Imitation and Artifice in Apes, Humans, and Machines

DEREK EDWARDS
Loughborough University, United Kingdom

> Mrs. Johns . . . was delighted with her bird's new home. "You could tell he was really pleased, he ran about rounding up his hens."
>
> —*The Guardian*, June 1, 1993
> (reprinted by permission)

This snippet from a human interest story concerns a noisy cockerel, "Corky," recently relocated to a remote farmyard to avoid disturbing a neighbor who had won a court case to get the bird silenced. Other neighbors were reported to have joined Mrs. Johns in what the judge called "some sort of crusade on behalf of oppressed chickens" (Sharrock, 1993). Note the similarity between Mrs. Johns's reported reaction and that attributed to her cockerel, and the reason given for that attribution. Try substituting *annoyed, anxious, unhappy* or other negative feelings for "pleased" and see if the sentence is any less convincing.

The concern here is with what happens when ordinary people, as well as behavioral scientists, contentiously attribute to animals (machines, infants, and other Others) the kinds of mental states or competencies that they attribute to each other, and the grounds on which they do so. In the more serious literature, the prime topics are language and intelligence, and the objects in question are often apes, computers, or human infants. My aim is not to resolve the question of nonhuman agency in any of these cases but to examine some ways in which that issue is generally constituted and decided. The focus is on how various criteria of agency (or "human-ness") work, such as imitation, description, intentionality, attribution, and comembership. Although this might look like a side step, a focus on the debate in preference to its topic, the argument here is that the operation of these criteria, both in ordinary discourse and in the scientific literature, are the substance of it all.

LOOKING IN THE MIRROR

The most common trope for these concerns must be the "Mirror to Man." We study apes, "those amazing creatures that can teach us so much about ourselves" (Goodall, 1971, p. 14). "Discovery of extraterrestrial life would utterly trans-

AMERICAN BEHAVIORAL SCIENTIST, Vol. 37 No. 6, May 1994 754-771

form our own view of ourselves as a species" (Professor Martin Rees, Director of Cambridge University's Institute of Astronomy, quoted in "SETI—the search," 1990). Technology in general may be seen as the Mirror (Woolgar, 1987, p. 312), or at least artificial intelligence (AI) may, and some projections of that enterprise produce reflections indistinguishable from the human original: "If machines do come to simulate all of our internal cognitive activities, to the last computational detail, to deny them the status of genuine persons would be nothing but a new form of racism" (Churchland, 1988, p. 120).

On the other hand, what the Mirror reveals, as Snow White's wicked stepmother discovered, may not be as welcome, or (it is hoped) as accurate or informative, as anticipated. Some folk would deny the trope or want the Mirror broken. Despite hopes for informative reflections, it may all be a flawed exercise in domination and hubris, a wrong analogy: "Western primatology is Simian orientalism" (Haraway, 1989, p. 10, drawing a further analogy to Said, 1979), a hall of dark and distorted mirrors where neither ape nor human being is clearly seen. It turns into an imposition of cultural assumptions, indeed a reification of them in the form of science, mere confirmation rather than transformation of our view of ourselves. Looking into Man's Mirror guarantees no clear vision, either of the object or of ourselves looking back. It is an issue not only for the contentious humanity of "marginal objects" (Turkle, 1984) such as apes and machines but for human studies too, and especially social anthropology, a more obviously cultural sort of mirror "for displacing the dulling sense of familiarity with which the mysteriousness of our own ability to relate to one another is concealed from us" (Geertz, 1973, p. 14).[1]

Turkle's notion of "marginal objects" invokes the sense of a to-be-defined set of protohuman, or "almost human" (Strum, 1987; Yerkes, 1925), beings at the margins of full membership with, let us say, us.[2] The computer "gives people new ways to think about themselves," it is "a new mirror," "a Rorschach, a projective screen for other concerns" (Turkle, 1984, pp. 24, 306, 320). Such objects or beings are both similar to and different from us, in ways to be defined, investigated, and *changed*. Apes are considered a good choice to start with in that they seem the most like us, yet sufficiently unlike us, to make the project interesting. They are the closest modern equivalents to the creatures in Arthur C. Clarke's (1968) *2001: A Space Odyssey*—on the verge of humanity, just needing a little external help. The (un)likeness varies across descriptions; the "anthropoid" apes may deceptively look like and imitate us, but what we cannot immediately see is the underlying neural and intellectual chasm that separates language-using humans from their signal-swapping cousins. Or perhaps we and they "do not *look* like sibling species" at all, but are closer beneath appearances, where "Man/Ape protein structure is practically identical" (Desmond, 1979, pp. 14-15).

The same sorts of contrary opposites are available for androids and humans; computers, brains and minds.[3] And so also for the other senses of "marginal," the moral outsiders, the not-us, the beings on which/whom efforts are expended or wasted, contentiously, to make them more like us. We put animals through

the conceptual categories of our own prehistory, from wild, to domesticated, to virtually human, "an epochal unfolding of animal to man, from instinct to consciousness" (Haraway, 1989, p. 60). Talking apes and simulating computers are of our making, whether by manufacture or by description, made in our image as God made us; monsters that may threaten, like Frankenstein's creation[4] and the wicked queen's reflection, our sense of our own unique worth. Yet there is the sense also of ourselves as cultural beings, inventions, artifacts, of our own if not of God's making; but in whose image? Artificial intelligence may yet be the only kind there is.

IMITATION

Imitation is a rich idea. In use, it resonates with contradictions. It is both the aim and the critique of simulation. Displaced by cognitivism from its role as behavioral explanation of how children become competent adults, it becomes available for dismissing cognitive inadequacies: superficial copies, mere imitations. Children's language is contrasted with Washoe's, Lana's, Nim's and Sarah's on the grounds that unlike theirs it is nonimitative, nonaped; it is produced creatively from grammar (Brown, 1973; Terrace, 1979). The exceptions are just that, exceptional, one-offs, capable of alternative, more parsimonious explanations;[5] they prove the rule. So ape language is like but unlike, too much of a copy yet not close enough. Similarly, however closely computers can be made to function like people, failure is guaranteed by the failures of imitation: "So long as they are not members of our society . . . they cannot imitate our intelligent activities" (Collins, 1990, p. 18). Of course, this begs the question of what it takes to be counted as a member and whether there is anything to that apart from being counted as.

