

Dear Author

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For **fax** submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style.
- Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL:

<http://dx.doi.org/10.1007/s12021-013-9182-5>

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information, go to:

<http://www.springerlink.com>.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us, if you would like to have these documents returned.

Metadata of the article that will be visualized in OnlineFirst

1	Article Title	Ontogenetic Ritualization of Primate Gesture as a Case Study in Dyadic Brain Modeling
2	Article Sub- Title	
3	Article Copyright - Year	Springer Science+Business Media New York 2013 (This will be the copyright line in the final PDF)
4	Journal Name	Neuroinformatics
5	Corresponding Author	Family Name Gasser
6		Particle
7		Given Name Brad
8		Suffix
9		Organization University of Southern California
10		Division Neuroscience Graduate Program and USC Brain Project
11		Address Los Angeles, CA, USA
12		e-mail bgasser@usc.edu
13	Author	Family Name Cartmill
14		Particle
15		Given Name Erica A.
16		Suffix
17		Organization University of Chicago
18		Division Department of Psychology
19		Address 5848 S. University Ave., Chicago 60637, IL, USA
20		e-mail
21	Author	Family Name Arbib
22		Particle
23		Given Name Michael A.
24		Suffix
25		Organization University of Southern California
26		Division Neuroscience Graduate Program and USC Brain Project
27		Address Los Angeles, CA, USA
28		Organization University of Southern California
29		Division Computer Science Department

30	Address	Los Angeles 90089, CA, USA
31	e-mail	
32	Received	
33	Schedule	Revised
34		Accepted
35	Abstract	<p>This paper introduces <i>dyadic brain modeling</i> – the simultaneous, computational modeling of the brains of two interacting agents – to explore ways in which our understanding of macaque brain circuitry can ground new models of brain mechanisms involved in ape interaction. Specifically, we assess a range of data on gestural communication of great apes as the basis for developing an account of the interactions of two primates engaged in <i>ontogenetic ritualization</i>, a proposed learning mechanism through which a functional action may become a communicative gesture over repeated interactions between two individuals (the ‘dyad’). The integration of behavioral, neural, and computational data in dyadic (or, more generally, social) brain modeling has broad application to comparative and evolutionary questions, particularly for the evolutionary origins of cognition and language in the human lineage. We relate this work to the neuroinformatic challenges of integrating and sharing data to support collaboration between primatologists, neuroscientists and modelers that will help speed the emergence of what may be called <i>comparative neuro-primatology</i>.</p>
36	Keywords separated by ' - '	Computational model - Primate - Gesture - Social learning - Ontogenetic ritualization - Neuroinformatics
37	Foot note information	

1
3
2

ORIGINAL ARTICLE

4
5

Ontogenetic Ritualization of Primate Gesture as a Case Study in Dyadic Brain Modeling

6

Brad Gasser · Erica A. Cartmill · Michael A. Arbib

7
8

© Springer Science+Business Media New York 2013

9

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

Abstract This paper introduces *dyadic brain modeling* – the simultaneous, computational modeling of the brains of two interacting agents – to explore ways in which our understanding of macaque brain circuitry can ground new models of brain mechanisms involved in ape interaction. Specifically, we assess a range of data on gestural communication of great apes as the basis for developing an account of the interactions of two primates engaged in *ontogenetic ritualization*, a proposed learning mechanism through which a functional action may become a communicative gesture over repeated interactions between two individuals (the ‘dyad’). The integration of behavioral, neural, and computational data in dyadic (or, more generally, social) brain modeling has broad application to comparative and evolutionary questions, particularly for the evolutionary origins of cognition and language in the human lineage. We relate this work to the neuroinformatic challenges of integrating and sharing data to support collaboration between primatologists, neuroscientists and modelers that will help speed the emergence of what may be called *comparative neuro-primatology*.

30
31
32

Keywords Computational model · Primate · Gesture · Social learning · Ontogenetic ritualization · Neuroinformatics

Q6

B. Gasser (✉) · M. A. Arbib
Neuroscience Graduate Program and USC Brain Project,
University of Southern California, Los Angeles, CA, USA
e-mail: bgasser@usc.edu

E. A. Cartmill
Department of Psychology, University of Chicago,
5848 S. University Ave.,
Chicago, IL 60637, USA

M. A. Arbib
Computer Science Department, University of Southern California,
Los Angeles, CA 90089, USA

Introduction

33 Q1

This paper is intended to highlight the promise of the emerging field of *comparative neuro-primatology* and to propose informatic tools and interdisciplinary directions that will open up new avenues of research for ethology and neuroscience. Computational modeling, and specifically *dyadic/social brain modeling*, can be used to integrate, extend, and test theories from both the neuroscientific and behavioral sciences. However, there are considerable theoretical and practical challenges to building realistic neuro-computational models of social behavior – both capturing the social elements of behavior, and making the most of the limited data that is currently available.

To illustrate the challenges of this integrative modeling approach, we propose a computational model based on the gestural communication of great apes. Gesture – in particular gestural acquisition – provides an excellent case study in social brain modeling because it raises issues that would be problematic for modeling any social behavior generally. For example, how do the brains of interacting agents process shared events differently? How do agents respond to behavioral changes in others, and how are these changes reflected in brain activations and/or adaptive synaptic wiring? Are there dedicated neural structures or pathways for social interaction, or do primates largely rely on domain-general regions for social cognition? Focusing on gesture also grants us empirical purchase as ape gestural behavior has long been studied, and manual action production and recognition systems in monkeys are fairly well characterized at the neural level. Finally, the added learning component of gesture acquisition forces us to consider both immediate and lasting changes in the neural organization of behavior.

The learning process we discuss – *ontogenetic ritualization* – has been proposed as a mechanism through which great apes may acquire new communicative gestures through the mutual shaping of action, resulting in a stable,

34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68

69 but non-arbitrary gestural form. Modeling the process of
70 ontogenetic ritualization provides several distinct challenges
71 that must be confronted. The model must be able to account
72 for the fact that ontogenetic ritualization is (i) a *dynamic*
73 *process* in which (ii) *multiple individuals* process and re-
74 spond to the interaction differently, while the interaction
75 itself (iii) *changes* and develops over time.

76 Constructing a model of the dyadic interactions of the
77 social brain requires integration of data across multiple
78 datasets, methodologies, and disciplines, and thus places
79 unique demands on informatics tools and resources. Data
80 management tools and techniques for integrating resources
81 must focus on efficiently navigating questions of homolo-
82 gies between species, finding the appropriate granularity of
83 data for modeling projects, and producing simulation results
84 specific enough to test existing frameworks and offer novel
85 hypotheses. We highlight the need for new and more inte-
86 grated resources for researchers operating in these highly
87 interdisciplinary fields, and offer new suggestions and chal-
88 lenges for the neuroinformatics community.

89 Primate Social Behavior

90 All animals face the challenges of finding and obtaining
91 food, water, shelter, and suitable mates while, at the same
92 time, minimizing injury from competitors or predators. For
93 social animals, these physical challenges arise in an abstract
94 network of social relationships that often impact an individ-
95 ual's success, and which must therefore be tracked, fostered,
96 and exploited. The need for such socio-cognitive abilities
97 likely provided a strong selection pressure that helped shape
98 both brain structure and cognitive skill in the primate line-
99 age (Byrne and Whiten 1988; Dunbar 1998; Sallet et al.
100 2011). The study of the primate brain has only begun to
101 explore the neural correlates of these socio-cognitive abili-
102 ties, but new developments in brain imaging and neurophys-
103 iological designs allow neural activity to be measured in
104 both human and non-human primates during social interac-
105 tion. For this paper, we focus on non-human primate –
106 henceforth “primate” – data, but recognize the substantial
107 insights that can be gained from human lesion and neuro-
108 imaging studies (Adolphs 2010; Amodio and Frith 2006;
109 Shilbach et al. 2012). Combining insights from neural and
110 behavioral studies promises to greatly increase the scope of
111 the questions that may be addressed.

112 Ape Gesture

113 Great apes – gorillas, bonobos, chimpanzees and orang-
114 utans – are proficient at copying manual actions, though
115 their skills in this area are limited relative to those of humans
116 (Byrne and Russon 1998; Dean et al. 2012). The ability to

117 acquire manual skills through observation of conspecifics
118 likely plays a role in the development of group-specific
119 behaviors in both wild and captive populations. ‘Local
120 traditions’ (behaviors restricted to particular populations or
121 subgroups of individuals) involving the presence or varia-
122 tion of manual actions, such as tool use, food processing,
123 and grooming, have been reported in both wild and captive
124 great apes (Byrne 2004; Hobaiter and Byrne 2010; van
125 Schaik et al. 2003; Whiten et al. 2001), and provide evi-
126 dence that apes can develop *cultures* of behavior (Whiten et
127 al. 1999).

128 Surprisingly, manual *gestures* do not show the same
129 levels of inter-group variability as manual actions do. Stud-
130 ies of ape gesture consistently report that the majority of
131 gestures are either “species typical” (i.e., used by members
132 of a species regardless of what geographic site they inhabit),
133 or are idiosyncratic and therefore produced by only a single
134 individual—and presumably recognized by at least one oth-
135 er (Genty et al. 2009; Hobaiter and Byrne 2011; Liebal et al.
136 2006; Pika et al. 2003). There have been some reported
137 differences in gesture form or use between apes at different
138 research sites, but the predominant pattern is one of similar-
139 ity across sites, with most of the gestures observed at site A
140 also observed at sites B and C. The gestural repertoires of
141 individual apes at different sites typically overlap as much
142 as those of apes at the same site (Call and Tomasello 2007).
143 Additionally, gestural repertoires typically vary more
144 strongly between age classes than between sites – with
145 juveniles displaying largely different repertoires from
146 adults. The proportion of apes using a particular gesture
147 does vary between site – at some sites a gesture will be used
148 by the majority of individuals, while at others it will only be
149 used by a few – but it does not vary much (Genty et al.
150 2009; Hobaiter and Byrne 2011). There are some exceptions
151 to the ‘species-typical or idiosyncratic’ characterization of
152 gesture use. A few group-specific gestures have been ob-
153 served in orangutans (Cartmill 2008; Liebal et al. 2006),
154 gorillas (Genty et al. 2009; Pika et al. 2003), chimpanzees
155 (Nishida 1980), and bonobos (Pika et al. 2005). These
156 studies suggest the existence of group-specific gestures that
157 may result from social learning (Arbib et al. 2008), though
158 some have argued that reports of relatively infrequent ges-
159 tures observed only in one group may simply be an artifact
160 of under-sampling (Genty et al. 2009). However, if a gesture
161 is used frequently in one group and rarely or never in
162 another, a strong case can be made for a local ‘gesture
163 culture.’

164 A significant problem with this approach of comparing
165 gesture use across sites is that unless all data are collected
166 and coded according to the same criteria, gestures at multi-
167 ple sites may vary in how they are defined and recorded.
168 This, in turn, may lead to inaccurate estimates of the reper-
169 toire overlap between groups. Comparing gestures across

170 multiple sites and species is crucial to understanding gestural
171 al communication in primates, and new informatics
172 approaches to integrating data gathered at different sites
173 are needed to make significant progress in this field. We
174 return to this issue in the final section.

175 Primate Neurophysiology

176 We are concerned with building a bridge between ape social
177 behavior and its underlying neural circuitry. However, while
178 there are data sets on single cell recordings in monkeys
179 (especially in macaques), no such data are available for apes
180 (although brain imaging data are now becoming available).
181 Thus, our strategy is indirect, extending our understanding
182 of brain processes in monkeys to construct a framework for
183 modeling the social behavior of apes. In this section we will
184 describe neurophysiological studies on macaques that have
185 linked neural activity to both manual behaviors and cogni-
186 tive abilities. In a later section we review a key set of
187 computational models linking vision and action and which
188 describe the neurophysiological data. We examine the
189 assumptions and limitations of these models and ask: “what
190 properties must be added to macaque models to support
191 brain modeling of ape (social) behavior?”

