CrossMark

THEORETICAL REVIEW

Formal models in animal-metacognition research: the problem of interpreting animals' behavior

J. David Smith[1] · Alexandria C. Zakrzewski[2] · Barbara A. Church[3]

Abstract Ongoing research explores whether animals have precursors to metacognition—that is, the capacity to monitor mental states or cognitive processes. Comparative psychologists have tested apes, monkeys, rats, pigeons, and a dolphin using perceptual, memory, foraging, and information-seeking paradigms. The consensus is that some species have a functional analog to human metacognition. Recently, though, associative modelers have used formal-mathematical models hoping to describe animals' "metacognitive" performances in associative-behaviorist ways. We evaluate these attempts to reify formal models as proof of particular explanations of animal cognition. These attempts misunderstand the content and proper application of models. They embody mistakes of scientific reasoning. They blur fundamental distinctions in understanding animal cognition. They impede theoretical development. In contrast, an energetic empirical enterprise is achieving strong success in describing the psychology underlying animals' metacognitive performances. We argue that this careful empirical work is the clear path to useful theoretical development. The issues raised here about formal modeling—in the domain of animal metacognition—potentially extend to biobehavioral research more broadly.

✉ J. David Smith
jsmith395@gsu.edu

[1] Department of Psychology, Georgia State University, 738 Urban Life Building, 140 Decatur St., Atlanta, GA 30303, USA

[2] Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY, USA

[3] Language Research Center, Georgia State University, Atlanta, GA, USA

Introduction

Metacognition is the awareness of one's mental states or cognitive processes (knowing, doubting, remembering, etc.) coupled with the ability to regulate behavior adaptively using that awareness. Metacognition in humans is central to thinking, memory, comprehension, and decision making (Dunlosky & Bjork, 2008; Flavell, 1979; Koriat & Goldsmith, 1994; Nelson, 1992). A large experimental and educational literature takes metacognition as its focus. Metacognition is a sophisticated cognitive capacity possibly linked to consciousness and self-awareness (Koriat, 2007; Nelson, 1996). This link is why metacognitive states are often personalized, as when we say: *I* don't know; *I* don't remember. Metacognition emerges late in human development (Balcomb & Gerken, 2008)—its earliest developmental roots have not been mapped. Indeed, metacognition might be so sophisticated a capacity that it is uniquely human.

Still, given metacognition's importance, it is a natural question whether nonhumans share aspects of this capacity (Kornell, 2009; Metcalfe, 2008; Smith, 2009). If they do, it could bear on their consciousness and self-awareness. It could affect theoretical debates in comparative psychology, and affect the interpretation of behavioral research. It could illuminate metacognition's evolutionary beginnings and possibly the earliest (nonverbal) forms it takes in very young human children (Balcomb & Gerken, 2008). It could point to nonverbal forms of cognitive regulation that might benefit children with special needs. Thus, researchers have actively explored animal metacognition, creating one of comparative

Springer

psychology's influential literatures (e.g., Basile, Hampton, Suomi, & Murray, 2009; Basile, Schroeder, Brown, Templer, & Hampton, 2015; Beran & Smith, 2011; Beran, Smith, & Perdue, 2013; Call, 2010; Couchman, Coutinho, Beran, & Smith, 2010; Foote & Crystal, 2007; Fujita, 2009; Kornell, Son, & Terrace, 2007; Paukner, Anderson, & Fujita, 2006; Roberts et al., 2009; Smith, Coutinho, Church, & Beran, 2013; Suda-King, 2008; Sutton & Shettleworth, 2008; Templar & Hampton, 2012; Washburn, Gulledge, Beran, & Smith, 2010; Washburn, Smith, & Shields, 2006; Zakrzewski, Perdue, Beran, Church, & Smith, 2014). Many researchers have contributed to this literature—our citations are only illustrative. Primates have shown diverse, seemingly metacognitive performances in tasks involving food search, psychophysical discrimination, information seeking, memory monitoring, and so forth. Perhaps some animals monitor their cognition, and know when they know or remember. It is widely agreed that this would be a singularly important conclusion about animal minds.

Associative cues and "metacognitive" performances

However, as in all domains of behavioral research, interpretative issues arise. A range of psychological interpretations could be given to a performance that seems metacognitive on the surface. One could emphasize lower-level associative-learning processes, or higher-level cognitive processes, or metacognitive processes that are perhaps reliant on something like working consciousness.

In particular, in animal-metacognition research, there has been a focus on associative explanations of "metacognitive" performances. These performances might be explained as low-level (reactive) conditioning phenomena if animals' "uncertainty" responses are cued by stimuli or conditioned by reinforcement. The associative-metacognitive issue has dominated the theoretical debate (Basile & Hampton, 2014; Basile et al., 2015; Carruthers, 2008; Hampton, 2009; Jozefowiez, Staddon, & Cerutti, 2009a,b; Le Pelley, 2012, 2014; Smith, 2009; Smith, Beran, & Couchman, 2012; Smith, Beran, Couchman, & Coutinho, 2008; Smith, Couchman, & Beran, 2012, 2014a, b; Staddon, Jozefowiez, & Cerutti, 2007). It has dictated the form and application of formal models in this domain. It has even framed the literature's article titles (e.g. Metacognitive Monkeys or Associative Animals?—Le Pelley, 2012).

The associative-metacognitive debate is sharpened because the field's methods do sometimes entangle procedural learning/associative responding with possible metacognitive monitoring. That is, researchers often give animals indeterminate stimuli to produce the uncertainty animals may monitor. Those stimuli will also produce errors and reduce rewards. Procedural-learning systems could sense these contingencies.

Avoidance responses could entrain to problematic stimuli, helping the animal avoid errors for responses made to those stimuli. These responses might seem metacognitive, but might not be. Thus, attributing higher-level cognitive monitoring to animals is difficult, though such monitoring could be present.

Yet strong empirical progress is being made. Acknowledging the difficulty of interpretation, researchers have introduced empirical approaches that address associative interpretations. One concern is that animals have often received tangible rewards for metacognitive responses (Foote & Crystal, 2007; Fujita, 2009; Hampton, 2001; Inman & Shettleworth, 1999; Kornell et al., 2007; Suda-King, 2008; Sutton & Shettleworth, 2008). As pointed out by Smith, Beran, et al. (2008), this approach might give those responses associative strength and appetitive attractiveness independent of their metacognitive basis. But researchers have addressed this concern repeatedly. Animals show adaptive metacognitive responses even when those responses earn no concrete reward (Beran, Smith, Redford, & Washburn, 2006; Couchman et al., 2010; Smith, Beran, Redford, & Washburn, 2006; Smith, Redford, Beran, & Washburn, 2010).

A second concern is that researchers often use first-order stimulus qualities to create metacognitive uncertainty (i.e., concrete, visible stimulus features like size, color, etc.). Error-causing (timeout-bringing!) stimulus features could become associatively aversive and avoided for this reason—not based on a metacognitive judgment. Research with monkeys has allayed this concern by showing adaptive metacognitive responses even when the task requires conceptual or memory judgments not linked to particular stimuli—even in some cases when no stimulus is visible (Hampton, 2001; Hampton, Zivin, & Murray, 2004; Kornell et al., 2007; Shields, Smith, & Washburn, 1997; Smith, Shields, Allendoerfer, & Washburn, 1998; Washburn et al., 2010).

A third concern is that researchers often give trial-by-trial reinforcement. If animals can associate consequences to the stimulus-response combinations that earned them, they might condition through low-level mechanisms to avoid problematic stimuli, with no metacognitive basis for avoidance. Research with monkeys has allayed this concern by showing metacognitive responses when reinforcement is deferred (i.e., presented only after each trial block) so that assigning credit for reinforcement to particular stimulus-response combinations is difficult and the normal pathways for procedural learning are blocked (Couchman et al., 2010; Smith et al., 2006).

