Title: Connecting primate gesture to the evolutionary roots of language: a systematic review

Short title: Primate gesture and language evolution

Authors: Evelina D. Rodrigues [1], António J. Santos [1], Flávia Veppo [1], Joana Pereira [2], Catherine Hobaiter [3]

Affiliations and Addresses:

[1] William James Center for Research, ISPA – Instituto Universitário, Portugal;

[2] Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

[3] School of Psychology and Neuroscience, St Mary's College, University of St Andrews, St Andrews, Scotland

Corresponding Author: Evelina D. Rodrigues

Email address: erodrigues@ispa.pt

Address: Rua Jardim do Tabaco, nº34, 1149-041 Lisboa

Telephone: +351 218 811 700

Abstract:

Comparative psychology provides important contributions to our understanding of the origins of human language. The presence of common features in human and non-human primate communication can be used to suggest the evolutionary trajectories of potential precursors to language. However, to do so effectively, our findings must be comparable across diverse species. This systematic review describes the current landscape of data available from studies of gestural communication in human and non-human primates that make an explicit connection to language evolution. We found a similar number of studies on human and non-human primates, but that very few studies included data from more than one species. As a result, evolutionary inferences remain restricted to comparison across studies. We identify areas of focus, bias, and apparent gaps within the field. Different domains have been studied in human and non-human primates, with relatively few non-human primate studies of *ontogeny* and relatively few human studies of gesture *form*. Diversity in focus, methods, and socio-ecological context fill important gaps and provide nuanced understanding, but only where the source of any difference between studies is transparent. Many studies provide some definition for their use of gesture; but definitions of gesture, and in particular criteria for intentional use, are absent in the majority of human studies. We find systematic differences between human and non-human primate studies in the research scope, incorporation of other modalities, research setting, and study design. We highlight 8 particular areas in a call to action through which we can strengthen our ability to investigate gestural communication's contribution within the evolutionary roots of human language.

Keywords: gestures, language evolution, domains, humans, non-human primates

1. Introduction

Human language is a rich system of communication often argued to be qualitatively distinct from other animal species' communication (Christiansen & Chater, 2015; Hauser et al., 2002; Pinker & Jackendoff, 2005). The thousands of modern human languages are distinct in their expression, so much so that being a proficient language user in one language allows access to, at best, just a handful of others. Despite this, across the spectrum of modern human languages, each one contains a similar capacity for language-like communication. How and when language emerged remains under significant debate (e.g., Corballis 2002; Richerson et al. 2010; Bolhuis et al. 2014); but it was likely built on precursors, some of which may be detected in the communication of our modern primate relatives.

The availability of fossil evidence to resolve these debates is limited (Hsieh & Plotnick, 2020; although cf. Arensburg et al. 1989; MacLarnon and Hewitt 1999); and comparative studies of modern primate species' communication provide valuable insight. The presence and absence of features within human and non-human primate communication suggest potential precursors to modern systems of communication, and the pattern of presence across primate phylogeny can offer a potential timescale for their emergence. Comparisons within and across species typically rely on comparison across studies. Diversity in study methodology – from study species to socio-ecological context and more – is critical to properly exploring the full landscape of primate gestural expression. However, such diversity makes subsequent comparison across studies challenging, particularly where definitions and conceptual boundaries are opaque. In this systematic review we investigate those studies of spontaneous gestural communication in human and non-human primates that articulate an explicit connection between gesture

and the evolutionary origins of modern human language. We incorporate a first use of Bourjade et al.'s conceptual framework for systematic comparison of gesture definitions (Bourjade et al., 2020) and investigate variation in different domains of research, in the study scope, in the inclusion of other signal sources (e.g., vocalizations), in the research setting, and in study design. We aim to provide an up-to-date description of the field, highlighting both what is understood and the areas in need of further research.

Language can be expressed in many forms, including spoken and signed: it is not the signal modality or channel of communication that defines human language so much as the way in which it is used. Many systems of communication across species encode sophisticated information, but non-human communication is typically broadcast irrespective of a recipient's attention, interest, or even presence (Rendall et al., 2009). Detecting language-like intentional communication is challenging because it depends not on the observable physical form of the signal but on the cognitive intention of the signaller. Imagine driving along a road and hearing another driver honking their horn as they approach you; there is no fixed information encoded in that signal. Unlike the acoustic structure of a monkey alarm-call (e.g., Seyfarth and Cheney 2003a, b), the uses of a car horn can mean very different things depending on what the signaller *intends* them to mean.

Intentional use, while a fundamental property of human language, remains apparently rare in communication of other species, including in many primate vocalizations (Rendall et al., 2009; Seyfarth & Cheney, 2003; Seyfarth & Cheney, 2003; although see Schel et al., 2013; Townsend et al., 2017). There is an exception: evidence for flexible intentional use is abundant in non-human ape (hereafter ape) gesture (Leavens, Russell, et al., 2005; Plooij, 1978; Tomasello et al., 1985), driving interest in

92 the evolutionary connections between ape gesture and human language, and 'gesture-
93 first' hypotheses of language evolution (Corballis, 2002; Hewes et al., 1973; Rizzolatti &
94 Arbib, 1998). More recent articulations recognise that language – like all animal systems
95 of communication – is multimodal, and likely derived from multimodal systems
96 (Gillespie-Lynch et al., 2014; Leavens, Russell, et al., 2010; Prieur et al., 2020;
97 Taglialatela et al., 2011) but may have included a transition in the role of the different
98 modalities, for example a shift in the vocal modality from supporting to carrying of
99 information (e.g., Fröhlich et al. 2019).

100 Comparative studies seeking to draw specific comparisons between primate
101 gesture and human language have been used to explore different aspects of primate
102 species' gesturing including the physical *form* (as compared to linguistic lexicons; e.g.,
103 Brentari et al., 2012) and *meaning* (as compared to language-like semantics) of gestural
104 signals (often through the study of message and context; e.g., Graham et al., 2018). The
105 *structure* of gestural communication (in sequences of gestures; as compared to
106 combinatorial structure and syntax in language; e.g., Hall et al., 2015), and the integration
107 of gestural signals with other signal types, such as vocalizations and facial expression
108 (*combination of sources;* e.g., Hobaiter et al., 2017a). From the perspective of the
109 signaller and recipient, researchers have investigated how gesture develops behaviourally
110 during *ontogeny* (e.g., Salo et al., 2018) and neurologically (*neural processes;* e.g., Biau
111 et al., 2016), and similarities between how gesture and language are deployed (for
112 example in brain or limb *laterality;* e.g., Meguerditchian et al., 2011).

113 One complication within gestural research, is that fact that researchers have no
114 direct access to cognitive states (of either non-human or human subjects), and instead
115 employ visible behavioural criteria to infer signaller intentions. These behavioural criteria

were first developed in studies of pre-verbal human infants' ability to capture the attention of others and manipulate their behaviour (Bates et al., 1975, 1979). Today, criteria include behaviour such as whether the signaller shows sensitivity to the attentional state or composition of the audience, whether they pause (wait) for a response, and if they persist or elaborate when the recipient fails to respond (Leavens, Russell, et al., 2005; Liebal et al., 2004; Tomasello & Call, 1997). However, as the study of non-human primate gesture developed there has been variation in how these criteria have been operationalised and employed (Bourjade et al., 2020; Fröhlich et al., 2018; Leavens, Russell, et al., 2005).

Our ability to reliably detect patterns of similarity and distinction across modern primate species' communication is central to our ability to make inferences about the evolutionary trajectory of language. Variation in research settings, methods, or contexts can represent a strength, allowing for robust exploration of a particular finding. However, for this to be the case it is key that diverse methods are transparent about sources of variation (Bourjade et al., 2020; Fröhlich & Hobaiter, 2018). Characteristics of our study sample such as social background, responsiveness, or prior experience impact the generalizability of our findings (c.f. STRANGE framework, Webster & Rutz, 2020), and the over-representation of particular species or populations distort our ability to make phylogenetic comparisons (for example WEIRD – Western, Educated, Industrialized, Rich, and Democratic – humans, or BIZARRE – Barren Institutional Zoo And other Rare Rearing Environment – chimpanzees; Henrich et al., 2010; Leavens, Bard, et al., 2010). Previous studies have highlighted how systematic species-differences in individual history, tasks, and testing environments are confounded with apparent species-differences in communicative or other socio-cognitive abilities, such as their ability to follow gaze or produce pointing (Boesch, 2020; Leavens et al., 2019).

Differences in methodology and context of study appear particularly profound when comparing human and non-human primate behaviour (Bard & Leavens, 2014; Leavens et al., 2019). Some of these differences may involve, for example, comparisons of institutionalized adult apes with non-institutionalized human children, or apes in caged environments with free-roaming human children (Bard & Hopkins, 2018; Leavens et al., 2019; Boesch, 2020). In some cases variation in our understanding across species is limited by what is both technologically feasible and/or ethical in non-human species, for example: the exploration of *neural processes* inside of living brains (c.f. Rizzolatti and Arbib 1998; Meguerditchian et al. 2010).

To explore meaningful patterns of similarity and distinction between human language and non-human gestural communication, we need to address apparent discrepancies in research approach and understanding. A crucial first step in this process is to better understand where any differences currently lie. A systematic assessment of the field allows us to better gauge the impact of any biases on our ability to develop clear hypotheses about the evolutionary trajectory of gesture and language. Bourjade and colleagues recently (2020) developed a framework to allow the systematic comparison of gestural definitions across primate studies – incorporating body parts, sensory modalities, social expression, and communicative and intentional properties. We include a first use of this tool, describing how primate species and study domains of gestural research vary in their concept of gesture, and then assessing how the species and study domains are differently represented in terms of study scope, the inclusion of additional sources such as vocalization or facial expression, and in research settings and study design. With this review we aim to 1) Identify both the areas of focus and apparent gaps within the field in studies that explore the connection between gestural communication and the evolution of

164 human language, and 2) Identify to what extent useful comparison can be made across
165 human and non-human studies at the present time and make recommendations for the
166 future.
167
168
169 **2. Methods**
170 In March 2020, we conducted a search of peer-reviewed articles and book chapters in
171 two search engines: Web of Science and PsycINFO. We used the SPIDER framework
172 (Cooke et al., 2013) as the search tool to define our question scope and organize and list
173 terms by the main concepts in the search question (Table 1).
174
175 **2.1. Search**
176 We used the Phenomenon of Interest and the Evaluation categories from the SPIDER
177 framework (Table 1) for our search string. We employed the largest time window allowed
178 (1900 to 2019) and used both Web of Science and PsychInfo databases. While our search
179 window extended back to 1900, more recent work is more thoroughly indexed in
180 electronic databases, and as a result our search procedure may have failed to detect some
181 earlier studies. Literature within Web of Science is systematically structured literature
182 from the 1950s and both Web of Science and PsychInfo include books and other material
183 within the Social Sciences and Humanities that are out of copyright. We did not apply
184 search terms related to the sample (e.g., human, non-human, or even primates) at this
185 stage, because research conducted on human (as compared to non-human) primates does
186 not typically specify taxonomic terms to define the sample. Similarly, the type of study
187 terms described in the design category (e.g., observational, experimental) is also often

188 omitted in human work, so we removed this criterion in the first selection phase. Although
189 no language restriction was applied at this stage, only studies with English abstracts or
190 keywords were returned by the search because the search terms were in English. In
191 PsychINFO the search term 'gestur*', AND 'evolutio*' OR 'origin*', AND 'languag*'
192 or 'communicat*' was used as a filter in the title (TI), abstract (AB), or keyword (KW).
193 In Web of Science, the same search terms were used as a topic (TS) filter (equivalent to
194 title, abstract and keywords in PsychInfo). The final search string used in Web of Science
195 was: (TS=gestur* AND TS= (languag* OR communicat*) AND TS= (evolutio* OR
196 origin*)).
197

198 **2.2. Inclusion and Exclusion criteria**

199 We included publications in the review if they (a) included data from primates, (b) had
200 gestures as a main focus, (c) make explicit the link between their study and language
201 evolution; and (d) relied on spontaneous communication. We included both journal
202 articles and book chapters. Whole books (as a single 'unit') were not included because
203 they typically include a range of differently structured studies, so considering them as
204 their individual chapters was more compatible with the data extraction for journal articles.
205 While all publications had to explore at least gestural signals, we also considered those
206 employing a "multi-source" approach (extracting the information on data from other
207 sources, such as vocalizations, for analysis).

