainly of peat swamp forest and show high orang-utan densities, with 7 individuals per km² at *Suaq* and 4 at *Tuanan*^{55,56}. Captive Bornean orang-utans were observed at the zoos of Cologne and Münster, and at Apenheul (Apeldoorn), while Sumatran orang-utans were observed at the zoo of Zurich and at Hellabrunn (Munich; see EEP studbook for details on captive groups; Becker 2016). While captive Sumatran orang-utans were housed in groups of nine individuals each, captive Bornean groups were smaller (on average four individuals, with the one in Apenheul including only a mother and her dependent and independent offspring). Signallers included in this study consisted of 33 Bornean (21 wild/12 captive) and 38 Sumatran orang-utans (20 wild/18 captive; see Tab. S8 for detailed information on subjects and group compositions).

Focal observations were conducted between November 2017 and October 2018 (*Suaq Balimbing*: November 2017 – October 2018; *Tuanan*: January 2018 – July 2018, European zoos: January 2018 – June 2018). At the two field sites, these observations consisted of full (nest-to-nest) or partial follows (e.g. nest-to-lost or found-to-nest) of mother-infant units, whereas in zoos 6-hour focal follows were conducted. Two different behavioural sampling methods were combined: First, presumable intra-specific communicative interactions of all observed social interactions of the focal either as signaller or receiver with all partners ($N = 7137$ acts), and among other conspecifics ($N = 888$ acts) present

were recorded using a digital High-Definition camera (Panasonic HC-VXF 999 or Canon Legria HF M41) with an external directional microphone (Sennheiser MKE600 or ME66/K6), enabling recordings of high-quality footage. In captive settings with glass barriers, we also used a Zoom H1 Handy recorder that was placed in background areas of the enclosure whenever possible. Second, using instantaneous scan sampling at ten-minute intervals, we recorded complementary data on the activity of the focal individual, the distance and identity of all association partners, and in case of social interactions the interaction partner as well as several other parameters. During ca. 1600 hours of focal observations, we video-recorded more than 6300 communicative interactions.

Coding procedure

A total of 2655 high-quality recordings of orang-utan interactions (wild: 1643, captive: 1012) were coded using the program BORIS version 7.0.4.⁵⁸. We designed a coding scheme to enable the analysis of presumably communicative acts directed at conspecifics (i.e. behaviours that apparently served to elicit a behavioural change in the recipient and were mechanically ineffective i.e. achieve a presumed goal without physical force; see also³³. Manual, bodily, and facial acts were defined and aligned (see Tab. S1) based on previous studies on orang-utan communication in captive^{26, 32, 33, 46} and wild settings^{35, 36, 59, 60}. Comparing our dataset to this literature, we then identified the subset of setting- and species-specific communicative acts. Although we also coded vocalizations based on field studies⁶¹, we did not include vocalizations in the analyses of repertoire and functional specificity as we could not equally pick up soft, low-frequency sounds in captive and wild settings, which hampered the fine-grained comparison across settings. For each communicative act, we coded the following “modifiers”: presumed goal following the distinction of³³, outcome, and other variables not directly relevant in this study (see Tab. S9 for levels and definitions of all coded variables). To ensure inter-observer reliability, we evaluated the coding performance of all observers

with alternating datasets using the Cohen’s Kappa coefficient ⁶² during an initial training period and at regular intervals afterwards. Trained observers (MF, NB, CF, CW, LM, MJ) proceeded with video coding only if at least a ‘good’ level ($\kappa \geq 0.75$) of agreement was found for signal type, articulator, sensory modality, context, and response. For our repertoire analyses, we plotted the cumulative number of communicative behaviours over the number of coded interactions for each study group (Fig. S1, S2) and for a subset of highly sampled individuals (Fig. S3), to estimate how many observations are necessary to grasp the repertoire of these groups/individuals as indicated by an asymptote. Communicative acts were counted as part of individuals’ repertoire only when observed at least twice per subject.

