



Comparative Approaches to Metacognition: Prospects, Problems, and the Future

Jonathon D. Crystal

Department of Psychological and Brain Sciences, Indiana University, Bloomington

*Corresponding author (Email: jcrystal@indiana.edu)

Citation – Crystal, J. D. (2019). Comparative approaches to metacognition: Prospects, problems, and the future. *Animal Behavior and Cognition*, 6(4), 254-261. <https://doi.org/10.26451/abc.06.04.05.2019>

Abstract – This article describes the author’s perspective on research in comparative approaches to metacognition. A brief review highlights the research that has informed this perspective. The prospects for making progress in comparative metacognition are outlined. Next, the article describes problems that any research program on comparative metacognition needs to address. Some recommendations for future research are offered.

Keywords – Metacognition; Comparative Metacognition, Uncertainty Monitoring, Metamemory, Quantitative Modeling

Introductory Perspective

I became interested in metacognition when I talked to Alastair Inman, who was a postdoctoral fellow at the University of Toronto in the laboratory of Sara Shettleworth, while he was conducting research on the costs and benefits of remembering. They published their research (Inman & Shettleworth, 1999) suggesting that pigeons may show signs of metacognition, although later research (Sutton & Shettleworth, 2008) has raised questions about this conclusion. My interest solidified when I read Rob Hampton’s initial work on metacognition (Hampton, 2001). Hampton’s (2001) approach (see Figure 1) came to be considered the gold standard for documenting metacognition in nonhumans, in part because the work carefully considered how simple alternative explanations could potentially explain putative metacognition data without positing metacognition. The field developed from earlier research (for an early review, see Smith, Shields, & Washburn, 2003) that primarily focused on the uncertainty response. Notably, Smith and colleagues’ seminal research showed that monkeys used the uncertainty response in adaptive ways, consistent with metacognition.

Hampton (2001) argued that an animal that is able to discriminate the presence of memory from the absence of memory would improve its accuracy if it was allowed to decline memory tests when the memory was forgotten. I will refer to the chosen-forced performance advantage as accuracy divergence. Like earlier research that focused on uncertainty monitoring, Hampton also noted that such an animal should decline tests most frequently when memory is attenuated or weak. According to the metacognition proposal, being forced to take a test will result in lower average accuracy because forced tests include trials that would have been declined had that opportunity been available.