Imitation is available as a category for discounting things. Apes' efforts at acquiring human language (or rather, humans' efforts to train them to use it) are dismissed as either too loose or too tightly constrained. They either babble without grammatical structure or they converse in routines, fixed by imitation and rehearsal (Brown, 1973). Rote learning produces routines without understanding, behavior without competence. Overlearned laboratory tricks are responsible for a semblance of grammar, the mere product of a long training session.[6] Otherwise, the semblance of productive spontaneity in less rigid ape-human interactions merely hides an imitative pattern revealed by close scrutiny of videotaped sequences (Terrace, 1979). Methods of quoting, presenting, and analyzing ape language are said to have exaggerated their human-likeness, by obscuring the routine and imitative context of their production and by "trimming" repetitive strings of signs, transforming them into "data" that "manufacture a human-like end product" (Desmond, 1979, p. 49).

The notion that imitation-based talk is not proper talk is also part of a dispute between behaviorism and cognitive psychology. Howard and Beatrice Gardner,

who trained Washoe in sign language (Gardner & Gardner, 1969) were explicitly engaged in an argument against Chomsky (1959) for the merits of behavioral methods and conceptions of human competence, including language. From their perspective, it was not an error or oversight that Washoe's performance was based on imitation and behavioral training but an important point of principle and practice. The imitative character of Washoe's signing was a product of how they shaped her hands into gestures and rewarded her for getting it right. Imitation was built in, a feature of the behavioral scientist's equation of control with understanding; laboratory animals are like artificial intelligence devices in this respect, "*designed* . . . in short, engineered, to answer human queries" (Haraway, 1989, p. 62, original emphasis), in the pursuit of the perfectibility of "man." Apes have to be *made into* mirrors.

Imitation works as a *descriptive* category; the making into mirrors can be a discursive construction, as well as an engineering one. Harvey Sacks describes how, in early American ethnographies, "again and again I found references to the activities of Negroes as 'imitating whites'." And they were characterized as being "marvelous imitators." Such reports are very similar to the way the behavior of children is characterized (Sacks, 1992, p. 70). Sacks (1992) makes sense of this in terms of how classes of persons are seen as entitled to behave:

> "Imitation" seems to involve a way of characterizing some action which somebody does when they are unentitled to do that class of action. And if you watch the way the Negro slaves got talked about, or the way the emerging Negro is talked about, you can see how marvelous a category "imitation" is, because it turns out that everything whites can do Negroes can imitate, but they can't *do* any of these things that whites can do. (p. 70, original emphasis)

So attribution can precede evidence; you categorize the actor, and that provides for how to interpret the actions. Imitation is a way of describing, and thus constituting as such, the actions of entities that are not proper, fully fledged, members or persons. That is not to say that it is a false category, or that some actions are not imitations; and this is not to say that some are. Rather, it is to say that it is a category, a form of description, and it has its attributional uses. As Sacks remarks, "Such categories as 'imitation' and 'phoney' . . . serve as boundary categories around the term 'Member' " (p. 70).

There are several linked senses of "imitation": There is the behavior itself, and also its ersatz product, such as imitation antique furniture. There is also an intentional sort of usage, invoked by "phoney," or "artifice," in which imitating is a kind of deceptive, goal-oriented action.[7] In critiques of what apes, infants, computers (and Negroes) can do, imitation is dismissive, a way of discounting some performance as not genuine, less than the original, not produced by the same mechanism, mere imitation. This is the sense used also in critiques of behavioral theories—that invoking imitation, like other behavioral routines, as explanations, leaves out what is definitive about human nature, that people can do things creatively and intentionally, as proper persons, and not merely by imitation.[8] However, if imitative apes were to be described as *passing themselves*

off as human, or computers to be *dis*-simulating, the implications would be fully reversed. Deception or artifice is taken to be a hallmark of sophistication, of mind, intelligence, even of "mind reading" (Leslie, 1987; Whiten, 1989), and a capacity not beyond the capacities of baboons, apes, and prelinguistic children. For example, the Gardners provide accounts of Washoe embarrassing new project members, new to sign language, by signing particularly slowly and carefully for them (Gardner & Gardner, 1974). These observations are, in the first instance, like Sacks's, matters of description.[9] Although distinctions between actions on the grounds of intentionality *can be made* between human and nonhuman, animal and machine, one animal and another, or one human action and another (Collins, 1990), these are attributive descriptions of the sort that participants in interactions might themselves deal with, as an issue of accountability (cf. Edwards & Potter, 1992). The problem for analysis would then be, how are actions *brought off as* intentional or otherwise?

Collins (1990) distinguishes a class of "behavior-specific actions" that persons can do and that machines are able to mimic. The category is defined by intention, being the things people are trying to perform in precisely the same way every time, such as parade ground drill. The trouble is that trying, not trying, not really trying, not trying enough, and so on are available as participants' descriptions for such actions, as are claims that particular performances were inadequate efforts (tryings or performances) compared to the standard, where the standard might even vary across persons and occasions. Intentionality is an idealized criterion when used by theorists in this *essentialist* way, outside of any study of situated descriptions and accounts. Signatures are an example of human actions where their "imitative" form defines how they work, as indexes of identity and authorship, where the more complex and interesting act would be forgery, a further but illegitimate imitation. Imitation here becomes the hallmark of agency. To list signatures and forgeries as imitative actions in the dismissive sense used for animals, machine simulations, and children would be presumptive indeed, as would any notion that some particular intended meaning can be legislated for them (Derrida, 1977).

ATTRIBUTION AND RAPPORT

My earlier distinction between discursive and engineering "construction," drawn for the making of mirrors, should now be reexamined. The distinction itself is a discursive construction and on the "engineering" side, when it comes to the "manufacture" of agents, whether humans, apes or computers, then the discursive, descriptive activity may figure as an important part. According to a strong tradition in developmental psychology, one of the ways in which human infants are thought to become agentive, enculturated, language-using persons is by being treated as such (e.g., Bruner, 1983; Lock, 1978; Vygotsky, 1987). This is proposed not as a treated-as, counts-as, descriptive sort of process, where

children are agents because some consensus claims they are but as an ontological, indeed ontogenetic, sort. People get constructed as agents in both senses, descriptive and engineering. The interesting point is how similar those senses are. It is the same descriptive, treated-as, counts-as, sort of procedure for inclusion into the club of agents that accomplishes the human engineering.

