



Can Comparative Psychology Crack its Toughest Nut?

Daniel J. Povinelli

Department of Biology, University of Louisiana

Corresponding author (Email: daniel.povinelli@louisiana.edu)

Citation – Povinelli, D.J. (2020). Can comparative psychology crack its toughest nut? *Animal Behavior and Cognition*, 7(4), 589-652. <https://doi.org/10.26451/abc.07.04.09.2020>

Abstract – What is the likelihood that humans will ever determine if other animals engage in higher-order thinking? In examining what has happened in the twenty years since the publication of our book, *Folk Physics for Apes*, I conclude that comparative psychologists, the academic stalwarts charged with making progress on this front, are stuck in a series of intractable, and largely unacknowledged, conceptual problems. Because higher-order mental states depend on the existence of first-order, perceptually-based representations of objects and events, and because those first-order representations are necessary and sufficient to explain current experimental and observational results, the approaches deployed by comparative psychologists are doomed to failure. I examine this *Asymmetric Dependency Problem* in detail and show how the failure to confront its implications leads to viciously circular arguments that cannot be fixed within the current paradigm of research. Next, I offer a seven-step method for isolating the common structural flaw in any given experiment, and work through several examples. Finally, I examine the central claims that my colleagues and I made in *Folk Physics* through the lens of the *Asymmetric Dependency Problem* and current research trends. Although the optimism we expressed that experimental approaches could implicate the presence of higher-order thinking in animals requires considerable dampening, the challenges we isolated remain as vital today as they were twenty years ago.

Keywords – Folk physics, Theory of Mind, Chimpanzees, Comparative Psychology, Higher-order Cognition

Candy's Coconut

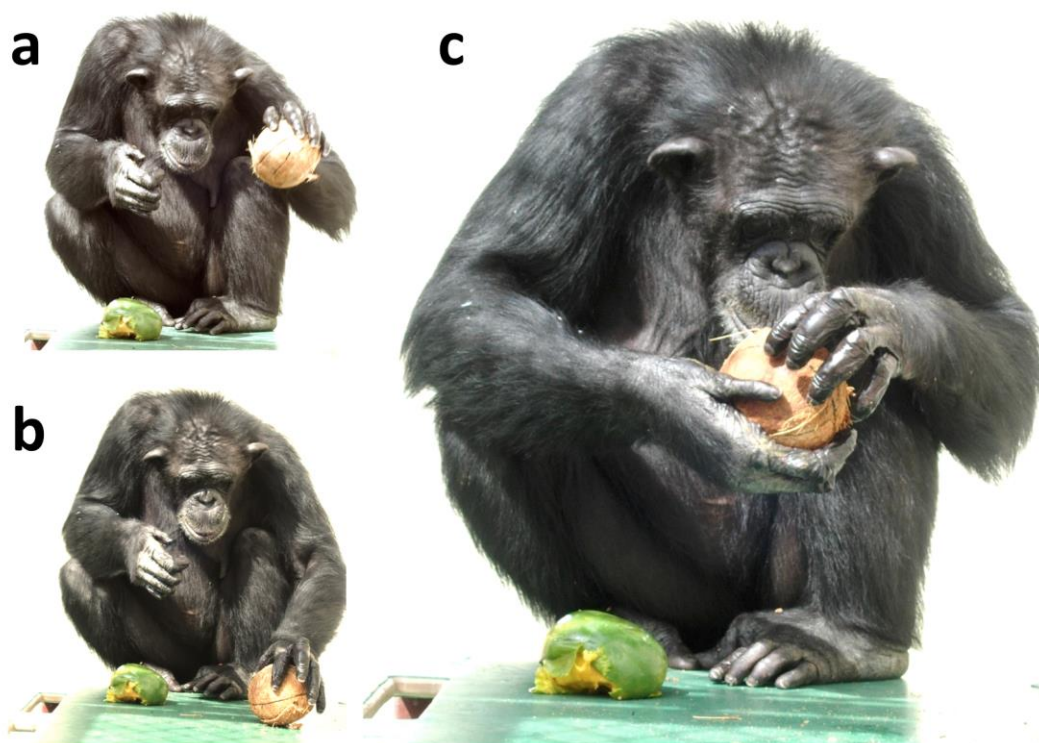
Consider Candy as she cracks a coconut and consumes its contents (see Figure 1). What kinds of mental events precede, accompany, and follow from her actions? Twenty years ago, *Folk Physics for Apes* offered what was, and (I surmise) still remains, a controversial answer. In that monograph, my colleagues and I argued that in order to make sense of Candy's coconut capers—in particular, using it as a hammer to crush the mango visible in the lower left corner of Figure 1—we need to fully embrace the idea that she is a *mental agent*. Explaining her behavior requires positing the presence of a mind able to engage in mental operations running the gamut from keeping track of how her body can interact with the objects strewn around her, to how those objects can interact with each another. What's more, her mental representations about the world are *temporally stable*: once formed, they do not disappear. What she knows about how a coconut can crush a mango does not vanish when she looks away to monitor a squabble brewing between Brandy and Apollo. When a bird flies overhead, briefly distracting her, she does not become confused about her local goals (how to hold the coconut, where to strike the mango, or whether to eat something else altogether). No, all the precious intelligences packed in her body and brain remain alert, invisible sentinels waiting to be deployed at a moment's notice, latent potentials hungry for action.

And so, when we see Candy reach for a coconut, we are also seeing a mind hard at work—a mental system rapidly updating the possible consequences of innumerable plans. The visible movements of her body belie the intricate invisible manifold that guides her actions, a mental system subtly re-sculpted with each new experience.

Admittedly, at almost twenty years of age, this is not Candy's first rodeo busting into the stony endocarp of a dry drupe (the squashed mango is the fleshy outcome of her latest ride). Nor, more generally, is it the first time she has reaped the rewards of accelerating one object into another. Indeed, *Folk Physics* (and its sequel, *World Without Weight*) carefully tracked how Candy and her six companions acquired numerous skills involving the manipulation of tools and other objects. *Folk Physics* and *Weight* reported the results of almost 60 experiments designed to target the nature of our apes' representations of specific object properties and their relations. My colleagues and I then walked Candy and her peers through them, each task programmatically building on the previous ones, the entire effort aimed at zeroing in on the mental skills that supported her developing mastery of using the objects around her. Seems simple enough.

Figure 1

Candy uses a Coconut as a Tool, then Cracks it Open



Note. Having already used a coconut to crush a mango, Candy (a) raises the coconut, (b) cracks it open, and (c) investigates the milk spilling from it.

Looking back, Candy's journey reminds us that not all of her functional potentials were up and running from the get-go. Given what we know about the slow and steady development of the

tool-using skills in wild-born chimpanzees (see Boesch & Boesch-Achermann, 2000; Boesch et al., 2019), this should not be the least bit surprising. To achieve mastery, Candy required *experience*. But her journey also underscores the patent fact that she was born into the world with a core set of tightly canalized developmental pathways already in motion, pathways that, as an infant, brought her into contact with objects other than her body. At first, these contacts were primarily buccal, but shortly thereafter they became more manual. These mouthing and grasping activities established dynamical starting points for hefting, banging, sliding, pulling, throwing, etc., as well as scaffolded her growing sensitivities to the effects that her actions were causing in the world around her. These experiences, then, cannot be thought of as distinct from Candy's natural range of behaviors relevant to tool use. *The experiences are themselves behaviors*. They are an inextricable part of her ever-expanding manifold of object-oriented behavior.¹

Importantly, Candy's behavioral potentials did not develop in isolation from her social world. From the beginning, her caretakers and peers played a strong role in her interactions with objects. The precise impact of these social inputs remains highly contested (see Bernstein-Kurtycz et al., 2020). Indeed, some of the most seemingly rock-solid folk psychological assumptions about the causal role of maternal input in the development of tool use in wild chimpanzees have collapsed under close empirical scrutiny (Estienne et al., 2019, and footnote 72). Nonetheless, somehow Candy's direct handling of objects, along with whatever she extracts by watching others do the same, come together to generate a competent tool-using animal.

By the time she was a year-and-a-half old, for example, Candy and her companions had invented familiar social routines involving balls and sticks and blankets and plastic blocks. Just like their wild counterparts, they used these objects to bait each other into playful bouts of chase-and-tickle until it inevitably devolved into the familiar rough-and-tumble horseplay so common among young chimpanzees everywhere (see Goodall, 1986). Thus, through the indivisible combination of her starting resources and her ongoing experiences, Candy was forming abstract representations of the role that objects and actions can play in orchestrating her goal-directed interactions with her peers and the world around her.

I do not write *abstract* whimsically. Any robust theory of action understanding—human or otherwise—must handle the *many-to-one* mapping of sensory inputs onto mental representations that can cope with the superficial differences of particular situations. Thus, as Candy lifts her coconut to smash her mango, she is bringing to bear the full weight of her abstracted set of concepts and operations that allow her to get to the heart of the matter. A novel dry drupe does not fool her. Simply put, after Candy learns to crack open a pecan, she is not left speechless confronting a walnut.

At this point, the reader may be confused: *Why is any of this controversial? Humans possess and operate on the basis of such well-studied cognitive resources all the time. Surely animals do as well.* Indeed, the mental abilities that *Folk Physics* postulated Candy needed to possess in order to accelerate the mass of a coconut enough to squash a mango are, quite frankly, pretty cosmopolitan stuff (see Povinelli, 2000, p. 297-328). So cosmopolitan, in fact, that I decided to present the true import of *Folk Physics* backwards, to help place the main conclusion of the project in full relief. The radical thing about the book had nothing to do with its claims about the cognitive resources *available* to Candy and her fellow chimpanzees. Instead, it excited controversy because of its claims about the cognitive resources that might *not* be available to them: namely, higher-order representations of things like *<gravity>*, *<mass>*, *<force>*, *<physical connection>*, *<shape>*, *<weight>*, etc. *Folk Physics* argued that despite their ability to deftly navigate the world

¹ For an enlightening perspective of how these activities lead to simple tool use in young children, I recommend the work of Jeff Lockman and his colleagues (e.g., Jung et al., 2015; Kahrs et al., 2014). For similar investigations with chimpanzees, see Inoue-Nakamura and Matsuzawa (1997) and Hayashi and Matsuzawa (2003). For New Caledonian crows, see Kenward et al. (2006).

of objects (including Candy's use of a coconut as a hammer), chimpanzees (and, by extension, other species) might not reason about inherently intangible things:

The results of our investigations have convinced us that although chimpanzees possess an excellent ability to reason explicitly about relations between objects and events that can be perceived, they appear to know very little (if anything) about phenomena that are, in principle, 'unobservable.' Chimpanzees appear to share with us a common set of . . . processing systems which cohere a common set of object properties (such as solidity and boundedness), whereas system which map unobservable causal descriptors onto these objects and their spatial relations may be a cognitive specialization of the human species. . . . [T]his would mean that humans and chimpanzees have access to the same kinds of perceptual information (both in terms of the kinds of objects in the world, as well as information of the statistical regularities that characterize their interactions), but that the two species interpret this information differently (p. 298).

