Methodological Challenges in the Study of Primate Cognition
Michael Tomasello, *et al.*
Science **334**, 1227 (2011);
DOI: 10.1126/science.1213443

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

have not yet encountered cases in which data and code were submitted for reproducibility review but results were not reproducible as claimed. It is encouraging that authors are taking advantage of the policy to make their work reproducible by others, but more work could be done to promote a broader adoption of the policy.

The fact that an analysis is reproducible does not guarantee the quality, correctness, or validity of the published results. The "R" kite-mark is meant to convey the idea that a knowledgeable individual has reviewed the code and data and was capable of producing the results claimed by the author. In cases in which questionable results are obtained, reproducibility is critical to tracking down the "bugs" of computational science. In cases with interesting findings, reproducibility can greatly facilitate building on those findings (*12*).

Perhaps the biggest barrier to reproducible research is the lack of a deeply ingrained culture that simply requires reproducibility for all scientific claims. Not unlike the culture of replication that persists across all scientific disciplines, the scientific community needs to develop a "culture of reproducibility" for computational science and require it of published claims. Another important barrier is the lack of an integrated infrastructure for distributing reproducible research to others. The current system is ad hoc with researchers in some fields having access to sophisticated central data repositories and researchers in other fields having few useful resources for sharing code and data. In many cases, a researcher does not have an obvious place to turn to make sure their work is reproducible and accessible by others. Journals' supporting online materials have some severe limitations, such as the inability to search and index available data.

Given the barriers to reproducible research, it is tempting to wait for a comprehensive solution to arrive. However, even incremental steps would be a vast improvement over the current situation. To this end, I propose the following steps (in order of increasing impact and cost) that individuals and the scientific community can take. First, anyone doing any computing in their research should publish their code. It does not have to be clean or beautiful (*13*), it just needs to be available. Even without the corresponding data, code can be very informative and can be used to check for problems as well as quickly translate ideas. Journal editors and reviewers should demand this so that it becomes routine. Publishing code is something we can do now for almost no additional cost. Free code repositories already exist [for example, GitHub (http://github.com) and SourceForge (http://sourceforge.net)], and at a minimum, code can be published in supporting online material. The next step would be to publish a cleaned-up version of the code along with the data sets in a durable nonproprietary format. This will involve some additional cost because not everyone will have the resources to publish data. Some fields such as genomics have already created data repositories, but there is not yet a general solution.

Last, the scientific community can pool its collective resources to create a DataMed Central and CodeMed Central, analogous to PubMed Central for all data, metadata, and code to be stored and linked with each other and with corresponding publications. Such an effort would probably need government coordination and support, but each would serve as a single gateway that would guide researchers to field-specific data and code repositories. Existing repositories could continue to be used and would interface with the gateway, whereas fields without existing infrastructure would be given access to these resources. The ultimate goal would be to provide a single place to which people in all fields could turn to make their work reproducible.

The field of science will not change overnight, but simply bringing the notion of reproducibility to the forefront and making it routine will make a difference. Ultimately, developing a culture of reproducibility in which it currently does not exist will require time and sustained effort from the scientific community.

References

1. B. Hanson, A. Sugden, B. Alberts, *Science* **331**, 649 (2011).
2. R. D. Peng, F. Dominici, S. L. Zeger, *Am. J. Epidemiol.* **163**, 783 (2006).
3. J. P. Madrid, D. Macchetto, *Bull. AAS*, arXiv:0901.4552 (2009).
4. M. Schwab, N. Karrenbach, J. Claerbout, *Comput. Sci. Eng.* **2**, 61 (2000).
5. C. Laine, S. N. Goodman, M. E. Griswold, H. C. Sox, *Ann. Intern. Med.* **146**, 450 (2007).
6. G. King, *PS: Polit. Sci. Polit.* **28**, 444 (1995).
7. R. Gentleman, *Stat. Appl. Genet. Mol. Biol.* **4**, article 2 (2005).
8. Yale Law School Roundtable on Data and Code Sharing, *Comput. Sci. Eng.* **12**, 8 (2010).
9. J. P. A. Ioannidis *et al.*, *Nat. Genet.* **41**, 149 (2009).
10. J. P. Mesirov, *Science* **327**, 415 (2010).
11. R. D. Peng, *Biostatistics* **10**, 405 (2009).
12. B. D. McCullough, K. A. McGeary, T. D. Harrison, *Can. J. Econ.* **41**, 1406 (2008).
13. N. Barnes, *Nature* **467**, 753 (2010).