The notion of anthropomorphism now becomes interesting, not as an error to be avoided when describing the accomplishments of animals and machines but as an essential, built-in feature of what makes persons. Latour (1988) notes that anthropomorphism is generally taken to be an interpretative error, a false "projection" of human attributes onto something nonhuman. However,

> The automatic groom [door-closer] is already anthropomorphic through and through . . . and in three senses: first, it has been made by men, it is a construction; second, it substitutes for the actions of people, and is a delegate that permanently occupies the position of a human; and third, it shapes human action by prescribing back what sort of people should pass through the door. (Latour, 1988, p. 303)

This is a deep kind of anthropomorphism, one inherent in human relations, rather than a matter of merely false attribution.[10]

All of this is becoming unfortunate for hopes of resolving the issue of agency in nonhuman creatures/creations. Among our most promising analytical tools, or descriptive categories, for sorting out that issue are the distinctions we might draw between the real thing and mere imitation, or reality and attribution. These dichotomies go together. When mere imitation is unmasked as such, the mistake of treating it as the genuine article becomes an attributional error. This can be especially persuasive when the unmasker is also the creator, as when Baron Frankenstein rejected his creation (and Mary Shelley authored it so) and when Joseph Weizenbaum (1976) dismissed his early interactive program ELIZA as "the simplest mechanical parody" (p. 6). Weizenbaum reports how he was "startled to see how quickly and how very deeply people conversing with DOCTOR [a version of ELIZA] became emotionally involved with the computer and how unequivocally they anthropomorphized it" (p. 6).[11] However, if anthropomorphism cuts deeply into what human agency is, then the distinction between reality and its attribution is blurred. The creator/author is making it so, not "merely" describing it so.

The kind of rapport that ELIZA's human interactants felt, which so astonished Weizenbaum, is routinely recommended by ape trainers as essential for establishing successful communications with apes. Gardner and Gardner (1969) stressed the need to develop close social relations with Washoe, so that she would have something to say and find them worth talking to. This view is repeated by Rumbaugh (1977) and his colleagues, by Patterson (1978) and others. One of the objections to Terrace's imitative chimpanzee Nim was that his "battery of caretakers and trainers" (Desmond, 1979, p. 47) failed to establish the right kind of rapport, and Nim was a disturbed animal.[12] The counter to any such insistence on rapport is, of course, the danger of subjectivity, of anthropomorphic attributions: "Enthusiastic observers of animals are constantly in danger of interpreting

their behavior in more complex terms than is necessary or correct" (Griffin, 1976, p. 72). Once again, this is not an issue confined to the study of nonhuman beings. It is, phrased in terms of cultural differences, the classic problem of cultural anthropology (Geertz, 1988).

For animal research, it has been suggested that the dangers of rapport might be lessened by employing an expert in interaction microanalysis (Sebeok & Umiker-Sebeok, 1980) to spot when false attributions are occurring.[13] This would be like the stage magician and fraud buster James Randi's (1975) recommendations for checking on Uri Geller's spoon-bending claims; Geller apparently also requires the establishment of rapport with his audience. With regard to ethnographies of human cultures the problem appears both less tractable and less worrying. Not only is such anthropomorphic rapport only natural and desirable (indeed, the notion of anthropomorphism seems silly in this context) but it is not clear who, for any culture yet unstudied, could do the checking. The distinction between animal work and human ethnography is interesting, not only because it establishes a difference between humans and others but because it merely reiterates that difference. In the first place, it is a sign of what sorts of attributions "we" are willing to countenance.

Echoing Sacks's discussion of imitation, Haraway (1989) also notes how, in colonial accounts and early anthropological works, "if a black person accomplished some exceptional feat of intelligence or daring, the explanation was that he (or she?) was inspired, literally moved, by the spirit of the master" (p. 52).[14] This is a kind of reattribution, a reapportioning of responsibility for some performance, where its ostensible producer (ape, savage, infant, machine) is a mere butt or conduit for the agentive work being done by someone else. This other is one of *us*, standing in front or behind; the agent who wrote, controls, made or designed it, or else the attributing perceiver, tricked by appearances.

Nevertheless, this also is a conceptualizing move, one that splits the producer of the message from its recipient. If we adopt an interactive reader/recipient perspective on language and semiosis, rather than the more standard "communication" model, a move common to both poststructuralism and conversation analysis (Heritage, 1984; Potter & Wetherell, 1987), then it becomes less convincing to make the distinction between what the ape, the machine, the infant, and so on can really do, and what it is interactionally treated as doing. If ape-human, infant-human, and machine-human interactions are seen as such, as interactions, then it may make as little sense to extract the nonhuman partner's contributions from that matrix as it would for those of the human partner. If participants' meanings are those received, treated as meaningful, oriented to, and taken up for their interactional value, we should stay with interactions, where uptakes provide analysts with "proof procedures" for meanings (Sacks, Schegloff, & Jefferson, 1974). The ape trainers, ELIZA's dialogical partners, and anthropomorphic human parents become constitutively right; the nonhumans took part and did the things they seemed to. Of course, we could refuse to go along with any of this and declare that there are no proper grounds for treating apes, babies, and machines in the ways we treat adult humans. However,

if we are to avoid circularity, we would have to find some other, noninteractional basis for that.

Collins (1990) argues that there is a genuine criterion of mere simulation, at least for artificial intelligence. The classic attributional definition of nonhuman competence in AI is the "Turing test" (Turing, 1950). Its criterion is the inability of a human interactant/interrogator to distinguish between the conversational output of a machine and a human being. In Turing's version, the machine and the person are situated behind a screen, and communication is via typescript. To demonstrate how any such test must ultimately fail, Collins suggests a parallel humans-only case, one that would be sure to unmask the imitator. The case is that of a British spy who is arrested and interrogated by the Russian State Security Committee, the KGB, having been trained in London to impersonate a native of Semipalatinsk. The KGB's task, like that of the interrogator in the Turing test, is to cross-examine and try to unmask him.