We do not prejudge here the correct theoretical interpretation of animals' "metacognitive" performances. But the sense of the literature is that metacognition research has moved beyond some associative hypotheses. The emerging consensus is that some species share some aspects of humans' metacognitive capacity. Sutton and Shettleworth (2008, p. 266) concluded that "metamemory, the ability to report on

memory strength, is clearly established in rhesus macaques (*Macaca mulatta*) by converging evidence from several paradigms." Fujita (2009, p. 575) concluded that "evidence for metacognition by nonhuman primates has been obtained in great apes and old world monkeys." Roberts et al. (2009, p. 130) concluded that "substantial evidence from several laboratories converges on the conclusion that rhesus monkeys show metacognition in experiments that require behavioral responses to cues that act as feeling of knowing and memory confidence judgments." Carruthers and Ritchie (2012, p. 76) concluded that "this body of work, taken as a whole, cannot be explained in low-level associationist terms, as involving mere conditioned responses to stimuli."

Associative models and "metacognitive" performances

Of course not everyone supports this consensus. One can question the weight to give the empirical demonstrations just described. One can question whether ruling out some associative hypotheses is decisive or not. One can propose additional associative mechanisms that have not yet been addressed. In this spirit, associative modelers have recently used formal models to criticize the animal-metacognition phenomena. That is, they have developed "associative" models that depict the reinforcement histories associated with error-causing stimuli that might cue avoidance (*not* metacognitive) responses. Staddon et al. (2007) and Jozefowiez et al. (2009a,b) used the *Behavioral Economic Model* to fit rats' metacognitive performance in Foote and Crystal (2007). Le Pelley (2012) asked whether an associative model could fit the metacognitive performances produced by macaques. He also assumed that stimulus-response registers were updated over trials so as to encode reinforcement histories and entrain response strategies to those histories. Below we describe one of the associative models in detail. To be fair and self-critical, one of us (Smith, Beran et al., 2008; Smith, Shields, & Washburn, 2003) initiated the formal-mathematical approach to animal metacognition.

These models have exerted a powerful theoretical pull in our field. Colleagues have shown they will reject a metacognitive interpretation of performance if an associative model fits the data. They have shown they will accept an associative interpretation if the associative model fits the data. Editors and reviewers have required these models to be incorporated in articles, recommended rejection of articles based on the outcome of modeling, and disallowed theoretical considerations of animal metacognition if associative models fit. A very distinguished colleague once told us that seeing associative models reproduce metacognitive phenomena had produced a conversion experience against the possibility of animal metacognition. If these interpretative and judgmental standards seem appropriate, then this article will be brightly illuminating.

This article is about the use of these models in the area of animal metacognition. It is about the logic of the use of formal models as mathematical fits for animals' performance. We will ask why researchers believe these models are explanatory. We will specify what models can never grant animal-metacognition research. We will consider their serious limitations. We will show that the current application of models in our field reflects poor scientific reasoning and potentially harms theoretical development of our field. We will consider the alternative ways in which researchers can advance the literature, given that these formal models do not presently contribute to that advance.

These issues have broad implications for biobehavioral research. Similar models are used similarly in many domains (e.g., numerosity, timing, foraging, gratification delay, decision-making). We do not assert that other models inevitably share the problems we point to—each field must make its own determination. But a general reassessment of how models are used and interpreted could be constructive. The animal-metacognition literature is an elegant case study in making this reassessment—one that springs from the work of many excellent researchers and modelers.

A target animal-metacognition performance

We begin by presenting a target data pattern from an animal-metacognition experiment. In this example (Smith et al., 2013), a rhesus macaque (*Macaca mulatta*) completed a sparse-uncertain-dense task. On each trial, he saw a box in the screen's top center filled with some number of randomly placed lit white pixels on a black background (Levels 1–60). The 30 sparsest and 30 densest trial levels, respectively, deserved the Sparse and Dense response. These responses received a food reward or a trialless timeout period if correct or incorrect. Trial levels near the discrimination breakpoint (Levels 30–31) were difficult and error producing. The macaque could also make an uncertainty response that produced the next trial without providing any feedback or reinforcement. Despite this neutral outcome, the uncertainty response was potentially useful. Through its judicious use—only on difficult trials—the macaque could fend these trials off, avoid errors and timeouts, and increase rewards in the task.

Figure 1 shows the macaque's performance over about 3, 000 trials. He showed a familiar pattern. He responded uncertain most for trials near the discrimination breakpoint. He somehow correctly evaluated that these trials were difficult and error producing, and he declined them selectively and adaptively. Humans often show identical data patterns. Interestingly, humans say their uncertainty responses are prompted by conscious metacognitive uncertainty.
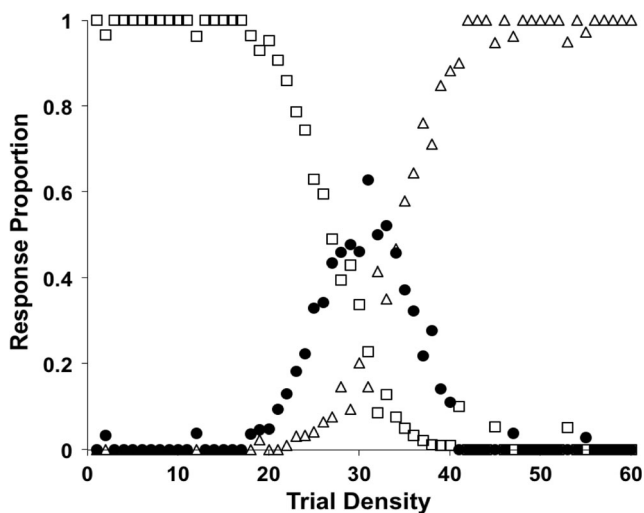
Fig. 1 Proportion of uncertainty responses (*solid circles*), sparse responses (*open squares*), and dense responses (*open triangles*) made by the macaque Murph in the sparse-uncertain-dense task of Smith et al. (2013). The *horizontal axis* indicates the objective density of the trial (Levels 1–30: Sparse; Levels 31–60: Dense)

Macaques say nothing! Their performance might be metacognitive or associative. Josefowiez, Le Pelley, Staddon, and their colleagues favor the associative interpretation. They apply associative formal-mathematical models to see whether they can describe associatively how animals perform in these tasks so that the metacognitive interpretation can be dismissed. We will illustrate these models next.

An "associative" model of animal metacognition

The macaque in the task just described received trial-by-trial feedback. Transparent reinforcement could let animals tabulate in memory the reinforcement histories attaching to different stimulus-response combinations. From a neuroscience perspective, we might say instead that immediate feedback would allow the updating of neural connections from visual cortex to motor cortex, perhaps with cells in the caudate nucleus as facilitating intermediaries, so that procedural learning ensued (e.g., Arbuthnott, Ingham, & Wickens, 2000; Calabresi, Pisani, Centonze, & Bernardi, 1996; Gamble & Koch, 1987; Hollerman & Schultz, 1998; MacDermott, Mayer, Westbrook, Smith, & Barker, 1986; Schultz, 1992; Wickens, 1993).

An illustrative associative model can instantiate this system. Its design is uncontroversial. Like all associative models in our area, it assumed that performance in uncertainty tasks is organized along a continuum of psychological representations of increasing strength (here, increasing density from sparse to dense). It assumed that objective stimuli were perceived with perceptual error (so that a Level 10 stimulus would create a perceptual impression from trial to trial in the range, say, of 8

to 12). The model assumed that simulated observers respond to the trial's subjective level or impression (not the objective stimulus level), just as a macaque must do. It assumed that animals tabulated in memory the reinforcement histories attaching to different stimulus-response combinations, and that they grew averse to responding Sparse or Dense to stimuli (and to particular responses made to them) proportionally to their errors. Figure 2 shows this response-strength function as it wanes toward the middle of the continuum containing error-prone levels. The function's steepness was controlled by a free parameter (*sensitivity*) in our model that governed the exponential decay of response strength.

