capacity to match across modality (Davenport and Rogers, 1970), so the match from tool to function is a reasonable possibility. This does not detract from the clever experiments of SR&B and of P&W, but I would have preferred more context from the rich literature on animal cognition.

*Griffin* presents no new data and omits much of the progress in animal cognition over the past twenty years. Also, he plays the intellectual game of shooting down a theoretical position which is already dead and inventing an imaginary field. This generalization obviously omits his own brilliant work on navigation in birds and echo-location in bats, and his more complete earlier attempt to raise the question of animal awareness (1976 *op. cit.* G, SR&B).

Behaviorism was never as monolithic as G's straw man. It lacked adequate technology, data, and theory to justify expending much time or effort in contemplating nonobservables when there were so many interesting things to observe, measure, and control. However, as technology advanced and data were collected, many theoretical structures associated with behaviorism had to be discarded. This resulted in a neobehaviorism with many ideas concerning animal cognition. Much of this work on animal cognition was done with the behavioristic penchant for careful control and was even conducted in alley mazes (Gleitman and Steinman, 1963) and Skinner boxes (Gleitman and Bernheim, 1963). Space does not permit full development of this argument, but a few examples may suffice to illustrate that comparative psychologists have been studying cognitive problems during the past twenty years and are much more in need of a theoretical structure than a fictional field such as cognitive ethology or sociobiology.

To illustrate that my own contention is not just an attack on another straw man, the following are areas in which psychologists have recently studied cognitive experiences by animals: (1) cross-modal transfer; (2) surgical and optical sensory recombination; (3) several hundred papers on memory and its development, including reinstatement, rehearsal, and the use of surprising cues; (4) transposition; (5) conditional reaction problems; (6) perceptual constancies; and (7) the perception of barriers and detours. Probably I should also have included Harlow's (1949) formation of learning sets, which bridges the gap between incremental learning and insight. This is a considerable commitment to a comparative and evolutionary analysis of cognition. We need to organize animal cognition into a coherent theoretical structure using the presently abundant data.

### REFERENCES


Darby, C. L., and Riopelle, A. J. Observational learning in the rhesus monkey. *Journal of Comparative and Physiological Psychology, 52,* 94–98, 1959.

Davenport, R. K., and Rogers, C. M. Intermodal equivalence of stimuli in apes. *Science, 168,* 279–280, 1970.

Gleitman, H., and Bernheim, J. W. Retention of fixed-interval performance in rats. *Journal of Comparative and Physiological Psychology, 56,* 839–841, 1963.

Gleitman, H., and Steinman, F. Retention of runway performance as a function of proactive interference. *Journal of Comparative and Physiological Psychology, 56,* 834–838, 1963.

Harlow, H. F. The formation of learning sets. *Psychological Review, 56,* 51–65, 1949.

Hebb, D. O., and Thompson, W. R. The social significance of animal studies. In G. Lindzey, (ed.), *Handbook of Social Psychology,* vol. 2, Reading, Mass., Addison-Wesley, 1954.

Michotte, A. *La perception de la causalité.* Louvain, Institut Supérieur de Philosophie, 1946.



*by* **Marian Dawkins**

***Animal Behaviour Research Group, Department of Zoology, University of Oxford, Oxford, England***


***The second time around.*** The behavioral sciences appear to be coming full circle. There was a time, many years ago, before the dominance of behaviorism, when questions about animal minds and emotions were considered to be quite legitimate, and indeed they are still seen as such by people not schooled in the behaviorist tradition. Then, in the early part of this century, the methodological self-discipline of behaviorism took over, and most ethologists and psychologists even up to the present day have fought shy of considering questions of consciousness and mental awareness on the grounds that they are not accessible to scientific investigation and should be left severely alone. Now times are again changing. Subjective phenomena are once more being seen as legitimate areas for scientific study. But whereas this appears to be simply a reversion to an earlier view, a counter-revolution to bring us back to a more enlightened time, it is in reality far more than this. The second time around for scientific consideration of subjective phenomena in animals is characterized by a rigor, a determination to consider alternatives, and above all by attempts to make predictions and test hypotheses, all of which was previously lacking.