It was this idea, *derived from an overall pattern in the empirical results*, that the project sought to highlight, and it was this idea that many scholars found unpalatable. For some, the fare was all the more unappealing because it was served up on top of our previous hypothesis that chimpanzees' social interactions may not be grounded in intangible ideas, either (i.e., each other's *<mental states>*).

To be fair, the central postulate of *Folk Physics* was situated inside a network of other arguments, all of which contributed (to greater or lesser degrees) to the project's rocky reception—a reception that fomented both important questions and, I believe, enduring misconceptions.² So, when the editors of *Animal Behavior and Cognition* invited me to offer my perspective on the August 2020 special issue devoted to *Folk Physics*, I realized it was an important opportunity to step back and assess that project with twenty years of hindsight. Several key questions arose immediately. First, is it possible to isolate more precisely why *Folk Physics* was so controversial? Second, which of its empirical and theoretical claims have withstood the test of time and which have not? Finally, and, most importantly, which of the thorny issues it raised continue to impede our attempt to characterize the mental lives of our fellow residents of Planet Earth?

Stepping Back to Move Forward

For the moment, let us direct our attention away from Candy and her friends (we will return to them shortly) and instead highlight *five core conceptual quagmires* that emerged in the wake of *Folk Physics* that have not been fully and precisely understood. For sake of ease, I dub them as follows:

Asymmetric Dependency Problem. For any representationally based view of cognition, a reasoning system that contains higher-order mental states must (by definition) be grounded

² The tone of *Folk Physics* may have impeded its reception. Jim Anderson (2001) certainly thought so: "Povinelli writes that he expects an 'extreme reaction' to this project. I suspect that any such reaction will not be so much to the experiments or the interpretations in the book (these can always be revisited and tested) as to the almost gladiatorial tone in which some points are made. National Geographic and the BBC are chastised for promoting, according to Povinelli, an erroneous picture of similarity between the mental lives of apes and humans; a similar swipe is taken at Darwin... An injection of humility might have been in order here. Views can change. Povinelli's own views about chimpanzee social cognition have changed fundamentally over the last decade. It is possible that his outlook on ape folk physics might also move in another direction in the years to come" (p. 1043).

in first-order, perceptually-based mental states (for greater detail, see Penn et al., 2008a). Briefly, first-order mental states consist in atomistic perceptual symbols that stably refer to objects and events that can be directly detected by an organism's primary senses. These first-order states are abstract in the sense that they pick out (refer to) the perceptual features of events and objects that are necessary to encode the relationships that are important to the organism in question. For example, after Candy learns that a stick can be used to pull an apple to within reach, presenting her with a stick of a different thickness, or a green apple instead of a red one, will not alter her first-order, perceptually-based relational reasoning ability: *lift stick, contact apple, pull*. Likewise, the perceptually-based categories of facial expressions and body postures of others that Candy carries with her, allow her to detect and reason about the consequence of encountering particular instances of those expressions in her peers. Higher-order relational reasoning, in contrast, involves reasoning over variables and thematic roles that cut across (and are not reducible to) particular perceptual symbol and relations. In the case of physical-causal reasoning, this would involve constructs such as *<force>*, *<gravity>*, *<mass>*, *<weight>*, *<time>*, *<space>*, etc.; in the case of social reasoning, this would include constructs such as *<emotions>*, *<intentions>*, *<desires>*, *<goals>*, *<belief>* and *<knowledge>*. The inescapable implication of these uncontroversial claims is that *higher-order mental states depend upon the existence of first-order, perceptually-based mental states, but not the reverse*. The problem arises in that first-order relational mental states are both *necessary and sufficient* to explain the results of the entire (and ever-expanding) genre of experiments that have been taken to implicate the presence of higher-order states in animals. Further, if this necessity and sufficiency claim is true, this means higher-order states are not necessary to explain the results. That is, *they do no causal work above and beyond that already specified by first-order states*. While each of the conceptual-methodological problems that follow are distinct, the *Asymmetric Dependency Problem* lies at the root of all of them.

Unprincipled Titration Paradox. Experimental strategies designed to escape the implications of the *Asymmetric Dependency Problem* (even when these strategies are not explicitly recognized as being designed for this purpose), typically take the form of an arrangement of elaborate control conditions. It can easily be shown, however, that such "controls" cannot eliminate the fact that the information presented to an animal must connect to stable, perceptually abstract, first-order representations that create the *Asymmetric Dependency Problem* in the first place. This reveals the underlying conceptual paradox embedded in both the explanations offered by the experimenters, as well as the experimental methodologies they use: Experimental stimuli must somehow engage an animal's existing first-order relational reasoning without allowing the animal to rely upon such reasoning to produce the allegedly diagnostic behaviors. In practice, *this drives experimenters to develop and deploy folk scientific practices* that attempt to titrate experimental stimuli in a decidedly unprincipled manner. The folk nature of this practice is both ensconced in and obscured by the scientific-sounding claim that the experiments in question deploy *novel stimuli*. As I explain below, *experimental stimuli are never novel in ways that matter to the debate*.

Experimental Necessity Dilemma. Attempts by experimentalists to escape the pervasive implications of the *Asymmetric Dependency Problem* typically fall prey to an unacknowledged dilemma. On the one hand, they can accept its claims and cease conducting experiments aimed at detecting higher-order reasoning that fall within the genre. On the other hand, they can reject the *conclusion* of the *Asymmetric Dependency Problem* by arguing their experiments are *consistent with* the presence of higher-order thinking, even though such states are *neither necessary nor sufficient* to account for the

results. But in doing the latter, they must also accept (a) that such experiments provide no unique evidence in favor of higher-order thinking, (b) that naturalistic data provide evidence that is equally convincing (after all, naturalistic data are typically rejected by experimentalists on the grounds that such behavior could have been generated by first-order relational reasoning), and, worse still, (c) that the data *are equally consistent with* the claim that the animals do not possess higher-order states. Here, I wish to highlight the bracing implication of (b): one accepts the existing naturalistic data as sufficient evidence for higher-order reasoning in animals, then one must also accept the claim that there was no need for experiments to settle this debate in the first place.

False Contrast Models Problem. As typically described, the alternative models tested by experimentalists (and naturalists), almost invariably pit *ideological camps* (e.g., behaviorism versus higher-order thinking) against one another, rather than pitting distinct cognitive models against one another. In doing so, they avoid confronting the implications of the *Asymmetric Dependency Problem*.

Conflation of Explanatory Levels. Experimentalists and naturalists alike frequently conflate functional- and representational-levels of analysis. From a functional-level perspective, everyone agrees that animals act *as if* they possess and are sensitive to certain higher-order representations and relations (see Dennett, 1987). The confusion arises when a researcher slips (without requisite evidence) into representational-level claims that argue or imply that the entities, variables and relations posited in the functional-level models "actually play a causally efficacious role in the computations and/or psychological dynamics of the subject's cognitive system" (see Penn & Povinelli, 2009, p. 25). This conflation creates confusion about the precise import of any given experiment, and thus creates the illusion that one has avoided the far-reaching scope of the *Asymmetric Dependency Problem*.

While previously unnamed, these problems were plaguing comparative psychology long before *Folk Physics*. But in the years since that book appeared, each of them (and their relation to each other) have crystalized as distinct problems. My understanding of these core conceptual problems has emerged from intimate collaborations with close colleagues. I am not sure if all of them would agree with the precise formulations offered here, but for my part, I now see them as a set of interconnected problems that must be resolved if humans are ever going to implicate the presence of *higher-order causal reasoning* in species other than our own—the very question that animated *Folk Physics*.

My path through this essay will be as follows:

- First, I will examine how the *Asymmetric Dependency Problem* arose in the context of debates about whether chimpanzees (or any other animals) possess higher-order social cognition. I show how the *Experimental Necessity Dilemma* and the *Unprincipled Titration Paradox* naturally emerge from the failure to confront the implications of the *Asymmetric Dependency Problem*. I further show that attempts to mischaracterize the core tenets of the *Asymmetric Dependency Problem* as a form of behaviorism, have created straw man models that still plague the literature.
- Second, I will present a 7-step method that can be used to analyze any experiment through the lens of the *Asymmetric Dependency Problem*. The purpose of this method is to provide a general method of determining if any given experiment within this genre can avoid the *Asymmetric Dependency Problem* and thus potentially assay higher-order

thinking.

- Third, I will outline the uncontroversial elements of animal cognition (i.e., first-order relational reasoning). I show that while all scholars assume these abilities to be present, they almost never explicitly describe how they must be operating in the context of their experiments (see Penn et al., 2008a; Povinelli & Vonk, 2004). This folk scientific practice (i.e., not explicitly sharing commitments about the causal power of first-order mental states) empowers researchers to present *False Contrast Models* and thus create the illusion of adjudicating between the operation of first- versus higher-order cognition.

After examining these foundational challenges, I then take a *forward-looking* glance at a number of the controversial claims made in *Folk Physics*. *Have experimentalists succeeded in generating a new genre of experiments that have overcome the Asymmetric Dependency Problem? Can naturalistic data be used to settle these debates? Do we understand enough about the causal linkages between our own higher-order thinking and our behavior to develop valid models for testing for higher-order reasoning in animals? Do chimpanzees (or other animals) ask, Why? And finally, are experiments on captive animals even relevant to the debate over higher-order reasoning in animals?* Throughout, I draw on the various contributions to the August 2020 special issue of this journal as examples of current practices in the field.