The model assumed that the third, avoidance response had a constant attractiveness across the continuum in accordance with its constant consequence (Fig. 2, horizontal line). This avoidance threshold's height on the *y*-axis (*threshold*) was also a free parameter in our model, with higher values producing more avoidance responding generally and more broadly across the continuum. Simulated observers made a Sparse or Dense discrimination response if the reinforcement-based response strength was greater. They made the avoidance response if its response strength was greater.

We varied the values for *sensitivity* and *threshold* to find those that best recovered the 180 proportions (three responses × 60 stimulus levels) the macaque showed. To quantify the model's fit, we found the sum of the squared deviations (SSD) across corresponding observation-prediction pairs. We minimized this measure to find the best-fitting parameters. They
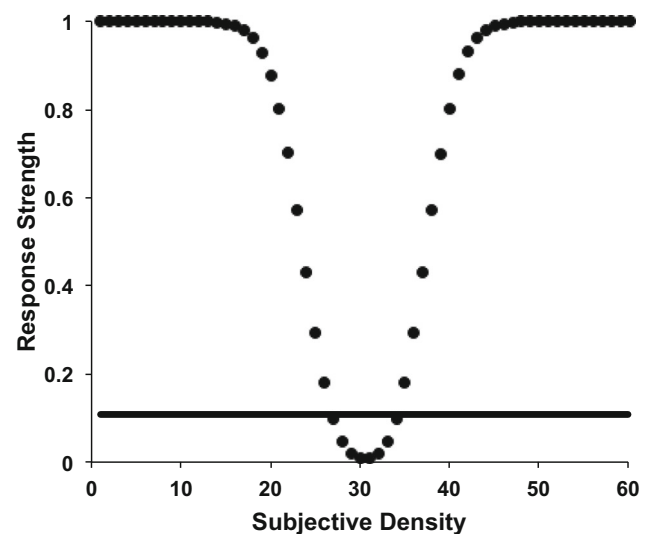

Fig. 2 A reinforcement-history portrayal of performance in a sparse-dense discrimination with a third response assumed to manage stimulus aversion and response avoidance. The *horizontal axis* indicates the subjective impression created by the objective stimulus on the trial. The *solid line* instantiates the idea that the third response could be the default option with a constant response strength that is selected when aversion or avoidance weakens the tendency to respond sparse or dense. The *dotted line* instantiates the idea that response strength for the sparse and dense responses would wane exponentially going inward as the frequency of errors increased

were .101 (*sensitivity*) and .108 (*threshold*). For these best-fitting parameters, we calculated an intuitive measure of fit—the average absolute deviation (AAD). This measure represents the average of the deviations between observed and predicted pairs (with the deviations always signed positively). Figure 3 shows that this model's predictions (lines) fit well the macaque's results (symbols). The SSD between corresponding points was .2019. The AAD per point was .0192; on average, the 180 observation-prediction pairs differed by less than .02.

That the associative model recovers "metacognitive" performance has a strong impact on animal scientists. We have seen many times the effect a fit like this has. A strong inference immediately follows that this model is the correct interpretation of the animal's performance, that the assumptions of associative-learning theory are justified regarding that performance, and that the target performance was therefore not metacognitive. The power of the fit of an associative model is extraordinary.

A "metacognitive" model of animal metacognition

In reality, an immediate obstacle blocks using an associative model to justify inferences like these. There is a rival model we also illustrate here. To be fair, we must let this model fit the same data first. This model assumes that the macaque monitored psychological signals of uncertainty and that he was able to place two confidence criteria along the Sparse–Dense
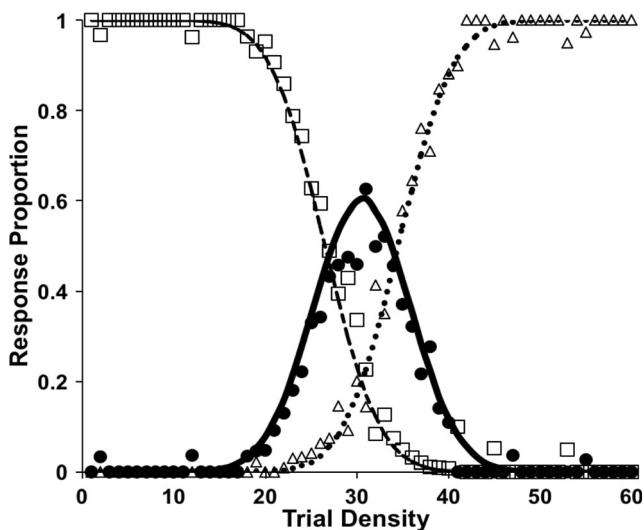

Fig. 3 Murph's performance in the sparse-uncertain-dense task of Smith et al. (2013), depicted with symbols as described in the legend to Fig. 1. Also shown are the best-fitting predictions produced by an "associative" model as it fit those observed data using a third response to manage stimulus aversion and response avoidance (*solid line* aversion-avoidance response; *dashed line* sparse responses; *dotted line* dense responses). Details of the model and model fitting are described in the text

continuum. These confidence criteria are central to all aspects of signal-detection theory (Macmillan & Creelman, 1991). These criteria offer an alternative descriptive framework. Using them, the macaque could use the Sparse and Dense responses for the easy and certain (i.e., high-confidence) trials outside the criteria and farther from the discrimination's breakpoint. He could reserve the uncertainty response for the difficult (i.e., low-confidence) trials between the criteria and nearer the breakpoint. Thus, the uncertainty response was modeled as arising from the application of a form of uncertainty monitoring or metacognition, though not necessarily the full human version of uncertainty monitoring and conscious metacognition.

The model assumed the same density continuum and perceptual error as the associative model. It assumed that one criterion (criterion sparse-uncertain, free parameter CSU) was placed to separate easier Sparse trials from Uncertain trials, and one criterion (criterion uncertain-dense, free parameter CUD) was placed to separate Uncertain trials from easier Dense trials. On each trial, the model assumed that the animal perceived an objective stimulus with perceptual error, yielding that trial's subjective impression. If that impression were below CSU, between CSU and CUD, or above CUD, respectively, the model chose the Sparse, Uncertain, or Dense response. If CSU and SUD were placed at 28 and 32, the uncertainty region would be narrow, and the animal would make uncertainly responses stingily for few trial levels. If the criteria were placed at 24 and 36, the uncertainty region would be wider and uncertainty responses more plentiful.

The metacognitive model did not include the sensitivity and threshold parameters of the associative model. They were replaced by CSU and CUD. The associative and metacognitive models in this article had the same number of free parameters—two. One must not suppose that the metacognitive model was more complex or parameter rich, or that the associative model is to be preferred because it was somehow simpler.

The metacognitive model was grounded in modeling techniques (e.g., perceptual error, discrimination thresholds, signal detection) that extend back many decades in experimental psychology. Thus, the associative and metacognitive models in this article used equally venerable and successful basic modeling assumptions. There is no basis for giving the associative model explanatory preference because of its historical depth.

The theoretical perspective of uncertainty monitoring and uncertainty responding have also had as long a history in human research as behaviorism has had in animal research (e.g., Fernberger, 1914; Jastrow, 1888). Along this dimension, too, there is no basis for preferring the associative model based on historical precedent.