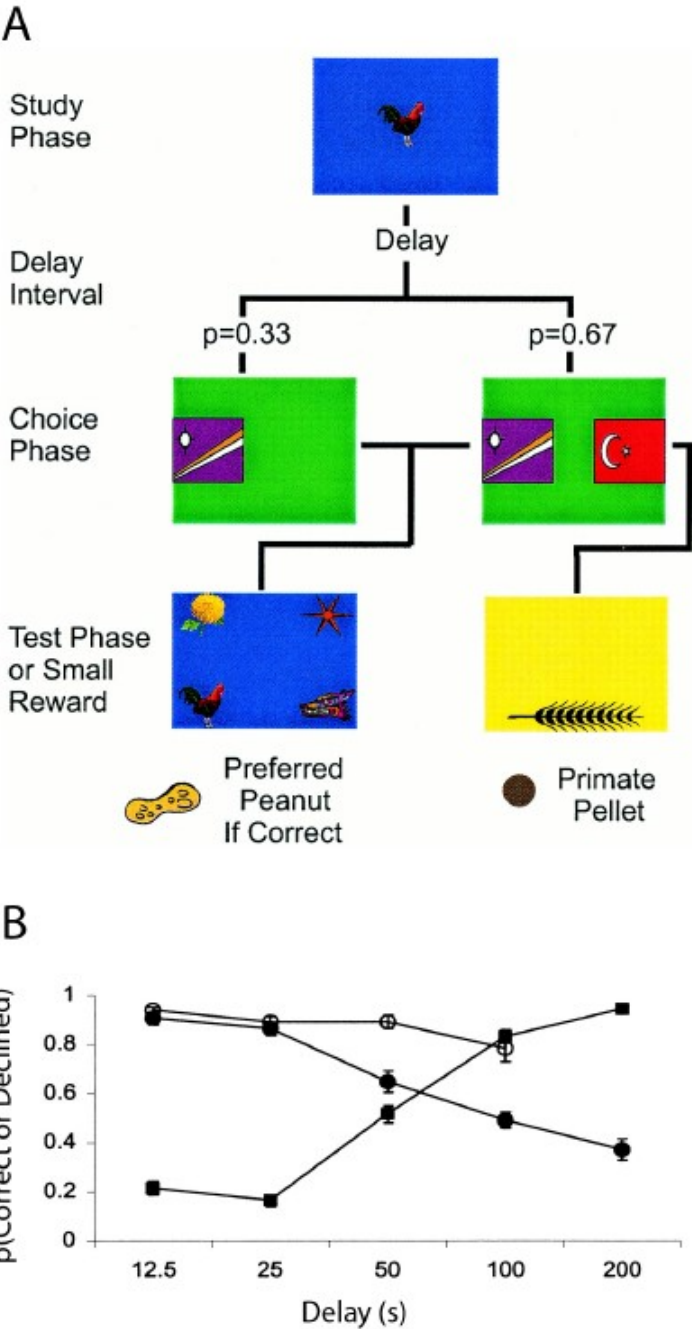


Figure 1. Schematic representation of design of study and data for monkeys. A. Procedure (top panel; Hampton, 2001): After presentation of a clip-art image to study and a retention-interval delay, a choice phase provided an opportunity for taking or declining a memory test; declining a test produced a guaranteed but less preferred reward than was earned if a test was selected and answered correctly (test phase); no food was presented when a distracter image was selected in the memory test. Items were selected by contacting a touch-sensitive computer monitor. B. Data (bottom panel; Hampton, 2001): Performance from a monkey that both used the decline response to avoid difficult problems (i.e., relatively long retention intervals) and had a chosen-forced performance advantage that emerged as a function of task difficulty (i.e., accuracy was higher on trials in which the monkey chose to take the test compared with forced tests, particularly for difficult tests). Filled squares represent the proportion of trials declined, and filled and unfilled circles represent proportion correct on forced and chosen trials, respectively. Error bars represent standard errors. (Adapted from Hampton, R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 5359-5362. © 2001 The National Academy of Sciences. Reprinted with permission.)

At the time my laboratory was conducting experiments on episodic memory in rats. The observation that rats may have episodic memory (Babb & Crystal, 2005, 2006a, b; for a recent review, see Crystal, 2018) raised the prospect that it might be possible to show that rats had metacognition. To that end, we adapted Hampton's (2001) experimental design with monkeys for an experiment with rats.

In our study (Foote & Crystal, 2007), trials consisted of three phases: study, choice, and test phases (Figure 2A). In the study phase, a brief noise was presented for the rat to classify as short (2 - 3.62 s) or long (4.42 - 8 s). Stimuli with intermediate durations (e.g., 3.62 and 4.42 s) are more difficult to classify as short or long than are more widely spaced intervals (e.g., 2 and 8 s). In the choice phase, the rat was sometimes presented with two response options, signaled by the illumination of two nose-poke apertures. On these choice-test trials, a response in one of these apertures (referred to as a take-the-test response) led to the insertion of two response levers in the subsequent test phase; one lever was designated as the correct response after a short noise, and the other lever was designated correct after a long noise. The other aperture (referred to as the decline-the-test response) led to the omission of the duration test. On other trials in the choice phase, the rat was presented with only one response option; accordingly, the rat was required to select the aperture that led to the duration test on these forced-test trials because the option to decline the test was not available. In the test phase, a correct lever press with respect to the duration discrimination produced a large reward of 6 pellets; an incorrect lever press produced no reward. A decline response (provided that this option was available) led to a guaranteed, but smaller, reward of 3 pellets.

We found (Figure 2B) that (1) the rate of declining to take the test increased as the difficulty of the discrimination increased and (2) accuracy declined as the difficulty of the discrimination increased, but this decline was greater when the rats were forced to take the test compared to trials on which the rats chose to take the test (Foote & Crystal, 2007).

Prospects

If an animal possesses knowledge about whether it knows or does not know the answer to a test, it would be expected to decline most frequently on difficult tests and show lowest accuracy on difficult tests that cannot be declined. Foote and Crystal's (2007) data provide evidence for both predictions in rats, which suggested that a non-primate has knowledge of its own cognitive state. The prospects for demonstrating metacognition appeared promising. At the time, we had planned to test for immediate generalization of the decline response in transfer tests, which had recently been demonstrated in monkeys (Kornell, Son, & Terrace, 2007; Washburn, Smith, & Shields, 2006). However, development of non-metacognition models (as described in the section below) undermines the prospects.