*Griffin* sets the stage by arguing forcefully for a cognitive ethology and attempting to meet head on some of the objections to it. For example, it is often argued by skeptics that postulating subjective experiences in animals does not lead to any specific predictions at all – it would not seem to make any difference to an animal's behavior whether or not it was conscious. In other words, the existence of subjective experiences is something for which there can be no rigorous test. G answers this objection in two ways. Firstly, he argues that open-minded agnosticism is more appropriate (and indeed more scientific) than flat denial, and that while it is true that we have no rigorous proof of the existence of subjective experiences in animals, neither are we able to claim rigorous proof in many other areas of biology (as is the case with the far-reaching and productive theories of sociobiology). G's second argument, which follows from this, is really the more exciting one, namely, that it *is* possible to formulate testable hypotheses about animal subjective awareness. The two papers on chimpanzees by P&W and by SR&B, which accompany G's in this issue, illustrate this point.

*Premack & Woodruff* set out to investigate something that appears at first sight to be quite out of bounds as far as respectable science is concerned, namely, whether chimpanzees ascribe mental states to other animals, such as human beings. The importance of this paper seems to lie not so much in whether it actually shows that chimpanzees do have a concept of mind as in the method of approach. There seem to be some objections to the evidence that is actually presented, but in thinking about these objections, one is forced to the conclusion that it would still be possible to devise an experiment which would be convincing. This is a far cry from saying that the phenomenon is one which is in principle *not open* to investigation at all.

In fairness to the authors I should be more specific. In the first experiment, the matching between the actor's stance in the video sequence and his stance in the "correct solution" photograph is acknowledged by the authors as being a confounding factor and considerably weakens the conclusion. This result is so crucial to the idea that the chimpanzees are really "assuming that the human actor wants the banana," as opposed to just completing a sequence, that it is a pity that it suffers from this objection. The later experiments with inanimate objects which are said to rule out physical matching suffer from the fact that although familiar sights cannot easily explain the results, the effects of familiar *sequences* do not seem to have been ruled out as completely. In any case, it is unfortunate to have to use a later experiment to justify the conclusions of an earlier one without simultaneous controls. But, as I said before, this is not the important point. The fact that controls could have been introduced and means could have been devised of more completely eliminating simpler explanations does not detract from the conceptual leap forward which the authors have taken in daring to ask their questions.

*Savage-Rumbaugh et al.'s* very valid criticisms of earlier experiments on chimpanzee communication (for example that there were often inadequate controls for experimenter bias) actually reinforce the point that it is in principle possible to devise rigorous experiments. Picking small holes in existing evidence is in fact a means of developing new and more watertight experimental paradigms, rather in the way that experiments on pigeon homing have been improved by the realization that more controls have to be introduced than was at first thought.

The second time around for the scientific assessment of animal consciousness is, as illustrated by the three papers in this special issue, characterized by adventurous questions and by the realization that it is possible to make determined attempts to answer them. It is not a time when "anything goes" and one interpretation is as good as any other. As G points out, the prospects for a cognitive ethology are much better now than they were in the days of Darwin and Romanes. At long last we can look forward to a real understanding of subjective awareness in animals, one of the profoundest mysteries in the whole of biology on this, the second time around.