10.1126/science.1213847

PERSPECTIVE

Methodological Challenges in the Study of Primate Cognition

Michael Tomasello* and Josep Call*

Laboratory studies of primate cognition face the problem that captive populations of a species are not always comparable, and generalizations to natural populations are never certain. Studies of primate cognition in the field face the problem that replications are expensive and difficult, and again different populations are not always comparable. To help remedy these problems, we recommend the creation of data banks where primary data and videotapes may be deposited (perhaps as a requirement of publication) to facilitate cross-examination, replication, and, eventually, the pooling of data across investigators.

The basic methodological premise in the study of primate, including human, cognition is that the same overt behavior may be underlain by different mechanisms. Cognitive mechanisms are prototypically those in which organisms cognitively represent nonperceived situations and draw inferences that go beyond immediate perception, which enables organisms to make flexible behavioral decisions even in novel environmental situations. Research in primate cognition is aimed at discovering these cognitive mechanisms by observing their effects in adaptive action.

Much of the research in primate cognition is conducted with captive populations in the laboratory, where investigators can best control conditions experimentally. Field researchers sometimes argue that, in this setting, subjects are faced with cognitive problems that are not well matched with the cognitive abilities that have evolved in their species-typical environments. [For more on the debate between laboratory and field researchers, see the paper by Ryan in this issue (*1*).] Giving captive populations novel tasks, however, sometimes reveals cognitive skills not observed in nature. For example, gorillas in captivity are quite proficient at using tools in a variety of extractive foraging tasks, although they have not been observed to use them in the wild (*2*). Because all of the other great apes also use tools proficiently in the laboratory, this suggests that the common ancestor to all apes used tools (*3*)—an hypothesis that could never be formulated if gorillas were not tested for tool use in captivity. The main methodological challenge of experimental research in primate cognition, therefore, is to design problems that are novel for individuals—so

Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.

*To whom correspondence should be addressed. E-mail: tomas@eva.mpg.de (M.T.); call@eva.mpg.de (J.C.)

that they are forced to use their cognitive skills flexibly and creatively—and yet still, in some sense, ecologically valid for the species.

Another challenge in experimental work is the comparability, or lack thereof, of different populations of individuals of the same species. Thus, it is never certain that an experiment with laboratory animals can be generalized to conspecifics in the wild because captive individuals in their human-designed environments may have developed a unique set of cognitive skills. (The reverse generalization from wild to captive populations is therefore always uncertain as well.) It is not even certain that results from one captive population may be straightforwardly generalized to another. The problem is that different captive populations may have had different experiences relevant to a particular cognitive task, and indeed some captive populations may even become "test savvy" due to the repeated testing of individuals over years of research [e.g., (*4*)]. Regularly comparing these populations to experimentally naïve ones can partially mitigate this problem and should be done whenever possible. In addition, many experimental studies of primate cognition involve only a relatively small number of subjects because the captive colony is small, making generalizations to the species as a whole problematic. A related problem is that some species of primates are relatively rare in captivity, so only one or a few researchers have access to them, making replication of experiments across different laboratories difficult.

