

Dolphins on the witness stand? The comparative psychology of strategic memory regulation

Morris Goldsmith and Asher Koriat

Department of Psychology, University of Haifa, Haifa, 31905 Israel.
 mgold@research.haifa.ac.il akoriat@research.haifa.ac.il
 http://iipdm.haifa.ac.il

Abstract: Smith et al. show that monkeys and dolphins can respond adaptively under conditions of uncertainty, suggesting that they monitor subjective uncertainty and control their behavior accordingly. Drawing on our own work with humans on the strategic regulation of memory reporting, we argue that, so far, the distinction between monitoring and control has not been addressed sufficiently in metacognitive animal research.

In their stimulating target article, Smith et al. argue that humans are not the only animals that metacognize. Seemingly, monkeys and dolphins also possess at least a rudimentary ability to monitor their own state of knowledge and control their behavior accordingly. In this commentary, we highlight the distinction between *monitoring* and *control* (a core distinction in metacognition; Nelson & Narens 1990), drawing on our own work with humans on the strategic regulation of memory accuracy performance (e.g., Goldsmith & Koriat 1999; Goldsmith et al. 2002; Koriat & Goldsmith 1994; 1996b; Koriat et al. 2001).

To draw an analogy between Smith et al.'s work and our own, consider the situation of courtroom testimony. In attempting to "tell the whole truth and nothing but the truth," witnesses must regulate what they report from memory, volunteering information that is likely to be correct and withholding information that is likely to be wrong. In our studies, we found that when allowed to withhold answers about which they feel unsure (*free-report* conditions), people can enhance substantially the accuracy of what they choose to report relative to forced report (Koriat & Goldsmith, 1994; 1996b; Koriat et al. 2001). Moreover, they exercise this option strategically, enhancing accuracy even further (by withholding more answers) when given stronger incentives to do so.

Could a monkey or a dolphin "witness" also regulate its memory reporting effectively? Apparently yes. Treating the "uncertain" response as a withheld ("don't know") response: (1) Monkeys and dolphins apparently can respond "don't know" when they feel uncertain. (2) When they choose to make a primary response (volunteer an answer), they are more likely to be right than when they opt for a "don't know" response. (3) Consequently, they can increase the accuracy of the chosen primary responses (volunteered answers) by selective use of the "don't know" response (see especially Monkey Baker's impressive performance in Smith et al. 1998).

Despite these similarities, however, before we can conclude that dolphins' and monkeys' metacognitive processes parallel those of humans, more refined distinctions need to be made than are afforded by the signal-detection-theory (SDT) approach endorsed in the target article.

In our model (Koriat & Goldsmith 1996b; for an extension, see Goldsmith et al. 2002), we assume that people first retrieve information, then assess the likelihood that the information is correct (monitoring), and then decide whether or not to report it (control). This decision depends not only on the monitoring output (confidence), but also on the perceived incentives for accurate reporting. This model is similar to the one implied in the target article. However, based on this model, we identified four distinct factors, alone or in concert, that contribute to free-report memory performance: (1) *memory retention* – the amount and quality of information accessible in memory, (2) *monitoring effectiveness* – the ability to discriminate between correct and incorrect information that comes to mind, (3) *control policy* (report criterion) – the liberality or strictness of the criterion that is adopted for volunteering or withholding answers, and (4) *control sensitivity* – the extent to which the control decision (volunteer/withhold) is in fact

based on the monitoring output. In testing the model (Koriat & Goldsmith 1996b), we developed an experimental paradigm and assessment methodology that allows each of the various cognitive and metacognitive factors to be measured independently (as well as free-report quantity and accuracy performance), by obtaining forced-report and free-report responses to each memory item and a confidence rating for each forced-report answer (see the QAP assessment procedure in Koriat & Goldsmith 1996a; 1996b).

Note that Smith et al.'s claims specifically concern monitoring – the ability to know that one knows. Thus, an important challenge facing students of animal metacognition is to distinguish between monitoring and retention, on the one hand, and between monitoring and control, on the other. Consider first the distinction between monitoring and retention. Clearly, one can have poor retention and yet be able to distinguish between correct and incorrect answers that come to mind. Conversely, one can remember quite a lot, yet be unable to screen out wrong information. In our work (Koriat & Goldsmith 1996b), we showed that output-bound memory accuracy (the proportion of volunteered answers that are correct) depends crucially on monitoring effectiveness, independent of retention.