My view about comparative approaches to metacognition is that we should start with the hypothesis that the animal has an array of psychological processes (e.g., working memory, executive function, episodic memory, associative learning, and others)¹. Next, we should compare that proposal with the hypothesis that the animal has the same array of psychological processes plus one other, namely metacognition. My reading of the literature at the time, based on (Hampton, 2001; Inman & Shettleworth, 1999), was that the field was in exactly such a position. The combination of increased use of a decline response and a corresponding accuracy divergence as difficulty increased was purported to be explained by metacognition (but not by an array of non-metacognition psychological processes). I came to believe that the field was not in such a position based on an article published by Smith, Beran, Couchman and Coutinho (2008). Smith et al. (2008) showed that the hallmark features of metacognition (decline rate and accuracy divergence) could be produced by a non-metacognition model.

¹ Of course, each psychological process is subject to empirical test. The point here is that the only way to determine if metacognition needs to be claimed is to assume an array of other processes alone vs. the same set of processes plus metacognition.

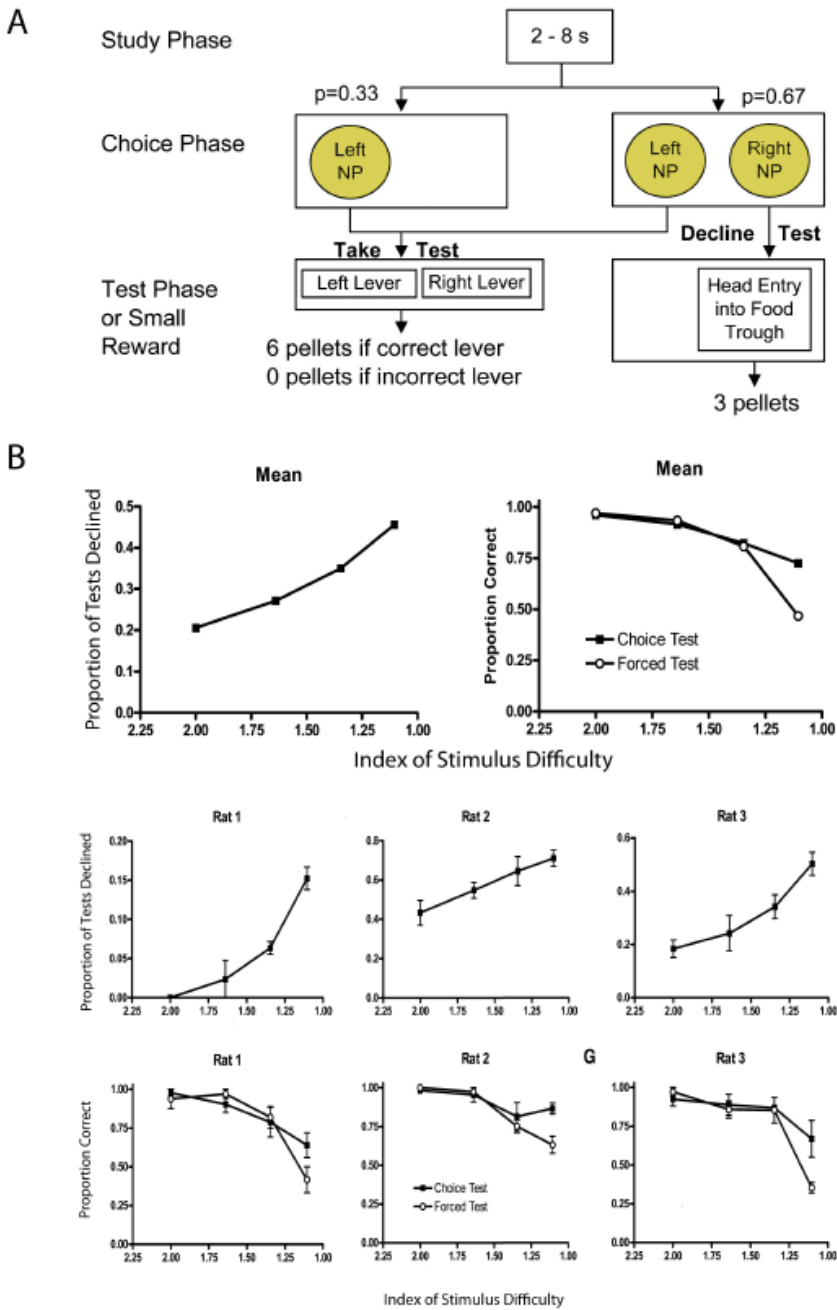
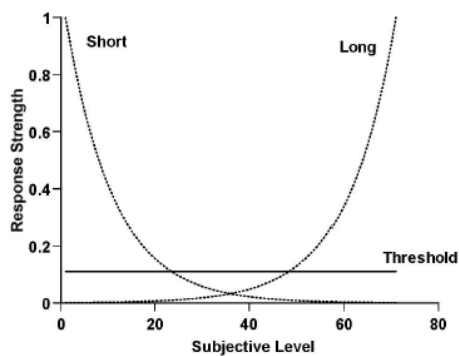