My central hope is that this re-examination of *Folk Physics* will serve as an antidote to the false choices about animal minds so prevalent in the current literature—a choice that frequently feels like a clash between cartoon-like heroes and villains.³

Folk Psychology Meets *Folk Physics*

At the time *Folk Physics* was published, I was up to my eyeballs in a controversy about a (seemingly) different higher-order question—namely, whether chimpanzees possess some form of theory of mind—that is, whether they reason about inherently unperceivable mental states such as *<beliefs>*, *<desires>*, *<seeing>*, *<attention>*; Premack & Woodruff, 1978). Over time, the thorny issues raised by this question forced me to confront some foundational, overlooked flaws both in the experimental methods my colleagues and I were using, as well as the theoretical frameworks within which those methods were being deployed. The widespread failure of comparative psychology to engage with the issues that arose as a consequence of these reflections (especially the *Asymmetric Dependency Problem*) not only continues to hobble humanity's attempt to say anything meaningful about whether chimpanzees (or other species) reason about *<mental states>*, but also equivalent efforts to determine if they engage in higher-order-thinking about *<force>*, *<weight>*, *<intrinsic connection>*, *<time>*, *<space>* and so on and so forth.⁴

³ My current favorite is the ongoing battle between *Complex Cognition Woman* and *Evil Doctor Contingency Learner* (see below, "Experimentalists React" and "Current Status of Animal Folk Physics"). Like other cartoon clashes in which the superheroes and their villains repeatedly square off, however, the final outcome is never really in doubt. Superheroes are constructed in such a way that their weaknesses are always more apparent than real. And vice versa for super-villains. This does not challenge the entertainment value of this genre. As long as these movies last, dopamine guarantees the journey will be intoxicating (cf. Christiansen, 2018; Jo et al., 2018).

⁴ Although some scholars (including myself) sometime divide such reasoning into distinct domains (social versus physical cognition), when it comes to human higher-order thinking, at least, I tend to see these divisions as more apparent than real, more about academic book-keeping than the way higher-order thinking is functionally organized. To be sure, there is a vast literature highlighting numerous domain-specific effects in information processing. But when it comes to deploying higher-order constructs such as *<transfer of force>*, *<weight>*, *<time>*, *<beliefs>*, *<desires>*, etc., these domain-specific effects may be at the level of

So, to begin, here is my perspective on what went wrong in that debate. After that, I will examine how structurally equivalent problems plague the debate over animal folk physics.⁵

Fooled by Folk Psychology

A turning point in my own thinking⁶ on the question of whether animals conceive of mental states *qua* <mental states>, occurred as the result of a series of studies Timothy Eddy and I presented in our early monograph, *What Young Chimpanzees Know About Seeing*, along with the longitudinal and cross-sectional follow-ups that followed (see Povinelli, 1996; Povinelli & Eddy, 1996a, b, c; Povinelli & Eddy, 1997; Reaux et al., 1999). Inspired by the already known uses of gaze information by social primates (not to mention other non-primate species), as well as our own direct observations, we designed one of our experimental protocols (the one initially reported in *Seeing*) to map the conditions under which our seven chimpanzees would deploy their visually-based begging gestures in order to request food.⁷ The situations we created involved numerous contrasts between two familiar caretakers, one who could see them, the other who could not.⁸

Critically, our aim was not to determine whether our chimpanzees could reason about the colloquial (folk) meaning of a speech act such as, "Candy knows that Megan can see the potatoes." This colloquial phraseology jumbles together too many distinctive cognitive operations. We already knew, for example, that our apes could skillfully navigate their way through the complex consequences that emerge from the bodily postures, facial direction, and eye movements of others. We wanted to know if, in addition, they interpreted what they were doing through the framework of a higher-order, inferentially coherent, higher-order conceptual system that included ideas about mental states such as <seeing> and/or <attention>.

Consider the most straightforward treatment—someone facing them versus someone facing away (i.e., *front vs. back*). Our chimpanzees *immediately and consistently* gestured to the

the first-order systems underwriting their deployment (see Clark & Thornton, 1997; Penn et al., 2008a). I tried to make this point by including five full pages of metaphorical epigraphs about "weight" in the front matter of *Weight*. That approach may have been too subtle (see Healy, 2012).

⁵ A note to the theory of mind fatigued, and/or those who think (hope?) the question is settled: the purpose of this section is neither about vindicating the theoretical grounding of our particular research program, nor about defending its results. Instead, I seek to illustrate the central problems our research raised, which, in the opinion of at least this former monkey mind doctor, remain unresolved.

⁶ I had initially interpreted my comparative research on theory of mind in chimpanzees, rhesus monkeys and human children as favoring the view that chimpanzees, but not other nonhuman primates were, in fact, reasoning about the <mental states> of others (Povinelli, 1991).

⁷ Only those interested in the (esoteric) debate over animal theory of mind could have ever entertained such a question. After all, our young chimpanzees spent all day attending to, responding to, and manipulating what we (the folk) would surely call "seeing" (see below, "Candy's Family"). But such *functional-level* descriptions are distinct from the *representational-level* questions of the mental operations that produce such behaviors. Answering the latter questions (arguably) requires experimental manipulation (see below, "Claim 2: Tool-Using Experiments").

⁸ An important point about these investigations: in every study, we carefully choreographed the eye direction of the two caretakers to whom the chimpanzees were responding, with elaborate checks during each of the tests. In *Seeing*, the experimenters fixed their respective gazes gaze directly in front of them where the chimpanzees would be responding. We adopted this approach for two reasons. First, we believed this procedure was more likely to tap into their *cold cognition* about visual perception, which we argued was more indicative of higher-order cognition about <seeing> (see Povinelli & Eddy, 1996a, p. 33-36). Second, this allowed us to separately investigate the influence of direct eye contact (see Povinelli & Eddy, 1996c). This purported distinction—between a (cold) cognitive appraisal of <seeing> versus a direct (hot) reaction to eye contact—may help to explain some, but not all of the contradictory results in the literature (*cf.*, Hostetter et al., 2001, 2007; Theall & Povinelli, 1999). In any event, this was clearly one of my many early encounters with the (at-the-time-unrecognized) *Asymmetric Dependency Problem*. See footnote 17.

person who could see them. No specific learning occurred during the testing. Was this good evidence that they understood <seeing>? Our answer was, no, this treatment did not have the power to distinguish between two possibilities. The first possibility was that our apes knew that the person facing them could <see> them and would therefore react to their begging gesture by handing over some food. The second was that they knew that someone facing them would hand them food when they gestured toward them (after all, there had to be some perceptual anchor points for their gestures). But, and this is a big but, as we proceeded to experimentally disentangle the factors influencing the deployment of this natural gesture, their behavior consistently surprised us. For example, when we confronted them with two people facing them, one with blindfolds over their eyes and the other with blindfolds over their mouth, they did not seem to get it—at least not right away. Indeed, in treatment after treatment, their responses consistently matched the predictions of a model we had developed of how they would behave if they were reasoning about postures and faces, and even eyes, as predictive stimuli, not internal, perceptual events like <seeing> or <attention>. Frankly, it was shocking. The complexity of their naturalistic observations, so powerfully confirmed by their reactions to the *front vs. back* procedure, did not prepare us for what followed. Our folk psychology of our chimpanzees' folk psychology let us down.

Some scholars have caricatured our studies in a way that suggests our apes never learned to navigate our experimental dissections of what nature typically (though not always) conjoined: the movements of the torso, face, and eyes. On the contrary: they most certainly did. What is more—in the lingo of the ongoing scientific folklore⁹ of comparative psychology—they did so *quickly*, typically in 4-20 trials (for a detailed data summary, see Reaux et al., 1999, Table 1). Importantly, many of the key findings of *Seeing* (and related projects) were confirmed in other labs (see, e.g., Kaminski et al., 2004).¹⁰

Thus, our interpretation that our apes were not reasoning about <seeing> was derived from how they behaved at each critical juncture in our longitudinal studies, not because of any *inability to learn* or even a particularly *slow rate of learning*. On the contrary, they learned pretty quickly. Thus, neither our hypotheses nor our interpretations of the data hinged on which relations they could learn, or how quickly they could learn them. Instead, it was the fact that *their specific signature of transfer* from one condition to the next matched the predictions of one model, not another. We can (indeed did) question whether these models were too weak to make the intended contrast (see Povinelli & Vonk, 2003, 2004). But in the present context, it is important to realize that *it was a positive empirical fact pattern about our chimpanzees' behavior*—a fact pattern that was robust, folk psychologically bracing, and ultimately replicated¹¹—*that drove our interpretations*.

⁹ For a discussion of the concept of *scientific folklore*, see Povinelli and Barker (2019). For a discussion of the case study of the scientific folklore surrounding the connection between the number of trials it takes an animal to learn a response on the one hand, and what that animal does or does not understand, on the other, see Barker and Povinelli (2019b).

¹⁰ Lurz et al. (2018) claim to have obtained results that are inconsistent with one of the conditions we used in *Seeing*, and later by Kaminski et al. (2004). This is false. In the Lurz et al. study, only a single experimenter was present. The researchers assume that the separate effects they obtained in two different conditions, each with one experimenter, and administered separately, would yield the same effects if they were combined into a single condition, with two experimenters present simultaneously. We have already shown that such folk psychological thinking can suffer sobering assaults when actually put to the test (for one particularly brutal set of results in the present context, see Reaux et al., 1999, Exp. 4).