208 We restricted our review of studies on primate gesture to publications that
209 included an explicit reference to the evolution (or origin) of language (or communication).
210 In order to assess the impact of this restriction on the literature returned we ran the same
211 search excluding this requirement.

To be considered 'spontaneous' gestural communication, gestures must have been produced without explicit training. We excluded signed languages because they represent language in its full modern form and, like spoken languages, they are taught to some extent. In contrast, home-sign often develops spontaneously in an environment in which one person does not have easy access to the language modality of others around them. Home-sign studies can provide a unique opportunity to investigate the emergence of a combinatorial linguistic system (Goldin-Meadow & Yang, 2017). While home-sign becomes formalised over time, it can incorporate the use of a range of spontaneous gesticulation, gestures, pantomime, and symbols and we retained these studies in this review. Our primary exclusion criteria for the publications were (a) publications with abstract, or full text not available; (b) publications written in languages other than English; publications not focused on (c) primates, (d) neuro-typical subjects, (e) gestures, (f) spontaneous communication, (g) publications without empirical data (e.g., theoretical); and (h) not about language evolution. We checked publications according to these criteria in the order described above, and excluded a publication as soon as they failed to fulfil any criterion (although in practice there may have been further additional reasons for their exclusion).

2.3. Selection of material

Figure 1 shows the selection process of the articles through the FLOW diagram. Each manuscript was independently and randomly assessed by at least two researchers in the two main stages of material selection: abstract screening and full text screening. When two people disagreed or both were unsure about its inclusion, we solicited a third opinion (abstract screening: 127 of 963 publications; full text screening: 78 of 221 publications).

We located 697 publications (692 articles and 5 book chapters) in Web of Science and 464 in PsychINFO (362 articles and 102 book chapters). We removed 198 duplicates, leaving 963 publications, including 862 articles and 101 book chapters (for a full list of all 963 publications see ESM1, all data are also available in an open repository at https://github.com/Wild-Minds/GestureStudies_SystematicReview); the oldest publication was published in 1975, and the most recent in 2019. Running the same searches but omitting the requirement for evolution/origin of language/communication from the string returned 6440 publications.

After abstract screening of the 963 publications, we removed a further 742 publications, leaving 221 publications for full text screening. In the full text screening phase, we excluded a further 58 publications because they did not fit the inclusion criteria (see detailed explanation in Figure 1). We excluded publications not focused on primates or focused on neuro atypical subjects during the abstract screening phase; however, the language evolution criterion was only applied at the full-text screening phase. Our final dataset included 163 publications (156 articles; 7 book chapters; for a full list of records see ESM2).

2.4. Fields for data extraction

Fields for data extraction and their categories (Table 2) were developed following systematic assessment of a pilot search of the literature by EDR (n=383 studies) and were then discussed to consensus between co-authors before being applied to the data set from the full systematic review. Please note that through the methods and results of the manuscript, capitalized terms refer to these fields and categories of information extracted for analysis.

We recorded which Species of primate(s) was included in the study. Primate species were grouped into two categories: Human and Non-human primate categories, and within Non-human primates we further distinguished the following groups: Great-apes (Chimpanzees, Bonobos, Gorillas, Orang-utans), Small-apes, Afro-Eurasian monkeys (also referred to as Old World), Monkeys of the Americas (also referred to as New World), and Strepsirrhines.

We described the proximity of the environment in which data were collected to the socio-ecological environment of evolutionary adaptation (EEA) for Non-human primate species. Given the extent of recent rapid anthropogenic change this can be challenging to define for any modern species; however, for Non-human primates we provide a cautious description of the environment as either Species-Typical (Wild, non-anthropogenically modified) or Species-Atypical (Man-made, anthropogenically modified). All captive Non-human primates were considered to be in Species-Atypical environments, wild primates were only considered to be in Species-Atypical environments where data collection occurred in anthropogenically modified habitats (e.g., in crop-fields, human settlements).

We do not feel that any similar distinction can be made for modern Human populations, both because there is no solid evidence for what our human EEA looks like, and because all humans today – whether urban city or hunter-gatherer communities – live in habitats that are substantially anthropogenically modified from that EEA. However, we provide an indication of local socio-economic structure as being WEIRD (Henrich et al. 2010), Non-WEIRD industrial, or Small-scale non-industrial. Please note that the category of subjects covered by “Rich” is a global one, and likely includes a range of economic groups in Western Industrialised Countries. These socio-economic categories

284 are not directly comparable to the Non-human primate ones of Species-Typical or
285 Atypical.

286 We defined the Research Domain(s) explored, asking which area(s) of gestural
287 communication were included in the study (for example: *form*, *structure*, or *ontogeny*;
288 Table 3 for full list and definitions).

289 We asked whether there was an explicit Definition included for Gesture (see Table
290 2). For the publications which provide an explicit definition of gesture, we used the
291 conceptual tool proposed by Bourjade et al (2020) to analyse the requirements for a given
292 behaviour to be categorized as a gesture. The authors provided 22 criteria covering 5 main
293 areas: the body parts used to gesture, the sensory modalities mobilized by the gesture, the
294 characteristics of its social expression, and its communicative and intentional properties
295 (see ESM 2 for full details).

296 Of those studies that provide an explicit definition of gesture, some include in this
297 definition a requirement that they must fulfil certain intentionality criteria (captured by
298 the 'communicative and intentional properties' aspect of Bourjade et al.'s framework).
299 However, other studies test for the presence or absence of these behavioural markers, and
300 do not include them within the definition of gesture (if provided). To capture these cases
301 we further extracted whether or not there was a definition of or criteria for Intentional
302 Communication, independently of their inclusion in any gesture definition. We also
303 extracted information on the Body parts involved in gesturing even if they were not
304 explicitly mentioned in the definition. We did not include Body parts recorded as involved
305 in non-gestural signals in this field, for example orofacial body parts used in facial
306 expressions. We marked a study as Multi-source if it included information on

307 Vocalizations, Facial expressions, and/or Lexigrams (arbitrary visual symbols
308 representing words) used in addition to the Gestures.

309 In the majority of the animal communication literature, including some primate
310 studies, the term 'multi-modal' has been employed to refer to the combination of
311 information from different sensory channels (e.g. visual, auditory, tactile; Partan and
312 Marler 1999; Partan 2002; Micheletta et al. 2013). However, within non-human ape
313 communication this term is sometimes used to refer to the combination of different signal
314 sources (e.g. gesture, vocalization, and/or facial expression; Pollick and de Waal 2007;
315 Waller et al. 2013; Wilke et al. 2017). To avoid confusion, here we follow the wider use
316 and employ the term 'multi-modal' to refer to the combination of sensory channels, and
317 the term 'multi-source' to refer to the combination of signal types.

318 We took into account the Gestural scope of the study. Here, we defined studies
319 that explored a specific context, or limited set of contexts (for example sexual solicitation
320 and consortship), specific gestures, or limited types of gesture (for example tactile
321 gestures, or specific gesture forms such as pointing) as Narrow. We defined studies that
322 explored a question across contexts and gesture repertoires without further specification
323 as Broad.

324 We then asked what Research setting data were collected from. We focused on an
325 individual's familiarity with the environment in which the study data were collected. We
326 defined two main categories: Familiar and Unfamiliar settings. These aimed to capture
327 whether or not the communication studied took place in an environment which was
328 similar to that in which the subjects spent their daily lives. In humans, Familiar settings
329 corresponded to studies conducted in their Homes and Day-care centres, and in Non-
330 human primates, corresponded to studies conducted in the areas of daily residence,

whether in: Wild, Semi-wild, or Captive settings. In both Humans and Non-human primates, Unfamiliar settings corresponded to studies conducted in Laboratories (for example, specific research-only rooms). Even when Non-human primates visited research-only rooms most days, the setting was classified as Unfamiliar since individuals only spent a small portion of their day in these rooms.

We defined Study design as either Experimental or Observational. In Observational studies, researchers did not manipulate or interact with the subjects and/or their socio-ecological environment (beyond being present). We considered any kind of socio-ecological manipulation across Research Settings as an Experimental design.

2.5. Data extraction

'Research domain' was the only field involving a potentially subjective judgement, so in all cases two of the authors extracted this field, and any discrepancies were argued until consensus between the two original raters. A third independent opinion was solicited (38 of 221 publications) when the two original raters could not reach consensus, or when the initial disagreement between raters involved more than one domain.

2.6. Data analysis

We compared the number of studies focused on the different research domains and conducted in Human and Non-human primates using Chi-square tests. The relative contributions of each cell in the computation of chi-square tests are reported through standardised residuals (Std. res.). For changes in the number of studies conducted in each domain over time (response variable) we used a Generalized Linear Model with a Poisson

error distribution and log link function. Test predictors included the interaction between domain and year. As an overall test of the effect of the year in the number of studies conducted in the different domains, we compared the full model's deviance with that of a null model comprising only the intercept, and the control predictor (domains). To test the differences across domains we compared the full model's deviance with that of a reduced model lacking the interactions (between domains and year). All statistical tests were two-tailed and run in R version 4.0.2.

3. Results

3.1. Primate species

Which primate species are represented in the review?

Within the studies of primate gesture that make an explicit connection to understanding the evolution of language we found a similar number included data on Human (N=80) or Non-human (N=87) primates, but only a small set (N=4) included empirical data on both Human and Non-Human primates. Fifteen of the Non-human primate studies included more than one species. Of the Non-human primate species studied, the majority focused on Great apes (N=75/87, 86%) with a strong bias towards Chimpanzees (N=56/87, 64%; see Figure 2). We found no studies of gesture that made an explicit connection to the evolution of language in Small apes, Monkeys of the Americas, or Strepsirrhines.

How does Socio-ecological environment vary within Non-human primates?

Of the 87 studies of Non-human primates, 57 (66%) were conducted in Captive anthropogenic Species-Atypical environments, and 30 (35%) were conducted within

Wild non-anthropogenic Species-Typical environments. No studies included data from Wild primates in anthropogenic Species-Atypical environments.

How does socio-economic culture vary within Humans?

The majority of Human participants were from WEIRD societies (N=68/80, 85%); ten studies included Human participants from Non-WEIRD industrial societies (N=10/80, 13%), and two studies included Human participants living in Small-scale non-industrial societies (N=2/80, 3%).

3.2. Research domains

Which research domains are studied and how does this change over time?

The majority of studies (N=153/163, 94%) included work on several of the seven research domains (total number of research domains recorded N=429; Table 4). The number of studies varied across these domains (Chi-square test: X2= 161.4, df=6, P<0.001). *Meaning* represented the most prolific domain of study (Std. res.=10.722; N=139/163, 85%), followed by *form* (Std. res.=2.582; N=80/163, 49%). Studies focused on *laterality*, *neural processes* and *structure* were less frequent (Std. res.=-3.213, -5.972, -3.765 respectively) and with *neural processes* being the least explored domain (N=18/163, 11%; Table 4).