Turning to the distinction between monitoring and control, people may adopt a strict or liberal report criterion independent of monitoring (confidence level), depending on the relative weight given to the (competing) incentives for complete versus accurate reporting (Koriat & Goldsmith 1994; 1996b). Moreover, it is not a foregone conclusion that the control decision will be based on the monitoring output: Although with undergraduate participants, control sensitivity is virtually at ceiling (Koriat & Goldsmith 1996b), ongoing research with elderly people (Pansky et al. 2002) and clinical populations (Danion et al. 2001; Koren et al. 2001), suggests that this factor too may be nontrivial. In general, then, monitoring (confidence) does not dictate specific report decisions, nor do such decisions exclusively reflect monitoring.

As discussed elsewhere (Koriat & Goldsmith 1996a; 1996b), the foregoing distinctions become blurred when the signal-detection framework is adopted. For example, in the "old/new" memory paradigm to which signal-detection methods are typically applied, "control" is isolated in terms of the parameter beta, but "memory retention" (overall memory strength) and "monitoring effectiveness" (the extent to which the participants' confidence distinguishes "old" from "new" items) cannot be operationally or conceptually separated: Both are equally valid interpretations of d' (see, e.g., Banks 1970; Lockhart & Murdock 1970). Moreover, control sensitivity is a nonissue: It is axiomatic in SDT that the control decision (e.g., old/new) is based on memory strength, which is generally equated with confidence. This may explain why, notwithstanding the cleverness and elegance of Smith et al.'s methodology, it only taps monitoring *via the control decision*, rather than independently (see also Higham 2002).

To support the independent status of subjective monitoring in animals and refute behavioristic interpretations, it is not sufficient to show that animals behave *as if* they monitor their own knowledge. In fact, in the absence of verbal measures of confidence, it is crucial to demonstrate that they can behave *otherwise*, that is, to demonstrate a dissociation between the monitoring and control functions. Notwithstanding the methodological challenges involved, Smith et al.'s (1995) study, described in the target article, offers room for optimism: If noninstrumental behavioral indices such as hesitancy, slowing, and wavering could be developed into reliable and sensitive measures of animal confidence, it would be possible to examine separately the effectiveness of memory monitoring and of control. For example, to isolate the control policy, one could manipulate the incentive for accurate reporting, and examine whether, as one would expect, the manipulation affects the animal's report criterion without affecting confidence. With regard to control sensitivity, one might ask whether the apparent absence of animal metamemory in Shields's (1999) prospective memory-monitoring task stems from an inability to monitor or, rather, from the animal's failure to control its responding on the basis of

its monitoring. More generally, one could investigate the conditions under which animals (including humans) override their monitoring, responding on other bases (e.g., desperation, compulsion, perseverance, drive). We believe that the dissociation between monitoring and control may offer another small window through which to examine issues concerning the role of conscious awareness in cognition, metacognition, and behavior.

ACKNOWLEDGMENT

We gratefully acknowledge support of this project by the German Federal Ministry of Education and Research (BMBF) within the framework of German-Israeli Project Cooperation (DIP).

Significant uncertainty is common in nature

Donald R. Griffin

Concord Field Station, Harvard University, Bedford, MA 01730.
griffin@fas.harvard.edu

Abstract: In animals' natural lives, uncertainty is normal; and certainty, exceptional. Evaluating ambiguous information is essential for survival: Does what is seen, heard, or smelled mean danger? Does that gesture mean aggression or fear? Is he confident or uncertain? If they are conscious of anything, the content of animals' awareness probably includes crucial uncertainties, both their own and those of others.

Our scientific thinking about uncertainty monitoring and metacognition in animals tends to be framed in terms of laboratory animals, whose lives have much more certainty and much less stress and danger than is usual under natural conditions. A broader consideration of the behavior and cognition of animals during their normal lives can add significant evidence about awareness of uncertainty. Smith et al. recognize that "Working consciousness is the perfect referee for life's close calls. . . . [and] . . . may thus have substantial phylogenetic breadth" (sect. 15, second last para.). A comparable view has been expressed by Dickinson and Balleine (2000) concerning the goal-directed action by laboratory rats.

The content of animal consciousness is doubtless very simple and limited in comparison with ours; but it can nevertheless be important to the animal for its survival.