To analyse functional specificity, we focused on goal-outcome matches ³³ or apparently satisfactory outcomes ³⁷, that is, whether the interaction outcome aligned with presumed goals identified by observers. We considered only those signal types that were produced at least three times towards a particular interaction outcome cf. ³⁷. We defined functional specificity depending on how often a communicative act was produced towards an apparently satisfactory outcome (ASO), adopting the definitions of Cartmill and Byrne ³³. Signal types that were used mainly towards a single interaction outcome, i.e. more than 70% of the time, were defined as having “tight meanings”. Signal types used frequently towards a single ASO, i.e. 50–70% of the time, were defined as having “loose meanings”. Finally, signal types that were used less than 50% of the time towards a single ASO were classified as having “ambiguous meanings”. For instance, “somersault” was exclusively produced to initiate “Play/affiliate” interactions (specificity value of 1, tight meaning), whereas “touch” was produced towards several different interaction outcomes, e.g. “Play/affiliate”, “Share food/object” and “Co-locomote” (specificity values < 0.7 , ambiguous or loose meaning).

Statistical analyses

We ran two separate linear mixed models LMMs;⁶³ with a Gaussian error structure and identity link function to examine sources of variation in (a) individual communicative repertoires (i.e. number of signal types used at least twice) and (b) specificity in signal function. We used LMMs rather than GLMMs in this study because it has recently been shown that linear models are more robust to violations of distributional assumptions^{64, 65}. We ran model (a) for two subsets of our data: first, including all individuals regardless of sample size; second, including only those individuals that contributed more than 50 communicative acts to the dataset (as the graphical inspection of asymptote plots in Fig. S3 suggested that this is a representative number to estimate individual repertoire sizes).

In model (a), which had individual repertoire size as response variable, we included research setting (2 levels: captive, wild) and orang-utan species (2 levels: Bornean, Sumatran) as our key test predictors. Because we assumed that the effect of research setting might depend on genetic predisposition (i.e. species), we included the interaction between these two variables into our model. Moreover, immature individuals often contribute the majority of both signal cases and types², for reviews see³¹ and the composition of study groups differed with regard to age-sex classes, hence we made sure that age class was taken into account in our analyses. We included the following variables as additional fixed effects (control predictors) into the models: subjects' age class (3 levels: "adult": females > 15 years, males > 16 years; "older immature": independent and dependent immature > 5 years of age, "younger immature": dependent immature < 5 years of age), sex (2 levels: female, male), the number of interaction outcomes the subject communicated for at least twice (range = 1–6, only four outcomes were coded in captive Borneans), and the number of observations (range = 5–467). To control for repeated measurements within the same sampling unit, group identity was

treated as random effect. To keep type 1 error rates at the nominal level of 5%, we also included all relevant random slopes components within group ID ⁶⁶.

In model (b), which had specificity in signal function as response variable, we included orang-utan species (two levels: Bornean, Sumatran), research setting (two levels: captive, wild) and dominant outcome (two levels: play, non-play), and the interaction between setting and species as our key test predictors. To control for confounding effects of sampling effort, the number of subjects contributing to the use of a signal type (range = 1–21) and the number of observations (range = 1–787) in the respective setting as additional fixed effects (i.e. control predictors). To control for repeated observations of the same signal types across settings, signal type was treated as random effect ⁶⁷.

All models were implemented in R (v3.4.1, ⁶⁸) using the function *glmer* of the package lme4 ⁶⁹. To control for collinearity, we determined the Variance Inflation Factors VIF; ^{70, 71} from a model including only the fixed main effects using the function *vif* of the R package car ⁷². This revealed no collinearity issues (maximum VIF = 2.8). Prior to running the models, we log-transformed the response variables and the control variables relating to sampling effort (number of observations/subjects), to achieve an approximately symmetrical distribution and avoid influential cases.). To test the overall significance of our key test predictors ^{73, 74}, we compared the full models with the respective null models comprising only the control predictors as well as all random effects using a likelihood ratio test ⁷⁵. Tests of the individual fixed effects were derived using likelihood ratio tests (R function *drop1* with argument “test” set to “Chisq”).