> The agent has learned the history and geography of Semipalatinsk from books, atlases, town guides, photographs, and long conversations with a defector who was once himself a native of the town. He has undergone long sessions of mock interrogation by this defector until he is word perfect in his responses to every question. His documents are in order. . . . The moment of crisis occurs for our hero when an interrogator enters who is himself a native of the town. . . . However good his training, we know that the spy will not survive cross-examination by a native of Semipalatinsk. (Collins, 1990, p. 6)

Collins invokes a form of "tacit" knowledge that can be acquired only from experience, and not from instruction, so that however much the spy's trainer had told him, it could never be fully programmed into him: "There is much more to this than can be explicitly described even in a lifetime . . . he could never know everything that the trainer knows, nor everything that the native interrogator knows" (p. 7). Sooner or later, the spy will be caught out, and so also must the programmed computer behind Turing's screen.[15]

There are two things that Collins's example underestimates. The first is an implication of his own remarks concerning how the personal nature of experience combines with the unending flexibility of description. It is not clear to what extent natives of Semipalatinsk are interchangeable with each other, so that even they, or their spy-catcher colleagues, could be sure of what exactly another native would absolutely have to know. Given the vagaries of description, there also ought to be as much uncertainty moving back from accounts to experiences as forwards; at what point would the spy's description be bound to fail?

The second problem is more severe. What one would like to know are the grounds on which the *true* native of Semipalatinsk got counted as such; presumably something more than that his papers were in order, that he claimed to come from there, knew a lot about the place, and so on. The production of some corroborating witnesses from Semipalatinsk, *real* natives, would only regress the problem. We can imagine the final scenes of the movie, when it turns out that the KGB interrogator was herself a British spy, a case of double agency, and another hall of mirrors. The trick here is to be the author of all this stuff (cf.

Edwards, Ashmore, & Potter, in press). Collins informs us that the successful interrogator really was what the spy could only contentiously claim to be, just as we are informed about what actually lies behind Turing's curtain. The Turing test requires that everything is obvious when we pull back the curtain and look; but how sophisticated an android are we allowed to find? We are back to an author's/creator's privilege, the usurpation by the author of the participants' problem of attribution. A more interesting issue is how people are able to make these judgments when nobody is writing their lines. Or could it be that writing the lines is the core phenomenon?

DESCRIPTION AND TRANSLATION: WHORF'S PARADOX

The issue of nonhuman agency hangs on how the actors/actants and their activities are described. Descriptions of animal, machine, indeed of ordinary human activities, are not simply the way things are but the way things are said and written. This is very easy to ignore, when we are writing texts and engaging in arguments, as we saw with the story about the native of Semipalatinsk. We ignore it when we accept the challenge of explaining Washoe's calling a duck a "water bird" (Fouts, 1974), when we accept that people mistook ELIZA for a sympathetic listener (Weizenbaum, 1976); or accept that a computer program counts as a piece of AI because it "fits the facts" of human psychology (Johnson-Laird, 1988, p. 26). It is not that these descriptions are wrong, but that they build into the ostensibly given world, the very thing that is at issue (cf. Woolgar, 1987). If we knew the facts of human psychology well enough so that they could stand as criterion for the output of computer simulations, then the claims of those simulations to be doing a superior sort of psychology would start to look circular.[16] The description of Washoe's behavior is already a translation of it. Lana is quoted as tapping onto the idiograms of her computer console the sentence "Please machine give fruit" (Rumbaugh, 1977); and Noam Chomsky cites a similar, ironic translation for the sequence of four colored stimuli that a pigeon, rewarded by food and subject to no claims about language, learned to tap with its beak (Chomsky, 1979; Straub, Seidenberg, Bever, & Terrace, 1979).[17]

Because the apes' performances are ones that use pictorial and gestural signs, these have to be rendered into English (or whatever written language) for the requirements of reporting and theorizing them as examples of language. Objections have been mounted not only to specific translations but to the general use of linguistic terms. Even for human children, the description of their early utterances as "language development" is a concession provided by the same anthropomorphic generosity that grants and anticipates their membership of the club of proper humans: "What goes before 'language' in development is only linguistic by courtesy of its continuity with a system which in fully elaborated form is indeed a language" (Brown, 1970, p. 37). According to this sort of

argument, a child and a chimpanzee might produce observationally identical kinds of actions, but only in the case of the child would we be looking at language. The ape researchers were unimpressed:

> Dispatches from the Gardners' front line suggested that linguists were having their cake and eating it, while denying ape researchers even a glance at the edibles. A child's variable word order argued its innate mastery of grammar; an ape's condemned it irretrievably. (Desmond, 1979, p. 44)

The claims on both sides are similar and are a classic form of scientific disputation. Each analyst claims that their opponent views some action, formulates a theoretical account of it, attributes that account to the actor/actant, and builds it into the action's description.

Let us call this "Whorf's paradox." One of the criticisms of Whorf's (1956) efforts to demonstrate the thesis of linguistic relativity was the fact that, to write papers on the subject and make his argument, he had to translate the various sentences of Hopi and Apache into peculiar-looking English and compare them to more ordinary-looking English versions. Eric Lenneberg (1953) and others pointed out the difficulties of doing this. Either Whorf's argument was vitiated by the success of his translations (so the meanings of one language can indeed be rendered in another), or else the translations were constitutive of the peculiarity attributed to Hopi and Apache, and what he should have done was to translate them colloquially, which would have ruined his argument. It is a problem for those who would define a nonlinguistic foundation for semantics (e.g., Lakoff, 1987; see Edwards, 1991), and for the ape researchers as well. When we are assessing the possibly agentive, intelligent, or linguistic productions of marginal objects, we are dealing with matters of translation.[18]

Desmond (1979) refers to "the distinct and distasteful possibility of cannibalising the living chimpanzee body for makeshift human counterparts" (p. 15); that is, organ transplants. The category "cannibalising" places humans and chimpanzees into the same species of creature. Similarly, when chimpanzees were reported to have complex relations with baboons, including living alongside, playing with, and hunting and eating them, Kortland (1975) refers to "a form of racism since, to chimpanzees, baboons are a sort of semi-conspecific" (p. 302). Jane Goodall's descriptions of wild chimpanzees are full of notions defined first for human practices such as murder, war, kidnapping, cannibalism, hunting, shyness, nurturing, sharing, and so on (Goodall, 1979, 1986). Desmond is more conscious of anthropomorphism when drawing distinctions between apes and humans, such as when questioning reports of Washoe's and Koko's "swearing," which seemed to him suspiciously modern looking: "The apes are too up to date. I wondered whether in fact we were not reading too much into their outbursts" (Desmond, 1979, p. 27). However, this is not a problem that can be cured by being more careful with particular descriptions and translations or by excluding specific cases of rendered ape language. It is intrinsic to the entire exercise.