192 Primate neurophysiology has been used to address some
193 aspects of manual and social behavior in monkeys, but the
194 designs have traditionally relied on ‘passive’ designs that do
195 not require interaction between the subject and another
196 individual. For instance, the research on ‘mirror neurons’
197 have always been *passively* social in that neural responses
198 could be elicited by observing the performance of other
199 individuals (di Pellegrino et al. 1992; Gallese et al. 1996)
200 rather than through interaction. Neurons were found in
201 premotor (and later in parietal) areas whose activity during
202 an individual’s own performance of a particular action was
203 found to be similar to the activity in response to observing
204 another individual – usually a human researcher –
205 performing a more-or-less similar action. In this way, mirror
206 neurons can be driven by social variables, but the experi-
207 mental designs do not require the monkeys to differentially
208 ‘use’ this information, and so cannot assess how these
209 responses affect downstream targets.

210 These passive designs can be contrasted with explicitly
211 social or ‘interactive’ designs, more recently developed, that
212 require the subject to directly interact with other entities,
213 whether computer agents (Lee et al. 2005; Seo et al. 2009;
214 Seo and Lee 2007) or conspecifics (Azzi et al. 2011; Chang
215 et al. 2012; Fujii et al. 2008; Yoshida et al. 2011, 2012).
216 These interactive designs have led to new insights into how
217 the brain is organized to process specifically social informa-
218 tion, how this information affects downstream targets, and
219 how interaction between a monkey and another agent places
220 unique demands on processing structures within the brain.

For instance, responses in medial parts of frontal cortex, in 221
and around pre-SMA, have been found to be ‘other’ respon- 222
sive neurons – instead of firing both when an action is done 223
by one’s self and when observing another, as in the above 224
‘mirror neurons’, these only fire during observation of 225
another’s actions (Yoshida et al. 2011). Orbito-frontal cortex 226
(OFC) neurons, recorded in monkeys playing interactive 227
games, revealed modulations encoding social influences on 228
motivation and reward processing (Azzi et al. 2011). OFC 229
and anterior cingulate (ACC) neurons, in a separate but 230
similarly ‘interactive’ study, were shown to differentially 231
process how rewards were allocated between others and 232
oneself, with ACC gyrus appearing important for processing 233
the ‘shared experience’ of rewards (Chang et al. 2012). 234
Together, these and other data demonstrate that social 235
behaviors are becoming increasingly accessible to neuro- 236
physiological study in monkeys, and not just in a ‘passively 237
social’ sense, but within tasks demanding back-and-forth 238
exchanges. Additionally, the emergence of functional mon- 239
itoring via PET, fMRI and other neuroimaging techniques 240
adapted to non-human primates is most promising. These 241
have the double advantage of being non-invasive, while 242
being of the same ‘format’ of the most used techniques in 243
humans, easing comparison of data across species. 244

For example, non-invasive EEG and ERP studies have 245
recently been applied to the production and comprehension 246
of vocal communicative behaviors in *apes* (Hirata et al. 247
2011; Ueno et al. 2008). These techniques complement the 248
emerging use of PET in functional brain monitoring in apes 249
(Parr et al. 2009; Tagliatela et al. 2011). And for eye- 250
tracking, an indirect measure for attentional processing, with 251
chimpanzees, see Kano and Tomonaga (2009). Combining 252
functional data with mathematical techniques to understand 253
these indirect measures of brain activity in terms of neural 254
firing allows researchers to ‘convert’ data between domains 255
of analysis, including making non-invasive functional data 256
more compatible with neuro-computational analysis (fMRI: 257
Arbib et al. 2000; PET: Arbib et al. 1995; ERP: Barrès et al. 258
2013). All the above methods can be combined with the use 259
of structural imaging techniques such as DTI (Hecht et al. 260
2012; Ramayya et al. 2010; Rilling et al. 2008), MRI (Sakai 261
et al. 2011), and other imaging, anatomic and cytoarchitec- 262
tonic methods comparing primate brains (Hopkins et al. 263
2010; Keller et al. 2012; Rilling et al. 2011; Schenker et 264
al. 2010). Given the difficulty in directly assessing brain 265
function in apes, it is necessary for researchers to use exist- 266
ing data from all available techniques to develop more 267
complete models of primate neural processing during social 268
behavior. 269

In this paper, we present the design of a brain-based 270
conceptual model – to be followed with a fully implemented 271
computational model elsewhere – aimed at testing a pro- 272
posed learning process through which great apes may 273

274 develop manual communicative gestures. Despite a dearth
 275 of direct neural data for gestural communication in great
 276 apes (Tagliatalata et al. 2011), we have several reasons to
 277 focus on gesture acquisition as a test case for modeling
 278 social cognition. Firstly, the proposed learning process –
 279 ontogenetic ritualization – rests on repeated interactions
 280 between pairs of individuals, thus demanding a direct treat-
 281 ment of social interaction. Secondly, computational model-
 282 ing of primate manual gesture intersects with a growing
 283 body of work on the mirror system and has implications
 284 for understanding the origins of human language. The Mir-
 285 ror System Hypothesis (MSH; Arbib 2010, 2005, 2008,
 286 2012) makes explicit claims about brain function evolution
 287 throughout the hominid line, and the concomitant capacity
 288 for social learning and flexibility in intentional communica-
 289 tion, that made the human brain ‘language-ready’. Although
 290 others have adopted a neuro-evolutionary approach to com-
 291 munication (e.g., Aboitiz 2012; Corballis 2002; Deacon
 292 1997), MSH is unique in explicitly grounding the evolution-
 293 ary account in the computational description of macaque
 294 neural processing (including ‘mirror neuron’ systems) and
 295 ape behavior (including gesture). It is within this MSH
 296 framework that we approach our case study, emphasizing
 297 the computational description of brain function to formalize
 298 hypotheses on gesture acquisition.

299 In order to properly contextualize our proposed model,
 300 we first outline the claims of ontogenetic ritualization, and
 301 then provide details on primate brain mechanisms known to
 302 be important for manual and social tasks (especially those
 303 formalized in computational models). We then describe our
 304 model – a conceptual analysis of the proposed process of
 305 ontogenetic ritualization – and discuss those features impor-
 306 tant for the field of ‘dyadic/social brain modeling’. Finally,
 307 we consider the impact of incorporating observational, ex-
 308 perimental, and computational approaches in the study of
 309 the social brain, and conclude with a discussion of issues
 310 related to data management and sharing that will support
 311 future interdisciplinary collaborations.

312 Ontogenetic Ritualization

313 Ontogenetic ritualization (OR) is the proposed process of
 314 ritualizing movements of ‘effective’ actions (those that di-
 315 rectly alter the behavior of other individuals) into commu-
 316 nicative signals aimed at eliciting particular responses in
 317 others (Tomasello and Call 2007; Tomasello and Camaioni
 318 1997). During this process of ritualization, a movement such
 319 as shoving another out of the way becomes ritualized over
 320 time into a ‘nudge’ as the actor learns that only the begin-
 321 ning of the movement is necessary to elicit the desired
 322 behavior in the recipient, and as the recipient learns to
 323 respond to the gesturer using only the initial movements of

the action. However, according to this process, the actor and
 recipient form different associations resulting from their
 respective roles in the interactions – the recipient may only
 be able to perceive but not produce the gesture, and vice-
 versa (Genty et al. 2009). The degree to which OR plays a
 role in the acquisition of ape gestures is debated (Genty et
 al. 2009; Perlman et al. 2012; Tomasello and Call 2007).
 Here, we do not take a strong stance on whether OR is the
 main acquisition mechanism for ape gesture, but we do note
 that OR could explain those species-typical (and not just
 idiosyncratic) gestures whose relation to species-typical
 actions is readily derivable through the OR process. We
 propose a model of the cognitive and neural changes that,
 we hypothesize, *could* support OR. It is our hope that such
 modeling work will make it possible to identify the condi-
 tions under which OR is a plausible mechanism for gesture
 acquisition, while simultaneously generating hypotheses for
 new behavioral and neuroimaging experiments that test
 social and communicative behaviors more broadly.

The process of ontogenetic ritualization is described by
 Call and Tomasello (2007) as proceeding in three steps
 (Fig. 1, left):

- (1) Individual A performs behavior X (not a communica-
 tive signal), and individual B consistently reacts by
 doing Y
- (2) Subsequently B anticipates A’s overall performance of
 X by starting to perform Y before A completes X
- (3) Eventually, A anticipates B’s anticipation and produces
 an initial portion of X in a ritualized form X^R in order
 to elicit Y.

Of particular relevance to social brain modeling is that
 this is a *dyadic* learning process – it requires *differential*
 learning in the brains of A and B, which reflects the chang-
 ing patterns of interaction between them throughout the
 ritualization process.

Ontogenetic ritualization is thought to underlie the devel-
 opment of some human gestures, but the process in human
 differs in some important ways from the ape process we
 focus on here (Fig. 2b). The palm-up ‘gimme’ gesture or the
 ‘arms up’ gesture in which an infant raises his arms to indicate
 a desire to be picked up are good examples of ritualized
 human gestures (Bruner et al. 1982; Clark 1978). Initially,
 these types of gestures occur only in the immediate context of
 the actions they are derived from – a 9-month-old infant will
 use the ‘arms up’ gesture only when an adult behaves as
 though she is about to pick him up (perhaps only following
 the adult’s contact under the arms of the child). Over time,
 however, the gestures become more removed from these
 particular contexts so that a 13-month-old infant might use
 the gesture according to his own desires to *request* rather than
 facilitate being picked up (Lock 2001). Thus, ‘arms up’
 emerges as a sign used with communicative intentions.

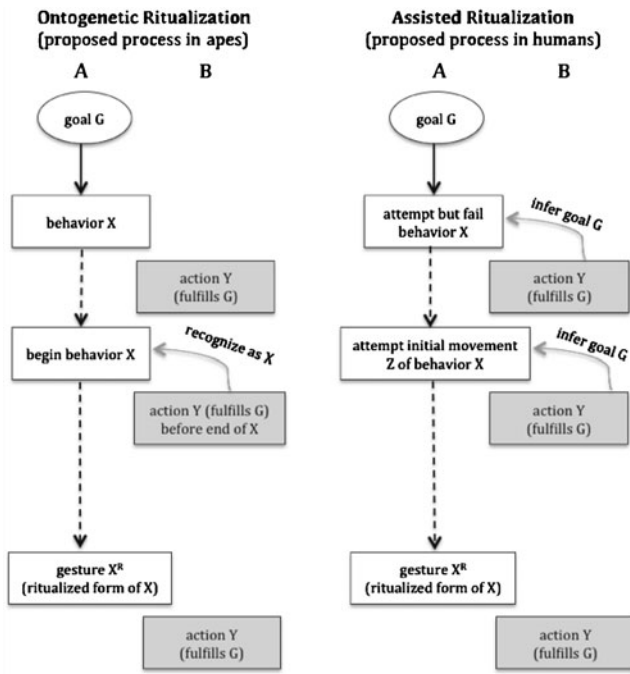


Fig. 1 Processes resulting in ontogenetic and assisted ritualization. (Left) Ontogenetic (naïve) ritualization yields a gestural form through the mutual shaping of behavior between individuals A and B. In each iteration, individual A begins with goal G, interacts in some way with individual B, and B fulfills the goal G through action Y (shaded box). Over time, B performs Y in response to shorter and shorter segments of X, resulting in A producing the ritualized gesture X^R (last boxes). (Right) Assisted ritualization is similar to ontogenetic ritualization, with the exception that individual B ‘guides’ the behavior of individual A, by inferring the goal G and modeling or facilitating the performance of X^R. Here, the shaping is primarily unidirectional (B shaping A), whereas at right, the shaping is bidirectional and makes fewer assumptions about the mental states of the interacting agents

376 In human infants, this ritualization process may be
 377 more accurately described as *assisted* ritualization
 378 (Zukow-Goldring and Arbib 2007) because the adult
 379 recipient is monitoring and reinforcing seemingly com-
 380 municative behavior in the infant, and in many cases the
 381 adult already has an idea of what the final gestural form
 382 should look like (X^R in Fig. 1, right) – because, for
 383 example, it already exists in the culture. Thus, the pro-
 384 cess and speed of ritualizing an action into a gesture
 385 becomes driven by the recipient.