¹¹ At a personal level, I have often wondered whether Kaminski and colleagues (2004) experienced the same sense of disbelief we did each time one of their chimpanzees gestured to someone who had their back and face turned away from them, instead of to the person who was looking directly at them. My curiosity is important only insofar as the category of results I have been describing as *negative evidence* are taken less seriously than those described as *positive evidence*. I unpack this issue in detail shortly (see below, "Positive Evidence or Negative Evidence?"). I should also note that Laurie Santos and her colleagues obtained

Double Fooled by Folk Psychology

If this was all there were to it, perhaps the only intelligent thing one might conclude is that (surprise, surprise) our folk psychology can trick us: *I guess our folk psychological interpretation of their natural behavior, coupled with their immediate experimental success on front vs. back trials, is just a bad map for what happens when you follow their development from 5 to 9 years of age and conduct twenty or more experimental variations of this procedure to try to pull apart and contrast all the relevant variables* (see Povinelli & Eddy, 1996a, c; Reaux et al., 1999). Maybe our folk psychological interpretation of what they are doing just is not up to the task of predicting their behavior—regardless of whether or not they possess their own folk psychology.

That would have been hard enough. But our folk psychology fooled us back again. Additional experiments in a gaze-following context (some interleaved inside the very experiments just described), revealed that Candy and her companions were robustly sensitive to both the eye and head orientations of others, easily following our gaze in response to both static and dynamic stimuli (Povinelli & Eddy, 1996a, Exp. 12; Povinelli & Eddy, 1996c, Exp. 1; Povinelli & Eddy, 1997; Povinelli et al., 2002). Additional experiments revealed that—get this—when our apes processed someone's dynamic gaze shift toward a barrier, *where they wound up looking depended on whether that barrier was opaque or transparent*—full stop (Povinelli & Eddy 1996b).¹² Also full stop: in an experimental set-up very similar to the original *Seeing* studies, our chimpanzees immediately preferred to gesture to someone who made direct (hot) eye contact with them versus someone who did not (Povinelli & Eddy, 1996c)—but (and again, this is a big *but*) a moving head oriented their way, with eyes closed, tended to have the same effect.

We obtained other full stops. For example, my colleagues and I arranged a study in which our chimpanzees sat across a table from a human cooperative partner to whom they could gesture for food (Povinelli et al., 2003). The table was divided lengthwise and two objects were placed just in front of their partner, out of the chimpanzees' reach. One object was a highly desirable food reward, the other was an undesirable distracter object. Next, we manipulated where their partner was looking, either at the food or at the distracter object. The results of this study definitively showed that when their partner's visual attention was on the distracter object, our chimpanzees modified the deployment of their begging gestures to accommodate the discrepancy between what they wanted (the food) and the object of their partner's attentional focus. One last full stop: in other experiments, when our chimpanzees entered the testing room and confronted two opaque cups, they immediately chose the one at which an experimenter was gazing (see Barth et al., 2005; Povinelli et al., 1999).

Critically, all of these *full stop* effects were obtained with the same chimpanzees and at roughly the same time—the same chimpanzees who were also gesturing just as often to someone with a blindfold over their eyes as someone with a blindfold over their mouth (see Povinelli & Eddy, 1996a; Reaux et al., 1999).

So, fooled by folk psychology—then *double-fooled* back again.

Positive Evidence or Negative Evidence?

Our initial interpretations of these (and related) data roughly matched up with how some other prominent scholars were interpreting the results emerging from their own investigations: chimpanzees might not, in fact, reason about mental states like *<seeing>* (see Tomasello, 1996).

strikingly different results when our *Seeing* tests were administered to rhesus monkeys, for whom direct eye contact is a threat (Flombaum & Santos, 2005; see also footnote 8 above).

¹² Conceptual replications of this finding have been reported for spider monkeys, capuchins, great apes, and ravens, but not, curiously, in northern bald ibises (see Amici et al., 2009; Bräuer et al., 2005; Bugnyar et al., 2004; Loretto et al., 2010; Met et al., 2014).

But this consensus quickly diverged when Hare et al. (2000, 2001) reported that, in situations of food competition with a dominant animal, subordinate chimpanzees tended (for example) to navigate toward opaque as opposed to transparent barriers. These (and other) findings were quickly promoted as *breakthrough studies*¹³—studies with the power to supplant the interpretations of all previous data sets (see Tomasello et al., 2003). A curious auxiliary hypothesis emerged to support this dismissal. Chimpanzees, these authors suggested, use their theory of mind only in competitive interactions (or at least the only way to detect their theory of mind prowess was in the context of such interactions: see e.g., Hare et al., 2000). Earlier tests were disregarded because, according to these scholars at least, those tests had been "cooperative."

Despite the obvious reasons to doubt this auxiliary hypothesis,¹⁴ the idea caught on and held sway. For a time, at least, it became part of the folk scientific narrative of the field.¹⁵ Coupled with excitement over the new results, this folklore characterized our *Seeing* results as *negative findings* (see Hare et al., 2001; Tomasello et al., 2003)—despite the immediate, intricate, and robust uses of all manner of things related to gaze that Candy and her companions had displayed across many experiments, not to mention their "rapid" patterns of learning in the others (see "Double Fooled" above). Nonetheless, the deadly epithet of *negative evidence* was pinned to a murky mixture of both the findings *and* the interpretations of the findings.

Jennifer Vonk and I have previously discussed the folk scientific role that the phrase "negative findings" began to play in the debate (see Povinelli & Vonk, 2003, 2004), but there are more foundational points to made here. First, as noted above, the phrase *negative evidence* was being deployed in a way that conflated at least two distinct connotations of the word *negative* (even though I suspect the distinctions obscured by this conflation were, and still remain, widely accepted). First, data may not implicate one theory or another (i.e., the experiment did not find evidence for the existence of, say, *the black swan*). Second, in a more narrow and statistical sense, it may be the case that no statistical difference is found between condition x and condition y. Nonetheless, in practice, the folk scientific equivocation proved devastating. After all, in the context of the *Asymmetric Dependency Problem*, the important issue should have been *the robustness of particular effects*, not whether they implicated one theory over another. This raises the troubling question of why some of the robust effects we reported (e.g., complicated gaze-

¹³ I believe that handsome dividends will be paid by an analysis of the extra-scientific role this term has played in obscuring the fundamental conceptual and methodological challenges in the study of higher-order cognition in animals.

¹⁴ In a lengthy discussion of this issue, Povinelli and Vonk (2004) offered a number of reasons to question this auxiliary hypothesis. To summarize: (1) the argument that co-operative contexts are unnatural for chimpanzees is contradicted by the fact that highly social species will possess evolved mechanisms that perfectly balance cooperative and competitive tendencies (see de Waal, 1986). Specifically, wild and captive chimpanzees both beg for food—from conspecifics and human counterparts; (2) our chimpanzees *immediately* used their begging gestures to request food from the person who (from a God's Eye perspective) could <see> them versus someone who could not (the *front versus back* condition), and rapidly learned to do so in many others (see Reaux et al., 1999, Table 1); (3) the *Asymmetric Dependency Problem* applies with equal force to designs involving competitive or cooperative interactions; (4) our chimpanzees excelled at many cooperative tests that involved gaze direction (see main text for a discussion of Povinelli et al., 2003; see also Barth et al., 2005). The best proof, however, is ultimately in the pudding: twenty years later after promoting this suspect idea, the scholarly team that most strongly promoted it has *experimentally shown* that the competitive versus cooperative contrast was—in the context of the genre of experiments that attempt (in our view, impossibly) to implicate theory of mind—a non-starter from the get-go (Grüneisen et al., 2017).

¹⁵ Admittedly, I paint with a broad brush here, but I think it's fair to say that the attitude of a sizable number of thoughtful researchers (some no longer in the field) can be summarized by Andy Whiten's (2001) reference to the Hare et al. (2000) results in his review of *Folk Physics*: "If Povinelli's gigantic prior analysis of chimpanzees' folk psychology can be overturned by an elegant experiment more intuitive for chimpanzees, what of the prospects for the current, equally voluminous onslaught on folk physics? Time will tell" (p. 133).

following abilities) were pursued full tilt by other laboratories,¹⁶ whereas others (deploying a begging gesture to someone who cannot see you), were not (even though they were replicated; see Kaminski et al., 2004). In practice, the phrase ‘negative evidence’ became a shorthand for: *These data are not consistent with the theoretical picture we prefer.*

I do not wish to paper over the importance of the internecine empirical debates that ensued about these (and related) studies. For example, key findings from the Hare et al. (2000) study proved difficult to replicate (see Karin-D’Arcy & Povinelli, 2002)—even by the original authors, who spent a great deal of time and effort determining precisely how far apart two bananas needed to be in order get a statistically reliable effect (see Bräuer et al., 2007). The larger point is this: coddling the folk scientific notion of ‘negative evidence’ allowed many comparative psychologists to dismiss certain robust findings, while simultaneously embracing others—the latter being easier to reconcile with the notion that animals possess higher-order mental states.

Thus, on the basis of a high-level interpretation of the so-called ‘breakthrough studies’ (and a widening net of related investigations), a consensus quickly emerged (among experimentalists, at least) that chimpanzees (later, ravens) conceive of at least some mental states as *<mental states>*.

But, negative evidence and reliable effects aside, could the underlying logic of any of these studies yield a strong inference about higher-order thinking? As I show next, the *Asymmetric Dependency Problem* suggests a clear answer: no, they cannot.

The Asymmetric Dependency Problem (aka the Logical Problem)

As we (and others) were measuring distances between bananas, I was slowly realizing something that should have been obvious from the outset: none of the tests we and others were so glibly deploying could ever uniquely implicate the presence of higher-order mental states such as *<seeing>*. After a lengthy series of discussions, Jennifer Vonk and I detailed the reasons why. We showed that an entire genre of experiments (of which ours were a prime example) could never provide unique evidence for theory of mind. Fortunately, the genre was easily definable: any experiments premised on requiring animals to predict the behavior of others and/or the consequences of such behavior (Povinelli & Vonk, 2003, 2004).¹⁷

Our argument was straightforward: for any representationally-based theory, higher-order representations such as *<seeing>*, depend (by definition) on the existence of robust, stable, and abstract perceptual representations of the behavioral categories that are purported to be recognized or processed as instances of *<seeing>*. Jennifer and I referred to these first-order representations of social behaviors as *behavioral abstractions*. Simply put, if chimpanzees (and other animals) do not possess stable, perceptually-based, abstract representations of the entities around them, then there is nothing for their purported representations of *<mental states>* to be based upon. More foundationally, their perceptual systems could never detect instances of them.

¹⁶ For example, see footnote 12.