The oldest study retained was published in 1975 (note that electronic databases used contain a limited number of publications prior to the 1950s and only 9 publications dated pre-2000 were retained following the abstract and full-text screening phases). Research domains were recorded with increasing frequency year-on-year (full-null model

401 comparison: χ^2=257.109, df=7, P<0.001), however the increasing rate didn't significantly
402 vary across domains (full-reduced model comparison: χ^2=7.224, df=6, P=0.301; Table 5)
403
404 *Which research domains are studied in Human and Non-human primates?*
405 Research effort across research domains was differently distributed in Human and Non-
406 human primates (Chi-square test: χ^2=27.204, df=6, P<0.001; Figure 3). *Meaning* was the
407 most frequently recorded domain in both Human (N=70/80, 88%) and Non-human
408 primate (N=73/87, 84%) publications and was similarly represented. *Ontogeny* was most
409 strongly biased towards Humans (Std. res.=3.051), and *form* was most strongly biased
410 towards Non-human primates when compared to Humans (Std. res.= 3.548, Figure 3).
411
412 **3.3. Defining gesture and intentional communication**
413 *Do studies include an explicit definition for Gesture and how does it vary with species*
414 *and domain?*
415 Although all publications had gesture as their main focus, over a third of the studies
416 (N=57/163, 35%) did not provide a formal definition for their use of Gesture (or any of
417 the specific forms of Gestures used). A definition of Gesture was more likely to be absent
418 in studies of Humans (N=41/80, 51%) than in studies of Non-human primates (N=16/87,
419 18%; Chi-square test: χ^2 =18.581, df=1, P<0.001). Only 42 studies explicitly included
420 mechanically ineffectiveness in their definitions of gesture. Studies of Non-human
421 primates included reference to mechanical ineffectiveness in their definitions of Gestures
422 (N= 37/87, 43%) more so than studies of Humans (N=5/80, 6%; Chi-square test:
423 χ^2=27.243, df=1, P<0.001).

Of the 106 publications that included a definition of Gesture(s), approximately half (N=52) defined only the specific gesture form considered in their studies (e.g. clap, defined as "One hand is brought forcefully downwards, palm facing down, to strike other hand, palm facing upwards; upper hand is recorded as dominant" in Fletcher, 2006) but provided no specific definition of what makes a Gestural signal. The proportion of studies defining gesture differed across domains (Chi-square test: χ^2=18.054, df=6, P=0.006; Table 4). Studies focused on gesture *form* were the most likely to define Gesture or subtypes of gestures used, and studies focused on *neural processes* were the least likely to include any definition of Gesture (Std. res.=2.852 and -2.376 respectively).

Of the 54 studies that provided an explicit definition of gesture, all gestures necessarily included a visual component in their Sensory Modality. Gestures were also defined as compounds that, in addition to visual information, also included an auditory or tactile component in the majority of studies (N=37/54, 69% and N=40/54, 74% respectively, Table 6). In terms of their Social Expression, gestures were defined as produced in the presence of an audience in 89% of studies (N=48/54); and as addressed to specific recipient(s) in 82% of studies (N=44/54); however, gestures were defined as produced while looking at the recipient in only 6 of the 54 studies (11%; Table 6). For gestures' Communicative Property, the majority of studies (N=51/54, 94%) defined gestures as a subcategory of communicative signals, while only three studies (6%) defined them as encompassing all communicative signals. Gestures were defined as mechanically ineffective in 69% of studies (N=37/54), and gesture was defined as being followed by a change in the recipient's behaviour in most cases (N=34/54, 63%; Table 6).

448 *Do studies include a specific definition of and criteria for intentional communication?*
449 *How does it vary with species and domain?*
450 Specific criteria for intentional communication were only provided in 52 of 163 articles
451 (32%; Table 4). Of the 54 studies that provided an explicit definition of gesture, 33 (61%)
452 included a requirement that gestures were necessarily an intentional communicative
453 signal. Response waiting and persistence or elaboration were included in approximately
454 half of the studies with a definition of gesture (N=24/54, 44%; and N=28/54, 52%
455 respectively). A more recent criterion (Cartmill & Byrne, 2010; Hobaiter & Byrne, 2014),
456 gestures were defined in association with an Apparently Satisfactory Outcome in only 9%
457 (N=5/54) of the studies. No studies specified that gestures should not be consistently
458 triggered by specific environmental conditions, and while information on functional
459 context was included in studies, no study required in their definition that gestures be
460 flexible across functions contexts (Table 6).
461 Across all studies, including those with no explicit definition of gesture, only a
462 very small proportion of publications with Human data provided explicit intentionality
463 criteria (n=5/80, 6%), while more than half of publications with Non-human primate data
464 did so (N= 48/87, 55%; Chi-square test: χ^2=43.812 df=1, P<0.001). Of the studies that
465 included intentionality criteria, persistence was the most common criteria provided
466 (N=34/52, 65%; see Table 7). The proportion of studies using intentional criteria differed
467 across domains (Chi-square test: χ^2=26.468, df=6, P<0.001). Research focused on *form*
468 was more likely to contain explicit intentional criteria than research focused on other
469 Domains (Std. res.= 3.921). No study focused on *neural processes* provided intentionality
470 criteria in their research (Std. res =-2.873; Table 4).
471

Which Body Parts are taken into account to produce Gestures? How does it vary with species and domain?

Across all studies we found a particular focus on Manual gestures: 54% of the publications (N=88/163) considered only Manual gestures, 4% considered only Non-manual gestures (N=7/163), 34% considered both Manual and Non-manual gestures (N=56/163), and 7% (N=12/163) did not specify the body part. However, within those studies with an explicit definition of gesture, 76% included movements of any body part (N=41/54; 5 of these definitions considered any body part including the face, and 36 of these definitions considered any body part but excluded the face; Table 6).

Across the full dataset, studies of Humans included more manual gestures than studies of Non-human primates (χ^2=10.774, df=2, P=0.005). Studies differed on the body parts included across all domains (χ^2 =21.432, df=12, P=0.044; Table 4). Studies on *form* were more likely to include both manual and Non-manual gestures, and studies on *neural processes* were more likely to include manual gestures (Std. res.=2.206 and 2.440 respectively).

3.4. Scope

Which type of Scope is represented in the studies included in the review? How does the Scope of the studies vary with species and domain?

Most studies were focused on specific gesture types or specific contexts and were classed as having a Narrow scope (N=112/163, 69%). Studies with a Broad scope were found less often in Human than in Non-human primate research (χ^2=18.914, df=1, P<0.001). Although all the domains were explored more often using a Narrow scope, the extent of this bias varied across domains (χ^2=13.948, df=6, P=0.030) with *form* and *laterality* least

biased (Std. res.=-1.606 and -0.954 respectively), as compared to *neural processes* which were exclusive studied with a Narrow scope (Std. res.= 2.842; Table 4).

3.5. Sources

Do studies incorporate additional Sources as well as Gesture? How does Source vary with species and domain?

Approximately half of the studies (N=87/163) considered additional sources to gesture. Vocalizations were considered in almost all multi-source studies (N=81/87, 93%), whereas Facial expressions and Lexigrams were described in 17% of multi-source studies (N=15/87). Nine articles considered two or more sources in addition to Gesture. Studies conducted in Humans (N= 53/80, 66%), as compared to Non-human Primates (N=37/87, 43%), were more likely to include a multi-source approach (χ^2=8.507, df=1, P=0.004). Studies of gesture in Humans including data on Vocalizations (N=51/80; 64%) more often than studies conducted in Non-human primates (N=33/87, 38%). The proportion of studies considering multiple sources varied across research domains (χ^2=69.984, df=6, P<0.001), with a multi-source approach always occurring in studies focused on *combination of sources* (Std. res.=7.635), and less frequently on studies focused on *form*, *laterality,* and *meaning* (Std. res.= -3.066, -2.378, and -2.443 respectively).

3.6. Research setting

Which Research settings are represented, and how does these vary with species and domain?

Most of the studies were conducted in Familiar settings (N=96/163, 59%). The frequency with which different Research settings were used varied across Human and Non-human

primates (χ^2=61.578, df=2, P<0.001). Studies of Non-human primates were conducted more often in Familiar settings (N=71/87, 82%), and studies with Humans were conducted more often in Unfamiliar settings (N=53/80, 68%; Table 8).

Research settings used also varied across Research domains (Chi-square test: χ^2=62.12, df=12, P<0.001). *Form* was studied most often in Familiar settings and *neural processes* were usually studied in Unfamiliar settings (Std. res. = 2.995, and 2.278 respectively, Table 4).

3.7. Study design

Which Study designs are represented and how does the Study design vary with species and domain?

Both Observational (n=79/163, 49%) and Experimental (n=79/163, 49%) designs were employed equally frequently in these studies; 5 studies incorporated both designs (3%). Studies with Humans applied Experimental designs more often (n= 58/80, 73%), and studies with Non-human primates applied Observational designs more often (n=61/87, 70%; Chi-square test: χ^2= 35.846, df=2, P<0.001).

The relative use of Observational and Experimental designs varied significantly across domains (Chi square test: χ^2=37.317, df=12, P<0.001). Studies analysing *form* used Observational designs most often, whereas studies focused on *neural processes* were always Experimental (Std. res.= 3.914 and 4.406 respectively; Table 4).

4. Discussion

Studies of gestural communication that make an explicit connection to the evolutionary origins of language exist in similar numbers for both human and non-human

544 primates; however, only four of the 163 studies included in this review incorporated data
545 from both human and non-human primates, and only 15 included more than one non-
546 human primate species. As a result, our ability to investigate species' similarities and
547 distinctions across primate taxa and infer an evolutionary trajectory for language from
548 this field, is almost entirely dependent on comparison across studies. By conducting a
549 systematic review we are able to describe to what extent current methods allow us to do
550 so reliably. We find substantial variation in the conceptual and methodological
551 approaches used. While variation allows for a diverse and robust examination of gesture
552 in this context, it presents particular challenges for the effective comparison across studies
553 and species on which the evolutionary approach depends.

554 There were limitations to the literature returned in our search process, for example,
555 older material (particularly pre-1950) is not systematically indexed in electronic
556 databases. However, perhaps the most important one was in our requirement for an
557 explicit reference to the evolution or origins of language or communication. We were
558 initially surprised that, in employing this restriction, we excluded work we regularly cite
559 as relevant to the evolutionary origins of human language, including our own studies.
560 Removing the requirement for the 'evolutionary' terms from our search returned around
561 6,500 results; whereas with them, our structured search returned around a thousand items
562 (with just over a sixth of that retained once the systematic selection criteria had been
563 applied). One explanation for the extent of these exclusions is that many of the empirical
564 studies on non-human primate gesture (including our own) avoid explicit discussion of
565 their potential connection to the evolution of human language – and in particular do not
566 do so in the title, keywords, or abstract; the fields most commonly indexed across
567 databases. In some cases, not doing so may be because the primary focus of the paper was

568 a description of the species communication – non-human primate gesture is of interest in
569 its own right, not just as a means of comparison to human communication. Nevertheless,
570 these studies may usefully inform our understanding of the evolutionary trajectory of
571 linguistic features. For example: some studies that explore the combination of gestures
572 into sequences (e.g., Tempelmann and Liebal 2012; McCarthy et al. 2013), relevant to
573 understanding similarities and differences to human language *structure,* or studies on
574 *neural processes* of homologous brain area activation in human and non-human primate
575 signalling (e.g., Hopkins et al. 2007, 2008) were excluded because they did not make an
576 explicit connection to the evolution of language. Similarly, some studies of human gesture
577 discussed its relationship to individual *ontogeny* of language, but did not explicitly
578 consider the relevance of the work to the evolution of language (e.g., Bates et al. 1979;
579 Iverson et al. 1994).

580 While widening our search to more broadly encompass primate gestural research
581 would successfully retain these studies, it would also add a very large literature that
582 provides limited insight into the evolution of language (for example, work on leaf-
583 clipping as a sexual solicitation in chimpanzees; Nishida 1980; or work on big loud
584 scratch as grooming solicitation; Nakamura et al. 2000) and such a large corpus risks
585 diluting our ability to provide a clear over-view of the field. There may also be a
586 justifiable reluctance to engage in 'just-so story-telling' in research that does not
587 explicitly test evolutionary hypotheses. Spurious statements about the relevance of any
588 non-human primate behaviour to human behaviour unhelpfully reinforce human-centric
589 approaches to the study of non-human behaviour, which risk us overlooking extraordinary
590 non-human species-specific capacities. Whereas carefully considered discussions can be
591 helpful, they require a substantial investment that may distract from the main aim of

592 research that did not set out to explicitly further a comparative approach. Nevertheless,
593 our theoretical papers often use the findings from these same empirical studies of primate
594 gesture as the foundation for the hypotheses we develop on the evolution of human
595 language. While it is certainly the case that studies of non-human primate gesture
596 contribute to a much wider range of questions beyond the possible evolutionary origins
597 of human language, it may be worth reflecting on the apparent scope in our field to more
598 explicitly test evolutionary hypotheses in a wider range of our empirical work. For
599 example: by establishing multi-species primate datasets that employ a coherent study
600 methodology, or allow for the extraction of like-with-like features for comparison, we
601 can test hypotheses that address how aspects of gesture are adapted to a particular species
602 socio-ecology (c.f. Prieur et al., 2020).