386 In the case study we consider, we restrict ourselves to the
 387 first interpretation of ontogenetic ritualization as a *naïve*
 388 interactive process through which a sign *emerges*, rather than
 389 a process in which a sign is shaped by a more knowl-
 390 edgeable individual. It may be the case that experienced primates
 391 play a more active role in shaping the behavior of others as
 392 humans are known to do (see Ferrari et al. 2009), but here
 393 we focus on the simpler, unassisted version of ontogenetic
 394 ritualization in which each participant is naïve as to what the
 395 final form of the gesture will be.

Action, Perception and Cognition in the Brain

396

To fully represent the changes in the brain of each partici- 397
 pant during ontogenetic ritualization, our model must mini- 398
 mally incorporate brain structures critical for (i) the visuo- 399
 motor control necessary for action and gesture, (ii) recog- 400
 nizing and responding to the actions of others, and (iii) 401
 motivating social interactions between conspecifics – as 402
 well as how learning affects each. We now review some 403
 known primate brain systems for visually-guided grasping, 404
 action-recognition, and decision-making, and outline their 405
 proposed computational properties. In a later section, we 406
 will suggest how these brain mechanisms supporting praxic 407
 actions directed at changing the physical state of an object 408
 can provide a basis for extended circuitry that also supports 409
 communicative actions (e.g., gestures) directed at changing 410
 the behavior of a conspecific. 411

Visually-Guided Grasping

412

The FARS model (Fagg and Arbib 1998) has been offered 413
 as a computational description of manual visuo-motor coor- 414
 dination in the macaque brain. FARS describes the fronto- 415
 parietal reach-to-grasp *production* circuitry macaques use 416
 when they manually grasp objects (so called reach-to- 417
 grasp actions), based on neurophysiological data. Briefly, 418
 parietal structures on the dorsal stream extract the ‘affordan- 419
 ces’ of the world relevant to the grasp (i.e., the physical and 420
 spatial properties of the object to be grasped) and forward 421
 these to premotor cortex for selection of an appropriate 422
 grasping action (Jeannerod et al. 1995). A ventral object- 423
 recognition path allows prefrontal structures to select an 424
 appropriate motor program when working memory or task 425
 structure provides relevant constraints. The model addition- 426
 ally invokes interaction between prefrontal cortex and the 427
 basal ganglia when a sequence of actions is required to 428
 complete the overall task. This computational description 429
 of monkey manual control – well supported by contempo- 430
 rary accounts of brain function (Cisek 2007; Cisek and 431
 Kalaska 2010) – can help frame our model of gesture 432
 learning. It is important to point out, however, that such an 433
 ‘affordance-driven’ description must be complemented with 434
 a description of the control structures participating in guid- 435
 ing hand motions without explicit physical targets, as would 436
 occur during the performance of *intransitive* gestures (as 437
 opposed to tactile gestures like the ‘nudge’ example dis- 438
 cussed previously). 439

Action-Recognition

440

The MNS, for Mirror Neuron System (Oztop and Arbib 441
 2002), and MNS2 (Bonaiuto et al. 2007) models build off 442
 of the computational description of manual *action* in FARS 443

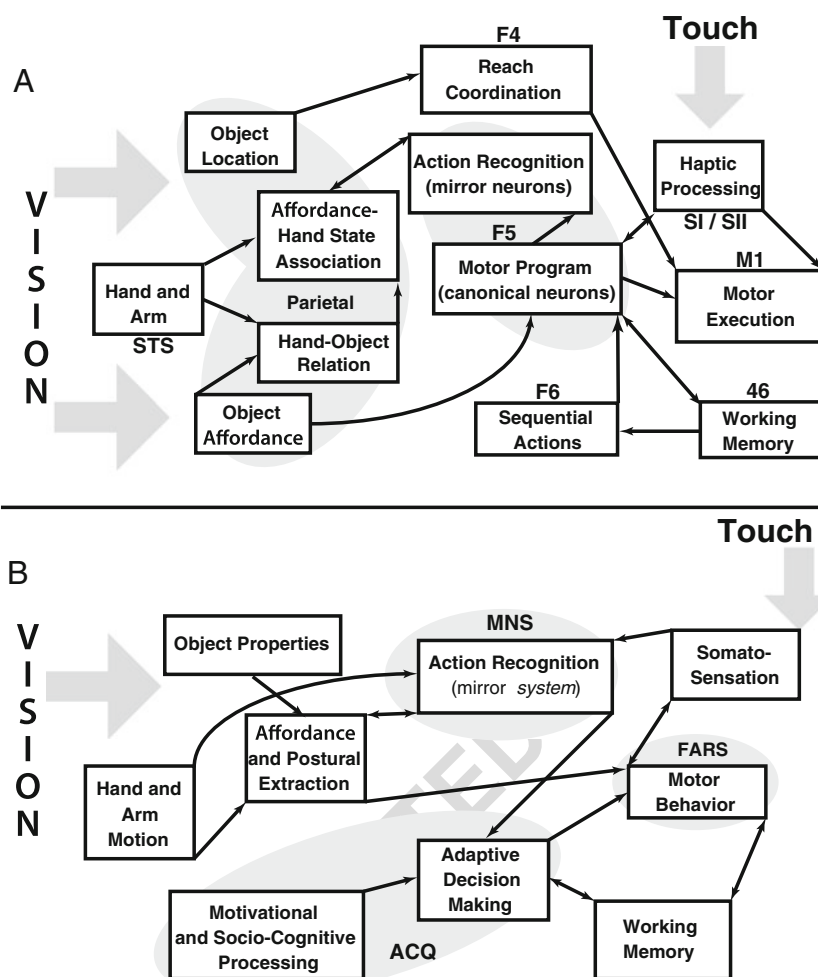


Fig. 2 Data-driven model development. Our proposed model is based on the functionality of many previous computational models, and of an analysis of where model integration is possible, and where model performance requires ‘extensions’ in its computational powers. **a** A schematic of primate manual control and recognition, based on the MNS (Oztop and Arbib 2002) model of action recognition, and the FARS (Fagg and Arbib 1998) model of action production, in the macaque. Note that mirror neuron responses are limited to grasping actions directed at objects, and manual control is similarly limited to

object-directed actions; the model would fail to respond to (simulated) intransitive movements. Shaded areas correspond to putative anatomical localization. **b** A schematic of simple ‘addition’ of models discussed in the text, including MNS, FARS, and ACQ (Bonaiuto and Arbib 2010), and of novel connections between modules, including greater postural (intransitive) and tactile-based action recognition, expanded postural control of limbs, and socially-motivated decision-making. Note that shaded regions correspond to primary architecture of previous models, and not anatomical localization

444 to describe the *recognition* component of the mirror neuron
 445 responses. These models have suggested that mirror neurons
 446 learn their property of ‘action parity’ – responding similarly
 447 for *production* and *recognition* – by learning the visual
 448 trajectory of the hand in relation to objects for actions
 449 already in the agent’s repertoire – combining signals of
 450 visual feedback during the course of generating an action
 451 with the efferent motor commands controlling that action
 452 (Oztop and Arbib 2002). These models formalize how neu-
 453 rons in parietal and premotor regions can learn to recognize
 454 a range of movements associated with a given reach-to-
 455 grasp action, and illuminate how action recognition in mac-
 456 aques may be supported by these parietal-premotor circuits.
 457 In our analysis of ape gesture learning, the MNS class of

458 models provides sufficient machinery for recognizing
 459 affordance-driven actions like reach-to-grasp – crucial, as
 460 we will see, for the learning that must occur in the initial
 461 stages of ritualization. However, again because of the mod-
 462 el’s emphasis on the relation of the hand to an explicit
 463 physical target, the MNS models (like the FARS model for
 464 action generation) is not flexible enough to account for the
 465 movements associated with known ape gestures. Thus, the
 466 MNS model for the macaque must be extended by addition-
 467 al visual-processing machinery to recognize movements not
 468 explicitly directed towards objects, and likewise for FARS
 469 (see Fig. 2). And while monkey reaching and grasping
 470 behavior has been long studied (Georgopoulos et al. 1981;
 471 Jeannerod and Biguer 1982; Rizzolatti et al. 1987; Taira et

472 al. 1990), non-human primate *gestural* control has not
 473 (Buxbaum et al. 2005; Petreska et al. 2007; Rothi et al.
 474 1991). This gap can be partially bridged by generating testable
 475 hypotheses derived from computational models (e.g., how
 476 do apes maintain approximate visual form when no explicit
 477 targets are available?).

478 **Decision-Making**

479 In order for an animal to adjust its actions to respond to the
 480 immediate environment, it must be able to evaluate contex-
 481 tual and motivational information and select an appropriate
 482 action from its repertoire on the basis of that information.
 483 For an animal to adapt its actions to environmental variables
 484 over time, the neural system must also be sensitive to the
 485 outcomes of its past actions. *Reinforcement learning* has
 486 been a successful framework for describing this type of
 487 adaptation – particularly when considering the decisions
 488 and actions of goal-directed, reward-driven agents (Sutton
 489 and Barto 1998). Crucially, estimates of the ‘value’ in
 490 performing particular actions in particular contexts are
 491 learnable, even when no explicit positive or negative rein-
 492 forcement is received until some time in the future – after
 493 the completion of further actions. These estimates of the
 494 value of particular actions predict how an agent will learn
 495 and act when it encounters similar circumstances in the
 496 future.

497 The Augmented Competitive Queuing model (Bonaiuto
 498 and Arbib 2010) places reinforcement learning mechanisms
 499 alongside MNS mechanisms, in the context of making deci-
 500 sions in the face of changing environments, goals and,
 501 crucially, skills. This allows actions to be evaluated in a
 502 particular context for *executability* – the availability of
 503 affordances that allow the given action – as well as *desir-*
 504 *ability* – the expected (future) reinforcement following that
 505 action (e.g., motivational components for decisions). This
 506 model predicts that actions will be chosen opportunistically:
 507 the next action chosen will be that which is most *desirable*
 508 among the set of *executable* actions. This separation of
 509 decision variables into cognitive and motivational compo-
 510 nents – and their ultimate integration – is supported by the
 511 neurophysiology of decision-making (Watanabe 2007;
 512 Watanabe and Sakagami 2007).