We varied the values for CSU and CUD to find those that best recovered the macaque's 180 response proportions. The

best-fitting values were 27.1 (CSU) and 34.2 (CUD). Figure 4 shows that the "metacognitive" model's predictions (lines) fit well the macaque's performance (symbols). The SSD between corresponding values was .126, i.e., a 37.5% reduction in the summed error of prediction compared to the "associative" model. The AAD per point in the two graphs was .014—a 25% better fit to the data than the "associative" model found (.019). The fit of the metacognitive model was at least as good as that of the associative model.

At this point, we have described two formal-mathematical models that reflect different psychologies. Both fit the data very well. One cannot simply accept the associative interpretation, because the metacognitive model certainly fits competitively. One cannot accept the metacognitive interpretation, though, because the associative model fits competitively, too. One cannot attack the metacognitive model, by claiming that its criteria need not be really metacognitive and that animals might actually be associative. For we did not attack the associative model similarly, by noting that its processes need not really be associative, and that animals might actually be metacognitive. Generally, we cannot approach the interpretative question using any bias toward associationism or cognitivism that we may have because then bias, and not science, would pre-decide the issue. Instead, we need a disciplined resolution between the perspectives behind the two models. Readers might consider which interpretation they might choose in this situation. So, instead of indicating to us the correct psychological interpretation of these performances,
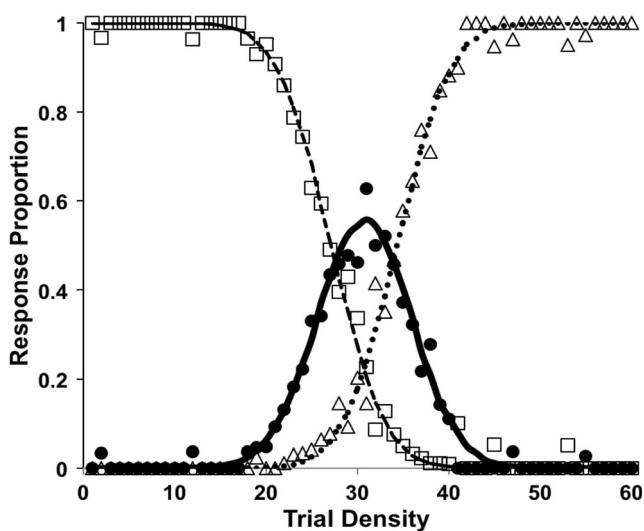
the models freeze us between equivalent associative and metacognitive interpretations. And now we will see that the interpretative situation regarding these two models is actually far worse than this.

Deep convergence, deep trouble

The deeper problem is that the associative and metacognitive models have a perfect mathematical correspondence. To illustrate, we took the predictions of the associative model that best fit the macaque's performance, and we used those predictions as the target to be fit by the metacognitive model. Both performance profiles are overlain in Fig. 5, as the metacognitive model (black symbols) tried to reproduce what the associative model had predicted (open symbols). If you do not see any open symbols, look very closely. That difficulty is the point here. These two data patterns are essentially identical. The SSD between the two performance patterns was .0011, indicating vanishingly small differentials. The AAD per associative-metacognitive data pair was .0013, indicating the same. Whatever the two nominal kinds of cognitive processing the models envision, in reality they are mathematically isomorphic.

We took one more step to confirm this. We created 17 simulated creatures using the metacognitive model, with CSU gradually decreasing from 30 down to 14, and with
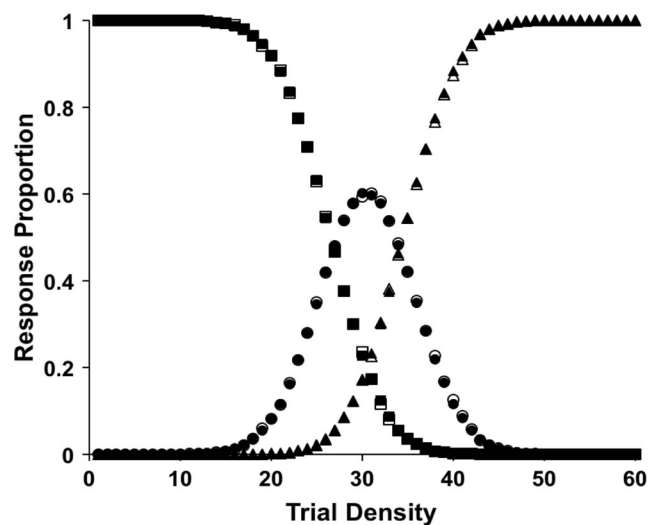

Fig. 4 Murph's performance in the sparse-uncertain-dense task of Smith et al. (2013), depicted with symbols as described in the legend to Fig. 1. Also shown are the best-fitting predictions produced by a "metacognitive" model as it fit those observed data assuming that the macaque monitored psychological signals of uncertainty and was able to place two confidence criteria along the Sparse-Dense continuum (*solid line* uncertainty responses; *dashed line* sparse responses; *dotted line* dense responses). Details of the model and model fitting are described in the text

Fig. 5 The convergence between the associative and metacognitive formal model. To make this graph, the best-fitting prediction of the associative model—as it fit the macaque's observed discrimination data (Fig. 1)—was used as the data-fitting target, and that target was then fit by the metacognitive model. The *horizontal* axis indicates the objective density of the trial (Levels 1–30, sparse; Levels 31–60, dense). Shown are the proportions of aversion-avoidance responses (*open circles*), sparse responses (*open squares*), and dense responses (*open triangles*) originally produced by the associative model, and the best-fitting proportions of uncertainty responses (*black circles*), sparse responses (*black squares*), and dense responses (*black triangles*) produced by the metacognitive model fitting the prediction of the associative model

CUD gradually increasing from 30 up to 46. The parameter configuration 30–30 would produce no uncertainty responses. The configuration 14–46 would produce exuberant uncertainty responding. Thus, we covered a range of "metacognitive" strategies from 0 to generous uncertainty responding given a wide uncertainty-response region that spanned more than half the stimulus continuum. We let each of these simulated "metacognitive" data patterns become the fitting target now for the "associative" model.

Figure 6 shows the result of this extensive simulation. The *x*-axis shows the width of the uncertainty region along the stimulus continuum for 17 versions of the metacognitive model (i.e., quantity CUD minus CSU). The *y*-axis shows the best-fitting threshold for choosing a Sparse/Dense response as the associative model fit each version. The remarkable result (filled symbols) is that these nominal metacognitive and associative parameter values trace a sigmoid of perfect equivalence.

Figure 6 (open circles) shows the result of fitting a logistic curve to the 17 points. The curves are nearly indistinguishable, with 99.93% of variance explained. As the "metacognitive" model's uncertainty region widens, the "associative" model's threshold for choosing a Sparse/Dense response heightens, in perfect mathematical lockstep.
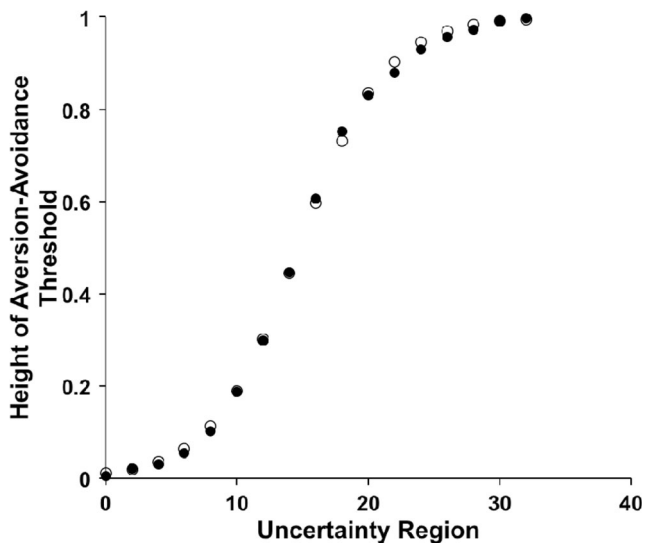