603 Developing hypotheses on the evolution of behaviour within primate history
604 requires that we have sufficient rich data across, and within, primate taxa. However, we
605 found that studies of non-human primate gestural communication were largely limited to
606 chimpanzees and bonobos. A number of studies in monkeys were excluded because they
607 involved training them to produce a particular gesture (e.g., Meunier et al. 2013; Defolie
608 et al. 2015) leaving only a handful of studies in Afro-Eurasian monkeys, and no studies
609 on monkeys of the Americas, small apes, or Strepsirrhines. While chimpanzees and
610 bonobos represent our closest living relatives, and these studies allow us to ask whether
611 or not a particular feature of language is uniquely human, they provide a more limited
612 scope for exploring the possible evolutionary trajectory of language over a longer period.
613 In addition, implicit hierarchies exist between species in their relevance to human origins
614 that may obscure the deeper roots of some features (Bourjade et al., 2020). We see this
615 illustrated in our data in the very limited number of human studies that explicitly test

616 whether the gestures explored meet criteria for intentional use – there is an assumption
617 that human behaviour always does. Our findings similarly highlight that there is often an
618 implicit assumption of the importance of ape behaviour for understanding human
619 evolution, whereas studies of monkey behaviour are required to more thoroughly
620 establish the grounds for comparison (Bourjade et al., 2020). Data on the spontaneous
621 gestural communication from a more diverse range of species – including direct
622 comparisons between non-human primates – are necessary for a deeper and more nuanced
623 understanding of how and when the capacities that underpin language evolved.

624 Our understanding of gestural signals' contribution to the evolutionary origins of
625 human language may also be compromised by the use of specific human and non-human
626 primate populations to represent species-specific characteristics. Every species' system
627 of communication – including primate gesturing – is in some way adapted to a species'
628 specific and distinct socio-ecological niche (Cheney & Seyfarth, 2018). The majority of
629 studies of non-human primate gesture in our review were conducted on groups living in
630 man-made anthropogenic environments that do not reflect the socio-ecological
631 environments to which their communication is adapted. Even among studies of wild
632 primates, a focus on a few specific groups or populations (e.g., Pika and Mitani 2006;
633 Hobaiter and Byrne 2011a; Roberts et al. 2012), likely impacts our understanding of
634 species-typical behaviour. Similarly, in human studies there was also a focus on specific
635 groups: as in many fields of study (Henrich et al., 2010), there was a strong bias towards
636 studies of human gesture in WEIRD socio-economic cultures. Our understanding of the
637 links between primate gesture and human language can be strengthened by more direct
638 testing of the impact of species socio-ecology and individual life-history characteristics

639 on gestural expression (e.g. Prieur et al., 2020), although doing so will take substantial
640 large-scale datasets.

641 Different environments may promote the use of certain gestures that are not
642 expressed in other environments (see Leavens et al. 2005 for the example of pointing).
643 The frequency and quality (context, interaction partner, and group membership) of
644 interactions seem to influence the frequency of gesture use and the size of gestural
645 repertoires, for example: higher interaction rates with non-maternal conspecifics and a
646 larger number of previous interaction partners are both related to more frequent gesturing
647 and the use of more gesture types (Fröhlich et al., 2017, 2018). Thus, it is particularly
648 important to complement the detailed data from captive settings with more diverse data
649 from primates living in their naturally structured social units (Cheney & Seyfarth, 2018;
650 Fröhlich & Hobaiter, 2018). Given the well-established presence of rich cultural variation
651 in behaviour (Boesch et al., 1994; McGrew et al., 1997; Whiten et al., 1999) a richer
652 understanding of the communicative abilities in diverse populations, in a range of
653 environments (Hobaiter & Byrne, 2011b, 2014), in other great apes (gorillas, bonobos
654 and orang-utans; Bard 1992; Genty et al. 2009; Schamberg et al. 2016; Knox et al. 2019),
655 and in other primate species (e.g., Japanese macaques, mandrills, pygmy marmosets,
656 capuchin monkeys, bonnet macaques; Itani 1963; Kudo 1987; De La Torre and Snowdon
657 2002; Wheeler 2010; Gupta and Sinha 2016) is necessary to better understand the
658 evolutionary trajectory of primate gestural communication and its relationship with
659 language.

660 The steady increase across most domains highlights the increasing interest in, and
661 relevance of gesture to, questions related to language evolution; however, research efforts
662 remain unequally distributed across domains and between species within domains.

663 *Meaning* was the most popular domain and was similarly explored in both Human and
664 Non-human primates. Studies of *form* were most biased towards Non-human primates
665 and *ontogeny* were most biased towards studied in humans.

666 Studies of *meaning* in gestural communication are now the most common focus.
667 Gestural research has sometimes employed the signaller's behaviour or the context of the
668 signal use as a proxy for 'meaning' (Bard and Leavens 2014; Tomasello et al., 1994).
669 Reflecting gesture's intentional use, recent studies employ a combination of signaller and
670 recipient behaviour in order to take a cognitive-linguistic perspective and infer the
671 signaller's intended meaning (e.g., Genty et al. 2009; Cartmill and Byrne 2010; Hobaiter
672 and Byrne 2014; Graham et al. 2017). The large number of studies of signal *form* in other
673 primates is likely because this domain includes the description of communicative
674 repertoires, a common focus when exploring the communication of any non-human
675 species (e.g., Berg 1983; Conner 1985; Edds-Walton and Edds-Walton 1997). However,
676 the lack of similar systematic descriptive studies of human gesture forms again makes
677 comparison with non-human primate research challenging (Kersken et al., 2018; Müller,
678 2005). Gestural *ontogeny* has been explored in humans for decades in the context of its
679 relevance for language development; however, it is only more recently that researchers
680 started to frame their results within the evolutionary puzzle or explore this domain in non-
681 human primates.

682 Both within and between human and non-human primates, studies used different
683 definitions of which movements and body parts constitute a gesture, and different criteria
684 to define their intentional use. Some studies defined gesture broadly to include body
685 postures, while other studies employed more restrictive definitions including specific
686 criteria, or specific body parts. Over half of the studies were restricted to manual gestures,

687 but a considerable number of studies (~40%) were more flexible, including movements
688 from the whole body or other body parts. None of this variation is necessarily
689 problematic; however, a significant cause for concern, given this variation in definition,
690 is that more than a third of the studies in this review didn't provide a definition for their
691 use of gesture at all. These differences in the fundamental basis of what 'a gesture' is may
692 have significant consequences for the study of comparative communication: particularly
693 where different studies across different species are directly compared. It would be
694 unnecessarily limiting to require a single definition of gesture, and, for example, there is
695 no particular problem with considering only manual gestures. However, given the
696 variation across the field, to allow effective comparison across species and populations,
697 it is imperative that studies define their specific usage.

698 One aspect of gesture *form* we did not directly compare were gestural repertoires.
699 Composed of graded movements (for example from shake to swing or reach to fling), the
700 identification of distinct gesture forms is typically done from the subjective and
701 anthropocentric-centric perspective of the researcher(s) (c.f. Hobaiter & Byrne, 2017b).
702 Even within a research group there may be variation between studies; for example, where
703 more data become available or new species comparisons suggest previously undetected
704 similarities or differences of *form* (c.f. Byrne et al., 2017; Genty et al., 2009; Graham et
705 al., 2017; Hobaiter & Byrne, 2011b; Knox et al., 2019). Descriptions of repertoires take
706 substantial study effort (Byrne et al., 2017; Hobaiter & Byrne, 2011b), and are vulnerable
707 to under-sampling (Knox et al., 2019) – making comparison across studies particularly
708 valuable. However, our current ability to do so remains very limited: under a third of the
709 studies we analysed provided any description of their specific gesture form(s). Even
710 where written descriptions are available these can be difficult to interpret. As open-access

711 online data-archiving becomes increasingly accessible, the use of video examples (e.g.
712 www.greatapedictionary.com) highlighting both the archetypical form and variation
713 around it, will make a substantial contribution to our ability to systematically compare
714 like-with-like within gesture *form*, without necessitating the use of universal – and
715 potentially restrictive – agreed repertoires.

716 The use of a clear framework for comparison allows us to highlight not only the
717 variation in the frequency with which criteria for intentional gesture use are applied, but
718 also differences in the types of criteria used between human and non-human primates
719 (Bourjade et al., 2020). Once again, across studies and across human and non-human
720 primates, we find substantial variation. Where definitions were available, Persistence was
721 the most popular criterion, with gestures considered intentional when the signaller persists
722 in producing further gestures if a response is not obtained from the recipient. The criteria
723 for defining intentional communication were originally developed in order to explore the
724 development of prelinguistic communication in human infants (Bates et al., 1975, 1979),
725 but their explicit use seems now largely restricted to non-human primate studies. Over
726 half of the non-human primate studies reviewed provided some criteria for defining
727 intentional gesture. However, while many required gesture cases to meet one or more
728 criteria for intentional use from a set, they typically did not specify which were met
729 (Genty et al., 2009; Hobaiter & Byrne, 2011b; c.f. Leavens et al., 2005). No one criterion
730 is a panacea for the challenge of identifying mental states from observable behaviour.
731 Audience-checking could simply reflect a shift in attention between objects of interest.
732 Response-waiting could reflect a brief rest in activity. Providing more detail on the
733 frequency and distribution of the different criteria within a study would allow for more
734 direct comparison of intentional gesture use across studies and species of non-human

735 primates and improve our ability to assess the extent to which particular criteria provide
736 robust, reliable, measures (e.g., Prieur et al., 2018).

737 In contrast to the relatively widespread use of criteria to define intentional gesture
738 in non-human primates, just four studies in humans and one on both human and non-
739 human primates, provided any criteria for intentional gesture use. While humans are
740 clearly capable of intentional communication, we are equally capable of producing fixed
741 non-intentional signals (for example an involuntary yelp, smile, or laugh; Provine 1992;
742 Kawakami et al. 2007). Moreover, mechanically ineffectiveness seems to be a criterion
743 often applied in non-human primate research to define gesture, but rarely seen in human
744 gestural research. Including all human gesture-like movements, irrespective of the
745 objective evidence for their communicative and intentional use, while limiting non-
746 human primate data to only those gestures used with evidence for intentional use, again,
747 impacts our ability to make meaningful comparisons between human and non-human
748 gesture. Doing so furthers the double standard too often applied in comparative research,
749 which sees systematic species-differences in testing conditions or criteria mistaken for
750 species-differences in cognition (Bard & Leavens, 2014; Leavens et al., 2019).

751 The majority of the studies reviewed employed a narrow scope of focus,
752 investigating specific gesture types or specific contexts. However, studies in non-human
753 primates were more likely to have a broad scope than studies with humans; for example:
754 they more often included descriptions of gestural repertoires rather than a specific *form*
755 such as pointing. Again, there is no intrinsic benefit to employing a narrow or broad
756 scope, but both are needed across species in order to compare like with like.

757 Almost half the studies in this review included other signal *sources* with their
758 gesture data; however, the integration of gestures, vocalizations, and facial expressions

759 remains under-studied in non-humans relative to humans, despite recent calls to
760 investigate it (e.g., Slocombe et al. 2011; Waller et al. 2013). Where signal sources are
761 combined in communication, for example gesture and facial expression, studying one in
762 the absence of the other may limit our interpretation of signal function (Wilke et al.,
763 2017). However, studying different signal types and sources in combination can be
764 methodologically challenging. For example, studies of gesture often focus on visual
765 information, and the signaller and recipient's visual attention; neither of which may be as
766 relevant to the production or receipt of vocal signals (Schel et al., 2013). The development
767 of methodologies that can be applied across sources will allow for more widespread multi-
768 source comparisons (Müller, 2005; Slocombe et al., 2011).