513 In the ACQ model, visual feedback analysis (mirror
 514 neuron system activity) of one’s own actions determine
 515 whether the action achieved its goal, and whether its execu-
 516 tion resembled some other action (the *apparent* action). On
 517 this basis, reinforcement learning can update the *executabil-*
 518 *ity* of the intended action and the *desirability* of the self-
 519 observed action (whether intended or apparent). In this way,
 520 an agent uses an evaluation of current context based on
 521 traces of past experiences to estimate the effectiveness of
 522 different possible actions. This, in fact, fits the observed

patterns of great ape gesturing, in which apes choose ges- 523
 tures based on their goals, the immediate social context, and 524
 their past interactions with their partner (Cartmill 2008; 525
 Hobaiter and Byrne 2011; Liebal et al. 2004a). 526

Given that computational models of neural circuitry for 527
 visually-guided grasping, action-recognition, and adaptive 528
 decision-making exist, our model of ape gesture acquisition 529
 need not be constructed de novo. The FARS, MNS, and 530
 ACQ models, along with insights drawn from other models 531
 – robotic (Chaminade et al. 2008; de Rengerv et al. 2010) – 532
 and neural (Bullock and Grossberg 1988; Caligiore et al. 533
 2010; Demiris and Hayes 2002) provide a circuitry frame- 534
 work upon which neuro-computational models of gesture 535
 acquisition may be based. Moving from simpler single- 536
 agent models into more complex, social brain modeling 537
 may necessitate a re-evaluation of previous models, and 538
 require ‘extensions’ to these models to more closely capture 539
 brain function. It is here that neuroinformatic tools could 540
 provide crucial insights into (i) model benchmarking – what 541
 can a model do or explain and what *can’t* it do? – (ii) 542
 representation of data at the appropriate ‘level’ – does it 543
 explain dynamic time-courses, or sequences of discrete de- 544
 cision events? – and (iii) comparing predictions derived 545
 from model simulations with empirical results from behav- 546
 ioral studies or neurophysiology. 547

548 **Case Study: Developing the Gesture ‘beckon’ via**
 549 **Ontogenetic Ritualization**

We now consider an analysis of the progressive changes in 550
 brain and behavior that would need to occur during the 551
 proposed process of OR. We do this by constructing a 552
 hypothetical sequence of interactions between a mother 553
 and child ape that could lead to the emergence of ‘beckon- 554
 ing’ as a gesture used by the child to get the mother to 555
 approach. This gesture, or variants of arm-extended ‘ap- 556
 proach’ gestures, has been observed in several ape species 557
 (Cartmill 2008; Pika and Liebal 2012; Pika et al. 2003; 558
 Tomasello and Call 1997), though it is not clear how (or 559
 whether) it is acquired. Our aim is not to claim that this 560
 specific gesture is learned in this way, but to use it to help 561
 clarify both the types of interactions and the neural changes 562
 that would be necessary to support the general transition 563
 from action to gesture via OR. Our model is *conceptual*, not 564
 a fully implemented *computational* model (though the latter 565
 is an ongoing research goal). The conceptual model serves 566
 to make general points about gestural acquisition through 567
 OR, and offer a framework to analyze existing behavioral 568
 data from a neuro-computational perspective. 569

Where Call and Tomasello (2007) describe ontogenetic 570
 ritualization with the above 3-step formula, we offer a finer- 571
 grained analysis using 6 stages to highlight the distinct 572

573 learning processes that we expect to occur in the Mother (M)
 574 and Child (C) as the child’s pulling action is ritualized into a
 575 beckoning gesture. We then walk through the neural
 576 changes that seem necessary to support the behavioral
 577 changes at each stage and discuss the challenges in model-
 578 ing the changes in mother and child at each stage.

579 **Proposed Behavioral Changes in Mother (M) and Child**
 580 **(C) During OR of a Beckoning Gesture**

- 581 Stage 1) C reaches out, grabs, and tugs on M, causing M
 582 to move towards C as a response.
- 583 Stage 2) C reaches out, grabs, and begins to tug on M, and
 584 M quickly moves towards C.
- 585 Stage 3) C reaches out and makes contact with M, and M
 586 quickly moves towards C.
- 587 Stage 4) C reaches out towards M, attempting to make
 588 contact, but M responds before contact is made.
- 589 Stage 5) C reaches part way towards M, and M responds
 590 by moving towards C.
- 591 Stage 6) C gestures towards M and M responds to this
 592 ritualized gesture by moving towards C.

593 It is our belief that such a finer-grained analysis, when
 594 pegged to behavioral and neural changes in each agent,
 595 presents a more appropriate framework with which to com-
 596 pare or benchmark a computational model, while still
 597 remaining consistent with Call & Tomasello’s description
 598 of the overall pattern of interaction.

- 599 Stage 1) Child reaches out, grabs, and tugs on Mother,
 600 causing Mother to move towards Child as a
 601 response.

602 Since our example is meant to illustrate the salient steps
 603 in all plausible cases of ontogenetic ritualization, the key for
 604 stage 1 is that the actor is able to achieve his desired goal
 605 directly through physical manipulation of the recipient. For
 606 this initial period of interaction, neither participant has any
 607 prior expectations of the others’ behavior.

608 *Child* In order for the child to achieve his goal, his only
 609 option is to physically manipulate the mother to bring her
 610 into physical contact with himself (that is, we assume no
 611 latent gestural form for this goal). He orients towards the
 612 mother, identifies appropriate surfaces for grasping to pull.
 613 He reaches out, grasps a part of her body, and initiates pulling
 614 on the mother. The pulling force begins the movement of the
 615 mother towards the child, and after enough tension, the mother
 616 complies and moves closer to the child. The neural machinery
 617 required to coordinate this sequence of actions can be fully
 618 described by the FARS model of visual control of grasping
 619 discussed above, which will serve as a benchmark for the
 620 child’s behavior in subsequent stages.

Mother Throughout this example, we will assume that the
 621 mother is motivated to complete her child’s request (not
 622 always the case!) and that her attention is appropriately
 623 oriented towards the child, allowing her to visually as well
 624 as haptically process the child’s actions. Assuming that the
 625 mother is attending to the child, her mirror neuron/action-
 626 recognition system would register the ‘reach-to-grasp’ fol-
 627 lowed by ‘pull’. A key property of mirror neuron firing,
 628 captured by the MNS models, is that mirror activity often
 629 signals recognition of the observed actions before the action
 630 is completed. However, the mother’s response to, as distinct
 631 from recognition of, the child’s ‘approach’ has not yet been
 632 established. Associative learning mechanisms establish this
 633 connection between the child’s action and the response
 634 ‘approach.’ Importantly, this association must be retained
 635 as the child’s action changes form over the OR process. We
 636 also note the need for ‘social’ motivation to fulfill the goals
 637 of the child or to prioritize physical proximity – a motivation
 638 shared by both agents.
 639

- 640 Stage 2) Child reaches out, grabs, and begins to tug on
 641 Mother, and Mother quickly moves towards
 642 Child.

643 In this step, both individuals experience an adaptive
 644 change in behavior in real-time and begin to alter their
 645 expectations of the other’s actions. For the child, (i) he need
 646 not pull as hard once the mother begins to comply, suggest-
 647 ing feedback modulation of his on-going action, and (ii) he
 648 forms the expectation that the mother will be increasingly
 649 compliant. For the mother, she learns that given contextual
 650 considerations – similar play conditions, perhaps, and/or
 651 perceived emotional state – and her child’s grasp-pull ac-
 652 tion, she is rewarded (socially) for moving herself to his
 653 side.

654 *Child* The child’s intention is to reach out, grab and phys-
 655 ically move the mother near him. However, following his
 656 grasp, his mother becomes more compliant and begins the
 657 movement towards his side. The child perceives the moth-
 658 er’s movement as beginning to satisfy his goal and acts less
 659 forcefully on the mother as she responds to his action. This
 660 further refines his expectations of his mother’s likely re-
 661 sponse. In future interactions, he will expect that less force
 662 is needed to achieve his goal.

663 At this stage, we encounter the problem of how recogni-
 664 tion of someone else’s actions can affect the ongoing exe-
 665 cution of one’s own actions – a general concern for social
 666 brain modeling. Here, the child, as in step 1, expects a full
 667 ‘reach-to-grasp-to-pull’ action is necessary to achieve his
 668 goal, but as he begins his tug on the mother, the mother
 669 responds by ‘completing’ the action. Recognition of the
 670 early success of the action must be able to modulate the
 671 child’s ongoing behavior in such a way that his action can be

672 modified either (i) by reducing the force he pulls on the
 673 mother, as in this step, or (ii) by interrupting and even
 674 extinguishing the action mid-trajectory (as we describe be-
 675 low). Such sensitivity to changing perceptual variables dur-
 676 ing grasping behavior has been explored in a computational
 677 model of how the reach, grasp, and their coordination may
 678 be affected by perturbations in the size and location of
 679 grasped objects (Hoff and Arbib 1993) which use on-line
 680 feedback to modulate what might otherwise have been a
 681 feed-forward movement. Thus, insights from other models
 682 may guide our own model development.

683 *Mother* Following contact, the mother moves towards the
 684 child, easing the tension on her arm. The association be-
 685 tween the child's action and the approach response is rather
 686 weak at this point, and can only be triggered by propriocep-
 687 tive contact and mechanical tension as a complement to the
 688 visual-form representation established by the mirror neuron
 689 system. As in Stage 1, the MNS models of monkey action
 690 recognition provide an explanation for the mother's recog-
 691 nition of the child's actions, but are unable to provide a clear
 692 description of the *effects* of this recognition – a problem we
 693 explore below. Future models of action recognition must
 694 address the role of multisensory integration in the recogni-
 695 tion process more thoroughly. The MNS2 model (Bonaiuto
 696 et al. 2007) characterized the audio-visual neurons seen
 697 experimentally in (Kohler et al. 2002), and showed how
 698 associative learning mechanisms may link acoustic cues
 699 with the visual form of actions. A key for a model of OR
 700 would be extending this to visual-*haptic* cues (see Fig. 2).

701 Stage 3) Child reaches out and makes contact with Mother,
 702 and Mother quickly moves towards Child.

703 *Child* The child's attempt at grasping and pulling the moth-
 704 er remains the same as stages 1 and 2, with the exception
 705 that he becomes increasingly sensitive to the mother's an-
 706 ticipatory response, having in the past two stages come to
 707 expect a 'completing' response. In stage 3, as he begins to
 708 make contact with the mother, the mother's response
 709 appears consistent to his expectation, and he aborts the
 710 second half of the action sequence: the pull on his mother.
 711 As we see in Fig. 3, however, such a process may be
 712 described at different levels of representation – discrete
 713 and continuous, or 'event' and 'trajectory'. Models of reach-
 714 ing and grasping (e.g., Bullock and Grossberg 1988; Fagg
 715 and Arbib 1998) emphasize the dynamic unfolding of the
 716 behavior and how certain elements (the positions of joints,
 717 perhaps) vary *continuously* in time. Models of learning and
 718 decision-making (e.g., Bonaiuto and Arbib 2010; Botvinick
 719 et al. 2009) on the other hand, emphasize the serial structure
 720 of decisions as *discrete* events. Both levels may be helpful
 721 in understanding brain function, and in fact the brain

appears to utilize both (see: Averbeck et al. 2002; Campos 722
 et al. 2010; Georgopoulos 2002; Sawamura et al. 2002). The 723
 challenge for neuroscientists is to understand how both may 724
 coordinate behavior, and how best to represent these 725
 descriptions in models. 726

Mother Visual recognition of the child's reach-to-grasp 727
 action, coupled with contextual cues and the proprioceptive 728
 contact as above, is sufficient for the mother to select an 729
 appropriate response consistent with the child's goals. This 730
 stage represents the terminal phase of proprioceptive cues 731
 involved in *training* the recognition-response linkage – in 732
 future stages visual recognition alone suffices to initiate the 733
 response. 734

Stage 4) Child reaches out towards Mother, attempting to 735
 make contact, but Mother quickly responds be- 736
 fore contact is made. 737

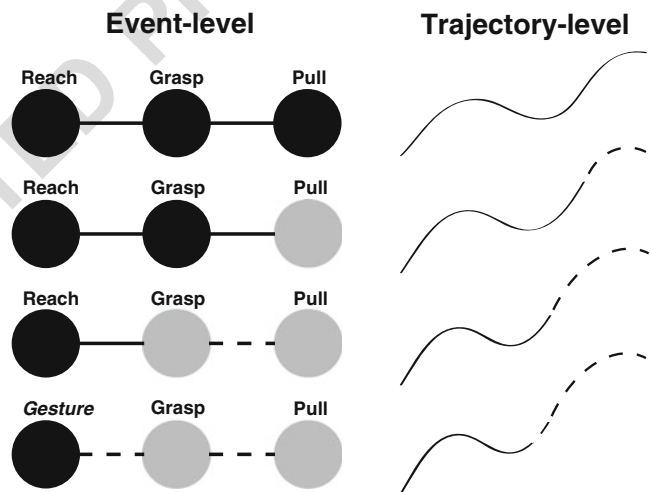


Fig. 3 Event- and trajectory-level representations in brain modeling. (Left) Event-level representations, treating actions and decisions as discrete units, emphasize higher-order representations and the sequential unfolding of distinct actions in series. Neurophysiological recordings show that the brain can maintain such state-by-state representations of sequences (Campos et al. 2010; Sawamura et al. 2002). (Right) Trajectory-level representations treat actions, both single actions and action sequences, as dynamic and emergent trajectories in 'action spaces', sensitive to idiosyncratic context and performance and the on-line modulation from feedback centers. Such a perspective is supported by behavioral and neurophysiological data (Jeannerod et al. 1995). From top to bottom, both levels of representation show the putative 'truncation' of an instrumental action, to that of a ritualized gesture. Dashed lines on the left indicate possible next-states in the action sequences (e.g., priming activation), while the shading indicates the sensitivity to feedback (e.g., 'grasp' may simply become a 'touch' if recipient responds quickly; see, for example, Stage 3 in the text). Dashed lines on the right similarly indicate possible next-states contingent on the performer's evaluation of the goal state (e.g., whether the recipient has responded appropriately). In both representations, then, we see that the original effective action/action sequence is not lost, and may be substituted for the gesture when more appropriate – for example, recipient not visually attentive (Liebal et al. 2004b)

738 Here, we have reached the point where both (i) the child
 739 has learned that a ‘reach-to-grasp-to-pull’ action is not nec-
 740 essary (though perhaps unsure about the extent to which he
 741 must contact and attempt to manipulate the mother), and (ii)
 742 the mother has learned that (attempted) grasps to her arm
 743 may signal an opportunity for social bonding. Note that
 744 whereas the child’s learning consists largely in tuning his
 745 forward expectations of the mother’s behavior, the mother’s
 746 learning consists in mapping the recognition of the child’s
 747 actions to behavioral responses that satisfy mutual goals.