Fig. 6 The convergence between the associative and metacognitive formal models. To make this graph, we produced 17 simulated performers who performed according to the predictions of the metacognitive model. They had Sparse-Uncertainty and Uncertainty-Dense confidence criteria, respectively, placed at Levels 30–30, Levels 29–31, Levels 28–32, and so forth out to Levels 14–46. As the width of the uncertainty region increased (e.g., 46 − 14 = width 32), the uncertainty response was used more generously. Each of 17 performance profiles was then fit by the associative model, so that we could assess the relationship between the width of the metacognitive uncertainty region in the metacognitive model (*x*-axis) and the height of the aversion-avoidance threshold in the associative model (*y*-axis). The two parameters— uncertainty-region width and aversion-avoidance threshold height— have a perfect mathematical correspondence (*solid symbols*). A simple logistic function recovered this relationship perfectly (*open symbols*)

The "metacognitive" and "associative" models are only alternative ways to parameterize performance space mathematically. The threshold parameter (associative model) and the width parameter (metacognitive model) accelerate the use of the uncertainty response in the same way. There might be six other isomorphic ways to parameterize performance space. There is no interpretative choice embodied by these models, or required by them, or allowed by them, because they are exactly the same mathematical thing. Indeed, the models are so hopelessly entangled mathematically that they cannot even validly or reasonably express the kinds of processes that the two theoretical ideas behind the models supposedly espouse. These are two arbitrary mathematical descriptions, two glorified factor-analyses of the data that are simple rotations of each other and therefore psychologically empty. The models are weak, inseparable, indistinguishable, and they are not up to the task of driving psychological interpretation or theory in either direction, in any direction. They must be removed from the theoretical discussion regarding animal (and human!) metacognition.

The trouble with "associative" and "metacognitive" models

That is the situation in the animal-metacognition literature. But the lessons there extend more broadly to other research areas. Therefore, here we offer some general observations about formal models as a general caution.

Models are pure mathematics. They are gradients, numerical transformations, decay functions, thresholds. One cannot read anything into a model beyond these transformations, or extrapolate them to any mind or cognitive system.

Models restate the empirical result. The models' output in Figs. 3 and 4 recovers the behavior of declining more trials near the midpoint of the continuum. We already observed the macaques doing this. Neither model adds value to this observation.

Models are post hoc rationalizations. Their parameters take on the values they do to reproduce a graph. The model's parameters are exhausted—that is, used up—re-expressing that graph. They have no additional interpretative or conceptual content.

Models only reparameterize the observations. For example, they turn the animal's frequency of uncertainty responding into a threshold-parameter value in the associative model. But they have only redescribed the original situation mathematically. They are a conceptually empty translation of the data into a new form. Moreover, it is likely, and it is devastating, that the translation is only one of many. For no one mathematical translation is better than another. None is closer to the animal's mind. None has any intrinsic tie to the animal's psychology.

Therefore, models are silent on the issue of psychological processing. Take the height of the aversion-avoidance threshold. It tells us only that we needed that value to recreate the animal's frequency of uncertainty responses. But the animal need not have performed the task in any way like that nominally specified by the model. So models do not and cannot express associative processing or metacognitive processing. They embody no psychological processes, only abstract mathematical transformations.

Finally, models may blur even the most fundamental distinctions about animals' minds. In our case, fully conscious (human!) metacognition and purely associative responding would produce the same data pattern as we have seen. Both underlying psychologies would then be fit by the same mathematics/model. Some have mistaken the fit of the associative model to assert that everything is associative. One could use the fit of the metacognitive model to assert that everything is metacognitive. Neither assertion is correct. Indeterminate models cannot make a processing determination.

Illustrating a definitive contrast between models

This criticism of animal-metacognition models will not extend to all research areas. Sometimes formal models can cut deeper to analyze the true structure of animal cognition. An example will show why animal-metacognition models do not cut deeper.

Smith, Redford, and Haas (2008) tested exemplar models of categorization. Exemplar theory holds that animals store category exemplars as separate, individuated memory traces spread out like a cloud in the mind's psychological space. They endorse new items into the category if—on comparison to these multiple, separated cognitive reference points—the items are similar enough to belong. Exemplar theory makes an elegant, obligatory prediction. If the animal stores individuated exemplars spread out in psychological space, then even a perfectly typical new item will not be maximally endorsed into the learned category. The reason is that the item can never be close to all the stored exemplars at once. It will always be near to some exemplars in psychological space but far from others (Smith, 2002; Smith & Minda, 2001, 2002).

Figure 7 (filled symbols) shows one macaque's endorsement gradient in a categorization task—that is, the proportion of times the animal endorsed items into the category from random items clearly outside the category to prototypical items at the category's center. His prototype endorsement level was as high as possible. The exemplar model (Es) predicted lower prototype endorsements and its predictions were consistently off by 10% per item type, a strongly disconfirming degree of error. The exemplar model is strong, clear and testable in this case, because it instantiated the organization of psychological space in the animal's mind as dictated by its
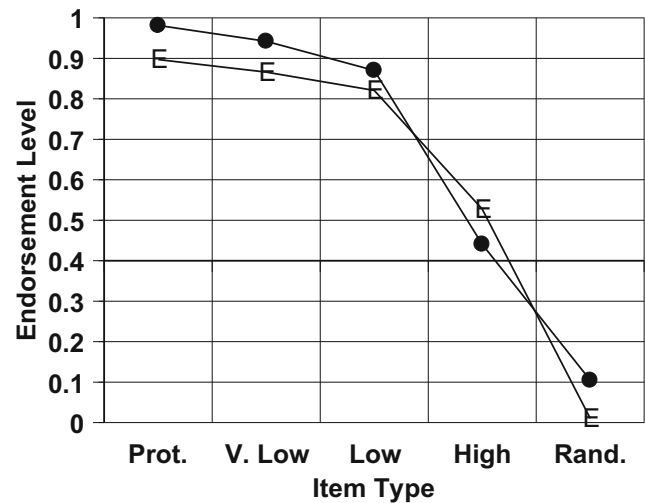

Fig. 7 The proportion of times a macaque endorsed into a learned category to-be-categorized test items that were outside the category (*rand.*), non-typical category members (*high*-level distortions), typical members (*low*-level distortions), highly typical members (*v. low*-level distortions), or prototypical members (*prot.*). Also shown is the best-fitting predicted profile (*E*) when a standard exemplar-based categorization model fit the macaque's performance as well as it could. From "Prototype abstraction by monkeys (*Macaca mulatta*)," by J. D. Smith, J. S. Redford, and S. M Haas, *Journal of Experimental Psychology: General*, 137, 390–401. Copyright 2008 by the American Psychological Association. Reprinted with permission

theory. This representational geometry enforced predictions that macaques strikingly disconfirmed.

The animal-metacognition models lack this strength. They define response regions of avoidance or uncertainty. But they are only surface mathematical descriptions of data patterns that could have any psychology underlying them.

There are doubtless many instances in biobehavioral research of models that cut deeper because they instantiate an inherent property of the animal mind that can be clearly tested. So, to be clear, we are not dismissing formal modeling. But researchers and modelers, and consumers of research and models, must be vigilantly evaluative about the uses to which models are put. Models that are only mathematical descriptions of behavioral patterns—that could have any psychology underlying them—must be given no ability to influence debate or theory in biobehavioral fields.