769 Ape cognitive and social development, including in their communicative
770 repertoires (Boesch, 2007; Leavens et al., 2019), is sensitive to a wide range of social and
771 environmental factors, and interactional experience has been shown to impact the
772 development of gestural use (e.g., Bard et al. 2014; Fröhlich and Hobaiter 2018). We
773 found that non-human primates were more often studied in environments that were
774 familiar to them and used observational research designs. In contrast, human research was
775 mainly conducted in unfamiliar environments, such as research laboratories, and used
776 experiment designs. Collecting data within a laboratory setting allows nuanced control of
777 specific variables; however, these methods are typically challenging to replicate in an
778 ethical manner with Non-human primates. Slocombe et al. (2011) previously highlighted
779 the lack of non-human primate gestural work in the wild (although see now for example:
780 Hobaiter and Byrne 2011a,b; Roberts et al. 2012; Graham et al. 2018), but it is similarly
781 noteworthy that very little gestural work on humans is done outside of unfamiliar
782 laboratory settings (c.f. Kersken et al. 2018).

Diversity in definitions and methods allows for robust testing of a particular line of research, with different perspectives and findings contributing to a deeper and more nuanced understanding of the question. As a result, it is critical that we allow researchers' sufficient flexibility in their definition of and exploration of primate gesture. Without this we would only stifle comparison – potentially missing crucial similarities and differences. However, where diversity is present, the absence of definitions and transparent descriptions of methods are a problem – particularly so as our ability to make explicit points about the evolution of a particular aspect of language often depends on comparison across species and studies.

We can summarise out findings in the following 8-point call to action for researchers interested in how gestural communication may inform our understanding of language evolution.

1. There is substantial scope in the wider gestural field to more explicitly test evolutionary hypotheses in our empirical work.
2. Data are needed on spontaneous gestural communication from a more diverse range of species, in particular outside of *Pan* ape species, and including direct comparisons between non-human primates.
3. Data are needed from more diverse populations in diverse environments that consider the impact of socio-ecology and socio-economy on the use of gesture.
4. There is particular scope for studies of gesture *forms* in humans, and studies of gesture *ontogeny* in non-human primates.
5. Given the variation across the field it is imperative that studies define their specific usage of gesture.

6. Providing more detail on the frequency and distribution of the different criteria for intentional use, in particular in humans, will improve our ability to assess the extent to which particular criteria provide robust, reliable, measures.
7. The development of methodologies that can be applied across sources will allow for more widespread multi-source comparisons.
8. In addition to studies of human gesture outside WEIRD populations, there is also substantial scope for studies of natural human gesture in familiar, non-laboratory, environments.

We hope that this review serves to highlight not only the challenges, but also the areas of particular promise for future research. A detailed understanding of human and non-human primate gesture will take more than one researcher's or research group's lifetime of study. Diversity in our study subjects and approach will provide a more nuanced understanding, but transparency and replicability in our methods are equally crucial to our ability to draw meaningful conclusions about gestural communication's role in the evolution of human language.

Acknowledgments

The structure of this paper was developed during the January 2019 ISPA Advanced Course on Scientific Writing, and we would like to thank P. McGregor for his guidance. We are grateful for the thoughtful comments raised in the review process that allowed us to incorporate important new analyses and discussion. The authors gratefully acknowledge the financial support provided by the Portuguese national funding agency for science, research and technology to the first author (SFRH/BD/138406/2018).

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest: The authors declare that they have no confict of interest.

Data Availability Statement: The data that supports the findings of this study are available in the supplementary material of this article and in a public repository on github.com/Wild-Minds/GestureStudies_SystematicReview

References

Arensburg, B., Tillier, A. M., Vandermeersch, B., Duday, H., Schepartz, L. A., & Rak, Y. (1989). A Middle Palaeolithic human hyoid bone. *Nature*, *338*(6218), 758–760. https://doi.org/10.1038/338758a0

Bard, K. A. (1992). Intentional Behavior and Intentional Communication in Young Free-Ranging Orangutans. *Child Development*, *63*(5), 1186–1197. https://doi.org/10.1111/j.1467-8624.1992.tb01688.x

Bard, K. A., Dunbar, S., Maguire-Herring, V., Veira, Y., Hayes, K. G., & Mcdonald, K. (2014). Gestures and social-emotional communicative development in chimpanzee infants. *American Journal of Primatology*, *76*(1), 14–29. https://doi.org/10.1002/ajp.22189

Bard, K. A., & Hopkins, W. D. (2018). Early Socioemotional Intervention Mediates Long-Term Effects of Atypical Rearing on Structural Covariation in Gray Matter in Adult Chimpanzees. *Psychological Science*, *29*(4), 594–603. https://doi.org/10.1177/0956797617740685

851 Bard, K. A., & Leavens, D. A. (2014). The Importance of Development for
852 Comparative Primatology. *Annual Review of Anthropology*, *43*(1), 183–200.
853 https://doi.org/10.1146/annurev-anthro-102313-030223
854 Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The*
855 *emergence of symbols*.
856 Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to
857 speech. *Merrill-Palmer Quarterly of Behavior and Development*, *21*(3), 205–
858 226.
859 Berg, J. K. (1983). Vocalizations and Associated Behaviors of the African Elephant
860 (Loxodonta africana) in Captivity. *Zeitschrift Für Tierpsychologie*, *63*(1), 63–
861 79. https://doi.org/10.1111/j.1439-0310.1983.tb00741.x
862 Biau, E., Morís Fernández, L., Holle, H., Avila, C., & Soto-Faraco, S. (2016). Hand
863 gestures as visual prosody: BOLD responses to audio–visual alignment are
864 modulated by the communicative nature of the stimuli. *NeuroImage*, *132*, 129–
865 137. https://doi.org/10.1016/j.neuroimage.2016.02.018
866 Boesch, C. (2007). What Makes Us Human (Homo sapiens)? The Challenge of
867 Cognitive Cross-Species Comparison. *Journal of Comparative Psychology*,
868 *121*(3), 227–240. https://doi.org/10.1037/0735-7036.121.3.227
869 Boesch, C. (2020). The Human Challenge in Understanding Animal Cognition. In L. S.
870 M. Johnson, A. Fenton, & A. Shriver (Eds.), *Neuroethics and Nonhuman*
871 *Animals* (pp. 33–51). Springer International Publishing.
872 https://doi.org/10.1007/978-3-030-31011-0_3

873 Boesch, C., Marchesi, P., Marchesi, N., Fruth, B., & Joulian, F. (1994). Is nut cracking
874 in wild chimpanzees a cultural behaviour? *Journal of Human Evolution*, *26*(4),
875 325–338. https://doi.org/10.1006/jhev.1994.1020
876 Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How Could
877 Language Have Evolved? *PLoS Biology*, *12*(8), e1001934.
878 https://doi.org/10.1371/journal.pbio.1001934
879 Bourjade, M., Cochet, H., Molesti, S., & Guidetti, M. (2020). Is Conceptual Diversity
880 an Advantage for Scientific Inquiry? A Case Study on the Concept of 'Gesture'
881 in Comparative Psychology. *Integrative Psychological and Behavioral Science*,
882 *54*(4), 805–832. https://doi.org/10.1007/s12124-020-09516-5
883 Brentari, D., Coppola, M., Mazzoni, L., & Goldin-Meadow, S. (2012). When does a
884 system become phonological? Handshape production in gesturers, signers, and
885 homesigners. *Natural Language & Linguistic Theory*, *30*(1), 1–31.
886 https://doi.org/10.1007/s11049-011-9145-1
887 Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017).
888 Great ape gestures: Intentional communication with a rich set of innate signals.
889 *Animal Cognition*, *20*(4), 755–769. https://doi.org/10.1007/s10071-017-1096-4
890 Cartmill, E. A., & Byrne, R. W. (2010). Semantics of primate gestures: Intentional
891 meanings of orangutan gestures. *Animal Cognition*, *13*(6), 793–804.
892 https://doi.org/10.1007/s10071-010-0328-7
893 Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate
894 vocalizations. *Proceedings of the National Academy of Sciences*, 201717572.
895 https://doi.org/10.1073/pnas.1717572115

896 Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: A usage-
897 based account of natural language recursion. *Frontiers in Psychology*, *6*, 1182.
898 https://doi.org/10.3389/fpsyg.2015.01182
899 Conner, R. N. (1985). Vocalizations of Common Ravens in Virginia. *The Condor*,
900 *87*(3), 379–388. https://doi.org/10.2307/1367219
901 Cooke, A., Smith, D., & Booth, A. (2013). Qualitative Health Research Beyond PIC :
902 The SPIDER Tool for Qualitative Evidence Synthesis. *Qualitative Health*
903 *Research*, *22*(10), 1435–1443. https://doi.org/10.1177/1049732312452938
904 Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton
905 University Press. https://doi.org/10.1525/jlin.2006.16.1.146
906 De La Torre, S., & Snowdon, C. T. (2002). Environmental correlates of vocal
907 communication of wild pygmy marmosets, Cebuella pygmaea. *Animal*
908 *Behaviour*, *63*(5), 847–856. https://doi.org/10.1006/anbe.2001.1978
909 Defolie, C., Malassis, R., Serre, M., & Meunier, H. (2015). Tufted capuchins (Cebus
910 apella) adapt their communicative behaviour to human's attentional states.
911 *Animal Cognition*, *18*(3), 747–755. https://doi.org/10.1007/s10071-015-0841-9
912 Edds-Walton, P. L., & Edds-Walton, P. L. (1997). Acoustic communication signals of
913 mysticete whales. *Bioacoustics*, *8*(1–2), 47–60.
914 https://doi.org/10.1080/09524622.1997.9753353
915 Fletcher, A. W. (2006). Clapping in chimpanzees: Evidence of exclusive hand
916 preference in a spontaneous, bimanual gesture. *American Journal of*
917 *Primatology*, *68*(11), 1081–1088. https://doi.org/10.1002/ajp.20308

Fröhlich, M., & Hobaiter, C. (2018). The development of gestural communication in great apes. *Behavioral Ecology and Sociobiology*, *72*(12), 194. https://doi.org/10.1007/s00265-018-2619-y

Fröhlich, M., Müller, G., Zeiträg, C., Wittig, R. M., & Pika, S. (2017). Gestural development of chimpanzees in the wild: The impact of interactional experience. *Animal Behaviour*, *134*, 271–282. https://doi.org/10.1016/J.ANBEHAV.2016.12.018

Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & Schaik, C. P. (2019). Multimodal communication and language origins: Integrating gestures and vocalizations. *Biological Reviews*, *94*(5), 1809–1829. https://doi.org/10.1111/brv.12535

Fröhlich, M., Wittig, R. M., & Pika, S. (2018). The ontogeny of intentional communication in chimpanzees in the wild. *Developmental Science*, *22*(1), e12716. https://doi.org/10.1111/desc.12716

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (Gorilla gorilla): Repertoire, intentionality and possible origins. *Animal Cognition*, *12*(3), 527–546. https://doi.org/10.1007/s10071-009-0213-4

Gillespie-Lynch, K., Greenfield, P. M., Lyn, H., & Savage-Rumbaugh, S. (2014). Gestural and symbolic development among apes and humans: Support for a multimodal theory of language evolution. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01228

Goldin-Meadow, S., & Yang, C. (2017). Statistical evidence that a child can create a combinatorial linguistic system without external linguistic input: Implications