Figure 2. Schematic representation of design of study and data for rats. A. Procedure (left panel; Foote & Crystal, 2007): After presentation of a brief noise (2-8 s; study phase), a choice phase provided an opportunity for taking or declining a duration test; declining a test produced a guaranteed but smaller reward than was earned if a test was selected and answered correctly (test phase). The yellow shading indicates an illuminated nose-poke (NP) aperture, used to decline or accept the test. B. Data (Foote & Crystal, 2007): Performance from three rats (bottom panels) and the mean across rats (top panels). Difficult tests were declined more frequently than easy tests; difficulty was defined by proximity of the stimulus duration to the subjective middle of the shortest and longest durations). The decline in accuracy as a function of stimulus difficulty was more pronounced when tests could not be declined (forced test) compared to tests that could have been declined (choice test). Error bars represent standard errors. (Adapted from Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. *Current Biology*, 17, 551-555. © 2007 by Elsevier Ltd. Reprinted with permission.)

Problems

Smith and colleagues (2008) proposed that reward of the decline response produces a low-frequency tendency to select that response independent of the stimulus in the primary discrimination. Accordingly, the decline response has a constant attractiveness across the stimulus continuum according to Smith and colleagues' proposal; constant attractiveness means that the tendency to produce the response is constant across varying situations. We refer to this class of threshold explanations as a stimulus-independent hypothesis to contrast it with a stimulus-response hypothesis, according to which the animal learns to select the decline response in a specific stimulus condition. The stimulus-independent hypothesis turns out to be a powerful explanation (Crystal, 2014; Crystal & Foote, 2009a, b, 2011; Foote & Crystal, 2012). For the primary discrimination, Smith et al. adopted standard assumptions about exponential decay of a stimulus (i.e., generalization decrements for an anchor stimulus in a trained discrimination). According to the proposal, the primary discrimination and the decline option give rise to competing response strengths. Smith et al. proposed a winner-take-all response rule (i.e., the behavioral response on a given trial is the one with the highest response strength). A schematic of the formal model appears in Figure 3A. Simulations with this quantitative model document that it can produce both increased use of the decline response and accuracy divergence as difficulty increases (Figure 3B). Use of the decline response effectively allows the subject to avoid difficult problems, and the development of superior performance on chosen trials, relative to forced trials, emerges as a function of task difficulty. Notably, both empirical aspects of putative metacognition data are produced by the simulation (Figure 3B) without the need to propose that the animal 'knows when it does not know' or any other metacognitive process.

A



B

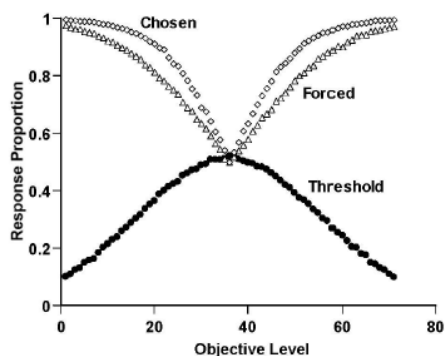


Figure 3. Schematic of Smith and colleagues' (2008) low-level, response-strength model and simulation.

A. Presentation of a stimulus gives rise to a subjective level or impression of that stimulus. Each response has a hypothetical response strength for any given subjective level. The schematic outlines response strengths for two primary responses in a two-alternative forced-choice procedure and for a third (i.e., decline or uncertainty) response (labeled threshold). Note that response strength is constant for the third response (i.e., it is stimulus independent). By contrast, response strength is highest for the easiest problems (i.e., the extreme subjective levels). Note also that the decline-response strength is higher than the other response strengths for the most difficult problems (i.e., middle subjective levels). Note that Smith and colleagues' (2008) specifically proposed the response-strength model as a low-level, non-metacognition model. Reproduced from Smith et al. (2008).