¹⁷ To my embarrassment, when the full force of this idea finally dawned on me, I was overtaken by a flashback from a conversation with a philosopher who had approached me after a lecture in Boston: “Your results are very interesting, and I follow your reasoning: the pattern of results you have obtained with your monkeys [sic] suggests they were tracking features of others that we wouldn’t use, which certainly weighs against the possibility they have a theory of mind like ours. But if I understood your bigger argument, wouldn’t the worry be that even if they *had responded* to your tests like you think humans would, we still couldn’t use those results to implicate the *presence* of a theory of mind?” Despite how obvious his point seems to me now, I merely blinked. I was clearly not prepared to grapple with his point. One instructive hypothesis as to why: some subterranean part of me knew how radically my daily professional practice would have to change if I did.

Why is this a problem? Once it is granted that higher-order representations depend on the existence of first-order, perceptually-based representations, there is no specifiable causal work left for any purported higher-order mental states *within this experimental genre*:

...[T]he design of these tests necessarily presupposes that the subjects notice, attend to, and/or represent, precisely those observable aspects of the other agent that are being experimentally manipulated. Once this is properly understood, however, it must be conceded that the subject's predictions about the other agent's future behavior could be made either on the basis of a single step from knowledge about the contingent relationships between the relevant invariant features of the agent and the agent's subsequent behavior, or on the basis of multiple steps from the invariant features, to the mental state, to the predicted behavior (Povinelli & Vonk, 2004, p. 8-9).

Thus, the first-order representations are both necessary and sufficient to account for the experimental results. This means that higher-order mental states are unnecessary to explain the results of these experiments. Furthermore, because of the asymmetry of the relationship between first- and higher-order mental states (the latter depend on the presence of the former, but not vice versa), higher-order mental states are not by themselves sufficient to account for the experimental results (see also Povinelli & Henley, 2020). Nor are they necessary. Having said this, it is important to underscore that Jennifer and I did not argue that higher-order mental states are causally inert. Such mental states may perform causal work in the behavior of both humans and other animals. Our point was narrow, but still vitally important: these *kinds of experiments* cannot provide unique evidence for the existence of such mental states.

Thus, even our initial, somewhat cursory specification of the first-order cognitive resources available to most organisms revealed a logical flaw not just in the Hare et al. (2000, 2001) studies, but in the entire (again: definable) genre of experiments that were being designed and deployed—including our own (a more detailed specification of these first-order mental resources can be found below, "Playing With a Full Cognitive Deck"). The claim that higher-order mental states depend upon lower-order states, but not vice versa, along with its devastating implications for experimental designs, quickly became known as *the logical problem* (Lurz, 2009). For purposes of precision, however, I prefer to call it the *Asymmetric Dependency Problem*.

The Experimental Necessity Dilemma

Jennifer Vonk and I (2004) made an additional argument that grew obscured in the ensuing debates: If (despite our analysis) one elected to accept the Hare et al. (2000, 2001) studies as evidence for theory of mind, then the behavior of Candy and her companions in our *Seeing* (and related) investigations must also be taken as evidence of theory of mind (see above, "Double Fooled"). There is no conceptual difference between those results, and the behavior of the chimpanzees in the Hare et al. (2000) studies.

More bracing, and more generally, this kind of evidence was available from *naturalistic observations* long before any of the experimental work (including our own) began (see Povinelli & Eddy, 1996a, p. 17-24; Tomasello et al., 1994; Whiten & Byrne, 1988). The natural behavior of chimpanzees and other animals clearly shows their functional-level sensitivity to the bodily postures, faces, and eyes of others—not to mention the overt manifestations of other mental states such as intentions, desires—even states of knowledge and false belief. Recall that it was the desire to answer the representational-level question of whether chimpanzees interpret these states as *<mental states>*, that drove researchers to argue that experiments were needed in the first place (e.g., Heyes, 1998; Povinelli, 1988; Premack, 1988).

This problem boils down to choosing between the horns of what I call the *Experimental Necessity Dilemma*: either Jennifer and I were correct that the current genre of experiments cannot

implicate higher-order reasoning in animals, *or we never needed them (including our own) in the first place*. As I show throughout this essay, this reasoning applies with equal force to the current generations of experiments.¹⁸

Any attempt to escape the *Experimental Necessity Dilemma* by choosing the first horn (i.e., that the experiments do not show that higher-order thinking is necessary to explain a given set of experimental results), but then countering that there are a lot of said experiments *and* they are at least *consistent with* the existence of higher order thinking, forces one to acknowledge that the same is true of naturalistically derived data. I should also add that any such move *de facto* acknowledges that first-order reasoning (alone) is both necessary and sufficient to explain both the experimental and naturalistic data, leaving no reason to posit the higher-order states to account for the results.¹⁹

Beyond the inferential roadblock specified by the *Asymmetric Dependency Problem*, and highlighted by the *Experimental Necessity Dilemma*, another important, albeit side issue arises: from whence do the all-important perceptually-based behavioral abstractions derive?

The Unprincipled Titration Paradox: Relational Reasoning *au Naturel*

Given that chimpanzees (like other social organisms) are biologically pre-prepared to learn myriad relations involving social dynamics and object interactions — coupled with the facts of their daily experiences — we can be assured that long before any wily experimentalists appear on the scene, Candy and her companions' minds are already deep at work developing their ever-expanding manifolds of behavioral propensities. Thus, it is beyond myopic for anyone to have imagined that the chimpanzees participating in our experiments, let alone those of Hare et al. (2000, 2001), had anything other than a *lifetime of experience* with the relevant abstracted set of relations on which they were being tested. The refrain that the *specific* objects used in testing *had never been seen before*, misses the point entirely: Perceptual abstractions are *abstract*²⁰ (that is, they are the perceptual invariants derived by organisms during their innumerable experiences with the specific relational exemplars, guided by the canalizing influences of their evolutionarily-defined developmental systems).²¹

¹⁸ It is no coincidence that the acronym for the *Experimental Necessity Dilemma* is END.

¹⁹ For some, this may spark the following idea: if the experimental data possesses such grave limitations, perhaps data derived from naturalistic observations could, with the right framing, allow for strong inferences about higher-order thinking. I examine this issue in the equivalent debates over tool use a bit later in this essay (see below, "Claim 1: Natural Tool Use Cannot Provide Strong Evidence for Higher-Order Thinking"). For now, suffice it to say that I do not believe this argument can get traction.

²⁰ See Cook et al. (2006). There are, of course, many scholars who argue that at least some taxa (e.g., primates) may, in addition, form higher-order concepts (see Zentall et al., 2008). Penn et al. (2008a) show why the evidence does not warrant such an inference (see also Glorioso et al., in press).

²¹ Another overlooked argument Jennifer Vonk and I made was that for any given animal, known relations based on such perceptual abstractions can (perhaps logically, must) be freely applied cross-species. This will continue until countervailing evidence (e.g., the heterospecific does not respond as expected) forces a new relation to be learned. Our effortless application of our own folk psychology to animals is dramatic proof of this fact. But if our argument is taken seriously, then human folk psychological attributions to animals are initially driven by human first-order mental machinery. Turning this around, this means that when Candy spontaneously and "appropriately" directs her social behavior to us—a fact that grounds her everyday interactions with us—we cannot herald this as independent evidence to support the claim that she possesses her own folk psychology. Curiously, Boesch (2007) has overlooked the depths of this problem altogether by arguing that experimentally derived evidence for theory of mind in chimpanzees is better when chimpanzees interact with other chimpanzees, not other humans. A careful examination of his claims reveals many fatal problems, including that, when a chimpanzee subject is allowed to receive input from another chimpanzee, there can be little control over the first-order variables in the experiments, introducing all manner of methodological inconsistencies (for a more detailed discussion, see Povinelli & Vonk, 2004, Appendix

Experimenters rarely (if ever) grapple with this issue. And yet, because the logic of such experiments do in fact, presuppose (*read*: require/assume) that animals already possess loads of first-order relational information directly related to the experimental procedures, a curious medley of scientific folklore guides the design and interpretation of the studies. Informal practices, perpetuated in daily discussions within individual laboratories, attempt to (mysteriously) balance the need to anchor the elements of any given task to an animal's pre-existing (or pre-trained) first-order reasoning systems, and the need to (somehow) prevent the network of relational reasoning attached to these elements from being the source of the allegedly diagnostic experimental outcomes. Earlier, I dubbed this the *Unprincipled Titration Paradox*. Povinelli and Henley (2020) describe the problem in the following way:

...all experimental protocols claiming to assay higher-order reasoning in animals rest upon an extremely suspect, and ultimately unprincipled, titration. Specifically, researchers implicitly assume that the tasks are perceptually *similar enough* to what their subjects have previously encountered to allow them to make sense of the problem, but perceptually *different enough* that the subjects must (somehow) rely upon higher-order reasoning to navigate their way through it (p. 392).²²

This paradox frequently interacts with other folk scientific practices for interpreting results that do not (seem to) confirm the presence of higher-order thinking. This so-called *negative evidence* (see earlier discussion) is frequently discounted as uninformative due to the titration being incorrectly balanced. This ignores the fact that admixture was unprincipled in the first place.

The *Unprincipled Titration Paradox* is more than a minor methodological nuisance. The practices and flawed inferential reasoning it supports allow researchers to ignore the core obstacle posed by the *Asymmetric Dependency Problem*: the attempt to control for first-order relational reasoning in a way that keeps it from poisoning inferences about the presence of higher-order reasoning (regardless of whether it is attempted through more complicated designs or more elaborate control conditions). Such titrations will never suffice within the current experimental genre. Why? Because while these carefully titrated elixirs can most assuredly isolate the *perceptually-based* mental representations that are *not* involved in generating the behaviors of interest, *they can never "control away" the perceptual representations upon which any higher-order representations depend* (see Povinelli & Vonk, 2004, p. 10-11). Worse yet, as shown above, because first-order representations are both necessary and sufficient to produce the behavioral responses, the *purported* higher-order ones are unnecessary. In the end, the *Unprincipled Titration Paradox* amounts to what I dreamily imagine as the intellectual equivalent of an ever-expanding carnival shell game.²³

1)—not to mention that his review is ruinously selective in its review of the evidence (see above, "Double Fooled", for examples of the dynamic and coherent responses of chimpanzees to human behavior; see also the additional refutations in Tomasello & Call, 2008). But the core, irreducible challenge is that it is the experimental paradigm, not the stoges acting within it, that constitutes the source of the *Asymmetric Dependency Problem*.