941 for language evolution. *Neuroscience & Biobehavioral Reviews*, *81*, 150–157.
942 https://doi.org/10.1016/j.neubiorev.2016.12.016
943 Graham, K. E., Furuichi, T., & Byrne, R. W. (2017). The gestural repertoire of the wild
944 bonobo (Pan paniscus): A mutually understood communication system. *Animal*
945 *Cognition*, *20*(2), 171–177. https://doi.org/10.1007/s10071-016-1035-9
946 Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo
947 and chimpanzee gestures overlap extensively in meaning. *PLoS Biology*, *16*(2).
948 https://doi.org/10.1371/journal.pbio.2004825
949 Gupta, S., & Sinha, A. (2016). Not here, there! Possible referential gesturing during
950 allogrooming by wild bonnet macaques, Macaca radiata. *Animal Cognition*,
951 *19*(6), 1243–1248. https://doi.org/10.1007/s10071-016-1012-3
952 Hall, M. L., Ahn, Y. D., Mayberry, R. I., & Ferreira, V. S. (2015). Production and
953 comprehension show divergent constituent order preferences: Evidence from
954 elicited pantomime. *Journal of Memory and Language*, *81*, 16–33.
955 https://doi.org/10.1016/j.jml.2014.12.003
956 Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is
957 it, who has it, and how did it evolve? *Science (New York, N.Y.)*, *298*(5598),
958 1569–1579. https://doi.org/10.1126/science.298.5598.1569
959 Henrich, J., Heine, S., Nature, A. N.-, & 2010, U. (2010). (Henrich, 2010) Most people
960 are not WEIRD. *Nature*, *466*(7302), 29.
961 Hewes, G. W., Andrew, R. J., Carini, L., Choe, H., Gardner, R. A., Kortlandt, A.,
962 Krantz, G. S., McBride, G., Nottebohm, F., Pfeiffer, J., Rumbaugh, D. G.,
963 Steklis, H. D., Raliegh, M. J., Stopa, R., Suzuki, A., Washburn, S. L., &
964 Wescott, R. W. (1973). Primate Communication and the Gestural Origin of

965 Language [and Comments and Reply]. *Current Anthropology*, *14*(1/2), 5–24.
966 https://doi.org/10.1086/201401
967 Hobaiter, C., & Byrne, R. W. (2011a). Serial gesturing by wild chimpanzees: Its nature
968 and function for communication. *Anim Cogn*, *14*, 827–838.
969 https://doi.org/10.1007/s10071-011-0416-3
970 Hobaiter, C., & Byrne, R. W. (2011b). The gestural repertoire of the wild chimpanzee.
971 *Animal Cognition*, *14*(5), 745–767. https://doi.org/10.1007/s10071-011-0409-2
972 Hobaiter, C., & Byrne, R. W. (2014). The Meanings of Chimpanzee Gestures. In
973 *Current Biology* (Vol. 24, Issue 14). https://doi.org/10.1016/j.cub.2014.05.066
974 Hobaiter, C., & Byrne, R. W. (2017a). What is a gesture? A meaning-based approach to
975 defining gestural repertoires. *Neuroscience & Biobehavioral Reviews*, *82*, 3–12.
976 https://doi.org/10.1016/j.neubiorev.2017.03.008
977 Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017b). Wild chimpanzees' use of
978 single and combined vocal and gestural signals. *Behavioral Ecology and*
979 *Sociobiology*, *71*(6), 96. https://doi.org/10.1007/s00265-017-2325-1
980 Hopkins, W. D., Dunham, L., Cantalupo, C., & Taglialatela, J. (2007). The Association
981 between Handedness, Brain Asymmetries, and Corpus Callosum Size in
982 Chimpanzees (Pan troglodytes). *Cerebral Cortex*, *17*(8), 1757–1765.
983 https://doi.org/10.1093/cercor/bhl086
984 Hopkins, W. D., Taglialatela, J. P., Meguerditchian, A., Nir, T., Schenker, N. M., &
985 Sherwood, C. C. (2008). Gray matter asymmetries in chimpanzees as revealed
986 by voxel-based morphometry. *NeuroImage*, *42*(2), 491–497.
987 https://doi.org/10.1016/j.neuroimage.2008.05.014

988 Hsieh, S., & Plotnick, R. E. (2020). The representation of animal behaviour in the fossil
989 record. *Animal Behaviour*, *169*, 65–80.
990 https://doi.org/10.1016/j.anbehav.2020.09.010
991 Itani, J. (1963). Vocal communication of the wild Japanese monkey. *Primates*, *4*(2), 11–
992 66. https://doi.org/10.1007/bf01659149
993 Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language
994 in two modalities. *Cognitive Development*, *9*(1), 23–43.
995 https://doi.org/10.1016/0885-2014(94)90018-3
996 Kawakami, K., Takai-Kawakami, K., Tomonaga, M., Suzuki, J., Kusaka, F., & Okai, T.
997 (2007). Spontaneous smile and spontaneous laugh: An intensive longitudinal
998 case study. *Infant Behavior and Development*, *30*(1), 146–152.
999 https://doi.org/10.1016/j.infbeh.2006.08.004
1000 Kersken, V., Gómez, J.-C., Liszkowski, U., Soldati, A., & Hobaiter, C. (2018). A
1001 gestural repertoire of 1- to 2-year-old human children: In search of the ape
1002 gestures. *Animal Cognition*. https://doi.org/10.1007/s10071-018-1213-z
1003 Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J. F., & Morrogh-
1004 Bernard, H. (2019). Gesture Use in Communication between Mothers and
1005 Offspring in Wild Orang-Utans (Pongo pygmaeus wurmbii) from the Sabangau
1006 Peat-Swamp Forest, Borneo. *International Journal of Primatology*, 1–24.
1007 https://doi.org/10.1007/s10764-019-00095-w
1008 Kudo, H. (1987). The study of vocal communication of wild mandrills in Cameroon in
1009 relation to their social structure. *Primates*, *28*(3), 289–308.
1010 https://doi.org/10.1007/BF02381013

1011 Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2010). BIZARRE chimpanzees do not
1012 represent “the chimpanzee.” *Behavioral and Brain Sciences*, *33*(2–3), 100–101.
1013 https://doi.org/10.1017/S0140525X10000166
1014 Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2019). The mismeasure of ape social
1015 cognition. *Animal Cognition*, *22*(4), 487–504. https://doi.org/10.1007/s10071-
1016 017-1119-1
1017 Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2005). Understanding the Point of
1018 Chimpanzee Pointing: Epigenesis and Ecological Validity. *Current Directions in*
1019 *Psychological Science*, *14*(4), 185–189. https://doi.org/10.1111/j.0963-
1020 7214.2005.00361.x
1021 Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005). Intentionality as Measured in
1022 the Persistence and Elaboration of Communication by Chimpanzees (Pan
1023 troglodytes). *Child Development*, *76*(1), 291–306.
1024 https://doi.org/10.1111/j.1467-8624.2005.00845.x
1025 Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2010). Multimodal communication by
1026 captive chimpanzees (Pan troglodytes). *Animal Cognition*, *13*(1), 33–40.
1027 https://doi.org/10.1007/s10071-009-0242-z
1028 Liebal, K., Call, J., & Tomasello, M. (2004). Use of gesture sequences in chimpanzees.
1029 *American Journal of Primatology*, *64*(4), 377–396.
1030 https://doi.org/10.1002/ajp.20087
1031 MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: The role of
1032 enhanced breathing control. *American Journal of Physical Anthropology*,
1033 *109*(3), 341–363. https://doi.org/10.1002/(SICI)1096-
1034 8644(199907)109:3<341::AID-AJPA5>3.0.CO;2-2

1035 McCarthy, M. S., Jensvold, M. L. A., & Fouts, D. H. (2013). Use of gesture sequences
1036 in captive chimpanzee (Pan troglodytes) play. *Animal Cognition*, *16*(3), 471–
1037 481. https://doi.org/10.1007/s10071-012-0587-6
1038 McGrew, W., Ham, R., White, L., … C. T.-I. J. of, & 1997, U. (1997). Why don't
1039 chimpanzees in Gabon crack nuts? *International Journal of Primatology*, *18*(3),
1040 353–374. https://doi.org/10.1023/A:1026382316131
1041 Meguerditchian, A., Molesti, S., & Vauclair, J. (2011). Right-handedness predominance
1042 in 162 baboons (Papio anubis) for gestural communication: Consistency across
1043 time and groups. *Behavioral Neuroscience*, *125*(4), 653–660.
1044 https://doi.org/10.1037/a0023823
1045 Meguerditchian, A., Vauclair, J., & Hopkins, W. D. (2010). Captive chimpanzees use
1046 their right hand to communicate with each other: Implications for the origin of
1047 the cerebral substrate for language. *CORTEX*, *46*(1), 40–48.
1048 https://doi.org/10.1016/j.cortex.2009.02.013
1049 Meunier, H., Fizet, J., & Vauclair, J. (2013). Tonkean macaques communicate with
1050 their right hand. *Brain and Language*, *126*(2), 181–187.
1051 https://doi.org/10.1016/j.bandl.2013.05.004
1052 Micheletta, J., Engelhardt, A., Matthews, L., Agil, M., & Waller, B. M. (2013).
1053 Multicomponent and multimodal lipsmacking in crested macaques (Macaca
1054 nigra). *American Journal of Primatology*, *75*(7), 763–773.
1055 https://doi.org/10.1002/ajp.22105
1056 Müller, C. (2005). Gestures in human and nonhuman primates: Why we need a
1057 comparative view. In *Gesture* (Vol. 5, Issue 1, pp. 259–283).
1058 https://doi.org/10.1075/gest.5.1.17mul

1059 Nakamura, M., McGrew, W. C., Marchant, L. F., & Nishida, T. (2000). Social scratch:
1060 Another custom in wild chimpanzees? *Primates*, *41*(3), 237–248.
1061 https://doi.org/10.1007/BF02557594
1062 Nishida, T. (1980). The leaf-clipping display: A newly-discovered expressive gesture in
1063 wild chimpanzees. *Journal of Human Evolution*, *9*(2), 117–128.
1064 https://doi.org/10.1016/0047-2484(80)90068-8
1065 Partan, S. (2002). Single and multichannel signal composition: Facial expressions and
1066 vocalizations of rhesus macaques (Macaca mulatta). *Behaviour*, *139*, 993–1028.
1067 Partan, S., & Marler, P. (1999). Communication goes multimodal. *Science*, *83*, 1272–
1068 1273.
1069 Pika, S., & Mitani, J. (2006). Referential gestural communication in wild chimpanzees
1070 (Pan troglodytes). *Current Biology*, *16*(6), R191–R192.
1071 https://doi.org/10.1016/j.cub.2006.02.037
1072 Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it?
1073 *Cognition*, *95*(2), 201–236. https://doi.org/10.1016/j.cognition.2004.08.004
1074 Plooij, F. X. (1978). Some basic traits of language in wild chimpanzees? In A. Lock
1075 (Ed.), *Action, gestures, and symbol: The emergence of language* (pp. 111–132).
1076 Pollick, A. S., & de Waal, F. B. M. (2007). Ape gestures and language evolution.
1077 *Proceedings of the National Academy of Sciences of the United States of*
1078 *America*, *104*(19), 8184–8189.
1079 Prieur, J., Barbu, S., Blois-Heulin, C., & Lemasson, A. (2020). The origins of gestures
1080 and language: History, current advances and proposed theories. *Biological*
1081 *Reviews*, *95*(3), 531–554. https://doi.org/10.1111/brv.12576

1082 Prieur, J., Lemasson, A., Barbu, S., & Blois-Heulin, C. (2018). Challenges Facing the
1083 Study of the Evolutionary Origins of Human Right-Handedness and Language.
1084 *International Journal of Primatology*, *39*(2), 183–207.
1085 https://doi.org/10.1007/s10764-018-0038-6
1086 Provine, R. R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs
1087 and smiles. *Bulletin of the Psychonomic Society*, *30*(1), 1–4.
1088 https://doi.org/10.3758/BF03330380
1089 Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal*
1090 *Behaviour*, *78*(2), 233–240. https://doi.org/10.1016/j.anbehav.2009.06.007
1091 Richerson, P., Biolinguistics, R. B.-, & 2010, U. (2010). Why Possibly Language
1092 Evolved. *Biolinguistics*, *4*(2–3), 289–306.
1093 Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. In *Trends in*
1094 *Neurosciences* (Vol. 21, Issue 5, pp. 188–194).
1095 Roberts, A. I., Vick, S.-J., & Buchanan-Smith, H. M. (2012). Usage and comprehension
1096 of manual gestures in wild chimpanzees. *Animal Behaviour*, *84*(2), 459–470.
1097 https://doi.org/10.1016/j.anbehav.2012.05.022
1098 Salo, V. C., Rowe, M. L., & Reeb-Sutherland, B. C. (2018). Exploring Infant Gesture
1099 and Joint Attention as Related Constructs and as Predictors of Later Language.
1100 *Infancy*, *23*(3), 432–452. https://doi.org/10.1111/infa.12229
1101 Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call
1102 combinations, vocal exchanges and interparty movement in wild bonobos.
1103 *Animal Behaviour*, *122*, 109–116. https://doi.org/10.1016/j.anbehav.2016.10.003