Under natural conditions animals must wonder whether a marginally visible movement of vegetation or a barely audible rustling means that a dangerous predator is present. But wind and other harmless events cause very similar rustlings. Seed-eating animals probably wonder whether a particular speck on the ground is a pebble or a partly buried seed. Social signals often have uncertain meaning: Does that slight movement of a furry shoulder mean he will attack, or does that brief glance mean that she will be friendly? Which of those infant distress calls is from my baby? Such uncertainties are often vitally important for the animal and are therefore likely to be consciously experienced and evaluated.

I suggest that we adopt the potentially testable working hypothesis that many animals are consciously aware of whatever is critically important in their lives, and that simple perceptual consciousness is a core function of central nervous systems. Selecting actions the animal believes will obtain what it wants or avoid what it dislikes or fears is an efficient way to use a central nervous system. This ability is adaptive because it makes appropriate decisions more likely and thus increases the animal's evolutionary fitness. If animals are aware of anything, the many uncertainties that are critical for survival must often require conscious attention.

A major obstacle to evaluating this hypothesis is the widespread opinion that it is impossible to determine with absolute certainty whether an animal is or is not conscious. Yet we seldom if ever demand perfect proof before evaluating imperfect evidence about other difficult questions in the behavioral sciences, so that this double standard is a form of paralytic perfectionism that discourages research. Furthermore, there is now abundant evidence of

nonhuman cognition and consciousness, as reviewed by Heyes and Huber (2000), Griffin (2001), and Bekoff et al. (2002). Considerable information about the content of consciousness is readily available. We make inferences about the conscious states of our human companions by interpreting their communicative behavior, both linguistic and nonverbal. This is increasingly feasible with animals as more is learned about the versatility of their communication. They often appear to be communicating their conscious experiences, which amounts to declarative consciousness, even though their communication systems differ from human language.

Smith et al. are concerned primarily with metacognitive thinking about uncertainty itself in contrast to thinking about alternative possibilities on the basis of imperfect information. I suspect that many animals are keenly aware of uncertainties about the meaning of sensory information that may or may not mean danger or opportunity; but whether they think about uncertainty as an abstract concept is much more difficult to judge. Perhaps we should search for communicative behavior that reports awareness of uncertainty itself. This might occur naturally, once we learn where to look for it, or it might be instilled by extensions of the types of experiments reviewed in the target article. In many challenging situations when animals are uncertain what to do, they actively seek better information, peering, listening, tasting, probing, or sniffing. Prey animals often show greater anxiety and caution when moving through thick vegetation where predators are more difficult to see than out in the open. And dangerous predators are sometimes inspected cautiously, apparently in search of an indication of the intention to attack. Sometimes this information quest entails exchanging communicative signals with other animals.

One example is the exchange of symbolic gestures by swarming honeybees recently reviewed by Griffin (2001). When it is necessary that the swarm find a suitable cavity into which they can move, some scout bees that have located reasonably suitable cavities report their distance, direction, and desirability by the same symbolic gestures ordinarily used to report the location of food sources. Some follow dances of a sister that describe a better cavity. Occasionally, the first bee then changes her dance message to that describing the better cavity about which she has learned as a dance follower. Sometimes this occurs *without first inspecting the second cavity herself*. This appears to be an example of the "substantial phylogenetic breadth" of working consciousness, recognized by Smith et al., applied in a situation where the bees are uncertain which is the better cavity and are seeking additional information before making this vitally important decision.

Metacognition as evidence for explicit representation in nonhumans

Robert Russell Hampton

Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, MD 20892. robert@ln.nimh.nih.gov

Abstract: Metacognition is either *direct*, as when information is recalled before making a confidence judgment, or *indirect*, as when the probability of successful future retrieval is determined inferentially. *Direct* metacognition may require an explicit mental representation as its object and can only be demonstrated under specific experimental circumstances. Other forms of metacognition can be based on publicly observable stimuli rather than introspection.

Metacognition requires two distinct components, an object-level mental process, such as a memory, and a meta-level, or executive, process that monitors the object-level process (Nelson & Narens 1996). In some cases, the meta process has direct access to an explicit representation at the object level (e.g., Koriati 1996), but in many other cases monitoring is indirect or inferential (Flavell 1979). Contrast the following two situations requiring a metacog-