Acknowledgements

General: We thank Erin Vogel and Suci Utami Atmoko (Tuanan), Kerstin Bartsch (Tierpark Hellabrunn), Claudia Rudolf von Rohr (Zoo Zürich), Alexander Sliwa (Kölner Zoo), Simone Sheka (Allwetterzoo Münster) and Thomas Bionda (Apenheul Primate Park) as well as all research staff and zoo keepers for a fruitful collaboration during this study. We gratefully acknowledge Clemens Becker for providing the EEP studbook, the Indonesian State Ministry for Research and Technology (RISTEK), the Indonesian Institute of Science (LIPI), the Directorate General of Natural Resources and Ecosystem Conservation – Ministry of Environment & Forestry of Indonesia (KSDAE-KLHK), the Ministry of Internal affairs, the Nature Conservation Agency of Central Kalimantan (BKSDA), the local governments in Central Kalimantan, the Kapuas Protection Forest Management Unit (KPHL), the Bornean Orang-utan Survival Foundation (BOSF) and MAWAS in Palangkaraya. Moreover, we thank Simone Pika and Eva Luef for providing some of the essential equipment, Santhosh Totagera for coding assistance, Uli Knief and Alex Hausmann for help with a customized jitter-plot function. **Funding:** MF was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, DFG, grant no- FR 3986/1-1) the Forschungskredit Postdoc (grant FK-17-106) and the A.H. Schultz Foundation of the University of Zurich, the Sponsorship Society of the German Primate Center (DPZ), the Stiftung Mensch und Tier (Freiburg) and the Christiane Nüsslein-Volhard Foundation. CvS acknowledges the support of the NCCR Evolving Language Program (SNF #51NF40_180888). **Author contributions:** MF and CvS conceived of the study. MF designed the project, collected, coded and analysed data. NB, CF, CW, LM, MJ helped to collect, curate and code data. TMS, CS, MvN and CvS provided resources. MF wrote the manuscript with inputs from CF, LM, CS, MvN and CvS. All authors approved the submission of the manuscript.

Competing interests: The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper and the R code can be found on GitHub: <https://github.com/MarlenF/repertoire-orang>.

References

1. Hockett CF. The origin of speech. *Scientific American* **203**, 88-97 (1960).
2. Liebal K, Waller BM, Burrows AM, Slocombe KE. *Primate Communication: A Multimodal Approach*. Cambridge University Press (2013).
3. Townsend SW, *et al.* Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biol Rev* **92**, 1427–1433 (2017).
4. Sievers C, Gruber T. Reference in human and non-human primate communication: What does it take to refer? *Anim Cogn* **19**, 759-768 (2016).
5. Maynard Smith J, Harper D. *Animal signals*. Oxford University Press (2003).
6. Bradbury JW, Vehrencamp SL. *Principles of animal communication*. Sinauer (1998).
7. Cheney DL, Seyfarth RM. Flexible usage and social function in primate vocalizations. *Proc Natl Acad Sci USA* **115**, 1974–1979 (2018).
8. Wegdell F, Hammerschmidt K, Fischer J. Conserved alarm calls but rapid auditory learning in monkey responses to novel flying objects. *Nature Ecology & Evolution* **3**, 1039–1042 (2019).
9. Fröhlich M, Hobaiter C. The development of gestural communication in great apes. *Behav Ecol Sociobiol* **72**, 194 (2018).
10. Fröhlich M, van Schaik CP. Must all signals be evolved? A proposal for a new classification of communicative acts. *Wiley Interdisciplinary Reviews: Cognitive Science* **11**, e1527 (2020).
11. Tomasello M. *Origins of human communication*. MIT press (2008).
12. Levinson SC, Holler J. The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **369**, 20130302 (2014).
13. Fröhlich M, Sievers C, Townsend SW, Gruber T, van Schaik CP. Multimodal communication and language origins: integrating gestures and vocalizations. *Biol Rev* **94**, 1809-1829 (2019).
14. Wich SA, *et al.* Call cultures in orang-utans? *PLoS ONE* **7**, e36180 (2012).

15. Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker RW. A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56-64 (2009).
16. Hopkins WD, Taglialatela JD, Leavens D. Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Anim Behav* **73**, 281-286 (2007).
17. Hardus ME, Lameira AR, Van Schaik CP, Wich SA. Tool use in wild orang-utans modifies sound production: a functionally deceptive innovation? *Proc R Soc Lond, Ser B: Biol Sci* **276**, 3689-3694 (2009).
18. Halina M, Rossano F, Tomasello M. The ontogenetic ritualization of bonobo gestures. *Anim Cogn* **16**, 653-666 (2013).
19. Fröhlich M, Wittig RM, Pika S. Should I stay or should I go? Initiation of joint travel in mother-infant dyads of two chimpanzee communities in the wild. *Anim Cogn* **19**, 483-500 (2016).
20. Tomasello M, Call J, Warren J, Frost GT, Carpenter M, Nagell K. The ontogeny of chimpanzee gestural signals: A comparison across groups and generations. *Evolution of Communication* **1**, 223-259 (1997).
21. Kummer H, Goodall J. Conditions of innovative behaviour in primates. *Philos Trans R Soc Lond, Ser B: Biol Sci* **308**, 203-214 (1985).
22. Fröhlich M, van Schaik CP. The function of primate multimodal communication. *Anim Cogn* **21**, 619-629 (2018).
23. Partan SR, Marler P. Communication goes multimodal. *Science* **283**, 1272-1273 (1999).
24. Kopp KS, Liebal K. Conflict resolution in socially housed Sumatran orangutans (*Pongo abelii*). *PeerJ* **6**, e5303 (2018).
25. Maple TL. *Orangutan behavior*. Van Nostrand Reinhold Co. (1980).
26. Zucker EL, Mitchell G, Maple T. Adult male-offspring play interactions within a captive group of orang-utans (*Pongo pygmaeus*). *Primates* **19**, 379-384 (1978).
27. Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature* **417**, 351-358 (2002).
28. Beecher MD, Brenowitz EA. Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution* **20**, 143-149 (2005).

29. Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J. Great ape gestures: intentional communication with a rich set of innate signals. *Anim Cogn* **20**, 755-769 (2017).
30. Fischer J, Hammerschmidt K. Towards a new taxonomy of primate vocal production learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20190045 (2020).
31. Call J, Tomasello M. *The gestural communication of apes and monkeys*. Lawrence Erlbaum Associates (2007).
32. Liebal K, Pika S, Tomasello M. Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* **6**, 1-38 (2006).
33. Cartmill E, Byrne R. Semantics of primate gestures: intentional meanings of orangutan gestures. *Anim Cogn* **13**, 793-804 (2010).
34. Cartmill EA, Byrne RW. Orangutans modify their gestural signaling according to their audience's comprehension. *Curr Biol* **17**, 1345-1348 (2007).
35. Knox A, *et al.* Gesture use in communication between mothers and offspring in wild orangutans (*Pongo pygmaeus wurmbii*) from the Sabangau Peat-Swamp Forest, Borneo. *International Journal of Primatology* **40**, 393-416 (2019).
36. Fröhlich M, Lee K, Mitra Setia T, Schuppli C, van Schaik CP. The loud scratch: a newly identified gesture of Sumatran orangutan mothers in the wild. *Biol Lett* **15**, 20190209 (2019).
37. Hobaiter C, Byrne Richard W. The meanings of chimpanzee gestures. *Curr Biol* **24**, 1596-1600 (2014).
38. van Schaik CP. The socioecology of fission-fusion sociality in orangutans. *Primates* **40**, 69-86 (1999).
39. Fröhlich M, *et al.* Social interactions and interaction partners in infant orang-utans of two wild populations. *Anim Behav* **166**, 183-191 (2020).
40. van Schaik CP. The primate origins of human nature. (ed[^](eds). John Wiley & Sons (2016).
41. Lehner SR, Burkart JM, van Schaik CP. An evaluation of the geographic method for recognizing innovations in nature, using zoo orangutans. *Primates* **51**, 101-118 (2010).
42. Marshall AJ, Wrangham RW, Arcadi AC. Does learning affect the structure of vocalizations in chimpanzees? *Anim Behav* **58**, 825-830 (1999).