1104 Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E.
1105 (2013). Chimpanzee Alarm Call Production Meets Key Criteria for
1106 Intentionality. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0076674
1107 Seyfarth, R., & Cheney, D. (2003). Meaning and emotion in animal vocalizations.
1108 *Annals of the New York Academy of Sciences*, 32–55.
1109 Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and Receivers in Animal
1110 Communication. *Annual Review of Psychology*, *54*(1), 145–173.
1111 https://doi.org/10.1146/annurev.psych.54.101601.145121
1112 Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: The need for
1113 multimodality in primate communication research. *Animal Behaviour*, *81*(5),
1114 919–924. https://doi.org/10.1016/j.anbehav.2011.02.002
1115 Taglialatela, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2011).
1116 Chimpanzee Vocal Signaling Points to a Multimodal Origin of Human
1117 Language. *PLoS ONE*, *6*(4). https://doi.org/10.1371/journal.pone.0018852
1118 Tempelmann, S., & Liebal, K. (2012). Spontaneous use of gesture sequences in
1119 orangutans. In *Developments in Primate Gesture Research* (pp. 73–92).
1120 https://doi.org/10.1075/gs.6.05tem
1121 Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford University Press, USA.
1122 Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Farrar, & Evans, A. (1985).
1123 The development of gestural communication in young chimpanzees. *Journal of*
1124 *Human Evolution*, *14*(2), 175–186.
1125 Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M.,
1126 Braga Goncalves, I., Burkart, J. M., Flower, T., Gaunet, F., Glock, H. J., Gruber,
1127 T., Jansen, D. A. W. A. M., Liebal, K., Linke, A., Miklósi, Á., Moore, R., van

Schaik, C. P., Stoll, S., … Manser, M. B. (2017). Exorcising Grice's ghost: An empirical approach to studying intentional communication in animals: Intentional communication in animals. *Biological Reviews*, *92*(3), 1427–1433. https://doi.org/10.1111/brv.12289

Waller, B. M., Liebal, K., Burrows, A. M., & Slocombe, Katie. E. (2013). How Can a Multimodal Approach to Primate Communication Help Us Understand the Evolution of Communication? *Evolutionary Psychology*, *11*(3), 147470491301100. https://doi.org/10.1177/147470491301100305

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, *582*(7812), 337–340. https://doi.org/10.1038/d41586-020-01751-5

Wheeler, B. C. (2010). Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (Cebus apella nigritus). *Behavioral Ecology and Sociobiology*, *64*(6), 989–1000. https://doi.org/10.1007/s00265-010-0914-3

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, *399*(6737), 682–685. https://doi.org/10.1038/21415

Wilke, C., Kavanagh, E., Donnellan, E., Waller, B. M., Machanda, Z. P., & Slocombe, K. E. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, Pan troglodytes schweinfurthii. *Animal Behaviour*, *123*, 305–316. https://doi.org/10.1016/j.anbehav.2016.10.024

Tables

Table 1. Search Tool SPIDER as applied to the current review

Search categories	Potential search terms

Sample	Human and Non-human primates
Phenomenon of **I**nterest	Language Evolution and Origins of Language
Design	Observational and Experimental studies
Evaluation	Gestural research domains of comparison
Research Type	Empirical Articles and Book Chapters

Table 2. Categories of each field extracted for analysis

Field	Categories
Species	Humans
	Non-human primates: Great-apes (Chimpanzees, Bonobos, Gorillas, Orang-utans), Small-apes, Afro-Eurasian monkeys, monkeys of the Americas, and Strepsirrhines.
Socio-ecology	Non-human primates: Species-Typical (Wild non-anthropogenic); and Species-Atypical (Wild anthropogenic, Captive anthropogenic);
Socio-economy	Humans: WEIRD, Non-WEIRD industrialised, Small-scale non-industrialised.
Research domains of comparison	Combination of sources, Form, Laterality, Meaning, Neural processes, Ontogeny, and Structure
Gesture definition	Defined (transcription provided in ESM2), Not defined
	Mechanical ineffectiveness: Mentioned, Not mentioned
Intentional communication criteria	Attention getting behaviour, Audience checking, Ceasing communication when goal is achieved, Change in recipient behaviour, Directed to recipient, Elaboration, Flexibility, Goal-directed, Initiating social interaction, Mechanically ineffective, Persistence, Response waiting, Sensitivity to the attentional state of the recipient
Body parts	Manual (hands, and upper limbs), Non-manual (head, orofacial, body), Both (manual and non-manual)
Gestural scope	Broad, Narrow
Multi-source	Vocalizations, Facial expressions, Lexigrams, None
Research setting	Familiar (Human: Home, Day-care centre; Non-human: Wild, Semi-Wild, Captivity); Unfamiliar (Laboratory); Both
Study Design	Experimental, Observational, Both

Table 3. Descriptions of the research domains of comparison

Research domains	Description

Combination of sources	Coexistence of two different signal categories (gesture, vocalization, and/or facial expression). Includes studies focused on the synchronization of the different sources.
Form	Physical form of the gesture(s). Includes studies that describe gestural repertoires.
Laterality	Predominance of one side of the body/brain involved in the execution of the communicative acts.
Meaning	Used in the broad sense without requirement for intentional use. Includes all informational and/or semantic content of gestures. Includes studies on iconicity, reference, and function.
Neural processes	Use of neuroimaging technology to study neural networks and their relation to communicative outputs.
Ontogeny	Development of communication and how it changes during the early stages of life.
Structure	Order and patterning of the communicative elements. Includes syntax.

Table 4. Number of studies in each domain according to the Gesture definition, Intentional criteria, Gestural scope, Source, Research setting, and Study design. N = number publications. Combination of sources (Comb), form (Form), laterality (Lat), meaning (Mean), neural processes (Neur), ontogeny (Ont) and structure (Struct). The sum of the studies specified by domain for each field differs from the total of studies considered in this analysis because a single study could include more than one domain of gestural research.

		N	Comb	Form	Lat	Mean	Neur	Ont	Struct
	Total	163	73	80	38	139	18	47	34
Gesture definition	Defined	106	44	63	24	93	7	32	16
	Not defined	57	29	17	14	46	11	15	18
Intentional criteria	Considered	52	17	39	8	47	0	10	10
	Not considered	111	56	41	30	92	18	37	24
Body parts	Manual	88	46	39	26	72	15	29	13
	Non-manual	7	4	2	2	4	2	2	1
	Both	56	21	36	9	52	0	14	12
Gestural scope	Broad	51	17	30	14	46	0	11	11
	Narrow	112	56	50	24	93	18	36	23
Sources	Gesture	76	0	44	22	67	5	15	18
	Multi-source	87	73	36	16	72	13	32	16
Research setting	Familiar	96	37	58	24	87	0	24	17
	Unfamiliar	60	31	19	9	47	11	21	16

	Both	7	5	3	5	5	7	2	1
Study design	Experimental	79	43	24	19	64	18	27	16
	Observational	79	27	54	17	71	0	19	17
	Both	5	3	2	2	4	0	1	1

1164

1165 **Table 6. Number of studies using the criteria proposed by Bourjade et al. 2020 for**

1166 **defining 'gesture'.**

Category of criteria	Specific criteria for defining Communicative gesture (CG)	Number of studies
Body part	Any body part including face	5
	Any body part excluding face	36
	Exclusively manual	13
	Exclusively facial	0
Sensory modality	CG has necessarily a visual component	53
	CG is visual or visual and auditory	37
	CG is visual or visual and vocal (mobilizing the vocal chords)	0
	CG is visual or visual and tactile	40
Social expression	CG is produced in the presence of an audience	48
	CG is addressed to specific recipient(s)	44
	CG is produced while looking at the recipient	6
Communicative property	CG encompasses all communicative signals	3
	CG is a subcategory of communicative signals	51
	CG is mechanically ineffective	37
	CG is followed by a change in the recipient's behaviour in most cases	34
Intentional property	CG is necessarily an intentional communicative signal	33
	CG is produced towards a visually attentive recipient	2
	CG is followed by response waiting (a pause in the gesturer's actions)	24
	CG is repeated, combined or replaced by other CGs if no response comes	28
	CG is not consistently triggered by specific environmental conditions	0
	CG is expressed in various functional contexts	0

Table 5. Results of the model (estimates, together with standard errors, confidence intervals, and significant tests). Domains were dummy coded with Combination of Sources being the reference category.

Terms	Estimate	SE	lower CI	upper CI	z-value	*P*
(Intercept)	-209.357	52.542	-324.425	-116.015		[1]
Domains Form	-100.185	80.641	-263.923	57.334		[1]
Domains Laterality	-61.053	77.288	-216.596	91.426		[1]
Domains Meaning	-39.745	69.151	-174.244	100.741		[1]
Domains Neural processes	-5.944	82.719	-176.749	154.343		[1]
Domains Ontogeny	3.194	69.615	-132.627	144.387		[1]
Domains Structure	-239.809	124.467	-514.930	-18.625		[1]
Year	0.104	0.026	0.058	0.161	3.981	<0.001
Domains Form : Year	0.050	0.040	-0.028	0.131	1.246	0.213
Domains Laterality : Year	0.030	0.038	-0.045	0.108	0.793	0.428
Domains Meaning : Year	0.020	0.034	-0.050	0.087	0.583	0.560
Domains Neural processes : Year	0.003	0.041	-0.077	0.088	0.068	0.946
Domains Ontogeny : Year	-0.001	0.035	-0.072	0.066	-0.042	0.966
Domains Structure : Year	0.119	0.062	0.009	0.256	1.926	0.054

[1] not shown because of having a very limited interpretation

Table 7. Different criteria used by the studies to define intentional gestural communication. Authors often used multiple criteria to define intentional communication. One of the studies using multiple intentional criteria was conducted in both Human and Non-human primates; as a result, the sum of the studies with humans and non-humans may exceed the total number of studies that considered each criterion.