Most field research in primate cognition is based on systematic observations. Systematic observations are to be distinguished from the kinds of anthropomorphized anecdotes often found in popular publications comprising one-time observations of individuals without any kind of background observations or context to constrain interpretation. Nevertheless, many experimentalists would claim that cognitive abilities are best (or even only) revealed in experimental settings, because, in principle, the same overt behavior may be underlain by very different cognitive processes, and these processes are only revealed if observations of the behavior take place under varied, ideally preplanned, conditions. Many methodological challenges to field approaches derive from the great time and expense involved. Thus, it is often the case that only one field worker can recognize all the individuals of a group—because it takes many months to acquire this skill—so interobserver observations and reliability estimates (cornerstones of laboratory behavioral research) are often not possible. In addition, because of the great time and expense involved, extensive corpora of systematic field observations are only rarely replicated at a later time on the same group by other researchers.

Other methodological challenges for field approaches to primate cognition emanate from the impossibility of controlling all relevant factors under "wild" conditions the way one might do in the laboratory. Thus, a persistent problem is that one

CREDIT: MAX PLANCK INSTITUTE FOR EVOLUTIONARY ANTHROPOLOGY

seldom knows in the field, as one typically does with captive individuals, the precise experiences individuals have had with problems relevant to the one being investigated, so novelty for individuals is almost always an open question. Field experiments are an obvious answer to many of the problems of natural observation, but for various reasons field experiments have been a critically underused strategy in the study of primate cognition, with experiments using techniques other than the playing back of recorded vocalizations to a social group [as in, e.g., the classic studies of (*5*)] being few and far between [see (*6*–*8*) for some exceptions]. More focus on field experiments is called for.

Cognitive mechanisms have physiological instantiations in the brain, of course, and investigating the physical bases of primate cognition is an important avenue of research as well. However, neurophysiological research has so far been conducted mostly with quite simple cognitive tasks involving basic processes of perception and memory, not with the more complex tasks used in experimental behavioral research. Moreover, neurophysiological research can only be done with some primates and in some ways because of ethical concerns. Intervening in the lives of wild animals experimentally may also, in some cases, raise ethical concerns. Research with individual great apes raised in human-like environments and taught human-like skills have produced some interesting, although not really systematic, results [e.g., (*9*)]—with investigations of this type unlikely to be pursued vigorously in the future, again due mainly to ethical concerns.

Despite a recent well-publicized case (*10*), scientific misconduct in the field of primate cognition would seem to be no more common than in other scientific fields. Still, the fact is that field observations are often not replicated because of the time and expense involved, and experimental investigations are sometimes difficult to replicate because of the rarity of some species or the incomparability of different populations. These issues could potentially make detecting misconduct more difficult in the study of primate cognition than in scientific disciplines in which replication is easier and more frequent. A mechanism that could help to correct this situation is a requirement by journals that researchers deposit copies of their primary data (including videotapes) on an institutionally operated Web site as a condition of publication, and then rules for data sharing could be established by consensus among scientists and journal editors.

References

1. M. J. Ryan, *Science* **334**, 1229 (2011).
2. M. Tomasello, J. Call, *Primate Cognition* (Oxford Univ. Press, New York, 1997).
3. W. C. McGrew, in V. Standen, R.A. Foley, Eds., *Comparative Socioecology: The Behavioural Ecology of Humans and Other Mammals* (Blackwell Scientific Publications, Oxford, 1989), pp. 457–472.
4. S. Inoue, T. Matsuzawa, *Curr. Biol.* **17**, R1004 (2007).
5. D. L. Cheney, R. M. Seyfarth, *How Monkeys See the World*. (Univ. Chicago Press, Chicago, 1990).
6. T. Matsuzawa, in T. Matsuzawa, M. Tomonaga, M. Tanaka, Eds. (Springer, Tokyo, 2006), pp. 3–33.
7. C. R. Menzel, *Anim. Behav.* **41**, 397 (1991).
8. E. Visalberghi *et al*., *Curr. Biol.* **19**, 213 (2009).
9. E. S. Savage-Rumbaugh *et al*., *Language Comprehension in Ape and Child: Monographs of the Society for Research in Child Development, 58, 3-4, No. Serie 233* (Wiley-Blackwell, Hoboken, 1993).
10. G. Miller, *Science* **329**, 890 (2010).

10.1126/science.1213443