*by* **Daniel C. Dennett***

***Department of Philosophy Tufts University, Medford, Mass. 02155***


***Beliefs about beliefs*** *[P&W, SR&B].* Because of its intrinsic interest – indeed its fascination – it is easy to lose track of the point of this kind of research. Getting a chimpanzee to talk takes on the aspect of sending a man to the moon. Suppose you succeeded. Then what? Presumably behaviorists would have to claim to be unimpressed, as unimpressed as they are by the verbal abilities of – themselves, for instance. So suppose we grant for the sake of a superannuated argument that *in principle* a suitably complex version of behaviorism can "handle" all ape behavior (and all human behavior too). That version of behaviorism will of course be scarcely distinguishable from mere mechanistic materialism – with micro-events in the brain being viewed as responses, for instance – and there is scant reason to oppose that creed, at least at this stage of our knowledge. The issue that remains is, on a first pass, how fancy a cognitive structure is required *in practice* to predict a chimpanzee's behavior. That is, granting that in practice it is desirable to intentionalize our account of chimpanzees (by attributing beliefs and desires, or belieflike states and desirelike states, Dennett, 1971, 1976), which beliefs and desires will it be useful, predictive, illuminating to attribute? In the present instance, will we find it valuable to attribute second-order beliefs and desires – beliefs and desires about the beliefs and desires of others? If so, then chimpanzees have a theory of mind in the requisite sense, for they use the concepts of belief and desire (or concepts importantly analogous) in their own action governance. If they turn out to have humanlike theories of mind, they will have use of even higher-order intentional attributions; they perhaps believe someone wants them to believe something, or want someone to believe they want something, and so forth. But how can these suppositions be put to the test?

I think the issue is analogous to the current controversy about mental images. What the growing literature on mental images shows is that whatever it is to which we may in the end "reduce" mental-image talk, there can be no doubt that there is *a* level of description of the phenomena at which imagistic characterizations are perspicuous because they are richly predictive of a surprisingly wide variety of behavioral effects. Talking of mental images may be a *façon de parler*, but it is no "mere" *façon de parler*, because taking the talk (quite) literally keeps on leading to confirmed predictions. This is undeniable even if it is also true that talking about mental images is itself in dire need of explanation – and even of ultimate elimination if one supposes that mental images cannot be taken *dead* literally as 3-D pictures in the brain.

What must be shown by *Premack & Woodruff,* analogously, is that imputing a theory of mind to chimpanzees (whatever that comes to literally, in the end) is richly predictive. As P&W note, any single test, however consonant its results with the theory-of-mind hypothesis, can be given a deflationary redescription by associationists *et al.* What one wants is a panoply of results elegantly predicted by the theory-of-mind hypothesis and only predictable with the aid of ad hoc provisions by its competitors. P&W do not yet have these results, as they grant, but while the experiments they are now undertaking would favor their hypothesis if the results were positive, they seem somehow slightly off target. P&W are searching for evidence that chimpanzees have expectations of the behavior of others that are better explained by supposing that they are (tantamount to) predictions derived from the chimpanzee's beliefs about the beliefs and desires of those others than from supposing that they are derived from either habits (of thought) or beliefs about other features of the world *(e.g.,* experienced regularities in the behavior of others). But the very training required to bring an animal into P&W's test situations seems to provide the relevant experience for engendering such alternate habits or beliefs. P&W are aware of this, and much of the complexity of the tests they have designed is dictated by their desire to make this alternative hypothesis less plausible. But in becoming so devious, the tests seem – to me – to sacrifice the most interesting hypothesis: it would be much more exciting to discover that chimpanzees normally have (naturally acquire in their lives) a theory of mind than to discover that chimpanzees can have a theory of mind dinned into them eventually. Bears can ride bicycles – a surprising fact of elusive theoretical interest. But when one tries (as I have, now, for several days) to dream up better experiments for P&W to run, one begins to appreciate that it is very hard to think up direct, natural, plausible tests. Why should this be?

Very young children watching a Punch and Judy show squeal in anticipatory delight as Punch prepares to throw the box over the cliff. Why? Because *they know Punch thinks Judy is still in the box.* They know better; they saw Judy escape while Punch's back was turned. We take the children's excitement as overwhelmingly good evidence that they understand the situation – they understand that Punch is acting on a mistaken belief (although they are not sophisticated enough to put it that way). Would chimpanzees exhibit similar excitement if presented with a similar bit of play acting (in a drama that spoke directly to their "interests")? I do not know, and think it would be worth finding out, for if they didn't react, the hypothesis that they impute beliefs and desires to others would be dealt a severe blow, even if all the P&W tests turn out positively, just because it can be made so obvious – obvious enough for four-year-old children – that Punch believes (falsely) that Judy is in the box.