Criteria for Intentional Gesture	Human	Non-human	Total
Persistence	2	33	34
Response waiting	1	29	30
Directed to recipient	3	24	27
Audience checking	2	23	24
Sensitivity to the attentional state of the recipient	1	23	24

Goal-directed	1	14	15
Elaboration	0	10	10
Change in recipient behaviour	0	7	7
Mechanically ineffective	1	5	6
Ceasing communication when the goal is achieved	0	5	5
Attention getting behaviour	1	3	3
Initiating social interaction	0	3	3
Flexibility	0	2	2

Table 8. Number of studies discriminated by Research setting

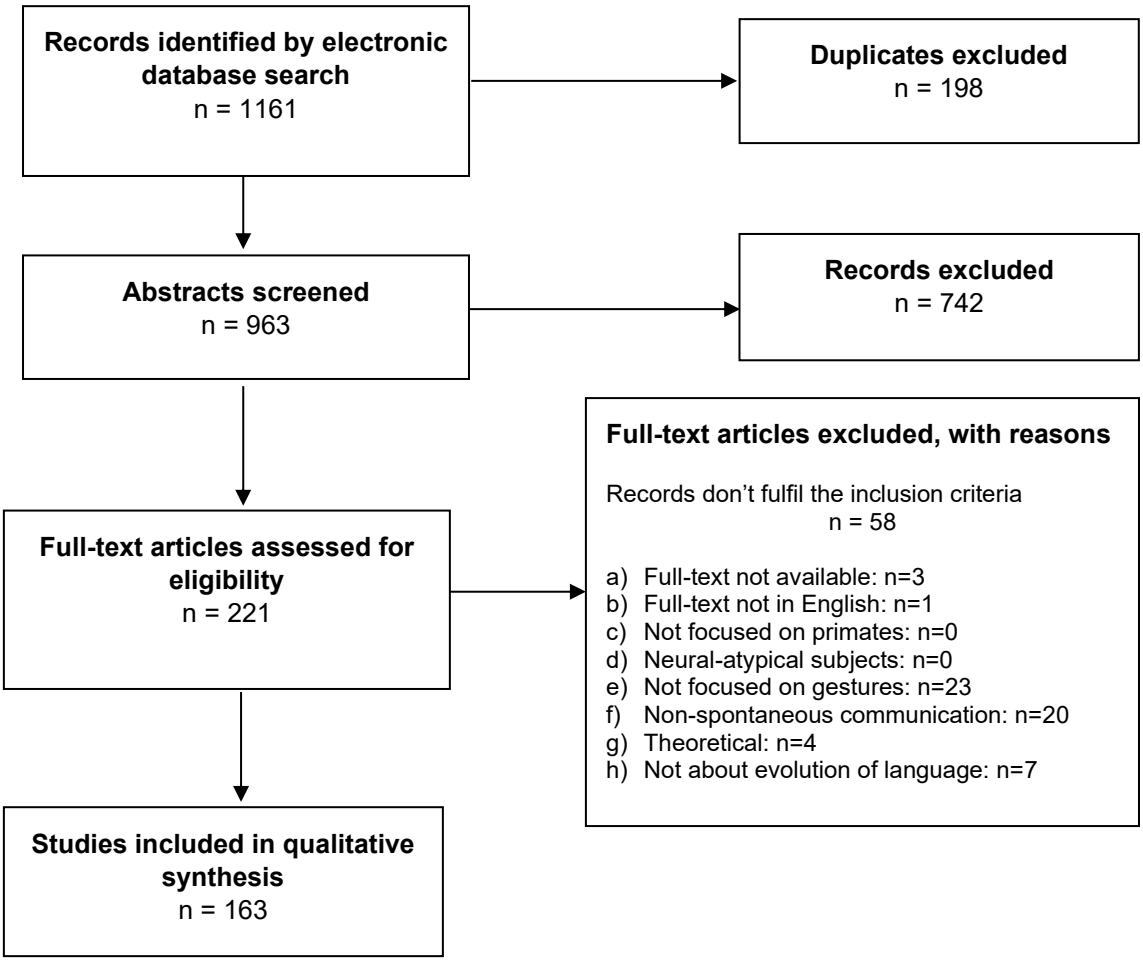
Research setting	Human	Non-human
Familiar		
Daycare	5	-
Home	19	-
Daycare + Home	1	-
Captivity	-	35
Wild	-	30
Captivity + Wild	-	1
Sanctuaries	-	5
Unfamiliar		
Laboratory	53	7
Both		
Captivity + Laboratory	-	7
Home (H)+ Laboratory (NH)	2	2

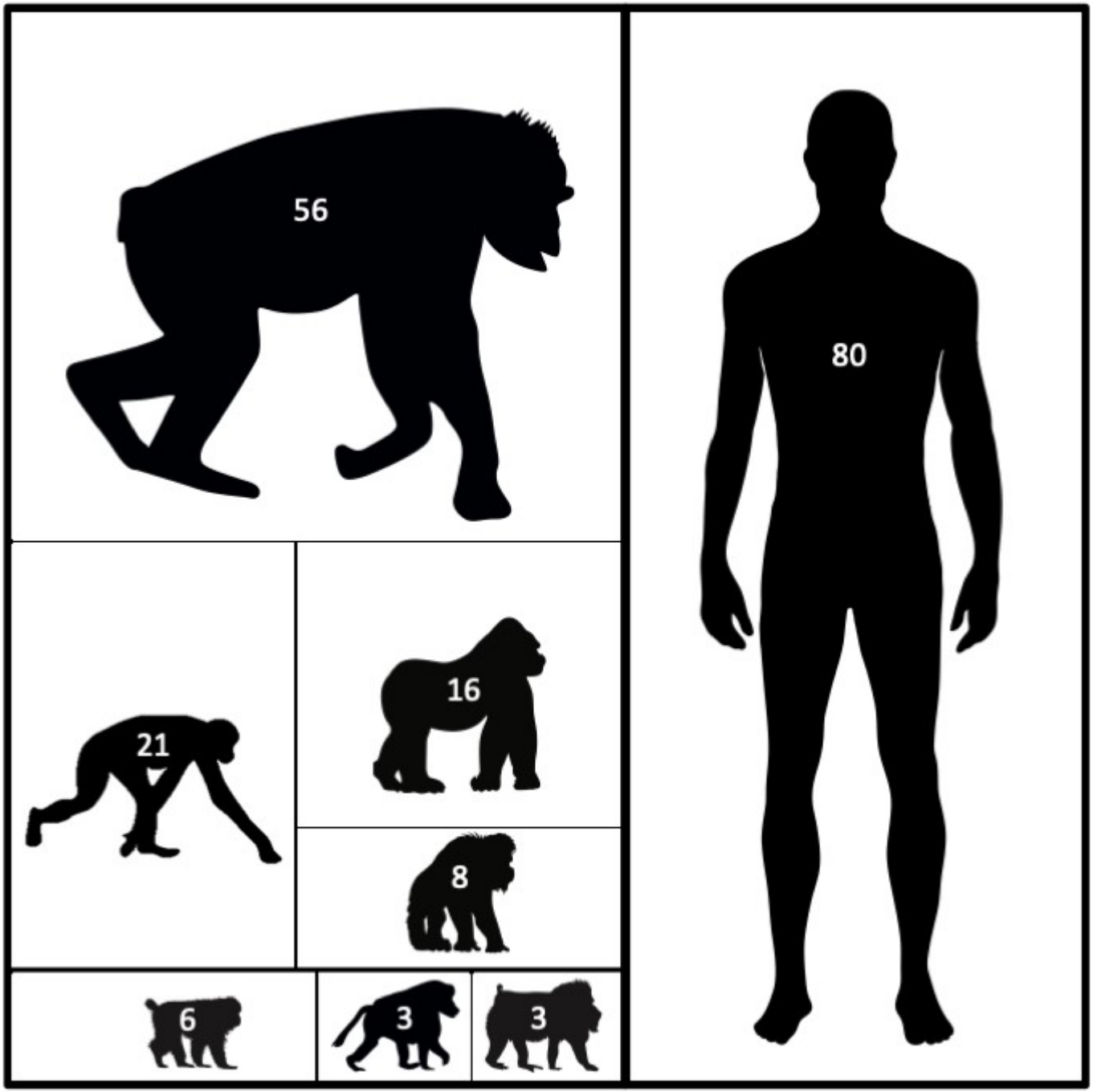
Figure legends

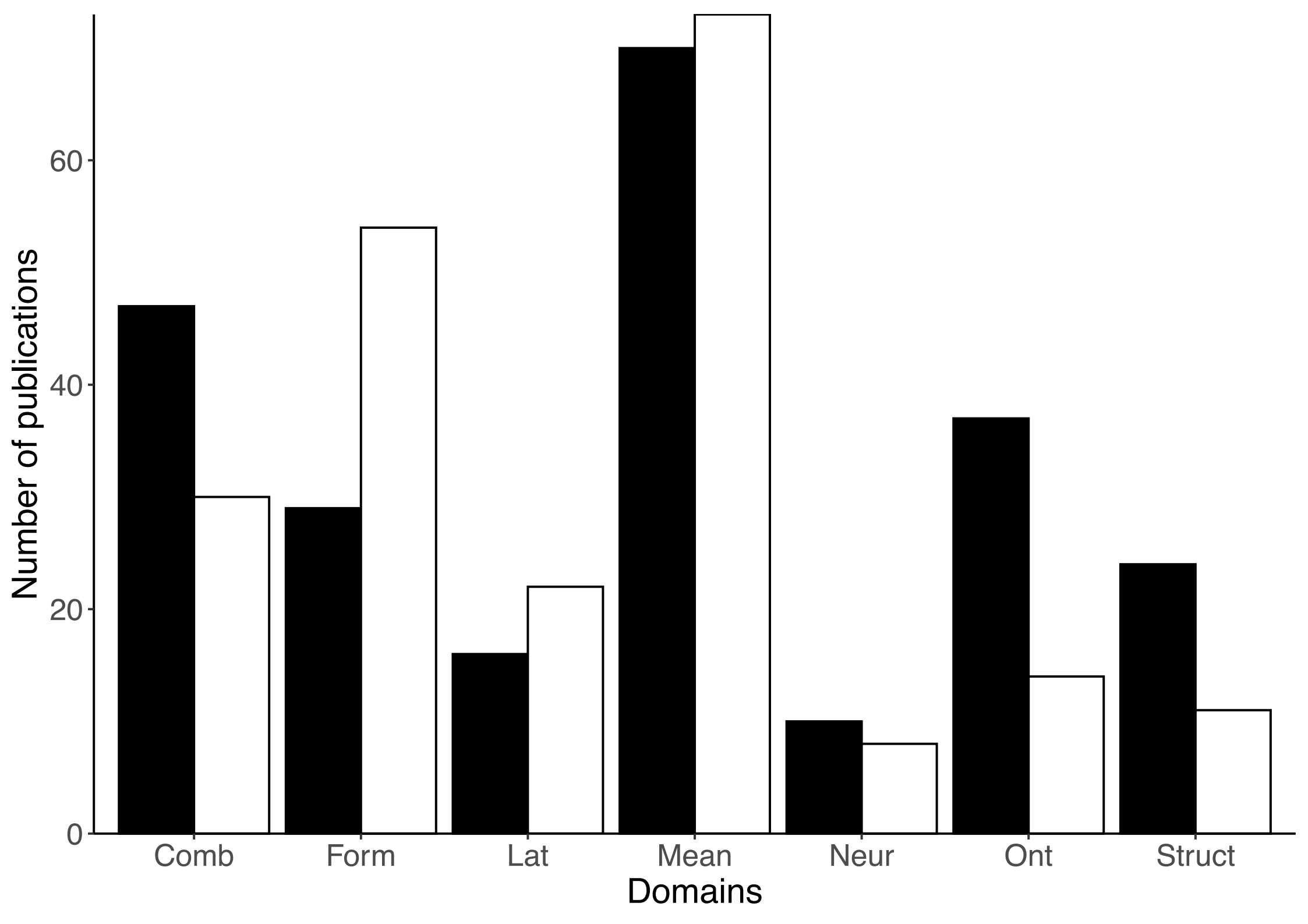
Figure 1. FLOW diagram. Number of studies identified, included and excluded and the reasons for exclusions through the different phases of the Systematic Review

Figure 2. Primate species represented in the review. Area represents the number of studies in the review including that species. Non-human primates (on the left) include Great apes (chimpanzees, bonobos, gorillas, and orang-utans), and Afro-Eurasian monkeys (macaques, baboons, and mandrills).

Figure. 3 Number of studies in Humans and Non-human primates across different research domains. Black bars indicate human subjects, white bars non-human subjects. Combination of sources (Comb), Form (Form), Laterality (Lat), Meaning (Mean), Neural processes (Neur), Ontogeny (Ont) and Structure (Struct). Note that some studies included more than one research domain.


Records identified by electronic database search
n = 1161
Duplicates excluded
n = 198
Abstracts screened
n = 963
Records excluded
n = 742
Full-text articles assessed for eligibility
n = 221
Full-text articles excluded, with reasons
Records don't fulfil the inclusion criteria
n = 58
a) Full-text not available: n=3
b) Full-text not in English: n=1
c) Not focused on primates: n=0
d) Neural-atypical subjects: n=0
e) Not focused on gestures: n=23
f) Non-spontaneous communication: n=20
g) Theoretical: n=4
h) Not about evolution of language: n=7
Studies included in qualitative synthesis
n = 163


56
80
16
21
8
6
3
3


Number of publications
0
20
40
60
Comb
Form
Lat
Mean
Neur
Ont
Struct
Domains