748 *Child* The child at this point has learned that incomplete
 749 versions of a ‘reach-to-grasp-to-pull’ action can be used to
 750 achieve his goal, and so only intends to make minimal
 751 contact. Here, the child’s attempted action should still be
 752 seen as *transitive*, directed at a surface. This will be the
 753 starting point for the last two stages, in which the actions
 754 become increasingly removed from orientation towards a
 755 specific surface, and instead the hand’s movement pattern
 756 in space becomes the most salient element, resulting even-
 757 tually in a ritualized, *intransitive* gestural form.

758 *Mother* By stage 4, the mother has robustly linked visual
 759 recognition of the child’s ‘reach-towards-body’ action with
 760 the ‘move-towards-child’ response, and can effectively ful-
 761 fill the child’s goal without even minimal haptic cues. The
 762 key to this stage of the ritualization process is that visual
 763 form alone is now sufficient for the mother to respond.
 764 Subsequent stages serve to train the mirror neuron/action-
 765 recognition system to recognize the now visual-only ‘proto-
 766 beckoning’ act in shorter and more reduced forms.

767 Stage 5) Child reaches part way towards Mother, and
 768 Mother quickly responds by moving towards
 769 Child.

770 *Child* This stage is unique in that now the child no longer
 771 intends to physically interact with the mother, but instead
 772 acts only in a way sufficient to elicit the appropriate re-
 773 sponse. Crucial for the child, though, is that his intransitive
 774 performance must appear visually similar enough to his
 775 transitive performance that the mother can recognize it and
 776 respond appropriately. This would suggest that propriocep-
 777 tive and visual feedback signals from the past transitive
 778 episodes trained a forward internal model – a model of an
 779 appropriate trajectory through space – that now can suffi-
 780 ciently control the limb in an apparent ‘proto-beckoning’
 781 action. Whereas previous models of object-directed motor
 782 control (e.g., FARS) neglected intransitive performance, we
 783 see that models of gesture acquisition may not. Instead, as
 784 Fig. 2 shows, additional machinery must be added to recog-
 785 nition and production processes, informed by analyzing
 786 behavioral and/or functional data.
 787

Mother The mother is able to recognize, in mid-trajectory, 788
 the motion of the child’s arm, and to respond appropriately. 789
 In this stage, the mother’s action-recognition system would 790
 need to begin to respond to smaller and smaller portions of 791
 the action. Just as the child must maintain visually similar 792
 performance in the absence of explicit targets (e.g., mother’s 793
 arm), the mother must be able to recognize the child’s 794
 actions absent such contextual cues. Without sufficient 795
 ‘overlap’ in trajectory, the putative visual training would 796
 be unable to maintain the link between recognition and 797
 response in the mother. 798

Stage 6) Child gestures towards Mother, and Mother 799
 responds to this ritualized gesture by moving 800
 towards Child. 801

A ritualized form of the gesture emerges. While previous 802
 stages have highlighted the changes that would allow both 803
 mother and child to progress towards using and recognizing 804
 an intransitive gestural signal, stage 6 represents a more 805
 stable communicative form that the child can continue to 806
 use in future interactions. 807

Child The child intends the mother to come to his side, and 808
 performs the ritualized gesture. As a result of continued 809
 interaction with his mother, he may learn that the gesture 810
 is more or less effective over various distances or in differ- 811
 ent contexts. The child maintains an association between the 812
 gesture and the original action while still able to substitute 813
 the action for the gesture if the gesture is not effective at 814
 achieving the child’s goal. 815

Mother The mother sees and recognizes his gesture, and 816
 responds in a manner that fulfills his goal. Again, we note 817
 that the mother would still respond accordingly to the orig- 818
 inal effective action. 819

Summary of the Ontogenetic Ritualization of ‘beckon’ 820

The ‘beckoning’ gesture that would result from this 6-stage 821
 process can be seen as a truncated (and modified) version of 822
 the instrumental act of reach-to-grasp-to-pull. While this 823
 case study illustrates general behavioral changes that would 824
 occur in the ‘naïve’ ontogenetic ritualization process as it 825
 has been discussed in the literature, we do not argue that this 826
 particular gesture is necessarily derived in this way. In fact, 827
 we note that any model of the acquisition of a gestural form 828
 is likely to have idiosyncratic features. What we have de- 829
 scribed may be seen as a generic analysis of ritualizing 830
 effective actions, and not wholly specific to ‘beckoning’. 831
 For example, ‘beckoning’ as described in the literature 832
 involves a sweeping of the hand or curling of the fingers 833
 towards the body. This requires additional machinery be- 834
 yond what we have described here – perhaps to anticipate 835

836 the recipient's movement towards oneself, or perhaps merely
837 to distinguish the form from other similar gestures (e.g.,
838 'palms-up give') – just as a 'nudge' gesture, or 'arms-up
839 play' gesture require unique mechanisms specific to each
840 gesture. Also, this transition from action to gesture may take
841 several forms – including phenotypic changes – neverthe-
842 less elements of the stages we propose are likely to be
843 central to any ritualization processes that yields gestural
844 forms.

845 Discussion

846 Computational models have been important in understand-
847 ing neural control of behavior, from object and action recog-
848 nition (Bonaiuto et al. 2007; Deco and Rolls 2004; Oztop
849 and Arbib 2002), to saccadic eye control (Dominey et al.
850 1995; Silver et al. 2012), and visual control of grasping
851 (Fagg and Arbib 1998). Such models have even made useful
852 contributions to our understanding of higher order cognitive
853 skills (O'Reilly and Frank 2006; Rougier et al. 2005). How-
854 ever, our proposed analysis of ontogenetic ritualization
855 presents several unique challenges to the brain modeling
856 community.

857 Challenges for Dyadic Brain Modeling

858 The first challenge for social brain models is simply being
859 *social*. Few brain models to date have incorporated explic-
860 itly social tasks that are central to primate behavior and
861 cognition. The Mental State Inference model (MSI) is per-
862 haps the first explicit instance of multiple brains in simula-
863 tion, and simulates the manual performance of one
864 individual being observed by another. The MSI model is
865 based in large part on the MNS models of action recognition
866 and suggests mirror responses are a single part of an ex-
867 tended 'internal model' that serves to decode the intentions
868 of others (Oztop et al. 2005). However, there is no explicit
869 interaction between the agents in the model, and the obser-
870 vations do not affect the subsequent behaviors of the ob-
871 server – that is, there is no 'task'. We are not able to predict
872 how observation would affect future performance, and we
873 are still lacking any consideration of interaction between the
874 brains. Thus, the MSI model is a 'passively social' model in
875 the same vein as the MNS and MNS2 models (and most of
876 the work on mirror neuron neurophysiology).

877 A few 'interactive' dyadic models have been put forth,
878 but they often lack the neural specificity found in models of
879 passively social or purely instrumental tasks. Taking an
880 interactive approach, Steels and colleagues have modeled
881 multiple interacting agents in 'language learning' games,
882 showing interesting results for ideas of grammar learning
883 and cultural transmission (Steels 2003). However, Steels'

multi-agent simulations lack 'brains', and instead describe
agents with simplified mechanisms that are highly task-
specific. The dynamics of the interactions are enlightening,
but say little about the specific brain processes involved in
interactions between primates. Other models simulate neural
dynamics between interacting agents, but do not engage
questions of the computational properties of detailed brain
circuitry during a specific task (Dumas et al. 2012).

Our analysis of ontogenetic ritualization diverges from pre-
vious work incorporating purely *observational* social dimen-
sions in that it is an explicitly *interactive* process of social
learning. Social learning encompasses all learning that is mod-
ulated by the actions of another individual (Galef and Laland
2005), though different kinds of learning may be distinguished,
for example, in the degree to which available environmental
information may be processed to influence future behavior
(Acerbi et al. 2011). Dyadic learning, like ontogenetic rituali-
zation, might be considered to be an 'interactive' form of social
learning in that the learner must interact with (rather than just
observe) another individual for learning to proceed. (Indeed,
both agents are 'learners' in ontogenetic ritualization.) In these
cases, brain models must show how distributed patterns of
neural activation in each individual affect their behavior and
how socially-influenced learning processes in the brain given
rise to adaptive changes in behavior.

A further challenge for social brain modeling concerns the
fact that the motivations underlying social behaviors may not
be the homeostatic motivations (hunger, thirst) that are often
used in simulations of behavior. During social interactions,
animals may perform the same task, but with differing moti-
vations – for example, preferring social information over food
rewards (Deaner et al. 2005; Klein et al. 2008). Brain model-
ing lacks serious consideration of these differing motivational
drives during behavior (especially social behavior), and rarely
incorporates reward-modulated or motivational responses in
neural network models (Arbib and Bonaiuto 2012; Guazzelli
et al. 1998). This failure to incorporate motivational elements
becomes more consequential when modeling social tasks,
which require both navigating the multi-dimensionality that
is 'motivation' (Berridge 2004) and incorporating the percep-
tion of motivations and intentions in partners, thought to be a
large driver of cognitive skill evolution in primates (Byrne and
Whiten 1988).

Another challenge is to address the debate over the extent to
which neural networks and cognitive modules can be said to be
'domain general' or 'domain specific', and how these systems
would interact, especially with respect to social cognition. As a
simple example, face-selective neurons have been described in
specific regions of temporal cortex (Barraclough and Perrett
2011). Using more complex interactive designs, (Yoshida et al.
2011) describe medial frontal neurons that appear to respond
exclusively to social variables (at least for the task studied), and
may suggest certain functional specialization within especially

937 frontal and prefrontal regions. Anterior Cingulate Cortex
 938 gyrus (ACCg) and Anterior Cingulate Cortex sulcus
 939 (ACCs) dissociate, in a lesion study, in their recruitment
 940 in response to social variables, with ACCg important for
 941 social valuation (Rudebeck et al. 2006) – though see
 942 (Chang et al. 2012) for a more complex result. Lastly,
 943 the LIP mirror neuron responses in monkeys suggest that
 944 ‘integration’ regions like PPC can represent both social and
 945 non-social information important for decision-making (Shep-
 946 herd et al. 2009). Together, these and other data must be
 947 analyzed to identify whether and which structures can be said
 948 to be ‘social-domain specific’, and how such regions would
 949 interact with wider neural systems (Fig. 4).