But suppose we are uncertain how to interpret the children's glee; how can we go about strengthening the hypothesis that they believe Punch believes . . . ? We can ask them questions, particularly "why questions," but others as well ("What do you think Punch would have done if . . . ?"). But are there nonverbal tests we can also employ? It is hard to think of any that would be decisive that wouldn't be too difficult for the children. This is because of the complexity of the "thought processes" one has to impute to any person or animal who acts on the basis of such a prediction from a theory of mind. So far as I can see, the *minimally* complex pattern has the following format:

1 C believes that E believes that p.

2 C believes that E desires that q.

3 C infers from his beliefs in (1) and (2) that E will therefore do x, and so, anticipating E's doing x,

4 C does y because

5 C believes that if E does x, then unless C does y, C won't get something C wants, or will get something C wants to avoid.

(This is the minimally complex pattern for doing something because you believe someone believes . . . ; doing something in order to get someone to believe something in order to get him to do something . . . has a different but equally complex scenario.)

The ideal experiment to establish the use of such an explanatory format will have the following features:

a. E's anticipated action x will be a (relatively) novel action, or at least an action that (arguably) could not be anticipated by C under the circumstances simply by virtue of being habitual for E or oft-repeated in just these circumstances. (An elegant way of accomplishing this is to ensure that the belief attributed to E in (1) is false *(cf.* Punch and Judy), for then E will be expected to act inappropriately to the circumstances, and hence, in all likelihood, not the way E has typically acted in the past.)

b. C's action y will also be an action as much as possible from C's natural repertoire, rather than a highly trained artificial response, for again, arduous training procedures almost inevitably provide grist for the associationist's mill (Dennett, 1976).

c. The perceived (by C) dependence of y on x should also be natural and obvious, so that C's belief in it (5) can be attributed to C on the basis of C's straightforward observation of a relatively novel circumstance, rather than on the basis of extensive training.

Trying to design experiments to meet these conditions soon reveals the difficulty. Conditions (1-3) are relatively easy to meet – one would think. For instance, suppose there is a key that E, the experimenter, uses to open the banana locker. One day two boxes, one red and one green, are placed in the scene, and C sees E put the key in the red box and leave the scene. Then C sees Sneaky Pete come in and move the key to the green box. When E returns to feed C, *ex hypothesi* C believes that E believes that the key is still in the red box, and hence C expects E to go to the red box (since he believes that E wants to get the key). But now, how can things be rigged so there is something C might see to do that is appropriate to C's expectation (meeting conditions (4) and (5))? P&W's solution at this step is to train C to perform a sort of proto-speech-act, a *prediction* by choice of photograph, with the assumption that, for predictions, truth is its own reward (thus satisfying C's desire in (5)). But this is gimmicky. One would prefer to have C's action y interact more meaningfully with E's action x, but reflection reveals that this is hard to set up without resorting to another sort of gadgetry: artificial dependencies created between x-type actions and y-type actions that C might be *trained* to recognize.

The conclusion that seems borne in on one is that unless there is a great deal of normal interaction – either competitive or cooperative – between C and E, there is just no way for C to come to perceive his own actions as meshing with E's in the tight way required at step (5). One can rig it up – *e.g.,* C could be taught that he will get a shock if E opens the red box unless C, anticipating this, moves to a particular location – but this requires training that removes the desired novelty listed in (a). This objection to gimmickry is not just aesthetic, of course; the more artificial the test circumstances, the more restricted the range of predictions available to the theory-of-mind hypothesis, and as noted at the outset, predictive fecundity is of the essence in this investigation.

It appears that except in tricky environments that require extensive training to produce familiarity, the only act-types that naturally meet the conditions are *communicative* acts, such as C's warning E, or requesting something from E, or asking E a question; and so the problem of the training factor now pertains to the training up of communicative act-types. In this regard *Savage-Rumbaugh et al.'s* format with Austin and Sherman looks much more straightforward and promising *if* the communicative mode of interaction between Austin and Sherman can be extended to relatively novel situations without (much) additional training. But still the conclusion that would follow success in such experiments would be at best that chimpanzees can be put in complex artificial environments (artificial for chimpanzees, not for people) in which they eventually develop a theory of mind. In their natural environments there seems to be no clear need for them to develop a theory of mind about each other, and hence no compelling reason to impute it to them. But perhaps further ingenious experiments will find a way of meeting the desiderata listed and make a believer out of me.