- 580
- 581 43. Leavens DA, Hopkins WD. Intentional communication by chimpanzees: A cross-sectional
- 582 study of the use of referential gestures. *Dev Psych* **34**, 813-822 (1998).
- 583
- 584 44. Leavens DA, Russell J, Hopkins W. Multimodal communication by captive chimpanzees
- 585 (Pan troglodytes). *Anim Cogn* **13**, 33-40 (2010).
- 586
- 587 45. Ashbury AM, *et al.* Why do orangutans leave the trees? Terrestrial behavior among wild
- 588 Bornean orangutans (*Pongo pygmaeus wurmbii*) at Tuanan, Central Kalimantan. *Am J*
- 589 *Primatol* **77**, 1216-1229 (2015).
- 590
- 591 46. Jantschke F. *Orang-Utans in Zoologischen Gärten*. Piper (1972).
- 592
- 593 47. Maestriperi D. Primate social organization, gestural repertoire size, and communication
- 594 dynamics. In: *The Origins of Language: What Nonhuman Primates Can Tell* (ed[^](eds King
- 595 BJ). School of American Research Press (1999).
- 596
- 597 48. Mitra Setia T, Delgado R, Utami Atmoko S, Singleton I, van Schaik CP. Social organization
- 598 and male-female relationships. In: *Orangutans: Geographic variation in behavioral ecology*
- 599 *and conservation* (ed[^](eds Wich SA, Utami Atmoko S, Mitra Setia T, van Schaik CP).
- 600 Oxford University Press (2009).
- 601
- 602 49. Taglialatela JP, Reamer L, Schapiro SJ, Hopkins WD. Social learning of a communicative
- 603 signal in captive chimpanzees. *Biol Lett* **8**, 498-501 (2012).
- 604
- 605 50. Tomasello M, Call J, Nagell K, Olguin R, Carpenter M. The learning and use of gestural
- 606 signals by young chimpanzees: A trans-generational study. *Primates* **35**, 137-154 (1994).
- 607
- 608 51. Moura ACdA. Stone banging by wild capuchin monkeys: an unusual auditory display. *Folia*
- 609 *Primatol* **78**, 36-45 (2007).
- 610
- 611 52. Garland Ellen C, *et al.* Dynamic Horizontal Cultural Transmission of Humpback Whale Song
- 612 at the Ocean Basin Scale. *Curr Biol* **21**, 687-691 (2011).
- 613
- 614 53. Grant BR, Grant PR. Cultural inheritance of song and its role in the evolution of Darwin's
- 615 finches. *Evolution* **50**, 2471-2487 (1996).
- 616
- 617 54. Perry S, *et al.* Traditions in wild white-faced capuchin monkeys. In: *The biology of*
- 618 *traditions: Models and evidence* (ed[^](eds Fragaszy DM, Perry S). Cambridge University
- 619 Press (2003).
- 620
- 621 55. Singleton I, Knott C, Morrogh-Bernard H, Wich S, van Schaik C. Ranging behavior of
- 622 orangutan females and social organization. In: *Orangutans: Geographic variation in*

- 623 *behavioral ecology and conservation* (ed[^](eds Wich SA, Utami-Atmoko SS, Mitra Setia T,
624 van Schaik CP). Oxford University Press (2009).
- 625
- 626 56. Husson SJ, *et al.* Orangutan distribution, density, abundance and impacts of disturbance. In:
627 *Orangutans: Geographic variation in behavioral ecology and conservation* (ed[^](eds Wich
628 SA, Utami-Atmoko SS, Mitra Setia T, van Schaik CP). Oxford University Press (2009).
- 629
- 630 57. Becker C. EEP studbook for zoo-housed orang-utans. (ed[^](eds) (2016).
- 631
- 632 58. Friard O, Gamba M. BORIS: a free, versatile open-source event-logging software for
633 video/audio coding and live observations. *Methods in Ecology and Evolution* **7**, 1325-1330
634 (2016).
- 635
- 636 59. MacKinnon J. The behaviour and ecology of wild orang-utans (*Pongo pygmaeus*). *Anim*
637 *Behav* **22**, 3-74 (1974).
- 638
- 639 60. Rijksen HD. *A fieldstudy on Sumatran orang utans (Pongo pygmaeus abelii, Lesson 1827):*
640 *Ecology, behaviour and conservation*. H. Veenman (1978).
- 641
- 642 61. Hardus M, *et al.* A description of the orangutan's vocal and sound repertoire, with a focus on
643 geographical variation. In: *Orangutans: Geographic variation in behavioral ecology and*
644 *conservation* (ed[^](eds Wich SA, Utami-Atmoko SS, Mitra Setia T, van Schaik CP). Oxford
645 University Press (2009).
- 646
- 647 62. Bakeman R, Quera V. *Sequential analysis and observational methods for the behavioral*
648 *sciences*. Cambridge University Press (2011).
- 649
- 650 63. Baayen RH. *Analyzing linguistic data*. Cambridge University Press (2008).
- 651
- 652 64. Schielzeth H, *et al.* Robustness of linear mixed[□]effects models to violations of distributional
653 assumptions. *Methods in Ecology and Evolution*, (2020).
- 654
- 655 65. Knief U, Forstmeier W. Violating the normality assumption may be the lesser of two evils.
656 *bioRxiv*, 498931 (2018).
- 657
- 658 66. Schielzeth H, Forstmeier W. Conclusions beyond support: overconfident estimates in mixed
659 models. *Behav Ecol* **20**, 416-420 (2009).
- 660
- 661 67. Pinheiro JC, Bates DM. Linear mixed-effects models: basic concepts and examples. *Mixed-*
662 *effects models in S and S-Plus*, 3-56 (2000).
- 663
- 664 68. R Development Core Team. R: A language and environment for statistical computing.
665 (ed[^](eds). R Foundation for Statistical Computing (2017).

69. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. *R package version 11-7*, (2014).
70. Quinn GP, Keough MJ. *Experimental design and data analysis for biologists*. Cambridge University Press (2002).
71. Field A. *Discovering statistics using SPSS*. Sage publications (2005).
72. Fox J, Weisberg S. *An R companion to applied regression*, 2 edn. Sage (2011).
73. Forstmeier W, Schielzeth H. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav Ecol Sociobiol* **65**, 47-55 (2011).
74. Mundry R. Statistical Issues and Assumptions of Phylogenetic Generalized Least Squares (PGLS). In: *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology* (ed[^](eds Garamszegi L). Springer (2014).
75. Dobson AJ. *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC (2002).

Figures and Tables

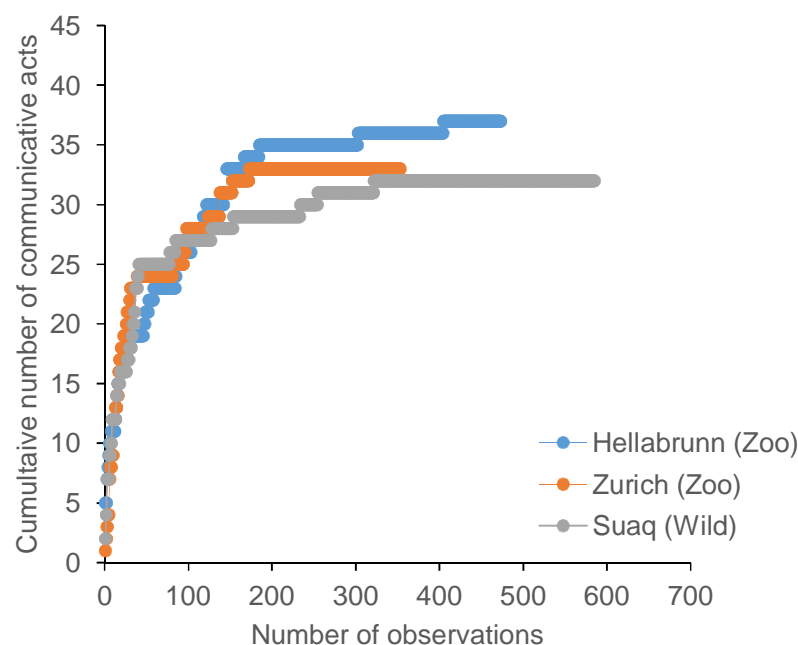


Fig. 1. Cumulative number of identified communicative acts over observation time. Asymptotes are depicted separately for the captive subjects (Hellabrunn, Zurich) and wild populations (Suaq) of Sumatran orang-utans in this study.