²² Having for the most part stepped away from public-facing sides of these debates many years ago, I recently had a chance to revisit this issue with Mike Tomasello in Salzburg, Austria. During my lecture, in which I was using the Hare et al. (2000) studies to illustrate the *Unprincipled Titration Paradox*, Mike interrupted (I had invited interruptions) and exclaimed: "You're missing the point! Our chimps were responding to *novel stimuli*. They had never seen our little transparent Plexiglas barriers before. So how could they possibly have a representation of them?"

²³ In a study involving physical cognition, Civelek et al. (2020) reveal the standard, unprincipled navigation of this problem. First, from their methods: "Chimpanzees in Leipzig Zoo had objects made of different materials in their outdoor and indoor enclosures (i.e., automatic metal feeders, tree logs, plastic buckets) and occasionally may hear the noises they make when they are hit/dropped. However, in comparison to children

This analysis exposes the soft underbelly of another tidbit of scientific folklore: namely, that *success on the first few trials in any given experiment (especially trial one)* holds some special status in implicating higher-order thinking.²⁴ When I first began my own research, I was as guilty as anyone in promulgating this theoretically untethered notion (see Barker & Povinelli, 2019a; Povinelli, 1988). But if the analysis Jennifer Vonk and I offered about the *Asymmetric Dependency Problem* is correct, and I believe it is, then an organism's behavior on trial 1 ought have no greater weight in implicating the presence of higher-order reasoning than the organism's behavior on, say, trial 22. Under our current theoretical umbrella, first-order behavioral abstractions are both necessary and sufficient to account for the behaviors in question, no matter when they appear.

I stress again: this conclusion applies with equal force to our own research, including, for example, our chimpanzees' *full-stop* understandings of: (a) the *front vs. back* distinction when begging for food (as well as their rapid understanding of many of the other seeing/not seeing conditions), (b) the geometric relationship between transparent/opaque barriers and a caretaker's gaze direction (Povinelli & Eddy, 1996b), (c) the mismatched attentional focus of themselves and a cooperative partner (see Povinelli et al., 2003), and (d) the relevance of a human's gaze direction in selecting a baited cup.²⁵

Experimentalists React to the *Asymmetric Dependency Problem*

Most experimentalists did not initially (and from what I can tell, still do not) recognize the depth and pervasive scope of the *Asymmetric Dependency Problem*. They have reacted to it, but rather than countering its premises, or challenging the strength of its conclusions, they have caricatured it as *modified behaviorism*, *inflexible perceptible cue learning*, *arbitrary perceptual cue learning*, *behavioral rule learning*, or as a hypothesis invoking *behavioristic principles of learning* (or some loosely equivalent terminology; e.g., see Call & Tomasello, 2008; Hare et al., 2006; Seed & Call, 2009; Tomasello et al., 2003; Tomasello & Call, 2006). Alas, the ideological alarm bells

we assumed their exposure to metal and wooden materials would be limited. Therefore, we prepared two sound-making training boxes: the "metal box" ...made from stainless steel and the "wooden box" ...made from plywood" (p. 9). Next, from their conclusion: "Three-year-old children and chimpanzees did not discriminate between the conditions, did not pass either of them and were more likely to be side biased. This could reflect a "true negative"...However, as with many negative findings, interpreting these results is not straightforward. *One explanation for the failure of chimpanzees could be that the initial training we implemented were not sufficient to build the necessary knowledge for solving the problem [italics added]*" (p. 14). Examples of chasing such controls abound in the literature, including one amusing protocol that was designed to test for knowledge attribution in chimpanzees while attempting to control for a so-called "evil eye" hypothesis—meanwhile, all the perceptual features that the subjects' required to respond appropriately were (by necessity) left intact (see Kaminski et al., 2008). See also Povinelli and Vonk (2012), "Case study: Levels of understanding floppy and rigid tools."

²⁴Ty Henley and I (2020) explore this issue in connection to higher-order causal reasoning in our contribution to the August 2020 special issue of *Animal Behavior and Cognition*.

²⁵ Consider the conclusion of our attentional mismatch project:

One interesting question raised by our research remains unanswered. Did our chimpanzees modify the location of their gestures as part of their construal of their partner as a psychological agent, or strictly as part of their understanding of the observable dynamics of their partner's observable posture and behaviour? In short, were the chimpanzees attempting to modify their partner's behaviour and attentional state, or strictly their behaviour alone? (Povinelli et al., 2003, p. 77).

But if Jennifer Vonk and I were correct, then even these "impressive" results had no bearing on whether our chimpanzees were reasoning about *<attention>*. This example underscores the point that from the early 2000s forward, the empirical fact pattern was not the (true) source of dispute over the question of theory of mind in animals. Understanding this historical episode in detail could be of great use in helping the next generation of comparative psychologists develop more productive lines of research.

raised by the charge of behaviorism, even if charitably read, are orthogonal to the problems picked out by the *Asymmetric Dependency Problem*. After all, the *Asymmetric Dependency Problem* exposes the *full cognitive commitments* of researchers and shows why, by their own (mostly unstated) assumptions, first-order reasoning is both necessary and sufficient to account for the steadily mounting experimental results.²⁶

The legacy of the mischaracterization of the *Asymmetric Dependency Problem* as a form of behaviorism has exacerbated another problem. Namely, it has heightened the *False Models Contrast Problem*—the practice of pitting naked higher-order thinking (i.e., higher-order thinking stripped of the first-order representational relational thinking upon which it depends) against unspecified models of associative learning that are so vague as to be meaningless.²⁷ This leaves the researchers shadow-boxing with Fred Skinner (see also, footnote 1).²⁸ Clearly, the foundational challenges posed by *Asymmetric Dependency Problem* are orthogonal to the parochial disputes involving behaviorist boogey-monsters (especially the creature colloquially known as "associative learning"; see Dacey, 2016; Penn & Povinelli, 2013; Povinelli, 2012). To emphasize: our theory of our *chimpanzees'* understanding of seeing,²⁹ not to mention Candy's understanding of her coconut,

²⁶ Some scholars (mistakenly) believe our position does, in fact, have strong ties to behaviorism. Marta Halina (2015), for example, argues that "[t]he logical problem [aka the *Asymmetric Dependency Problem*] constitutes a sort of second-order behaviorism—behaviorism not with respect to our scientific understanding of agents, but with respect to nonhuman animals' understanding of other agents" (p. 488). This characterization rests on a deeply problematic equivocation of the term *behaviorist*. Behaviorists defend a meta-methodological claim that *denies the need* to posit mental states in order to gain a complete theory of behavior. Nowhere do we claim that *chimpanzees deny the need* to posit <mental states> to explain the behavior of others. That would be a strange theory, indeed. Rather, our argument is merely that animals need not possess higher-order mental states in order to perform the behaviors in question. We make two claims that should not be run together. First, we argue that *in order to account for their behavior, chimpanzees must possess first-order mental representations*. Second, we argue that these *first-order mental representations are necessary and sufficient to account for their behavior*. Thus, our cognitive model explicitly defends the theoretical and empirical necessity of positing that chimpanzees (and other animals) possess a wide variety of mental states—namely, they possess atomistic and temporally stable perceptual symbols that are compositionally recombined to support goal-directed action (see Penn et al., 2008a). Thus, Halina's (2015) observation of a (very) loose linguistic resemblance between our claims and Skinner's (false) argument that all mental state variables can be dispensed with at no explanatory or predictive cost, has absolutely no bearing on (1) whether the *Asymmetric Dependency Problem* constitutes a distinct challenge for experimental designs, or (2) whether it will turn out that higher-order <mental states> are uniquely human. For a further discussion of the problem of equivocation in the use of the term *behaviorist*, see Povinelli and Eddy (1996a, p. 139).

²⁷ I recommend Mike Dacey's (2016) masterful examination of the history and contemporary usage of the term "association" in psychology. He concludes: "...[A]ssociation should...be seen as a highly abstract filler term, standing in for causal relations between representational states in a system. Associations, so viewed, could be implemented by many different mechanisms. I outline the role that this view gives associative models as part of a top-down characterization of psychological processes of any kind and of any complexity" (p. 1). At the risk of redundancy, I stress that whatever concerns one might have about how association-based psychology intersects with cognitivism, the *Asymmetric Dependency Problem* still holds and is unaffected by those worries.

²⁸ Although I have no sympathy for Skinner's radical behaviorist ontology or epistemology, the persistent pugilism directed his way occasionally makes me feel sorry for him. Even in his prime, it would not have been a fair fight: his hobbies were indoor gardening, old movies, French poetry, and Agatha Christie novels (The Harvard Crimson, 1990). And besides, he's been RIP for thirty years.

²⁹ To my ever-lasting regret, in *Seeing*, Tim Eddy and I (1996a) used Premack and Woodruff's (1978) labels for the models we were testing: namely, *behavioristic* versus *mentalist*. We dutifully noted that the behaviorist model was not intended to imply that chimpanzees were devoid of mental states (viz. they could most definitely "process and use information" p. 26), but rather to determine if there was any reason to believe they wielded the higher-order mental states of <seeing> and/or <attention> (Povinelli & Eddy, 1996a, p.

was (and remains) a *mental* theory grounded in uncontroversial constructs in the *cognitive sciences*.³⁰

Perfect Experimental Storms

The challenges I have outlined in the previous few sections invariably come together to create perfect experimental storms—and that is not a good thing. A typical case is the seminal attempt by Call et al. (2004) to demonstrate that chimpanzees think about *<intentions>/<desires>/<goals>*. At the level of analysis that matters for the *Asymmetric Dependency Problem*, this study falls squarely within the genre identified by Jennifer Vonk and I: take some apes, present them with some carefully choreographed behaviors (in this case, human behaviors), and then measure how they react. Study how Call et al. (2004) flail in the confluence of the *Asymmetric Dependency Problem*, the *Experimental Necessity Dilemma*, the *Unprincipled Titration Paradox* (along with its equally undisciplined disciple, *novelty*), and the *False Models Contrast Problem* as they attempt to explain their results:

Perhaps chimpanzees had learned from their previous experience to expect that certain actions usually result in them receiving food and certain actions usually result in them not receiving food . . . For example, normally after humans drop a piece of food on the way to giving it to the chimpanzee, they pick it up and give it to the chimpanzee, whereas normally when humans are eating they do not share their food with the chimpanzee... However, if chimpanzees were using their previous experience of E's actions to decide how to react, they would have had to have *a separate learning history for each of the five conditions* [emphasis added] in which they discriminated successfully. This is unlikely because some conditions, at least, arguably were *novel* [emphasis added] to the chimpanzees and because these chimpanzees had little experience with experimenters or testing situations in general because they were new to the facility... [Thus] We believe that chimpanzees were using the actions of the experimenter *not just as superficial discriminative cues but as a way to determine his goal* (p. 496-497).