The need for interpretative symmetry

In the animal-metacognition literature, when an apparently metacognitive model fits macaques' performance, this has no persuasive scientific impact. When the apparently associative model fits, this has a shaping influence. This asymmetry was expressed in Smith, Beran et al. (2008), Staddon et al. (2007), Josefowiez et al. (2009a,b), and Le Pelley (2012). Previous sections explained why a belief in associative

responding based on modeling is unjustified. Now we explain why it is troubling.

The logical positions in the two scenarios are symmetrical. Each model fits. Each indicates a nominal kind of information processing. Each is mathematically abstract and psychologically neutral. In reality, neither model has standing as a description of the animal's information processing. By favoring the associative interpretation, we are proceeding by pure bias without scientific grounding. The reason why the associative model and the associative interpretation carries weight is that it brings our underlying associative bias to the fore. We accept the associative account because we prefer and believe it. The decision is not evidence based. The models' fits tell us nothing supporting this decision. As we have said (Smith, Beran, & Couchman, 2012, p. 294), "The models do not specify concrete cognitive representations, processes of interest in mind, or regions of interest in brain. They do not specify levels of intentionality or awareness. They are psychologically empty because they are mathematically neutral. They do not point toward a high-level or low-level description of the data."

In our view, the resolution to this issue is that interpretative *symmetry* must hold. We must treat the metacognitive and associative models equivalently. When someone asserts that a metacognitive model confirms metacognition, or that an associative model confirms associative processing, we must reject both assertions equally sharply. Neither interpretative step is justified, because one simply cannot send mathematics to do psychology's job. Not in our domain or in any other area of biobehavioral research.

Inappropriate logic in modeling situations

In fact, both assertions from modeling exemplify poor scientific inference and logic. The problem is that one can derive multiple models to reproduce a data set. We confirmed this possibility here. The animal might be responding to associative strengths, or metacognitive doubts, or reinforcement history, or risk aversion, or stimulus-specific exemplar memorization, and produce the same graph in every case. So, there is nothing intrinsically true about a model's specific inner boxology. This is a fiction—a creation of the modeler. One cannot reify that boxology by inserting it into the animal's mind. For which of these five possible processing psychologies would one so insert?

If you do reify a model's processes, you have committed the basic logical error of affirming the consequent. That is, you have asserted a conditional (If the animal's performance is associative, my associative model will fit), you have affirmed the consequent (my associative model fits), and you have concluded that the antecedent premise is true (the animal's performance is associative).

An intuitive example shows why this conclusion is a fallacy. I can assert the conditional: if it's raining, I'll have an umbrella. You may affirm the consequent—I have an umbrella. This does not confirm the rain. It might be really hot; I might have had a skin cancer diagnosed; I might be on antibiotics; I might have had my eyes dilated. An affirmed consequent never allows backward inference to a particular cause, because many causes could have led to the umbrella or the animal's performance curve. Yet this improper form of logic has been common in the animal-metacognition literature and in other research areas, too, as associative modelers have interpreted their models' fits.

It is extraordinary that we all know about the indefensibility of affirming the consequent, yet we readily enable it when associative modelers treat animal metacognition and other phenomena. Our area, perhaps other areas, too, must raise its game beyond the current logical/scientific standards of the use of formal models.

Limiting our conclusion

The associative-metacognitive debate plays out on different levels in our field. We must delimit what we are and are not concluding.

First, there is the level of the formal models that describe mathematically animals' behavior. This is where our criticism lies. The models—mathematically shallow, inferentially weak—have no real tie to associative or metacognitive processing assumptions. The models do not distinguishably reflect their underlying theories and they cannot further theoretical discussion in this area.

Second, there is the level of associative and metacognitive processing assumptions. We are not criticizing this level of the debate, or the principled ideas of associative-learning theory. Those ideas are elegant and challenging. We are not favoring the constructs of animal metacognition. These ideas are new and developing. Neither psychology is yet ruled out or in. There are important issues of process and representation still to be resolved. To be clear, we think this level of the debate is scientifically strong and highly productive.

Third, there is a philosophy-of-science level to the debate that can be summarized as follows. Perhaps the "bias" favoring associative models has a basis. Associative models have been successful, applied to many phenomena. The constructs of associationism have naturally gained currency and popularity. Isn't it the way of science that popular and useful constructs are preferred? These constructs are also familiar and comfortable to comparative researchers, another possible reason to grant them descriptive privilege. Associative learning is a simple processing idea producing a simple model, another possible reason for preference. And some animal researchers would apply Morgan's Canon or Ockham's Razor to animal-

metacognition findings, again giving associative processing descriptive privilege.

Our conclusions about modeling do not reach this philosophical level. However, reviewers suggested we give our view on these issues to benefit readers and the broader dialog.

First, the idea of uncertainty-based decisional criteria joins associative-learning theory in having great historical depth, empirical/theoretical success, and wide-spread familiarity to researchers broadly, so these rationales for the associative-learning preference fall away in our research literature.

Second, a preference based on simplicity falls away as well. It is dispiriting that associative theorists go shopping for alternative low-level interpretations as the principal ones are disconfirmed. By turns, they have suggested stimulus aversion/avoidance, reinforcement history, reward maximization, latency, associative connections to dithering behavior, and so forth. There is nothing simple about constantly revisiting the associative-apps store, or trying to guess which of many possible associative cues the animal might be responding to in any specific case, or determining why and how it could continually switch among them. In contrast, if one grants animals a simple uncertainty-monitoring system, one explains performance in many tasks. A generalized uncertainty state will apply to tasks of perception, memory, foraging, numerosity, timing, and so forth. This is easily as simple an interpretation as the associative description provides.

Third, a preference based on parsimony falls away too. Humans and animals produce nearly identical graphs in some uncertainty tasks (e.g., Shields et al., 1997). It is unparsimonious to interpret humans' performance metacognitively but animals' performance associatively—it multiplies mechanisms inelegantly. In no other case we know of, be this younger/older children or younger/aged adults, would this sharp divergence follow from the same data pattern. Instead, one would naturally interpret similar performances similarly. This is even truer because monkeys and humans share evolutionary histories and homologous brain structures. The parsimonious interpretation when monkeys perform uncertainty tasks similarly to humans is that the psychological processes are similar (De Waal, 1991; Smith, Couchman et al., 2012; Sober, 2012). We think it is quite plausible that evolution would have given multiple species the adaptive capacity to manage uncertainty.

Thus, our view is that it is a poor choice to use any historical or simplicity consideration to tie-break models or to guide psychological interpretation. This is especially true in a new research area. Animal-metacognition research is about (potentially) opening a new window on animals' minds. Why would you pre-judge what you will see through that window? To let the popularity of associative models be the tie-breaker is a worse choice. This would let faddism guide modeling and interpretation in animal-metacognition research, and we believe this is not the way of science. All of

these choices would deny empirical research the space to itself guide theoretical interpretation, when it—and not our modeling preference—should be determinative.

We would not ever base our psychological interpretation of an animal's behavior on abstract precedence or preference, especially a behavior reflecting a new facet of animal cognition (metacognition). If one understands the processes and representations the creature is using, then one does not need to rely on precedence/preference. If one does not know, then surely waiting to interpret is the proper course. Waiting, and actively exploring the matter empirically. The only issue in comparative science is which processing assumption is true, never which assumption is venerable or popular.

The experimentalist's toolkit

This article would represent a disheartening assessment if we were doubting the broad potential of the animal-metacognition field to make theoretical progress. Emphatically, we are not. Only the application of the formal models has definitively failed in this area. But the understanding of animals' uncertainty-monitoring performances does not depend on that application (Smith et al., 2014a, b). By letting the models go, one can see clearly that researchers have developed many experimental tools and paradigms for furthering that understanding.