REFERENCES

Dennett, D., Intentional systems. *Journal of Philosophy.* 68:87-106. 1971.
Conditions of personhood. In: A. Rorty (ed.), *The Identities of persons.* Berkeley: U. Cal. Press, 1976.

EDITORIAL NOTE

°Received too late for a Response from P&W or SR&B. See Continuing Commentary. [Ed.]

***by* William Orr Dingwall**

*Linguistics Program, University of Maryland, College Park, Md. 20742*

***Animals and the rest of us: Descartes versus Darwin.*** Topics as complex, ill-defined, and open-ended as those to which this special issue of *BBS* is devoted cannot possibly be resolved, if at all, within the compass of a few thousand words. One can at best select those aspects of the problems that have personal appeal, relying on differences of interest as well as opinion to provide the needed breadth of coverage. Since Griffin, in his contribution and earlier book (1976 *op. cit.* G, SR&B), elaborates the kinds of questions that are partially addressed by the experiments in the other two contributions, it seems logical to begin with a discussion of his views.

*Are the meadows of cognitive ethology really greener [G]?* Since the views of Darwin, a firm believer in continuity, are often cited as forming one of the major pillars of modern ethology (cf. Eibl-Eibesfeldt, 1975; see also Eibl-Eibesfeldt: "Human Ethology" *BBS* 2(1) 1979), it seems a bit odd that G feels the need to sound a battle cry against Cartesian discontinuity – a doctrine that would appear to have been laid to rest in an earlier age (at least as far as ethology is concerned). Indeed, Darwin's brilliant defender, T. H. Huxley, discusses this issue at length using arguments from neurological continuity very similar to G's to affirm continuity of consciousness (1874). To imagine a genetic saltation suddenly resulting in a phenomenon as complex as consciousness (or human communication for that matter) appeared as nonsensical to Huxley as it does to the British philosopher, Mary Midgley, who, in a recent work, points out that it requires us to assume "a quite advanced point in animal evolution when parents who were merely unconscious objects suddenly had a child which was a fully conscious subject (1978, p. 217)." Indeed, it appears that ethologists do not deny the continuity of subjective states per se, only that one can say anything specific about them (cf. Eibl-Eibesfeldt, 1975). Thus, continuity does not seem to be the question; rather it is whether such vague terms as *cognition* and *consciousness* can be defined with sufficient explicitness to allow for the formulation of testable hypotheses.

What one will find, I believe, is that such terms do not designate entities that an organism either possesses in their entirety or not at all, but rather a mosaic of structures, skills, and knowledge that, like human communication, does not develop in children as a whole, does not disappear in cases of pathology as a whole, and undoubtedly did not evolve as a whole. But what is consciousness? Is it to be contrasted with unconsciousness or subconsciousness? Is it the same as thinking, self-awareness, ability to learn, volition, perception of relationships, a private world of mind? Each of these definientia yields different results in terms of the continuity question. If self-awareness is chosen as a distinguishing characteristic, perhaps only hominoids can be said to be conscious; if learning is chosen, then even one-celled organisms would appear to evince consciousness. It is not clear to me that terms like *cognition* and *consciousness* even define a coherent network of functions in the way I assume a term such as *communicative behavior* does (cf. Dingwall, in press). To postulate a neurophysiological basis for consciousness in terms of neuronal cell number and degree of connectivity as Rose (1976) has recently done is not based on any firm neurobiological evidence [see also Puccetti & Dykes: "Sensory Cortex and the Mind-Brain Problem" *BBS* 1(3) 1978] and implies a *scala naturae* of consciousness with the primates at its apex, which is already known to be unsupported in the case of some indices of learning (Hodos, 1970).