25-26). Peter Hobson (1996) charitably recognized this fact in his accompanying commentary. Nonetheless, our use of these labels became part of the legacy of the *False Models Contrast Problem* that continues to plague comparative psychology, our dogged efforts to dispel it, notwithstanding (see especially Penn & Povinelli, 2013). This underscores the need for researchers to develop and use of a shared set of formal notations to track the purported causal work of first- versus higher-order representations in their explanations of animal behavior (for a step in this direction, see below, "A Seven-Step Program to Recovery").

³⁰ The characterization of our theory as a form of behaviorism has allowed researchers to sidestep the *Unprincipled Titration Paradox* and design, conduct, and publish studies that create the illusion they are testing their chimpanzees in situations they have never experienced, measuring behaviors the chimpanzees have never produced before. One of my all-time favorite examples is a still-widely cited study by Hare et al. (2006) in which investigators sought to determine if chimpanzees would hide from human competitors in situations of food competition. The researchers offer their readers the choice between imagining the chimpanzees operate on the basis of an "inflexible reliance on contextual or behavioral cues alone" (p. 497) or a *theory of mind*. This was an obvious straw man at the time and remains one today. Seriously, who could possibly accept a theory positing that chimpanzees were inflexible! Alas, the descendants of this false dichotomy are alive and well. Worse, this false contrast allowed the researchers to ignore the obvious fact that such deceptive behavior occurs every day among captive chimpanzees. This directly raises the issue of what it means for a situation, object or relation to be *novel* in the first place—an issue that *Folk Physics* grappled with at length (Povinelli, 2000, Chapter 12). Indeed, the issue of what *novelty* is supposed to pick out about objects or behavior and/or the relations among them, continues to weigh heavily upon comparative psychology, including several of the experimental contributions to the August 2020 special issue of this journal (viz. the *Unprincipled Titration Paradox*).

Because they fail to engage with the core (and well-grounded) premises of the *Asymmetric Dependency Problem*, these scholars (a) wield undefined (folk) notions of novelty, (b) fail to distinguish between first-order behavioral abstractions versus higher-order representations of <goals>, (c) falsely contrast *superficial discriminative cues* with (apparently) higher-order representations of <goals> and (d) argue that their titration of first-order reasoning was *just right*. All of this creates a cyclone of vicious circularity. The storm begins by not acknowledging that any higher-order representation of <goals> depends upon corresponding abstractions of instances of such as understood by the humans who invented the test (see Povinelli, 2000; Povinelli & Vonk, 2003, 2004). Innumerable other studies can be analyzed in the same manner (a few widely-cited studies that are a delight to examine through this lens include Karg et al., 2015; Kano et al., 2019; Melis et al., 2006; for more examples, see below, "Philosophers React"; see also Perspective Pieces 1, 2 and 3, (this issue) and Povinelli & Henley, 2020, Appendix).

Those who either outright dismiss, or have not yet grappled with the far-reaching implications of the *Asymmetric Dependency Problem* are likely to be incredulous that I would spend so much time dissecting an experiment soon to be two decades old: *The field has moved on, Daniel. We now have scores of experiments using dozens of new protocols. Collectively, they provide converging (or better, or more diagnostic) evidence that chimpanzees, other great apes, ravens, and probably dolphins, elephants and dogs, have some kind of theory of mind. The debate is settled.* As an erstwhile experimentalist, I understand this reaction. As a former monkey mind doctor, though, I note that while the first half of this reaction is indisputably true (the field *has* moved on), the second half, alas, is not.

Philosophers React: Experimental Escapes from the *Asymmetric Dependency Problem*?

In direct contrast to the sentiment of many experimentalists, philosophers reacted to the *Asymmetric Dependency Problem* with direct engagement. Trained to confront the best possible version of an argument (and thus avoid undue violence to hapless straw men), they recognized that our hypothesis (1) was an explicitly *mental* theory and (2) that it presented a severe challenge for interpreting all current (and future) experiments of a type. These facts spawned a healthy debate about how to overcome it using modified experimental designs (Andrews, 2005, 2016; Halina, 2015; Lurz, 2011; Lurz & Krachun, 2011; Lurz et al., 2014; Roche, 2013; Sober, 2016; cf. Povinelli & Henley, 2020, Appendix). Indeed, in following these debates, I became far less sanguine about the proposals Jennifer Vonk and I had made for overcoming the problem (e.g., Lurz, 2011; see discussion in Gallagher & Povinelli, 2012).

In an exciting turn of events, the engagement of philosophers has led to the development of several specific experimental proposals that have been offered as escaping the obstacles laid out by the *Asymmetric Dependency Problem*. But have they, actually? Just to explore one example, Robert Lurz and colleagues (2018) attempted to determine whether chimpanzees could reason about <seeing> above and beyond keeping track of first-order relations such as the unobstructed geometric relation between the chimpanzee subject and the experimenter's face. They confronted their apes with a person seated on a stool and just out of reach. They then measured the number of begging gestures the chimpanzees made in what amounted to three informative conditions in which the person either:

- (a) faced away from the ape, but their face was visible in a mirror, looking back at the ape;
- (b) faced away from the ape, but instead of a mirror, a photo of the experimenter's face was pasted where the mirror was in (a);
- (c) faced away from the ape, but the person was looking back over their shoulder toward the ape.

Because the apes gestured more in (a) and (c) than in (b), Lurz et al. (2018) concluded that the chimpanzees were reasoning about <seeing>, as opposed to just the direct geometric relation between themselves and a face. This is a fallacious conclusion.

First, in both (a) and (c), the stimuli projected on the chimpanzee's retinae is a highly realistic face, complete with depth information, oriented toward them. Hence, the geometric relation between the face and the chimpanzee *is the same in both cases*. In (c) it is direct, in (a) it is mirror-mediated. Lurz et al. (2018) acknowledge this irrefutable fact, but attempt to dismiss its significance:

It's ingenuity notwithstanding, we do not find this mirror-mediated line of gaze hypothesis to be more plausible than the seeing hypothesis. As already noted, our chimpanzees did not have any experience prior to our tests with a contingency existing between humans feeding them and humans having a mirror-mediated line of gaze to them. Consequently, there is no independent reason to expect that chimpanzees would treat a mirror-mediated line of gaze of an experimenter in a feeding context as either (1) significantly different from no direct line of gaze...or (2) as equivalent to direct line of gaze (as they do in experiment 2) (p. 247).

But by Lurz et al.'s (2018) own reasoning, regardless of whether the chimpanzees possess a higher-order theory of <seeing>, they must compute the geometric relation between a face (whether in a mirror or otherwise) and themselves. *It is only by assuming ahead of time the chimpanzee possesses a higher-order understanding that connects <seeing> and "looking in mirrors"* (that is, that the mirrored face is at once *not* the experimenter's real face, but rather a reflection of it, and that this reflection is connected to the real experimenter via the experimenter's ability to <see>, and, by extension, that the experimenter's ability to <see> extends along a geometric line to the chimpanzee's current location) *that allows their inference that chimpanzees understand <seeing> to follow from the experimental results*. But if so, the premises assume the existence of the thing they sought to demonstrate. This is viciously circular.³¹

Lurz et al. (2018) next argue that the fact that the chimpanzees displayed higher levels of begging in (a) and (b), as compared to the photo condition, licenses the conclusion that the chimpanzees were reasoning about <seeing>. But by the same token, their chimpanzees would have gestured less to a cardboard cutout of a human body than to a real human facing them. Why? Because, the more realistic a stimulus, the more gestures it will elicit (*cf.* Tinbergen & Perdeck, 1951). Now, before one dives down the rabbit-hole of thinking that more controls can save the day, recall the core problem isolated by the *Asymmetric Dependency Problem*: There is no way to control away the very perceptual stimuli that are purported to connect the animal to the task. To hammer home why this is so, imagine eliminating (*read*: controlling for) the very perceptual information a subject needs to detect in order to generate a higher-order interpretation: *in this case, even a subject known to be capable of generating the relevant higher-order thought would be unable to do so, and hence unable to make the intended behavioral response*. Thus, no amount of making the photo more life-like (or, conversely, degrading the quality mirror image) will suffice. Eventually, the perceptual features anchoring the chimpanzees' behaviors (whether additionally

³¹ Lurz et al. (2018) provide an elaborate discussion of how their chimpanzees' previous experiences with mirrors might (or might not) connect to what the chimpanzees did (or did not) expect to happen when they gestured to the mirror. For purposes of simplicity, the analysis offered here assumes everything Lurz et al. state about the chimpanzees' prior experiences with mirrors is true, but then shows why these arguments neither support nor refute the <seeing> hypothesis. This issue is thus orthogonal to the challenges posed by the *Asymmetric Dependency Problem*.

linked to higher-order representations or not) will meet, and the same level of gesturing will be obtained.³²