950 **Linking Neuroinformatics to Gestural and Behavioral**
 951 **Datasets**

952 The dyadic brain approach to social behavior challenges the
 953 neuroinformatics community to provide resources to inte-
 954 grate data from neurophysiological and behavioral studies
 955 of primates in a way that could provide new insights into the
 956 cognitive and structural changes underlying the evolution of
 957 primate (and human) communication. We review discipline-
 958 wide concerns for managing data, including:

- 959 • Primate behavioral data
- 960 • Primate brain imaging data

- Macaque neurophysiological data 961
- Comparative neuroanatomical data 962
- Model simulation results 963

Behavioral Data Management 964
965

966 Researchers in biological and biomedical sciences have
 967 made significant advances in constructing searchable
 968 databases and have tackled the challenges of standardiz-
 969 ing and archiving data in a range of fields, including
 970 within neuroscience (see companion articles). Though
 971 these approaches could not be transferred verbatim to
 972 data in comparative cognition, the challenges inherent
 973 to linking studies and identifying patterns in data are
 974 common to all integrative databases and should be used
 975 to inform future efforts to consolidate data across studies
 976 of primate cognition.

977 Tomasello and Call (2011) raised concerns about the
 978 isolation of individual studies in primate cognition, particu-
 979 larly in relation to gesture studies in the great apes. Differ-
 980 ences as to what qualifies as a ‘gesture’ and how gestures
 981 are coded and defined lead to significant differences be-
 982 tween the conclusions of different studies of the same spe-
 983 cies, and can reflect local traditions – in research groups
 984 (Cartmill and Byrne 2011). The potential of drawing erro-
 985 neous conclusions based on single studies or only studies
 986 from a single research group underscores the importance of
 987 establishing a consistent ontology of primate social behavior

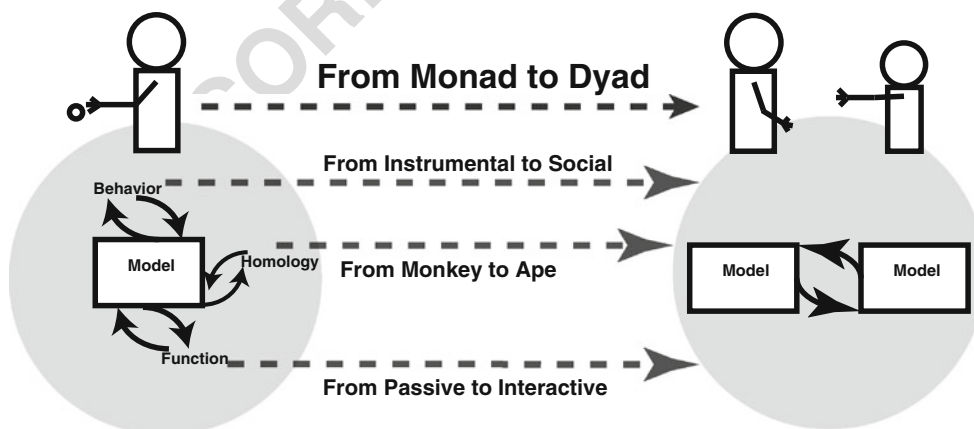


Fig. 4 Interactions between experimental and theoretical disciplines. Modeling can be a source of anchoring insights across experimental conditions, including anatomical, physiological and behavioral. A robust, computational model of brain systems that (i) are anatomically-based, (ii) compute with biologically-plausible models of neurons or populations of neurons, and (iii) generate patterns of overt behavior can both formalize findings in a unified framework, and support hypothesis generation to inspire new experiments or techniques of analysis. This back-and-forth between the experimental and theoretical disciplines – facilitated by informatics tools (*shading*) – is and has been highly profitable. (*Left*) Models in the past have been successful at engaging experimental findings, especially those that have relied on

instrumental behavior in monkeys – and in utilizing insights from monkey single-unit recordings during instrumental or ‘passively social’ task conditions. Informatics tools and resources assist in developing, testing and benchmarking computational models. (*Right*) We propose to similarly engage this back-and-forth between models and experiments, while moving each ‘cycle’ into an arena of novel questions as the arrows from left to right indicate. In particular, we seek to move from models of single agents engaged in instrumental tasks, to models of ‘dyads’ that interact directly with each other. Informatics tools, while providing important resources for this research venture, must be expanded to handle the new challenges that will result from this novel modeling approach

988 and developing resources for managing, integrating and
989 sharing behavioral data.

990 The level of comparison or ‘granularity’ of the searchable
991 data between studies is particularly important. Definitions of
992 behaviors or functional responses differ between studies and
993 these differences often make it difficult to directly compare
994 the results of studies without explicitly accounting for differ-
995 ences in methodologies. For example, one paper that sur-
996 veyed multiple groups of gorillas found a total repertoire of
997 33 gestures, most of which were shared between zoos (Pika
998 et al. 2003). Another paper reported 102 gestures, also from
999 a survey of multiple zoos (Genty et al. 2009). The differ-
1000 ences between the repertoire sizes reported in these two
1001 studies are not the result of group-specific gestures or cul-
1002 tural variation between sites. Rather, they result from gran-
1003 ularity differences in the researchers’ gesture definitions.
1004 The first study defines gestures by the predominant move-
1005 ment involved, but does not typically draw distinctions
1006 between gestures based on the limb or hand shape used.
1007 Thus all examples of hitting a surface with a hand would be
1008 considered a “slap ground” gesture. In the second study,
1009 however, the limb and hand shape used are considered part
1010 of the gesture definition so “knock object,” “punch object,”
1011 “slap object 1-handed,” and “slap object 2-handed” are all
1012 recorded as separate gestures. These differences in the gran-
1013 ularity of definition could result in the same set of observations
1014 yielding drastically different summary results (a potential
1015 problem not just for behavioral, but also neural datasets).

1016 On the one hand, it is necessary to accurately record
1017 those methodological differences that make direct compari-
1018 son of results between studies difficult at the moment. On
1019 the other hand, it is not practical to fully recode primary
1020 source data from different studies according to the same
1021 guidelines so that it can be easily pooled. Allowing individ-
1022 ual variation in the coding systems not only removes a
1023 substantial barrier to contributing data to a collective data-
1024 base, it also allows coding systems to be appropriately
1025 tailored to the differences between species’ communication
1026 systems. For example, one frequently coded behavior is
1027 whether an ape waits for a response from the recipient
1028 before giving up or attempting another gesture. This mea-
1029 sure of ‘response waiting’ is used as an indication of inten-
1030 tional communication since it signals that the gesturer
1031 expects a particular response from the recipient. Since pri-
1032 mate species differ in temperament and energy levels, the
1033 length of time that suggests waiting for a response is likely
1034 to differ. The amount of time thought to indicate response
1035 waiting in a low-energy species might be far too long for a
1036 high-energy species with a shorter attention span. In this
1037 case, it would be better to ignore the differences in definition
1038 of response waiting between studies since those differences
1039 account for temperament variation and allow the same be-
1040 havior to be measured across species. Though variation of

1041 definitions within a species is likely to cause problems (as in
1042 the gorilla gesture example), allowing definitions to vary by
1043 species facilitates direct comparison between studies by
1044 bringing the cognitive ability rather than the temperamental
1045 differences to the forefront.

1046 Longitudinal data are especially valuable because they
1047 allow us to ask direct questions about the development of
1048 gestures over time, but longitudinal studies in apes are rare
1049 and time intensive. Indirect questions about gesture devel-
1050 opment may be asked by comparing individuals of different
1051 age classes between different sites to identify developmental
1052 trajectories in gesture use. Incorporating longitudinal data of
1053 gesture in the same individuals into a cross-study/cross-site
1054 database would be invaluable to the field because the inte-
1055 gration of cross-sectional and longitudinal data would allow
1056 researchers to ask more sophisticated questions about devel-
1057 opment within a species, and facilitate comparative studies.

Neural Data Management 1058

1059 Resources for managing functional and neuroanatomical
1060 data provide a strong backbone for research in social brain
1061 modeling. BrainMap <<http://www.brainmap.org/>> and
1062 Brede <<http://neuro.imm.dtu.dk/services/jerne/brede/>> are
1063 tools for managing and performing meta-analyses for function-
1064 al neuroimaging data (and see companion articles in this vol-
1065 ume), and resources like BrainLiner <<http://brainliner.jp/>>
1066 offer a platform for managing and standardizing neurophysi-
1067 ological data. As non-invasive functional brain monitoring
1068 in apes becomes more available, specific resources may
1069 have to be developed tailored to the needs of researchers.
1070 For neuroanatomical data, the NeuroHomology DataBase
1071 (Bota and Arbib 2001), for example, was developed to offer
1072 researchers the tools to investigate the relationship between
1073 similar brain structures in different species. However, newer
1074 techniques like diffusion tensor imaging (DTI) now allow
1075 researchers to ask novel questions in a non-invasive design.
1076 Recent comparative DTI analyses, for instance, suggest
1077 significant differences in fiber pathways linking regions
1078 in parietal, temporal and frontal areas between modern
1079 primates – specifically between macaques, chimpanzees
1080 and humans (Hecht et al. 2012). The results suggest an
1081 increase in connectivity between STS and inferior parie-
1082 tal regions – moving from macaques, then to chimpan-
1083 zee, and finally to humans with robust connectivity –
1084 together processing the visual form of movements. These
1085 and other neuroanatomical studies may support, for ex-
1086 ample, model hypotheses regarding connectivity between
1087 kinematic-processing structures and action recognition
1088 and other structures (see Fig. 2). As these data become
1089 more prevalent, efficient ways to handle and link these
1090 data with functional and neuro-homology databases
1091 becomes more important.

1092 Model Result Management

1093 Software designed for computational neuroscience are
 1094 widely available (e.g., NEURON; [http://www.yale.edu/](http://www.yale.edu/neuron)
 1095 [neuron](http://www.yale.edu/neuron)) and code repositories like ModelDB ([http://](http://senselab.med.yale.edu/modeldb)
 1096 senselab.med.yale.edu/modeldb) offer researchers ways
 1097 to share code. These resources and others can often be linked
 1098 or ‘federated’ to offer access to data from other systems, as the
 1099 Brain Operations DataBase (BODB) does. BODB ([http://](http://bodb.usc.edu/bodb)
 1100 bodb.usc.edu/bodb; and see companion articles) currently
 1101 allows linkages to data sources ranging from neuroanatomical
 1102 datasets of monkey and human, to functional imaging sets like
 1103 those offered by BrainMap. BODB also offers tools for man-
 1104 aging Summaries of Empirical Data (SEDs) with the goal of
 1105 facilitating work in computational neuroscience. The SED
 1106 format is designed to be at the appropriate ‘level’ to offer
 1107 challenges to existing ideas for brain function, and flexible
 1108 enough to be understood both in relation to other data, and in
 1109 relation to specific models of the brain, allowing direct com-
 1110 parison between model simulation results and existing (or
 1111 future) empirical work against which the simulations can be
 1112 benchmarked. However, as the above analyses have shown,
 1113 model benchmarking becomes much harder when the behav-
 1114 ior studied – gesture, for example – has differing operational
 1115 definitions and levels of description.

1116 Ideally, integration should be possible at multiple levels of
 1117 representation. One study may want to ask how manual ges-
 1118 tures are used in different age groups and integrate this infor-
 1119 mation with what is known about primate brain systems
 1120 involved in action recognition. Another study may focus on
 1121 the ability to respond to the gaze of potential recipients by
 1122 using a visual vs. tactile gesture, and may be concerned only
 1123 with the neural representation of gaze awareness and not the
 1124 gesture type. Flexibility for future integration and expansion is
 1125 key. For example, BODB currently offers tools for managing
 1126 behavioral data, and has the possibility of integrating its
 1127 functionality with other, more specific databases. It would be
 1128 possible then to create links between BODB and a future
 1129 database of primate gesture research, thus enabling a platform
 1130 to manage behavioral and neuroscientific data. Still, existing
 1131 resources within neuroscience need more structuring, as the
 1132 examples of non-human primate brain imaging suggest.