Even when we are defining the apes' nonlinguistic actions (as Goodall does) or what they do as *non*human, we tend to deploy terms that describe alternative sorts of human actions and meanings:

> True, gorillas and chimps *do* readily learn the word "sorry." But again, we must shout *beware*: "sorry" signed by an ape . . . does not necessarily imply remorse, as does *sorry* said in English. . . . It transpires Nim signed "sorry" most frequently while under imminent threat of punishment . . . a ritual device to thwart an attack, not a symptom of deep remorse. . . . [But] I dare say that with a master-stroke of conditioning something *resembling it* might be attained. (Desmond, 1979, pp. 201-203, emphasis in original)

The trouble here is that saying *sorry* without remorse is not only an imitative shadow of what humans do but a perfectly recognizable human option, indeed a variety of options, redescribable in ways ranging from insincerity or deception through to the normal etiquette of ritualized "remedial interchanges" (Goffman, 1971). Given these possibilities, the attribution of a mental state such as remorse is transferred from its status as the analyst's theory of human meaning, to the kind of thing participants in interactions may perform for each other, apes and ape trainers included. Once analysts resist mapping talk onto putative internal states, even for human discourse (Coulter, 1990; Wittgenstein, 1953/1958), we are left with a domain of talk-in-interaction (Schegloff, 1989), where the legitimacy of taking anybody's individual utterances as diagnostic of mind is questionable. Therefore, this applies to apes, infants, and computers too, whether it is done by proponents or critics of their essential humanlike competence.

Anthropomorphism is arguably at its most effective when we are displaying caution, doing critique and objectivity. It is plain to see in descriptions such as the following: "impressionable adolescents or young adults setting out to make their mark on chimpanzee society," "these shy creatures," "as baboons became bolder," "political chicanery of community life" (Desmond, 1979, pp. 225-228). But it is present also here, even with the scare quotes:

> Melissa, partially crippled by polio and brutally beaten by Passion and Pam, extending her hand to Passion for "forgiveness" and reassurance—even as Passion eats her still-warm infant—is not an act we readily understand. (Desmond, 1979, p. 244, drawing on descriptions by Goodall)

It is, however, a description written out in words that we readily understand. Further, the effect of empathy or of alienation from the activities and minds of others is not unknown in descriptions of purely human conduct, where descriptions may either unite us all in a common sensibility (cf. Geertz, 1988, on the anthropology of Evans-Pritchard) or emphasize difference (Said, 1979, on "orientalism"). Even the details of Melissa's[19] alien behavior invite comparison with human experiences, such as the requirement for Victorian children to thank their fathers and schoolmasters for a beating, and descriptions of the feelings of concentration camp victims for their guards (Bettelheim, 1943).

Descriptions of what apes mean, whether language trained or wild, cannot avoid Whorf's paradox. To try to avoid it completely would be to avoid

translation altogether, running into a further conundrum: "If a lion could speak, we would not understand him" (Wittgenstein, 1953/1958, p. 223). The idea that humans and animals might mean different things, however, may be only a more severe, not a qualitatively different problem from that which besets humans themselves. According to Wittgenstein, we do not know, except by playing the same public language games, what *each other* means. But the fact that we are able to play those games provides no guarantee, as if one were needed, of sharing the inner life of the mind; it is, rather, what that sharing amounts to. So, although sharing a form of life with the lions is not altogether out of the question (cf. the movie, *Born Free*, based on the Adamsons's experience of doing just this), the various rich interactions between humans and apes might be considered ape-human forms of life, with their languages. The problem is that of representing those forms of life within the forms of another—human discourse, academic reports, and debates—where only one (human) partner plays both games and represents the first as a diagnostic test for the ape's possible membership of the second.

CONCLUSION

The arguments presented in this article, and encapsulated below, suggest the futility of trying to resolve the issue of nonhuman agency as a problem of ontology. Rather, it has to be understood as a matter of social, and particularly, descriptive practices. Further, these descriptive practices include not only those of participants in interactions but also those of the writers of research reports, and all constructors of lay and academic dichotomies and distinctions.

First, there is Whorf's paradox: the reflexive problems of description. Accounts of, and data on, human/nonhuman communications are themselves examples of discourse, of the thing under analysis: the capacity of nonhumans to render themselves in human language, or otherwise make themselves intelligible. These kinds of descriptions become available for analysis (for further description) like any other text, as forms of social action. For example, descriptions of nonhumans as agentive or nonagentive could figure in justifications for treating them as such (cf. Haraway, 1989, 1991, on the intimate relations between primate studies, eugenics, medical research, racism, and constructions of gender). Meanwhile, we must not forget that analysis of human-human interactions is founded on the same competence as participation (Heritage, 1984), and both are constituted in the "write-up" (Ashmore, 1989; Geertz, 1988).[20] Not only ape-human and machine-human interactions but the entire research process is a hall of mirrors.

Second, there is Malinowski's paradox (Geertz, 1988), the attribution/anthropomorphism quandary: the need both for rapport and for objectivity. This is not only a problem for how to do participant observation but a deeper issue of what kind of things participation and its description are. The ontology of us and them dissolves into text and description. Anthropomorphic attribution is as

plausible about the constitution of human relations and the formation of persons and their proper description, as it is a matter of truth and error. Its application to nonhumans is as much a matter of social practice, a participants' making-it-so procedure, as is its application to us. It marks "us" out.