B. Simulation of schematic shown in A. Simulation of Smith and colleagues' (2008) response-strength model with a constant threshold produces apparently functional use of the decline response (i.e., intermediate, difficult stimuli are declined more frequently than easier stimuli). The choice-forced performance advantage emerges as a function of stimulus difficulty despite the absence of metacognition in the model. Reproduced from Smith et al. (2008). (From Smith, J. D., Beran, M. J., Couchman, J. J., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review*, 15, 679-691. © 2008 Psychonomic Society. Reprinted by permission.)

The non-metacognitive model proposed by Smith and colleagues (2008) may be used to predict the patterns of data reported by Hampton (2001). The account is similar to that depicted in Figure 3A, except a single exponential function is used to model the strength of decaying memory traces (Crystal & Foote, 2009a). Moreover, the performance of Hampton's monkeys (that were tested with novel retention interval delays and omission of the to-be-studied item) is also predicted by Smith and colleagues' non-metacognition model (Crystal & Foote, 2009a). In addition, efforts to avoid direct reward of the uncertainty response and obscure feedback by delaying and scrambling the order of rewards for the primary task (e.g., Beran, Smith, Redford, & Washburn, 2006; Smith, Beran, Redford, & Washburn, 2006) can also be explained by application of Smith and colleagues' (2008) non-metacognition model. For an extended discussion of these issues, see Crystal and Foote (2009a).

In addition to Smith and colleagues' (2008) non-metacognition model, other non-metacognition models have been proposed. Some approaches (Jozefowicz, Staddon, & Cerutti, 2009; Staddon, Jozefowicz, & Cerutti, 2007) are tightly linked to the procedures used to produce putative metacognition data (Foote & Crystal, 2007). Other non-metacognition models are more similar to the Smith et al. (2008) model in that they are more general purpose in orientation (rather than being tightly linked to specific procedures). In particular, the non-metacognition model described by Le Pelley (Le Pelley, 2012, 2014) is based on basic properties of associative learning, and it can be applied to diverse experimental designs. Notably, Le Pelley (2012, 2014) showed that this type of non-metacognition model can produce the hallmark patterns of putative metacognition data published by Foote and Crystal (2007), Couchman and colleagues (Couchman, Coutinho, Beran, & Smith, 2010), Hampton (2001), and others. It is noteworthy that Le Pelley's model is not meant to describe an ideal associative learning model. Accordingly, although the identification of an ideal model of associative learning is of interest (primarily for students of associative learning), the issue for comparative models of metacognition is to validate claims that hallmark features of putative metacognition data can, in principle, be produced by a non-metacognition mechanism.²

The Future

What are the prospects for the future? Some reactions to the concerns noted above and similar concerns are focused on finding flaws in the specific non-metacognition proposals (e.g., Smith & Church, 2018; Smith, Couchman, & Beran, 2012, 2014). Although there is merit to optimizing specific proposals, the purpose of models such as Smith and colleagues' (2008) and Le Pelley's (Le Pelley, 2012) are not to work out the specifics of non-metacognition psychological processes. Instead, the value of these non-metacognition models is to highlight that metacognition need not be proposed to produce putative metacognitive data. Similarly, a productive approach to resolving the controversy will not focus on asserting that a consensus has emerged for the existence of metacognition in some animals.

Other reactions focus on clarifying and testing specific alternative proposals (Basile, Schroeder, Brown, Templer, & Hampton, 2014). I believe that this is a more productive approach for resolving the controversy. Here, I offer a complementary approach.

General-purpose non-metacognitive models (Le Pelley, 2012; Smith et al., 2008) are available to test any proposed new method of documenting metacognition; doubt would be registered for the new test of metacognition if a non-metacognition model can produce the putative metacognition data. The field of comparative metacognition would benefit from the development of computational models of

² The issue is not resolved by claiming that it is not necessary to deflect claims from other alternative perspectives. It is a basic requirement of experimental design that when a confounding variable is identified (i.e., alternative explanation), it is necessary to test such explanations as the route to vindicating the initial claim. The Smith et al. (2008) model was offered by the authors to show that indeed simpler explanations could produce putative metacognition data. Similarly, Le Pelley (2012) used basic principles of associative learning to produce putative metacognition data. Moreover, the issue is not resolved by noting that some species pass a test of metacognition but that other animals do not despite significant effort to provide opportunities to pass such tests; the absence of evidence is not evidence of absence, and innovations may change long-held conclusions (e.g., Beran, Perdue, Church, & Smith, 2016; Bramlett, Perdue, Evans, & Beran, 2012).

metacognition; that is, explicit identification of the stages of information processing that are proposed to implement genuine metacognition. Computational modelers, armed with a selection of metacognition and non-metacognition models, could then help the field to identify new experimental procedures that can be uniquely explained by metacognition (without also being explained by a non-metacognition model). The goal is to have a well-described model of metacognition to compare with available models of non-metacognition. Thus, a head-on-head comparison could be used to select between the competing models.