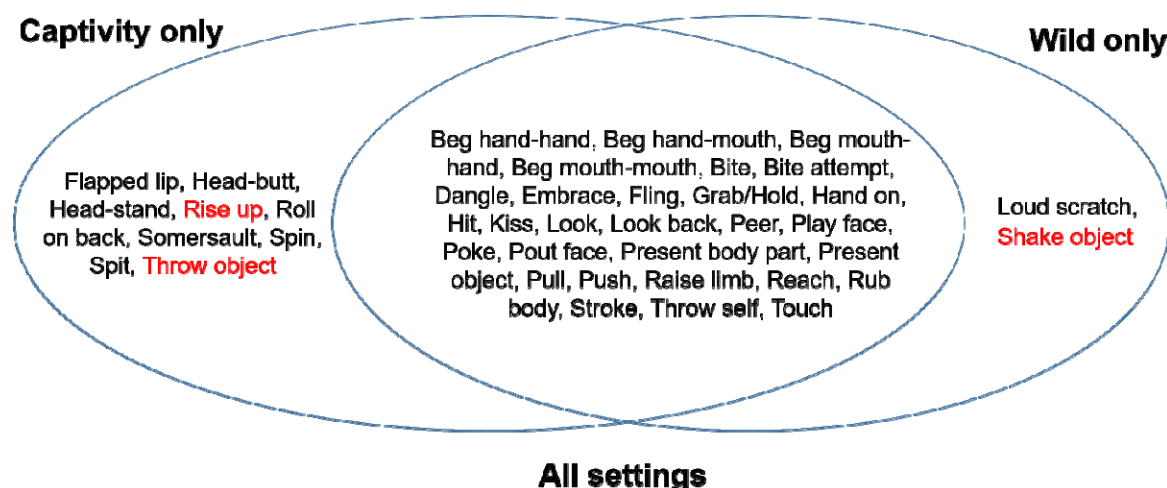
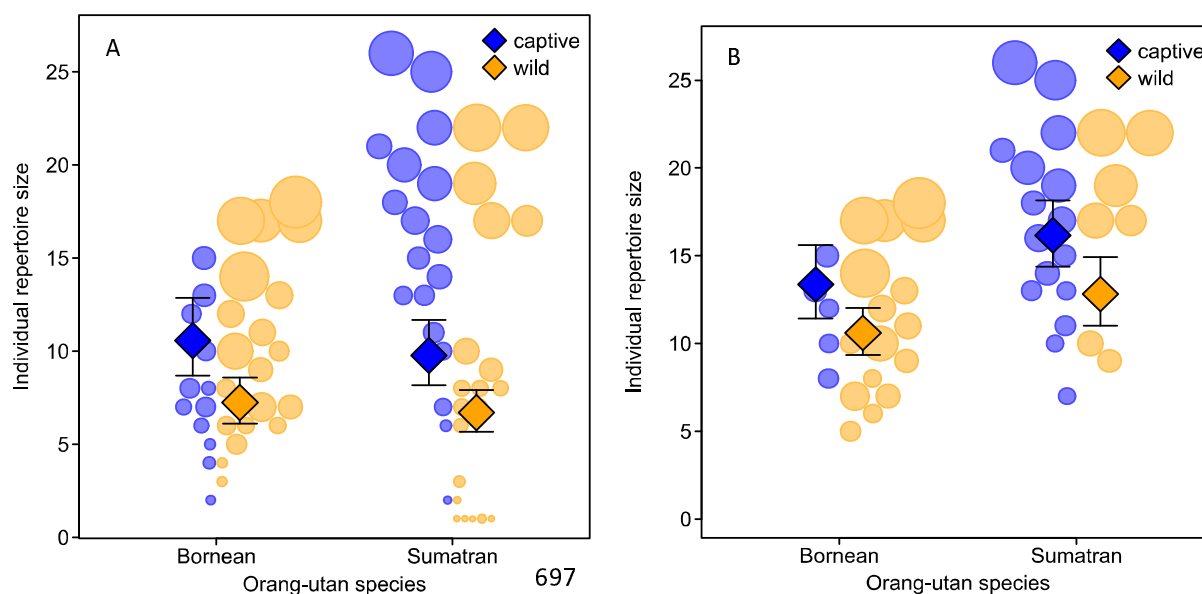


Fig. 2. Overview of communicative acts observed only in different research settings.