Third, there is the futility of what we might call the "impossibility argument," that there is no way *in principle* that nonhumans can ever cross the divide. The impossibility argument sustains, by definition, an ontological distinction between what humans and nonhumans can do. Partly, the futility stems from the Whorf and Malinowski conundrums, where reality reduces to what counts as. What counts as can change. Although we might presently doubt the capacity of machines and animals to enter into human "physiognomic language games" (Harré, 1988), there is no reason to insist that humans and nonhumans (to retain that provisional distinction) *cannot* engage together in forms of life, or interaction, nor that some kind of language could not serve to mediate them.

The impossibility argument, for example, refuses to countenance or take seriously Haraway's category of "boundary objects," called "cyborgs" (Haraway, 1991). These are ways of describing us, not merely hoaxes or fictional inventions.

> A cyborg is a hybrid creature, composed of organism and machine . . . entities made of, first, ourselves and other organic creatures in our unchosen "high-technological" guise as information systems, texts, and ergonomically controlled laboring, desiring, and reproducing systems. The second essential ingredient in cyborgs is machines in their guise, also, as communications systems, texts, and self-acting, ergonomically designed apparatuses. (Haraway, 1991, p. 1)

Haraway's description decomposes the cyborg into organism and machine, but only to put them together again. The argument emphasizes their fusion, as with Latour's (1987) "actants," where the division of labor is finally obscured between subject and object, knower and known, agent and artifact.

The argument of this article is not one that ends with the primacy of attribution. That would merely shift the explanatory focus from object to attributer, with mind and agency displaced and reapportioned from one to the other. Rather, it is an argument for the further step of dissolving these categories (object and attributer) into those of interaction and description, where attribution and its attendant problems are participants' concerns (manifested in descriptions, accusations, claims, error accounts, etc.) just as much as imitation and agency are. We are not so much investigating the agentive nonhuman object, nor the human agent that attributes and produces it, but we are taking part in the discourse that constitutes and deconstructs them both.

NOTES

1. It has been noted that, for human out-groups as well as for apes, the most forceful distinctions are often drawn closest to home. With regard to apes, "man's closest and homeliest relative has in history been the most vilified of all creatures: the wretch of creation" (Desmond, 1979, p. 18).

2. This is at least my fourth use of this pronoun ("us") so far, but here it is provisional and ironic. It is, in a sense, the point at issue: "The pronouns embedded in sentences about contestations for what may count as nature are themselves political tools, expressing hopes, fears, and contradictory histories. Grammar is politics by other means" (Haraway, 1991, p. 3). Note, for example, Sherry Turkle's use of "themselves" in the quotation that follows.

3. See, for examples, Dreyfus (1979), Hofstadter (1979), Searle (1980), Dennett (1971).

4. "That nightmare about the crushing failure of the project of man" (Haraway, 1989, p. 31).

5. Various examples have been offered of nonimitative, creative language production by chimpanzees and gorillas, including Washoe's *waterbird* for a duck, and *rock berry* for a Brazil nut (Fouts, 1974); Lana's *finger bracelet* for a ring (Rumbaugh, Gill, von Glasersfeld, Warner, & Pisani, 1975); Lucy's *cry hurt food* for radishes (Temerlin, 1975); and Koko's *white tiger* for a zebra (Patterson & Linden, 1981). Rejections of such evidence are based on their untypicality, selective reporting, and indeterminacy of situational reference (Desmond, 1979). I am reminded of the *iron horse* of the Wild West movies, the kind of thing the "Indians" would say, strangely but inventively, although in this case an emblem of difference rather than sameness to "us."

6. Brown's (1973) version of this "rote learning" objection is delightfully mechanical, in the context of nonhuman agency in general, and a change from his usually highly readable prose: "possible dependence of terminal accomplishments on specific atomic preliminary programs" (p. 48). For examples of what the criticism may apply to, see Premack and Premack (1972) and Rumbaugh (1977).

7. Imitation is also contrasted with originality, especially in a literary context. As Weinsheimer (1984) shows, in a wise and deeply reflexive work, literary originality, as a prime authorial norm, is quite recent; imitation of the masters was once the goal that was strived for, and could be again, if Weinsheimer's highly original *Imitation* is heeded.

8. A fruitful place to look for other uses and ironies of imitation might be in European and American comparative representations of Japanese education and manufacturing since the 1950s, and Japan's replacement in that role by the more recent economic emergence of Taiwan and Malaysia. Whereas Far Eastern businesses have "copied" or "cloned" the West's products, the reverse process seems more a rational adoption of efficient business practices and global new technologies. It is not insincerity or the production of counterfeit; on the contrary, it is the sincerest form of flattery.

9. It is claimed even for the chimpanzee Nim, whose imitative behavior became the basis for rejecting his and other apes' spontaneous grammaticality, that he displayed a humanlike orientation to intersubjectivity: "He was alive to new discoveries, and just had to share them; he would drag his flatmate [Laura Petitto] outside to point out planes or birds" (Desmond, 1979, p. 231). Again, this is in the first place a description, a motive-attributing narrative account in which the activity and its interpretation are conjoined. Thorpe (1972) cites evidence of "prevarication" in birds, such as feigning a damaged wing to lure a predator away from its eggs. How far such activities are intentional or automatic begs the question, of course:

> When Herb Terrace concluded that Nim was just imitating his trainer, he was implying that it's a very low-level ability. Experiments like these [by Sue Savage-Rumbaugh with "Kanzi"] have shown that imitation, understanding the goal and achieving it by the same method as the demonstrator, is a very sophisticated skill. (Jones, 1993, p. 22)

Terrace's study was itself sophisticatedly imitative (or not)—it was a replication of earlier studies.

10. Latour (1988) offers the distinction "figurative/nonfigurative" (p. 303), the degree of figuration being the extent to which artifacts or machines are made to directly resemble, personify, or embody human characteristics, rather than merely substituting for human actions as the door closer does. Therefore, the ape language research, the artificial intelligence (AI) program, and the way we bring up (socialize) our children are efforts at practical anthropomorphism, at increasing their figurative qualities as recognizable, accountable human actors.

11. This was considered a product neither of the particular gullibility of the persons concerned, nor of the program's lack of sophistication. Weizenbaum (1976) noted the "enormously exaggerated

attributions an even well-educated audience is capable of making, even strives to make, to a technology it does not understand" (p. 7). He also observed the same process at work for more sophisticated AI products: "The subsequent, much more elegant, and surely more important work of Winograd . . . is currently being misinterpreted just as ELIZA was" (p. 7).