For example, Basile, Hampton and their colleagues explored the associative cues that might underlie macaques' uncertainty performances, starting with a theoretical article (Hampton, 2009). Basile and Hampton (2014) also outlined this cue-based scientific investigation. Basile et al. (2015) evaluated several cues using a computerized task of macaques' memory monitoring. They found no evidence to support the hypotheses of behavioral cue association, rote response learning, expectancy violation, response competition, generalized search strategy, or postural mediation. Instead, they consistently found evidence for the metacognitive hypothesis. This research joins many other findings showing that animals' uncertainty monitoring transcends the associative dimensions of reinforcement, associable stimuli, and so forth (Hampton, 2001; Kornell et al., 2007; Shields et al., 1997; Smith et al., 1998; Washburn et al., 2010). The uncertainty-monitoring performances of macaques and apes especially have risen above what one could comfortably call associative responding. These species have answered negatively the associative-responding question in this domain. One sees that comparative psychologists are quietly writing the psychology of animal metacognition through empirical research, with no necessary contribution from mathematical models.

Smith and his colleagues have taken a complementary approach by asking about the cognitive level at which animals monitor and respond adaptively to uncertainty. For example, Smith et al. (2013) added a secondary task requirement to the

ongoing uncertainty-monitoring performance of macaques. The secondary task acted as a concurrent cognitive load. Smith et al. suggested that perceptual-classification responses (e.g., Sparse, Dense) would be stimulus-based and associative in the traditional sense, making few demands on working memory or executive attention. Then they would be barely affected by the load. Smith et al. suggested that uncertainty responses might be dependent on working memory or executive attention. Then they would be strongly affected by the load. In fact, in the Sparse-Uncertain-Dense task already described, concurrent tasks disrupted macaques' uncertainty responses far more than their Sparse or Dense responses. This result complements research in which humans performed memory tasks while reporting metacognitive states (Schwartz, 2008). Here, too, memory loads strongly affected metacognitive judgments, sharply decreasing tip-of-the-tongue experiences. Schwartz concluded that working memory and metamemory use similar processes, a conclusion supported by Smith et al. (2013).

Associative modelers have no way to explain the response dissociation produced by a concurrent load. One cannot pursue any interpretation that portrays the uncertainty response as an associative reaction to stimuli just as the Sparse and Dense responses are. Any such interpretation fails because the responses behave qualitatively differently when cognitive resources are occupied. Associative models even lack any way to distinguish different levels of executive and nonexecutive cognition because they are nonspecific stimulus-response models. To incorporate these results, one must grant that uncertainty responses in animals reflect some different psychological organization.

To be fair, though, one need not claim that the results show full-fledged metacognition as in humans, including consciousness and self-awareness of doubt and uncertainty. Smith et al.'s results do not grant anywhere near so generous a license.

But in this failure of reach, one sees one of the most salutary aspects of empirical research in this area. That is, current empirical work has grounded the distinctive theoretical premise that metacognition is not all-or-none. There can be a constructive theoretical middle ground wherein one grants organisms a basic uncertainty-monitoring capacity without overinterpreting that capacity. In this middle ground may lie the evolutionary roots of human metacognition and its (nonverbal) ontogenetic roots in human children. Then one sees that the behavioral animal paradigms expand the range of metacognition paradigms available for testing young human children. These paradigms might also be used to explore the metacognitive capacities of children with language-delay, autism, or mental retardation. There might be more basic forms of cognitive regulation (more implicit; less language-based) that could be preserved or fostered in children who are challenged in the highest-level introspective aspects of metacognition. In all these ways, the middle ground of theory integrates comparative in an appropriate way into the mainstream of cognitive science and human psychology. Of course the middle ground demotes and sidelines the sharp metacognitive-associative debate that has not really served us that well. It takes away any need for us to force onto animal-metacognition results the associative prejudgments of last century's philosophies of what is simple or parsimonious. Finally, most relevant to the present article, it naturally lets us rid ourselves of the apparently sharply contrastive, really completely entangled, formal models of metacognitive performances.

It is not our purpose to force through some resolution of the associative debate about animal-metacognition. But it is necessary to point out that the approaches just described have distanced many of animals' metacognitive performances from available stimulus or associative cues in any traditional sense.

It is our purpose to say that the approaches just described are strong theoretical-empirical approaches for exploring the phenomena of animal metacognition. This positive statement applies equally to researchers favoring associative or metacognitive perspectives. These empirical approaches trump the weak and unscientific practice of fitting an "associative" model and claiming it fits. This approach may be psychology's dullest tool. Its use in some cases may spring from the hope to prejudge important scientific issues, a poor path forward for science. Its use in some cases may reflect the wish to gloss over empirical findings that are uncomfortable for one's preferred theory, a worse path forward for science. Therefore, this formal-modeling approach has the potential to do substantial harm in slowing the progress of this area's theoretical development, and it must be carefully reconsidered. A similar reconsideration of formal models in other areas of biobehavioral research may also be warranted.

References

Arbuthnott, G. W., Ingham, C. A., & Wickens, J. R. (2000). Dopamine and synaptic plasticity in the neostriatum. *Journal of Anatomy, 196,* 587–596. doi:10.1046/j.1469-7580.2000.19640587.x

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science, 11,* 750–760. doi:10.1111/j.1467-7687.2008.00725.x

Basile, B. M., & Hampton, R. R. (2014). Metacognition as discrimination: commentary on Smith et al. (2014). *Journal of Comparative Psychology, 128,* 135–137. doi:10.1037/a0034412

Basile, B. M., Hampton, R. R., Suomi, S. J., & Murray, E. A. (2009). An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12,* 169–180. doi:10.1007/s10071-008-0180-1

Basile, B. M., Schroeder, G. R., Brown, E. K., Templer, V. L., & Hampton, R. R. (2015). Evaluation of seven hypotheses for metamemory performance in rhesus monkeys. *Journal of Experimental Psychology: General, 144,* 85–102. doi:10.1037/xge0000031

Beran, M. J., & Smith, J. D. (2011). Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition, 120,* 90–105. doi:10.1016/j.cognition.2011.02.016

Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 111–119. doi:10.1037/0097-7403.32.2.111

Beran, M. J., Smith, J. D., & Perdue, B. M. (2013). Language-trained chimpanzees (*Pan troglodytes*) name what they have seen but look first at what they have not seen. *Psychological Science, 24,* 660–666. doi:10.1177/0956797612458936

Calabresi, P., Pisani, A., Centonze, D., & Bernardi, G. (1996). Role of Ca^{2+} in striatal LTD and LTP. *Seminars in Neuroscience, 8,* 321–328. doi:10.1006/smns.1996.0039

Call, J. (2010). Do apes know that they could be wrong? *Animal Cognition, 13,* 689–700. doi:10.1007/s10071-010-0317-x

Carruthers, P. (2008). Meta-cognition in animals: A skeptical look. *Mind & Language, 23,* 58–89. doi:10.1111/j.1468-0017.2007.00329.x

Carruthers, P., & Ritchie, J. B. (2012). The emergence of metacognition: Affect and uncertainty in animals. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of metacognition* (pp. 76–93). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199646739.003.0006

Couchman, J. J., Coutinho, M. V. C., Beran, M. J., & Smith, J. D. (2010). Beyond stimulus cues and reinforcement signals: A new approach to animal metacognition. *Journal of Comparative Psychology, 124,* 356–368. doi:10.1037/a0020129

De Waal, F. B. (1991). Complementary methods and convergent evidence in the study of primate social cognition. *Behaviour, 118,* 297–320. doi:10.1163/156853991X00337

Dunlosky, J., & Bjork, R. A. (2008). *Handbook of metamemory and memory.* New York: Psychology Press.