Communicative acts observed in the contrasting setting by other studies are marked in red.

696



697

698 **Fig. 3. Number of observed types of communicative acts per individual as a function of**
699 **research setting and orang-utan species, with (A) all subjects included, and (B) restricted to**
700 **subjects with > 50 samples. Circles represent different individuals with area corresponding to sample**
701 **size, Diamonds depict model estimates with 95% confidence intervals (all other variables centered to**
702 **a mean of zero).**

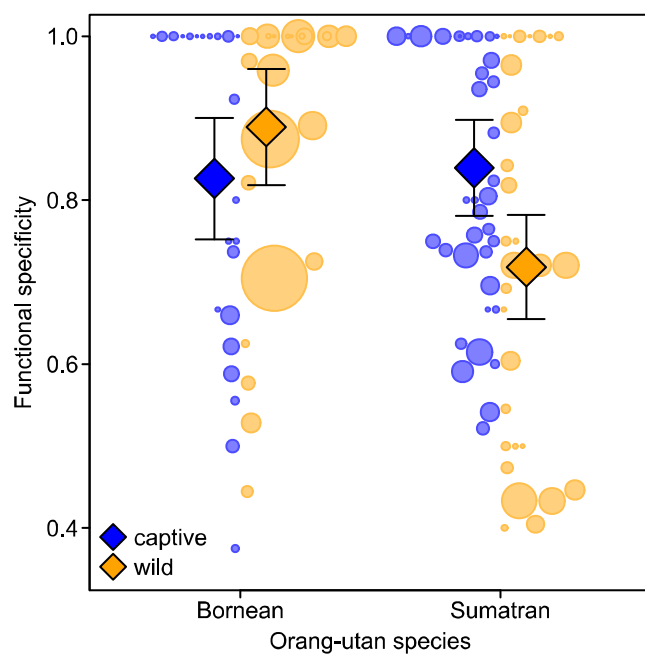


Fig. 4. Functional specificity of communicative acts as a function of research setting and orang-utan species. Circles represent different signal types with area corresponding to sample size, diamonds depict model estimates with 95% confidence intervals (all other variables centered to a mean of zero).

Tab. 1. Effects of research setting, orang-utan species, and control variables on (a) repertoire size of individuals (N = 44), and (b) functional specificity of communicative acts (N = 114), derived using LMMs with a Gaussian error structure and identity link function. Significant effects ($P < 0.05$) are depicted in italics.

| (a) Repertoire size | Estimate | SE | χ^2_1 | P |
|-----------------------------------|---------------|--------------|---------------|------------------|
| Intercept | 0.453 | 0.227 | - | - |
| <i>Setting [wild]</i> | <i>-0.29</i> | <i>0.069</i> | <i>12.727</i> | <i><0.001</i> |
| <i>Species [Sumatran]</i> | <i>0.148</i> | <i>0.061</i> | <i>6.641</i> | <i>0.010</i> |
| <i>Age class [young imm.]</i> | <i>0.314</i> | <i>0.076</i> | <i>14.571</i> | <i><0.001</i> |
| Age class [old imm.] | 0.243 | 0.138 | 3.396 | 0.065 |
| Sex [male] | -0.039 | 0.064 | 0.223 | 0.637 |
| No. functions | 0.058 | 0.03 | 3.654 | 0.056 |
| <i>Log (no. observations)</i> | <i>0.361</i> | <i>0.05</i> | <i>39.426</i> | <i><0.001</i> |
| (b) Functional specificity | | | | |
| Intercept | 0.898 | 0.06 | - | - |
| Setting [wild] | 0.059 | 0.043 | - | - |
| Species [Sumatran] | 0.002 | 0.039 | - | - |
| <i>Dominant outcome [play]</i> | <i>0.095</i> | <i>0.038</i> | <i>5.911</i> | <i>0.015</i> |
| Log (no. observations) | 0.004 | 0.019 | 0.043 | 0.836 |
| <i>Log (no. subjects)</i> | <i>-0.072</i> | <i>0.034</i> | <i>4.578</i> | <i>0.032</i> |
| <i>Setting x species</i> | <i>-0.175</i> | <i>0.055</i> | <i>10.395</i> | <i>0.001</i> |