12. This description of Nim can be contrasted with other ways of discounting Terrace's claims, noted above, in which Nim's imitations are described as highly sophisticated sorts of actions.

13. The classic debunking was performed a century ago by Pfungst (1965) on the horse Clever Hans. Apparently, the horse was able to perform feats of mental arithmetic by tapping out answers with his hoof. Pfungst showed that Hans was unable to do this when the questioner did not know the answer. Otherwise, he was skeptical or inattentive and reasoned that gestural cues were inadvertently being given.

14. Haraway (1989) cites several examples, including this from Mary Jobe Akeley, where attribution, empathy, and the descriptive classes of savages and animals are rolled together to bring off a piece of well-meant membership exclusion from the class of proper persons:

> Now with few exceptions our Kivu savages, lower in the scale of intelligence than any others I had seen in Equatorial Africa, proved kindly men . . . how deeply their sympathy affected me! As I think of them, I am reminded of the only playmate and companion of my early childhood, a collie dog. (p. 52)

15. Collins's argument parallels Harré's (1988) insistence on the importance for human language and intelligence of possessing a human body, on which human "physiognomic language games" are predicated (cf. Dreyfus, 1979; Hintikka & Hintikka, 1986; Lakoff, 1987). Weizenbaum (1976) makes a similar case, adding that "however much intelligence computers may attain, now or in the future, theirs must always be an intelligence alien to genuine human problems and concerns" (p. 213).

16. Compare Churchland's (1988) description of AI, that "generally speaking, the system proposed must do what the creature at issue succeeds in doing, or what its selected faculty does" (p. 93). But how are those criteria and successes defined, except by noncomputational criteria? AI would seem dependent on the same common sense or conventional psychological accounts it often seeks to replace.

17. In another retranslation study, Lenneberg (1975) trained a couple of schoolchildren to use the same sorts of magnetized plastic symbols for things by which "Sarah" (Premack & Premack, 1972) demonstrated impressive grammatical prowess. The children performed even better than Sarah; they were, however, unable to translate the "sentences" into English, having formed the impression that it was a kind of puzzle game rather than a kind of language.

18. Whorf's paradox enters the description not only of linguistic and gestural performances but of their domains of reference and usage. With children, for example, an argument for the prelinguistic structuring of infant activity and interaction, as a foundation for language, is much aided by the analyst's description of those nonlinguistic matters in terms derived from the analysis of language and speech acts (Bruner, 1983; Edwards, 1973). Similarly, one of the tropes for animal language research is the idea of bringing up an ape just like a child, as far as possible. This is an issue of substance, inserted as method. If we could, precisely (and this is where the "as far as possible" bites), bring up an ape as we bring up a child and make the ape's interactional environment identical to that of a child, then ape and child would be interactionally indistinguishable. The ape has virtually passed the Turing test in the report's Method section, before we get to see the results.

19. The naming of apes is a feature of their transformation into protohumans, just as it is for human infants. For the apes observed in the wild, these remain observers' names. Washoe and other trained apes learned them as labels for themselves.

20. In Geertz's (1988) view, Malinowski's legacy to ethnography is "not, as so often thought, a research method, 'Participant Observation' (that turns out to be a wish not a method), but a literary dilemma, 'Participant Description' " (p. 83). It is the schism in Malinowski, between "High Romance and High Science, seizing immediacy with the zeal of a poet and abstracting from it with the zeal of an anatomist, uneasily yoked" (p. 79). For Geertz, *both* High Science and High Romance

are subordinated to writing, which, rather than coming afterwards as a form of reporting, is primary and constitutes the sense of the others. The coming afterwards is itself a representation, which constructs those senses, both of reporting the intimacies of experience and of doing disengaged science.

REFERENCES

Ashmore, M. (1989). *The reflexive thesis: Wrighting sociology of scientific knowledge*. Chicago: University of Chicago Press.

Bettelheim, B. (1943). Individual and mass behavior in extreme situations. *Journal of Abnormal and Social Psychology, 38*, 417-452.

Brown, R. W. (1970). *Psycholinguistics: Selected papers by Roger Brown*. New York: Free Press.

Brown, R. W. (1973). *A first language: The early stages*. London: Allen & Unwin.

Bruner, J. S. (1983). *Child's talk*. Oxford: Oxford University Press.

Chomsky, N. A. (1959). Review of "Verbal Behavior" by B. F. Skinner. *Language, 35*, 26-58.

Chomsky, N. A. (1979). Human language and other semiotic systems. *Semiotica, 25*, 31-44.

Churchland, P. M. (1988). *Matter and consciousness: A contemporary introduction to the philosophy of mind* (rev. ed.). Cambridge, MA: MIT Press.

Clarke, A. C. (1968). *2001: A space odyssey*. New York: NAL/Dutton.

Collins, H. M. (1990). *Artificial experts: Social knowledge and intelligent machines*. Cambridge, MA: MIT Press.

Coulter, J. (1990). *Mind in action*. Oxford: Polity.

Dennett, D. (1971). Intentional systems. *Journal of Philosophy, 68*, 87-106.

Derrida, J. (1977). Signature event context. *Glyph, 1*, 172-197.

Desmond, A. (1979). *The ape's reflexion*. London: Blond & Briggs.

Dreyfus, H. (1979). *What computers can't do: The limits of artificial intelligence*. New York: Harper & Row.

Edwards, D. (1973). Sensory-motor intelligence and semantic relations in early child grammar. *Cognition, 2*(4), 395-434.

Edwards, D. (1991). Categories are for talking: On the cognitive and discursive bases of categorization. *Theory and Psychology, 1*(4), 515-542.

Edwards, D., Ashmore, M., & Potter, J. (in press). Death and furniture: The rhetoric, politics and theology of bottom line arguments against relativism. *History of the Human Sciences*.

Edwards, D., & Potter, J. (1992). *Discursive psychology*. London: Sage.

Fouts, R. S. (1974). Language: Origins, definitions, and chimpanzees. *Journal of Human Evolution, 3*, 475-482.