Fernberger, S. W. (1914). The effect of the attitude of the subject upon the measure of sensitivity. *The American Journal of Psychology, 25,* 538–543.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34,* 906–911. doi:10.1037/0003-066X.34.10.906

Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. *Current Biology, 17,* 551–555. doi:10.1016/j.cub.2007.01.061

Fujita, K. (2009). Metamemory in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12,* 575–585. doi:10.1007/s10071-009-0217-0

Gamble, E., & Koch, C. (1987). The dynamics of free calcium in dendritic spines in response to repetitive synaptic input. *Science, 236,* 1311–1315. doi:10.1126/science.3495885

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 5359–5362. doi:10.1073/pnas.071600998

Hampton, R. R. (2009). Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? *Comparative Cognition and Behavior Reviews, 4,* 17–28.

Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition, 7,* 239–246. doi:10.1007/s10071-004-0215-1

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience, 1,* 304–309. doi:10.1038/1124

Inman, A., & Shettleworth, S. J. (1999). Detecting metamemory in non-verbal subjects: A test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 389–395. doi:10.1037/0097-7403.25.3.389

Jastrow, J. (1888). A critique of psycho-physic methods. *The American Journal of Psychology, 1,* 271–309.

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. T. (2009a). Metacognition in animals: How do we know that they know? *Comparative Cognition & Behavior Reviews, 4,* 29–39. doi:10.3819/ccbr.2009.40003

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. (2009b). Reinforcement and metacognition. *Comparative Cognition & Behavior Reviews, 4,* 58–60. doi:10.3819/ccbr.2009.40007

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge: Cambridge University Press.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123,* 297–315. doi:10.1037/0096-3445.123.3.297

Kornell, N. (2009). Metacognition in humans and animals. *Current Directions in Psychological Science, 18,* 11–15. doi:10.1111/j.1467-8721.2009.01597.x

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science, 18,* 64–71. doi:10.1111/j.1467-9280.2007.01850.x

Le Pelley, M. E. (2012). Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 686–708. doi:10.1037/a0026478

Le Pelley, M. E. (2014). Primate polemic: Commentary on Smith, Couchman, and Beran (2014). *Journal of Comparative Psychology, 128,* 132–134. doi:10.1037/a0034227

MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J., & Barker, J. L. (1986). NMDA-receptor activation increases cytoplasmic calcium concentration in cultured spinal cord neurones. *Nature, 321,* 519–522. doi:10.1038/321519a0

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Metcalfe, J. (2008). Evolution of metacognition. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 29–46). New York: Psychology Press.

Nelson, T. O. (1992). *Metacognition: Core readings.* Needham Heights, MA: Allyn and Bacon.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51,* 102–116. doi:10.1037/0003-066X.51.2.102

Paukner, A., Anderson, J. R., & Fujita, K. (2006). Redundant food searches by capuchin monkeys (*Cebus apella*): A failure of meta-cognition? *Animal Cognition, 9,* 110–117. doi:10.1007/s10071-005-0007-2

Roberts, W. A., Feeney, M. C., McMillan, N., MacPherson, K., Musolino, E., & Petter, M. (2009). Do pigeons (*Columba livia*) study for a test? *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 129–142. doi:10.1037/a0013722

Schultz, W. (1992). Activity of dopamine neurons in the behaving primate. *Seminars in Neuroscience, 4,* 129–138. doi:10.1016/1044-5765(92)90011-P

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. *Memory & Cognition, 36,* 9–19. doi:10.3758/MC.36.1.9

Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental*

Psychology: General, 126, 147–164. doi:10.1037/0096-3445.126.2.147

Smith, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science, 13,* 437–442. doi:10.1111/1467-9280.00477

Smith, J. D. (2009). The study of animal metacognition. *Trends in Cognitive Sciences, 13,* 389–396. doi:10.1016/j.tics.2009.06.009

Smith, J. D., & Minda, J. P. (2001). Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 984–1002. doi:10.1037/0278-7393.27.4.984

Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 800–811. doi:10.1037/0278-7393.28.4.800

Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, W. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General, 127,* 227–250. doi:10.1037/0096-3445.127.3.227

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26,* 317–339. doi:10.1017/S0140525X03000086

Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. *Journal of Experimental Psychology: General, 135,* 282–297. doi:10.1037/0096-3445.135.2.282

Smith, J. D., Beran, M. J., Couchman, J. C., & Coutinho, M. V. C. (2008a). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review, 15,* 679–691. doi:10.3758/PBR.15.4.679

Smith, J. D., Redford, J. S., & Haas, S. M. (2008b). Prototype abstraction by monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: General, 137,* 390–401. doi:10.1037/0096-3445.137.2.390

Smith, J. D., Redford, J. S., Beran, M. J., & Washburn, D. A. (2010). Rhesus monkeys (*Macaca mulatta*) adaptively monitor uncertainty while multi-tasking. *Animal Cognition, 13,* 93–101. doi:10.1007/s10071-009-0249-5

Smith, J. D., Beran, M. J., & Couchman, J. J. (2012a). Animal metacognition. In E. A. Wasserman & T. R. Zentall (Eds.), *The Oxford handbook of comparative cognition* (2nd ed., pp. 282–304). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780195392661.013.0016

Smith, J. D., Couchman, J. J., & Beran, M. J. (2012b). The highs and lows of theoretical interpretation in animal-metacognition research. *Philosophical Transactions of the Royal Society B, 367,* 1297–1309. doi:10.1098/rstb.2011.0366

Smith, J. D., Coutinho, M. V. C., Church, B. A., & Beran, M. J. (2013). Executive-attentional uncertainty responses by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: GeneraI, 142,* 458–475. doi:10.1037/a0029601

Smith, J. D., Couchman, J. J., & Beran, M. J. (2014a). Animal metacognition: A tale of two comparative psychologies. *Journal of Comparative Psychology, 128,* 115–131. doi:10.1037/a0033105

Smith, J. D., Couchman, J. J., & Beran, M. J. (2014b). A tale of two comparative psychologies: Reply to commentaries. *Journal of Comparative Psychology, 128,* 140–142. doi:10.1037/a0034784

Sober, E. (2012). Anthropomorphism, parsimony, and common ancestry. *Mind & Language, 27,* 229–238.

Staddon, J. E. R., Jozefowiez, J., & Cerutti, D. T. (2007). Metacognition: A problem not a process. *PsyCrit.* http://www.psycrit.com/Targets/2007-Foote/2007-03-15_Staddon.

Suda-King, C. (2008). Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Animal Cognition, 11,* 21–42. doi:10.1007/s10071-007-0082-7

Sutton, J. E., & Shettleworth, S. J. (2008). Memory without awareness: Pigeons do not show metamemory in delayed matching to sample. *Journal of Experimental Psychology: Animal Behavior Processes, 34,* 266–282. doi:10.1037/0097-7403.34.2.266

Templer, V. L., & Hampton, R. R. (2012). Rhesus monkeys (*Macaca mulatta*) show robust evidence for memory awareness across multiple generalization tests. *Animal Cognition, 15,* 409–419. doi:10.1007/s10071-011-0468-4

Washburn, D. A., Smith, J. D., & Shields, W. E. (2006). Rhesus monkeys (*Macaca mulatta*) immediately generalize the *uncertain* response. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 185–189. doi:10.1037/0097-7403.32.2.185

Washburn, D. A., Gulledge, J. P., Beran, M. J., & Smith, J. D. (2010). With his memory magnetically erased, a monkey knows he is uncertain. *Biology Letters, 6,* 160–162. doi:10.1098/rsbl.2009.0737

Wickens, J. (1993). *A theory of the striatum.* New York: Pergamon Press.

Zakrzewski, A. C., Perdue, B. M., Beran, M. J., Church, B. A., & Smith, J. D. (2014). Cashing out: The decisional flexibility of uncertainty responses in rhesus macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Learning and Cognition, 40,* 490–501. doi:10.1037/xan0000041