1133 Establishing links between collections of neural and behav-
 1134 ioral data and allowing searches to span and connect data in
 1135 different fields would transform our ability to ask questions
 1136 about the evolution of cognition, brain, and behavior. Though
 1137 the creation of integrative databases holds great promise for
 1138 researchers, barriers to participation in a collective database
 1139 must be minimized. The simplicity and power of the built-in
 1140 tools for adding and managing data in a database greatly
 1141 impact the likelihood that people will contribute data and use
 1142 the database to conduct research. The behavioral and cognitive
 1143 sciences are becoming increasingly interdisciplinary and

advances in our understanding are more likely to be made by 1144
 comparing across studies and disciplines than by individuals 1145
 working on isolated datasets using a unidirectional approach. 1146

1147 In this paper we have argued that social brain modeling is a 1148
 promising field with potential to combine and extend the 1149
 insights gained from the neural and behavioral sciences. We 1150
 use gesture, and specifically the proposed learning of gesture 1151
 via ontogenetic ritualization, as a test case for the construction 1152
 of this integrative modeling approach. We focus on gesture 1153
 because it incorporates social features that are problematic for 1154
 modeling (e.g., different processes in the signaler and receiver, 1155
 goal attribution, recognition of social variables like attention, 1156
 and flexible deployment), but also because it allows us to build 1157
 on existing models of the production and perception of manual 1158
 action. As our proposed model of ontogenetic ritualization 1159
 illustrates, integrating ethological data with models grounded 1160
 in neural detail offers the possibility to ask interesting ques- 1161
 tions about social learning and cognition and to make testable 1162
 predictions about behavioral outcomes – and ultimately to help 1163
 unravel questions about development and evolution. However, 1164
 substantial challenges remain. We believe that many of these 1165
 challenges require innovative new informatics approaches, like 1166
 the construction of searchable databases that would allow 1167
 integration of data across studies, fields, and methodologies. 1168
 We call for a concerted interdisciplinary effort between prima- 1169
 tologists, neuroscientists, and computational modelers to con- 1170
 sider new collaborative approaches to the integration and 1171
 maintenance of both raw and summarized data. Even small 1172
 steps into this interdisciplinary terrain promise to transform the 1173
 research landscape from isolated studies to richly collaborative 1174
 conversations, and to open up powerful new approaches to 1175
 very old questions.

Information Sharing Statement 1176

This article was made feasible through various online 1177
 resources, including university-maintained journal access. 1178

Acknowledgments We thank the participants of the Workshop on 1180
 Action, Language, and Neuroinformatics at the University of Southern 1181
 California for their helpful discussion and insights and thank J. G. 1182
 Foster for his comments on the manuscript. Part of this work was 1183
 funded through NSF grant number 0924674, Michael A. Arbib, 1184
 Principal Investigator. 1185
 1186

References 1187

Aboitiz, F. (2012). Gestures, vocalizations and memory in language 1189
 origins. *Frontiers in Evolutionary Neuroscience*, 4. 1190
 Acerbi, A., Tennie, C., & Nunn, C. L. (2011). Modeling imitation and 1191
 emulation in constrained search spaces. *Learning & Behavior*, 39, 1192
 104–114. 1193

1194 Adolphs, R. (2010). Conceptual challenges and directions for social
1195 neuroscience. *Neuron*, *65*, 752–767.

1196 Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial
1197 frontal cortex and social cognition. *Nature Reviews Neuroscience*,
1198 *7*, 268–277.

1199 Arbib, M. A. (2005). From monkey-like action recognition to human
1200 language: an evolutionary framework for neurolinguistics. *The*
1201 *Behavioral and Brain Sciences*, *28*, 105–124. discussion 25–67.

1202 Arbib, M. A. (2008). From grasp to language: embodied concepts and the
1203 challenge of abstraction. *Journal of Physiology, Paris*, *102*, 4–20.

1204 Arbib, M. (2010). Mirror neurons and language. *Annual Review of*
1205 *Anthropology*, *40*, 110301094607049–49.

1206 Arbib, M. A. (2012). *How the brain got language: The mirror system*
1207 *hypothesis*. New York: Oxford University Press.

1208 Arbib, M. A., & Bonaiuto, J. J. (2012). Multiple levels of spatial
1209 organization: world graphs and spatial difference learning. *Adap-*
1210 *tive Behavior*, *20*, 287–303.

1211 Arbib, M. A., Bischoff, A., Fagg, A. H., & Grafton, S. T. (1995).
1212 Synthetic PET: analyzing large-scale properties of neural net-
1213 works. *Human Brain Mapping*, *2*, 225–233.

1214 Arbib, M. A., Billard, A., Iacononi, M., & Oztop, E. (2000). Synthetic
1215 brain imaging: grasping, mirror neurons and imitation. *Neural*
1216 *Networks: the Official Journal of the International Neural Net-*
1217 *work Society*, *13*, 975–997.

1218 Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization,
1219 gesture, and the evolution of human language. *Current Anthro-*
1220 *pology*, *49*, 1053–1076.

1221 Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P.
1222 (2002). Parallel processing of serial movements in prefrontal
1223 cortex. *Proceedings of the National Academy of Sciences of the*
1224 *United States of America*, *99*, 13172–13177.

Q2 1225 Azzi, J.C.B., Sirigu, A., & Duhamel, J-r. (2011). Modulation of value
1226 representation by social context in the primate orbitofrontal cortex.

1227 Barraclough, N. E., & Perrett, D. I. (2011). From single cells to social
1228 perception. *Philosophical Transactions of the Royal Society of*
1229 *London. Series B, Biological Sciences*, *366*, 1739–1752.

1230 Barrès, V., Simons, A., & Arbib, M. (2013). Synthetic event-related
1231 potentials: a computational bridge between neurolinguistic mod-
1232 els and experiments. *Neural Networks: the Official Journal of the*
1233 *International Neural Network Society*, *37*, 66–92.

1234 Berridge, K. C. (2004). Motivation concepts in behavioral neurosci-
1235 ence. *Physiology & Behavior*, *81*, 179–209.

1236 Bonaiuto, J., & Arbib, M. A. (2010). Extending the mirror neuron
1237 system model, II: what did I just do? A new role for mirror
1238 neurons. *Biological Cybernetics*, *102*, 341–359.

1239 Bonaiuto, J., Rosta, E., & Arbib, M. (2007). Extending the mirror
1240 neuron system model, I. Audible actions and invisible grasps.
1241 *Biological Cybernetics*, *96*, 9–38.

1242 Bota, M., & Arbib, M. A. (2001). The NeuroHomology database.
1243 *Neurocomputing*, *38–40*, 1627–1631.

1244 Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically
1245 organized behavior and its neural foundations: a reinforcement
1246 learning perspective. *Cognition*, *113*, 262–280.

1247 Bruner, J., Roy, C., & Ratner, N. (1982). The beginnings of request.
1248 *Children's Language*, *3*, 91–138.

1249 Bullock, D., & Grossberg, S. (1988). Neural dynamics of planned arm
1250 movements: emergent invariants and speed-accuracy properties
1251 during trajectory formation. *Psychological Review*, *95*, 49–90.

1252 Buxbaum, L. J., Kyle, K. M., & Menon, R. (2005). On beyond mirror
1253 neurons: internal representations subserving imitation and recog-
1254 nition of skilled object-related actions in humans. *Brain Research.*
1255 *Cognitive Brain Research*, *25*, 226–239.

1256 Byrne, R. W. (2004). Culture in the great apes: using intricate com-
1257 plexity in feeding skills to trace the evolutionary origin of human
1258 technical prowess. *Philosophical Transactions of the Royal Soci-*
1259 *ety, B*, *362*, 577–585.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: a
1260 hierarchical approach. *The Behavioral and Brain Sciences*, *21*,
1261 667–684. discussion 84–721.

1262 Byrne, R., & Whiten, A. (1988). *Machiavellian intelligence: Social*
1263 *expertise and the evolution of intellect in monkeys, apes, and*
1264 *humans*. Oxford: Clarendon Press.

1265 Caligiore, D., Borghi, A. M., Parisi, D., & Baldassarre, G. (2010).
1266 TRoPICALS: a computational embodied neuroscience model of
1267 compatibility effects. *Psychological Review*, *117*, 1188–1228.

1268 Call, J., & Tomasello, M. (2007). *The gestural communication of apes*
1269 *and monkeys*. New York: Lawrence Erlbaum Associates.

1270 Campos, M., Breznen, B., & Andersen, R. A. (2010). A neural repre-
1271 sentation of sequential states within an instructed task. *Journal of*
1272 *Neurophysiology*, *104*, 2831–2849.

1273 Cartmill, E. (2008). *Gestural communication in orangutans (Pongo*
1274 *pygmaeus and Pongo abelii): A cognitive approach*. St.
1275 Andrews: University of St. Andrews.

1276 Cartmill, E. A., & Byrne, R. W. (2011). Studying non-human primate
1277 gesture from the perspective of the individual In G. Stam, M.
1278 Ishino (Eds.), *Integrating gestures* (pp. 15–26). John Benjamins

1279 Chaminade, T., Oztop, E., Cheng, G., & Kawato, M. (2008). From self-
1280 observation to imitation: visuomotor association on a robotic
1281 hand. *Brain Research Bulletin*, *75*, 775–784.

1282 Chang, S.W.C., Gariépy, J-F., & Platt, M.L. (2012). Neuronal refer-
1283 ence frames for social decisions in primate frontal cortex. *Nature*
1284 *Neuroscience*.

1285 Cisek, P. (2007). Cortical mechanisms of action selection: the afford-
1286 ance competition hypothesis. *Philosophical Transactions of the*
1287 *Royal Society of London. Series B, Biological Sciences*, *362*,
1288 1585–1599.

1289 Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting
1290 with a world full of action choices. *Annual Review of Neurosci-*
1291 *ence*, *33*, 269–298.

1292 Clark, E.V. (1978). Strategies for communicating. *Child Development:*
1293 *953–59*.

Q5 1294 Corballis, M. C. (2002). *From hand to mouth, the origins of language*.
1295 Princeton: Princeton University Press.

1296 de Rengerv, A. D, Boucenna, S., Andry, P., & Gaussier, P. (2010).
1297 Emergent imitative behavior on a robotic arm based on visuo-
1298 motor associative memories. *Intelligent Robots and Systems*
1299 *(IROS), 2010 IEEE/RSJ International Conference on*.

1300 Deacon, T. W. (1997). *The symbolic species: The co-evolution of*
1301 *language and the brain*. New York: WW Norton.

1302 Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N.
1303 (2012). Identification of the social and cognitive processes under-
1304 lying human cumulative culture. *Science*, *335*, 1114–1118.

1305 Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per
1306 view: adaptive valuation of social images by rhesus macaques.
1307 *Current Biology: CB*, *15*, 543–548.

1308 Deco, G., & Rolls, E. T. (2004). A Neurodynamical cortical model of
1309 visual attention and invariant object recognition. *Vision Research*,
1310 *44*, 621–642.

1311 Demiris, Y., & Hayes, G. (2002). In K. Dautenhahn, C. Nehaniv (Eds)
1312 *Imitation as a dual-route process featuring predictive and learn-*
1313 *ing components: A biologically-plausible computational model*
1314 (pp. 327–61): MIT Press.

1315 di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G.
1316 (1992). Understanding motor events: a neurophysiological study.
1317 *Experimental Brain Research*, *91*, 176–180.

1318 Dominey, P. F., Arbib, M. A., & Joseph, J. P. (1995). A model of
1319 corticostriatal plasticity for learning oculomotor associations
1320 and sequences. *Journal of Cognitive Neuroscience*, *7*, 311–
1321 336.

1322 Dumas, G., Chavez, M., Nadel, J., & Martinerie, J. (2012). Anatomical
1323 connectivity influences both intra- and inter-brain synchroniza-
1324 tions. *PLoS One*, *7*, e36414–e14.