Gardner, B. T., & Gardner, R. A. (1974). Comparing the early utterances of child and chimpanzee. In A. Pick (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 8, pp. 3-23). Minneapolis: University of Minnesota Press.

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science, 165*, 664-672.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Geertz, C. (1988). *Works and lives: The anthropologist as author*. Oxford: Polity.

Goffman, E. (1971). *Relations in public: Microstudies of the public order*. Harmondsworth: Penguin Books.

Goodall, J. (1971). *In the shadow of man*. Boston: Houghton Mifflin.

Goodall, J. (1979). Life and death at Gombe. *National Geographic, 155*(5), 592-621.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.

Griffin, D. R. (1976). *The question of animal awareness: Evolutionary continuity of mental experiences*. New York: Rockefeller.

Haraway, D. (1989). *Primate visions: Gender, race and nature in the world of modern science.* London: Routledge & Kegan Paul.

Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature.* London: Free Association Books.

Harré, R. (1988). Wittgenstein and artificial intelligence. *Philosophical Psychology, 1*(1), 105-115.

Heritage, J. (1984). *Garfinkel and ethnomethodology.* Cambridge: Polity.

Hintikka, M.B.C., & Hintikka, J. (1986). *Investigating Wittgenstein.* Oxford: Blackwell.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid.* New York: Basic Books.

Johnson-Laird, P. N. (1988). *The computer and the mind: An introduction to cognitive science.* London: Fontana.

Jones, J. (1993). *Chimp talk: Text adapted from the programme transmitted 21st June 1993.* London: BBC/BSS Publications.

Kortland, A. (1975). Discussion. In R. H. Tuttle (Ed.), *Socioecology and psychology of primates* (pp. 298-304). The Hague: Mouton.

Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Latour, B. (1987). *Science in action.* Milton Keynes, England: Open University Press.

Latour, B. (writing as Johnson, J.) (1988). Mixing humans and non-humans together: The sociology of a door-closer. *Social Problems, 35,* 298-310.

Lenneberg, E. (1953). Cognition in ethnolinguistics. *Language, 29,* 463-471.

Lenneberg, E. (1975). A neuropsychological comparison between man, chimpanzee and monkey. *Neuropsychologia, 13,* 125.

Leslie, A. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review, 94,* 412-426.

Lock, A. J. (Ed.). (1978). *Action, gesture and symbol: The emergence of language.* New York: Academic Press.

Patterson, F. (1978). Conversations with a gorilla. *National Geographic, 154,* 438-465.

Patterson, F., & Linden, E. (1981). *The education of Koko.* New York: Holt, Rinehart & Winston.

Pfungst, O. (1965). *Clever Hans (the horse of Mr. von Osten)* (R. Rosenthal, Ed.). New York: Holt, Rinehart & Winston.

Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour.* London: Sage.

Premack, A. J., & Premack, D. (1972). Teaching language to an ape. *Scientific American, 227,* 92-99.

Randi, J. (1975). *The magic of Uri Geller.* New York: Ballantine.

Rumbaugh, D. M. (Ed.). (1977). *Language learning by a chimpanzee: The Lana project.* New York: Academic Press.

Rumbaugh, D. M., Gill, T., von Glasersfeld, E., Warner, H., & Pisani, P. (1975). Conversations with a chimpanzee in a computer-controlled environment. *Biological Psychiatry, 10,* 627-641.

Sacks, H. (1992). *Harvey Sacks: Lectures on conversation* (Vols. 1 & 2, G. Jefferson, Ed.). Oxford: Blackwell.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Said, E. W. (1979). *Orientalism.* New York: Random House.

Schegloff, E. A. (1989). Harvey Sacks—Lectures 1964-1965: An introduction/memoir. *Human Studies, 12,* 185-209.

Searle, J. R. (1980). Minds, brains and programs. *The Behavioral and Brain Sciences, 3,* 417-424.

Sebeok, T. A., & Umiker-Sebeok, J. (Eds.). (1980). *Speaking of apes.* New York: Plenum.

SETI—the search for extra-terrestrial intelligence. (1990, July 29). *Observer.*

Sharrock, D. (1993, June 1). Corky is cock-a-hoop as he finds new life in wide open spaces. *The Guardian.*

Straub, R. O., Seidenberg, M. S., Bever, T. G., & Terrace, H. S. (1979). Serial learning in the pigeon. *Journal of the Experimental Analysis of Behavior, 32,* 137-148.

Strum, S. (1987). *Almost human: A journey into the world of baboons.* New York: Random House.

Temerlin, M. K. (1975). *Lucy: Growing up human: A chimpanzee daughter in a psychotherapist's family*. Palo Alto, CA: Science & Behavior.

Terrace, H. S. (1979). *Nim, a chimpanzee who learned sign language*. New York: Alfred A. Knopf.

Thorpe, W. H. (1972). The comparison of vocal communication in animals and man. In R. A. Hinde (Ed.), *Non-verbal communication* (pp. 27-47). Cambridge: Cambridge University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*, 433-460.

Turkle, S. (1984). *The second self: Computers and the human spirit*. New York: Simon & Schuster.

Vygotsky, L. S. (1987). *Thought and language* (A. Kozulin, Ed.). Cambridge, MA: MIT Press.

Weinsheimer, J. (1984). *Imitation*. London: Routledge & Kegan Paul.

Weizenbaum, J. (1976). *Computer power and human reason*. New York: Freeman.

Whiten, A. (1989, April). *Pretence and mindreading: Steps in human ontogeny and phylogeny*. Paper presented at the symposium on "The Emergence of Mindreading: Evolution, Development and Simulation of Second Order Representations," University of St. Andrews, Scotland.

Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). Cambridge, MA: MIT Press.

Wittgenstein, L. (1958). *Philosophical investigations* (2nd ed.; G.E.M. Anscombe, Trans.). Oxford: Blackwell. (Original work published 1953)

Woolgar, S. (1987). Reconstructing man and machine: A note on sociological critiques of cognitivism. In W. E. Bijker, T. Hughes, & T. Pinch (Eds.), *The social construction of technological systems* (pp. 311-328). Cambridge, MA: MIT Press.

Yerkes, R. M. (1925). *Almost human*. New York: Century.