Conclusions

In principle, evidence for metacognition can be obtained in animals. I outline the standards that should be applied in evaluating putative evidence for metacognition. I will emphasize that the concerns noted here are not meant to stifle progress in the field. Instead, I believe that a clear-eyed assessment is needed to determine what is lost by the advent of general models of non-metacognition that predict putative hallmark features of metacognition. Ultimately, recognizing the problems faced in the field of comparative metacognition may unleash new creative approaches that resolve the controversy.

Acknowledgments

Supported by NIA grant AG053524.

References

- Babb, S. J., & Crystal, J. D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning & Motivation, 36*, 177-189.
- Babb, S. J., & Crystal, J. D. (2006a). Discrimination of what, when, and where is not based on time of day. *Learning & Behavior, 34*, 124-130.
- Babb, S. J., & Crystal, J. D. (2006b). Episodic-like memory in the rat. *Current Biology, 16*, 1317-1321.
- Basile, B. M., Schroeder, G. R., Brown, E. K., Templer, V. L., & Hampton, R. R. (2014). Evaluation of seven hypotheses for metamemory performance in rhesus monkeys. *Journal of Experimental Psychology: General, 144*, 85-102.
- Beran, M. J., Perdue, B. M., Church, B. A., & Smith, J. D. (2016). Capuchin monkeys (*Cebus apella*) modulate their use of an uncertainty response depending on risk. *Journal of Experimental Psychology: Animal Learning and Cognition, 42*, 32-43.
- 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.
- Bramlett, J. L., Perdue, B. M., Evans, T. A., & Beran, M. J. (2012). Capuchin monkeys (*Cebus apella*) let lesser rewards pass them by to get better rewards. *Animal Cognition, 15*, 963-969.
- Couchman, J. J., Coutinho, M. V., 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.
- Crystal, J. D. (2014). Where is the skepticism in animal metacognition? *Journal of Comparative Psychology, 128*, 152-154.
- Crystal, J. D. (2018). Animal models of episodic memory. *Comparative Cognition & Behavior Reviews, 13*, 105-122.
- Crystal, J. D., & Foote, A. L. (2009a). Metacognition in animals. *Comparative Cognition & Behavior Reviews, 4*, 1-16.
- Crystal, J. D., & Foote, A. L. (2009b). Metacognition in animals: Trends and challenges. *Comparative Cognition & Behavior Reviews, 4*, 54-55.
- Crystal, J. D., & Foote, A. L. (2011). Evaluating information-seeking approaches to metacognition. *Current Zoology, 57*, 531-542.
- Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. *Current Biology, 17*, 551-555.
- Foote, A. L., & Crystal, J. D. (2012). "Play it again": A new method for testing metacognition in animals. *Animal Cognition, 15*, 187-199.

- 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.
- Inman, A., & Shettleworth, S. J. (1999). Detecting metamemory in nonverbal subjects: A test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 389–395.
- Jozefowicz, J., Staddon, J. E. R., & Cerutti, D., T. (2009). Metacognition in animals: How do we know that they know? *Comparative Cognition & Behavior Reviews*, 4, 29–39.
- Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, 18, 64–71.
- 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.
- Le Pelley, M. E. (2014). Primate polemic: Commentary on Smith, Couchman, and Beran (2014). *Journal of Comparative Psychology*, 128, 132–134.
- Smith, J. D., Beran, M. J., Couchman, J. J., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review*, 15, 679–691.
- 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.
- Smith, J. D., & Church, B. A. (2018). Dissociable learning processes in comparative psychology. *Psychonomic Bulletin & Review*, 25, 1565–1584.
- Smith, J. D., Couchman, J. J., & Beran, M. J. (2012). The highs and lows of theoretical interpretation in animal-metacognition research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1297–1309.
- Smith, J. D., Couchman, J. J., & Beran, M. J. (2014). Animal metacognition: A tale of two comparative psychologies. *Journal of Comparative Psychology*, 128, 115–131.
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, 26, 317–373.
- Staddon, J. E. R., Jozefowicz, J., & Cerutti, D. (2007). Metacognition: A problem not a process. *PsyCrit*, April 13, 2007.
- 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, 266–82.
- 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.