1325

- 1326 Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary*
1327 *Anthropology: Issues, News, and Reviews*, 6, 178–190.
- 1328 Fagg, A. H., & Arbib, M. A. (1998). Modeling parietal-premotor
1329 interactions in primate control of grasping. *Neural Networks:*
1330 *the Official Journal of the International Neural Network Society,*
1331 *11*, 1277–1303.
- 1332 Ferrari, P. F., Paukner, A., Ionica, C., & Suomi, S. J. (2009). Reciprocal
1333 face-to-face communication between rhesus macaque mothers
1334 and their newborn infants. *Current Biology: CB*, 19, 1768–1772.
- 1335 Fujii, N., Hihara, S., Nagasaka, Y., & Iriki, A. (2008). Social state
1336 representation in prefrontal cortex. *Social Neuroscience*: 1–12.
- 1337 Galef, B. G., & Laland, K. N. (2005). Social learning in animals:
1338 empirical studies and theoretical models. *BioScience*, 55,
1339 489–89.
- 1340 Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action
1341 recognition in the premotor cortex. *Brain*, 119, 593–609.
- 1342 Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural
1343 communication of the gorilla (*Gorilla gorilla*): repertoire, inten-
1344 tionality and possible origins. *Animal Cognition*, 12, 527–546.
- 1345 Georgopoulos, A. P. (2002). Cognitive motor control: spatial and
1346 temporal aspects. *Current Opinion in Neurobiology*, 12, 678–683.
- 1347 Georgopoulos, A. P., Kalaska, J. F., & Massey, J. T. (1981). Spatial
1348 trajectories and reaction times of aimed movements: effects of
1349 practice, uncertainty, and change in target location. *Journal of*
1350 *Neurophysiology*, 46, 725–743.
- 1351 Guazzelli, A., Corbacho, F. J., Bota, M., & Arbib, M. A. (1998).
1352 Affordances, motivation, and the world graph theory. *Adaptive*
1353 *Behavior*, 6, 435–471.
- 1354 Hecht, E. E., Gutman, D. A., Preuss, T. M., Sanchez, M. M., Parr, L.
1355 A., & Rilling, J. K. (2012). Process versus product in social
1356 learning: comparative diffusion tensor imaging of neural systems
1357 for action execution-observation matching in Macaques, Chim-
1358 panzees, and Humans. *Cerebral cortex (New York, N.Y.: 1991)*.
- 1359 Hirata, S., Matsuda, G., Ueno, A., Fuwa, K., Sugama, K., et al. (2011).
1360 Event-related potentials in response to subjects' own names: a
1361 comparison between humans and a chimpanzee. *Communicative*
1362 *& Integrative Biology*, 4, 321–323.
- 1363 Hobaiter, C., & Byrne, R. W. (2010). Able-bodied wild chimpanzees
1364 imitate a motor procedure used by a disabled individual to over-
1365 come handicap. *PLoS One*, 5, e11959–e59.
- 1366 Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild
1367 chimpanzee. *Animal cognition*.
- 1368 Hoff, B., & Arbib, M. A. (1993). Models of trajectory formation and
1369 temporal interaction of reach and grasp. *Journal of Motor Behav-*
1370 *ior*, 25, 175–192.
- 1371 Hopkins, W. D., Coulon, O., & Mangin, J. F. (2010). Observer-
1372 independent characterization of sulcal landmarks and depth asym-
1373 metry in the central sulcus of the chimpanzee brain. *Neuroscience*,
1374 171, 544–551.
- 1375 Jeannerod, M., & Biguer, B. (1982). Visuomotor mechanisms in reach-
1376 ing within extra-personal space. In D. J. Ingle, R. J. W. Mansfield,
1377 & M. A. Goodale (Eds.), *Advances in the analysis of visual*
1378 *behavior* (pp. 387–409). Cambridge: The MIT Press.
- 1379 Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995).
1380 Grasping objects: the cortical mechanisms of visuomotor trans-
1381 formation. *Trends in Neurosciences*, 18, 314–320.
- 1382 Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures:
1383 a comparative eye-tracking study. *Proceedings. Biological Scien-*
1384 *ces/The Royal Society*, 276, 1949–1955.
- 1385 Keller, S. S., Deppe, M., Herbin, M., & Gilissen, E. (2012). Variability
1386 and asymmetry of the sulcal contours defining Broca's area ho-
1387 mologue in the chimpanzee brain. *The Journal of Comparative*
1388 *Neurology*, 520, 1165–1180.
- 1389 Klein, J. T., Deaner, R. O., & Platt, M. L. (2008). Neural correlates of
1390 social target value in macaque parietal cortex. *Current Biology:*
1391 *CB*, 18, 419–424.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297, 846–848.
- Lee, D., McGreevy, B. P., & Barraclough, D. J. (2005). Learning and decision making in monkeys during a rock-paper-scissors game. *Brain Research. Cognitive Brain Research*, 25, 416–430.
- Liebal, K., Call, J., & Tomasello, M. (2004a). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, 64, 377–396.
- Liebal, K., Pika, S., & Tomasello, M. (2004b). Great ape communicators move in front of recipients before producing visual gestures. *Interaction Studies*, 5, 199–219.
- Liebal, K., Pika, S., & Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture*, 6, 1–38.
- Lock, A. (2001). Preverbal communication. In J. G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 379–403). Oxford: Blackwell.
- Nishida, T. (1980). The leaf-clipping display: a newly discovered expressive gesture in wild chimpanzees. *Journal of Human Evolution*, 9, 117–128.
- O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, 18, 283–328.
- Oztop, E., & Arbib, M. A. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87, 116–140.
- Oztop, E., Wolpert, D., & Kawato, M. (2005). Mental state inference using visual control parameters. *Brain Research. Cognitive Brain Research*, 22, 129–151.
- Parr, L. A., Hecht, E., Barks, S. K., Preuss, T. M., & Votaw, J. R. (2009). Face processing in the chimpanzee brain. *Current Biology: CB*, 19, 50–53.
- Perlman, M., Tanner, J. E., & King, B. J. (2012). In S. Pika & K. Liebal (Eds.), *A mother gorilla's variable use of touch to guide her infant: Insights into iconicity and the relationship between gesture and action* (pp. 55–73). Amsterdam: John Benjamins Publishing Company.
- Petreska, B., Adriani, M., Blanke, O., & Billard, A. G. (2007). Apraxia: a review. *Progress in Brain Research*, 164, 61–83.
- Pika, S., & Liebal, K. (Eds.). (2012). *Developments in primate gesture research*. Amsterdam: John Benjamins Publishing Company.
- Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): gestural repertoire, learning, and use. *American Journal of Primatology*, 60, 95–111.
- Pika, S., Liebal, K., & Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): Repertoire and use. *American Journal of Primatology*, 65, 39–61.
- Ramayya, A. G., Glasser, M. F., & Rilling, J. K. (2010). A DTI investigation of neural substrates supporting tool use. *Cerebral Cortex (New York, N.Y.: 1991)*, 20, 507–516.
- Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., et al. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11, 426–428.
- Rilling, J.K., Scholz, J., Preuss, T.M., Glasser, M.F., Errangi, B.K., & Behrens, T.E. (2011). Differences between chimpanzees and bonobos in neural systems supporting social cognition. *Social cognitive and affective neuroscience*.
- Rizzolatti, G., Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., & Ponzoni-Maggi, S. (1987). Neurons related to goal-directed motor acts in inferior area 6 of the macaque monkey. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, 67, 220–224.
- Rothi, L.J.G., Heilman, K.M., Affairs, V., & Centre, M. (1991). A Cognitive Neuropsychological Model of Limb Praxis Disconnection Model of Apraxia. 8
- Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: 1457

- 1458 rules without symbols. *Proceedings of the National Academy of*
 1459 *Sciences of the United States of America*, 102, 7338–7343.
- 1460 Rudebeck, P. H., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S.
 1461 (2006). A role for the macaque anterior cingulate gyrus in social
 1462 valuation. *Science (New York, N.Y.)*, 313, 1310–1312.
- 1463 Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., et al.
 1464 (2011). Differential prefrontal white matter development in chimpanzees and humans. *Current Biology*, 21, 1397–1402.
- 1465 Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X.,
 1466 et al. (2011). Social network size affects neural circuits in macaques. *Science (New York, N.Y.)*, 334, 697–700.
- 1467 Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation
 1468 for action in the parietal cortex of the monkey. *Nature*, 415, 918–
 1469 922.
- 1470 Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A. R.,
 1471 Stimpson, C. D., et al. (2010). Broca's area homologue in chimpanzees (Pan troglodytes): probabilistic mapping, asymmetry, and comparison to humans. *Cerebral Cortex (New York, N.Y.: 1991)*, 20, 730–742.
- 1472 Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 27, 8366–8377.
- 1473 Seo, H., Barraclough, D. J., & Lee, D. (2009). Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 29, 7278–7289.
- 1474 Shepherd, S. V., Klein, J. T., Deaner, R. O., & Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Science*, 106, 9489–9494.
- 1475 Shilbach, L., Timmermans, B., Reddy, V., Costall, A., & Bente, G. et al. (2012). Toward a second-person neuroscience. *Behavioral & Brain Sciences*.
- 1476 Silver, M. R., Grossberg, S., Bullock, D., Histed, M. H., & Miller, E. K. (2012). A neural model of sequential movement planning and control of eye movements: item-order-rank working memory and saccade selection by the supplementary eye fields. *Neural Networks*, 26, 29–58.
- 1477 Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7, 308–312.
- 1478 Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.
- 1479 Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PloS One*, 6, e18852–e52.
- 1480 Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, 83, 29–36.
- 1481 Tomasello, M., & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.
- 1482 Tomasello, M., & Call, J. (2007). Intentional communication in non-human primates. In J. Call & M. Tomasello (Eds.), *The gestural communication of apes and monkeys* (pp. 1–15). Mahwah: Lawrence Erlbaum Associates.
- 1483 Tomasello, M., & Call, J. (2011). Methodological challenges in the study of primate cognition. *Science*, 334, 1227–1228.
- 1484 Tomasello, M., & Camaioni, L. (1997). A comparison of the gestural communication of apes and human infants. *Human Development*, 1997, 7–24.
- 1485 Ueno, A., Hirata, S., Fuwa, K., Sugama, K., Kusunoki, K., et al. (2008). Auditory ERPs to stimulus deviance in an awake chimpanzee (Pan troglodytes): towards hominid cognitive neurosciences. *PloS One*, 3, e1442–e42.
- 1486 van Schaik, C. P., Ancrenaz, N., Borgen, G., Galdikas, B., Knott, C. D., et al. (2003). Orangutan cultures and the evolution of the material culture. *Science*, 299, 102–105.
- 1487 Watanabe, M. (2007). Role of anticipated reward in cognitive behavioral control. *Current Opinion in Neurobiology*, 17, 213–219.
- 1488 Watanabe, M., & Sakagami, M. (2007). Integration of cognitive and motivational context information in the primate prefrontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 17(Suppl 1), i101–i109.
- 1489 Whiten, A., Goodall, J., McGrew, C., Nishida, T., Reynolds, V., et al. (1999). Cultures in chimpanzees. *Nature*, 399, 682–685.
- 1490 Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., et al. (2001). Changing cultural variation in chimpanzees. *Behaviour*, 138, 1481–1516.
- 1491 Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2011). Representation of others' action by neurons in monkey medial frontal cortex. *Current Biology: CB*, 21, 249–253.
- 1492 Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2012). Social error monitoring in macaque frontal cortex. *Nature Neuroscience*, 15, 1307–1312.
- 1493 Zukow-Goldring, P., & Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: caregivers and the directing of attention. *Neurocomputing*, 70, 2181–2193.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

- Q1. Please check if the section headings are assigned to appropriate levels.
- Q2. Please provide the JournalName, VolumeNumber and PageNumber for this reference item.
- Q3. Please provide Location of the PublisherName for reference item Cartmill & Byrne (2011); Demiris & Hayes (2002)
- Q4. Please provide the VolumeNumber and PageNumber of reference items Chang et al. (2012); Hobaiter & Byrne (2011)
- Q5. Please provide the VolumeNumber for reference item Clark (1978); Fujii et al. (2008); Rilling et al. (2011); Shilbach et al. (2012)
- Q6. Please check affiliations if captured and presented correctly.
- Q7. Missing citation of Figure 4 was inserted here. Please check if appropriate.

UNCORRECTED PROOF