At this point, I'm inclined to agree with the position of ethologists such as Eibl-Eibesfeldt concerning the vagueness of these terms. It is not the case, one should hasten to point out, that ethologists have shied away altogether from postulating theoretical constructs, often in terms of hypothetical neural structures (e.g., *motor coordination center, releasing mechanism,* etc.); indeed, as Konishi (1971) has pointed out, many of these constructs have been shown to have neurophysiological bases. It is the program of research which involves careful observation of behavior, (either in the field or in the laboratory), isolation of critical aspects of this behavior which can be provided with plausible neurological correlates, followed by neurophysiological testing for such correlates – it is in such a program (which might be termed: *neuroethology*) that I would place my hopes for progress, rather than in a *cognitive ethology,* which may well reveal itself a chimera. Incidentally, the use of animal surrogates in the investigation of communicative behavior, which, in my view, is likely to be one of the most productive of G's proposals, is clearly compatible with the program outlined above.

*No end of wonders [P&W].* Premack & Woodruff, who show themselves to be well aware of the need for careful definition of terms as well as for the careful evaluation of alternative explanations of experimental results, have provided a fascinating demonstration of how much remains to be learned about the mental states of chimpanzees (as well as of other animals). It will be most interesting to discover the level of performance of other populations mentioned (viz., normal and retarded children) on these same tasks. They may well prove to be significantly lower than those of such a sophisticated subject as Sarah.

Rather than raising quibbles concerning details of the experiments reported, it seems to me more important to address a theoretical point brought up at the very beginning of the paper. Since neither associationism nor rule systems of the type proposed by linguists have provided a viable theory of language, on what basis, I wonder, do P&W assume that these will prove adequate for a theory of something as poorly defined as mind?

*In the beginning was the word. [SR&B].* If P&W's contribution shows us how much remains to be learned about the chimpanzee's concept of problem, then the contribution by SR&B shows us how much remains to be learned about this great ape's communicative capacities. The sophisticated analysis of the role of function in word-learning and of levels of wordness constitutes a major contribution of this paper and provides interesting parallels with similar research involving children. As in P&W's experiments, it would be interesting to have direct evidence of normal children's performance in the tasks involved. The keyboard-off control condition provides an important supplement to the work of Menzel [q.v.] on the efficacy of the chimp's "natural" communicative behavior in transmitting information.

It is perhaps important to note at this point that the gap between what is known of the great apes' communicative behavior in the wild (which is very little as yet) and the languagelike capacities demonstrated in the laboratory is often cited as evidence for the discontinuity view. This involves a misunderstanding of evolutionary theory, for it is not the minimal abilities of an animal that are important in the struggle for existence but rather their behavioral potential for dealing with new or unusual situations.

There are a number of minor yet important statements made in the course of SR&B's contribution which may be misleading. Let me mention three. (1) It is an error to suggest that Myers (1976 *op.cit.* SRB) or anyone else who has investigated the neurological bases of nonhuman primate nonverbal (*sic*) communication has failed to notice that the limbs are under volitional control and differ significantly in this regard from the substrates of such communicative behaviors as vocalization and certain emotion-linked facial expressions. (2) The statement that the elements that make up the lexigrams of Yerkish are analogous to phonemes should not lead one to believe that such a level of analysis in the processing of Yerkish has been empirically demonstrated in either humans or chimps, which is, as far as I know, not the case. (3) Finally, the link between tool use and the emergence of language is tenuous at best; the only evidence cited by SR&B in support of this view, namely, the putative constraints on apes' manipulation of objects, is considerably weakened by the recent demonstration that an orangutan is quite capable of making and using flaked stone tools (Wright, 1978).

REFERENCES

Dingwall, W. O. The evolution of human communication systems. In Whitaker, H. and Whitaker H. A. (eds.), *Studies in neurolinguistics, vol 4.* New York: Academic Press, in press.
Eibl-Eibesfeldt, I. *Ethology.* New York: Holt, Rinehart and Winston, 1975.
Hodos, W. Evolutionary interpretation of neural and behavioral studies of liv-