Playing with a Full Cognitive Deck

Shortly after Jennifer Vonk and I published our analysis of the *Asymmetric Dependency Problem*, another colleague, Derek Penn, recruited me and Keith Holyoak to work with him to flesh out the domain-general format of the behavioral abstraction hypothesis—in essence, to lay all our cognitive cards on the table (see Penn et al., 2008a). Derek's work picked up where *Seeing, Folk Physics* and the behavioral abstraction hypothesis left off, outlining how any representationally-based theory of animal cognition, by necessity, must embrace the fact that organisms (including humans) possess fundamental atomistic perceptual symbols that can be productively recombined. These (and related) capacities, in turn, give rise to powerful forms of goal-directed behavior driven by first-order, perceptually-based relational reasoning. In humans, at least, such first-order capacities interact with *additional capacities* that allow disparate perceptual relations to be grouped under common thematic or argumentative roles—a hallmark signature of higher-order, role-based, analogical reasoning. This is what allows humans to deploy higher-order constructs such as <gods>, <ghosts>, <gravity>, <belief>, <time>, <space>, <weight>, etc. (see also Povinelli 2000, 2012; Vonk & Povinelli, 2006). In reviewing the evidence—across every domain of cognition—Derek, Keith and I examined what we believed were the best available experimental results available to support claims for higher-order reasoning in animals. In each case, we discovered that first-order, perceptually-based relational reasoning—the stuff of Candy and her coconut—was (using the language adopted here) both necessary *and sufficient* to account for the obtained results.³³

A Seven-Step Program to Recovery

The growing complexity of the experimental designs (devised and mobilized by both traditional experimentalists and philosophers alike), coupled with the problems isolated above, makes it tedious (and hence less fun for experimentalists) to isolate how, precisely, in each case, any new (or for that matter, old) experimental designs do (or more to the point, do not) overcome the *Asymmetric Dependency Problem*. Fortunately, for anyone who is interested, there is a rather straightforward, and more-or-less formal, method for doing so.

Because it may seem a bit heavy-handed (academics, including myself, typically do not like to be told how to approach a problem), let me stress that I am sure there are many eclectic ways to expose the hidden premises in the informal, verbal arguments that are typically presented in empirical journal publications. Therefore, to be clear at the outset, I endorse any and all methods

³² The actual situation is far worse for their arguments than I have outlined. To understand why, I invite the reader to write out the steps of the inductive (or abductive) argument they are ultimately making. This will require formal notations to keep track of the role of first-order perceptual representations and relations, and their connection to higher-order ones. This would minimally include the following: Face_{mirror}, Face_{real}, Face_{photo}, the geometric relations among each of them, and how the higher-order variables of <seeing> and <reflection> are mapped onto those relations. For examples of how to do this, see below, "A Seven-Step Program to Recovery".

³³ Expanding upon the arguments made in *Seeing, Folk Physics* (and later, *Weight*), Penn et al. (2008a) made the *additional* argument that the existing evidence strongly supports the claim that the ability to cognize over higher-order, role-based analogical relations is a uniquely human capacity, cutting across every domain of cognition. As explained earlier, this claim does not contradict the more foundational argument of the *Asymmetric Dependency Problem*. After all, if one were to reject the argument that experimental results can implicate the *absence* of higher-order thinking, one would be doing so by *accepting* the implications of the *Asymmetric Dependency Problem*.

for achieving the outcomes I aim at here. That being said, here is the method I have found most effective:

(1) Identify (and develop a notation for labeling) all stimuli, and relations among stimuli, presented to the subjects in the main experimental conditions that they must keep track of in order to participate. *To avoid circularity, do not include folk psychological descriptors at this point because this is what is being tested for.*

(2) Identify any unstated assumptions the researchers may be making about the stimuli that inadvertently usher in the very mental states for which they are testing. Elevate these unstated premises to an explicit level.

(3) Examine all critical dependent measures collected by the researchers and describe them as in (1).

(4) Examine all control conditions or critical contrast cases the researchers are using to leverage their interpretations and describe them as in (1).

(5) Identify all previous contexts in which exemplars of the perceptual relations have been experienced by the animals. These may have been encountered in training or testing or in the ongoing wash of their everyday lives, or, more likely, all of the above. Also, are there known facts about the evolutionary ecology of the organisms that leads one to strongly suspect the (abstract) perceptual relations in question are tightly canalized in development? If so, make these explicit.

*(6) Construct the best possible (inductive or abductive) argument (using premises and conclusions) that connect the first-order perceptual representations that are necessary for the organisms to possess (the ones that were identified in steps 1-4), to the researchers' conclusion that the specific higher-order representations are causally operative (e.g., necessary) to produce the results of the experiment.

(7) Delete all higher-representations in (6) and ask: Are the lower-order representations necessary and sufficient to explain the results? If yes, it follows that higher-order thinking is unnecessary.

Shannon Kuznar, Mateja Pavilic and Gabrielle Glorioso and I (Perspective Piece 1, this issue) use this general method to work through Bugnyar et al.'s (2016) study that claimed to have demonstrated theory of mind in ravens. In Perspective Pieces 2 and 3, Ty Henley and I (this issue) offer a head start on dissecting Kano et al.'s (2019) study involving chimpanzee theory of mind, as well as Jelbert et al.'s (2019) recent investigation of crow folk physics. Other, detailed, but less formal deconstructions can be found elsewhere (see Glorioso et al., in press; Penn et al., 2008a). Even if you think you already have reason to doubt the interpretations of these particular findings, I urge you to study these examples. It is only by working a sufficient number of examples that the common, fatal flaws of all such studies become apparent. True, every study has its own suite of minor-to-major methodological limitations. Analyses focused on such details are important to normal science. But the Seven-Step Program to Recovery has a more overarching (and inconvenient) aim: to show why *all studies of the identified genre are doomed by a common conceptual problem.*

I further encourage readers to work through other examples from the literature that they may find the most compelling to them. Suffice it to say that I have encountered none that survive step 7. In each case, the logical problem Jennifer Vonk and I described fifteen years ago stubbornly

clings to any and all experiments *of a type*. Until either a principled refutation of this argument appears (which may or may not be possible), or procedures are developed that fall outside the genre, no amount of wishing otherwise will change this (admittedly troubling) conclusion.³⁴

If everything I have said up to this point applied solely to investigations of animal theory of mind, this would be troubling enough. However, as I have been hinting at all along, the *Asymmetric Dependency Problem* applies with equal force to all higher-order representations, including <gravity>, <force>, <connection>, <shape>, <weight>, <time>, <past>, <present>, <future>, and so on and so on (see also Povinelli & Penn, 2011; Povinelli & Henley, 2020). But because *Folk Physics* appeared several years before Jennifer and I shared our analysis of the *Asymmetric Dependency Problem*, the polemics of the two debates were never fully integrated.

So, in what follows, I integrate the foundational challenges outlined in the first half of this essay, with key developments that have occurred across the ensuing two decades of research on the use and manufacture of tools in chimpanzees and other animals. I will use the central claims made in *Folk Physics* as the framework for assessing these developments.

Why the Central Claims of *Folk Physics* Still Matter

Claim 1: Natural Tool Use Cannot Provide Strong Evidence for Higher-Order Thinking³⁵

The opening argument of *Folk Physics* was that programmatic experimental strategies were needed to disentangle distinct *representational-level* hypotheses about the cognitive underpinnings of the *functional-level* descriptions of tool use in the wild. After all, if the functional competences displayed by wild chimpanzees were sufficient to infer whether they possessed higher-order representations, *Folk Physics* would have been unnecessary:

[This] project was designed to use experimental techniques to explore what chimpanzees understand about why tools produce the specific effects they do. In doing so, the project begins with a clear recognition that chimpanzees naturally make and use simple tools in the wild, and that, in captivity, these activities may be even further elaborated and refined (Povinelli, 2000, p. 2).

Furthermore, as a physical anthropologist, I knew that many field researchers would think our project was unnecessary, as tool use and manufacture by wild chimpanzees was historically,

³⁴ This method (or a suitably similar one) is so foundational to recognizing the common denominator among all experiments in this genre that two questions frequently arise in my mind: (1) Why have more analyses of this type not been applied to the scores of experiments that claim to seal the deal about higher-order thinking in animals? and, (2) Should researchers be required to produce such formalisms as part of their publications and/or the pre-registration of their experimental designs? The latter suggestion can be seen as complementary to the concerns of publication bias addressed by Farrar et al. (2020). In offering concerned graduate students a tool for doing the work themselves, my hope is that every new study will be analyzed in this manner before untold hours are invested collecting undiagnostic data. I believe there is a fishing proverb applicable to this situation. I am well aware this may not satisfy some scholars who favor detailed reviews of scores of studies.

³⁵ Some experimentalists may be tempted to skip this discussion, already convinced that data from the field should not be used in this way: *Strong inferences about the presence or absence of higher-order reasoning in animals can never be drawn from such observations. Their sole utility is for generating hypotheses to be tested in the lab.* This may be how things usually go, but I urge caution. Specifically, I argue it is vital to examine whether this conclusion is logically true, and if it is, how this reasoning impacts the current experimental approaches. There are two additional reasons why all parties should take this discussion seriously: (1) field researchers will sharpen their understanding of what is *representationally* at stake in the debates, and (2) experimentalists can avoid designing studies that simply re-demonstrate what we already know from functional-level descriptions of field-derived data (viz. the *Experimental Necessity Dilemma*).

functionally, and visually impressive. I imagined a skeptical reaction because the main skeptic was, in fact, myself:

'Look . . . isn't it obvious that chimpanzees and other great apes understand the physical principles governing simple tool use in just about the same manner that we do? Haven't we all seen enough National Geographic specials to know that chimpanzees make and use tools spontaneously and naturally? They crack nuts open using hammer stones and make simple fishing wands to extract termites from their mounds. So why do we need to bring them into the laboratory and test them on their ability to use tools? . . .' Starting with this skeptical voice may seem odd to some readers, suggesting we have adopted a defensive posture from the outset. Nothing could be further from the truth. After all, on our very best days as scientists, this skeptical voice repeats over and over . . . reminding us of the fact that our project is a difficult one indeed (Povinelli, 2000, p. 2).

Thus, in calling for a critical investigation of *representational-level* claims about the cognitive basis of tool use, *Folk Physics* began by fully embracing *functional-level* explanations of tool use as a vital part of what a chimpanzee is: an evolved organism situated in a range of natural ecologies, using a range of objects (often intentionally modified) to aid in their efforts to extract resources from their environments.³⁶

So, how has the claim the naturalistic data cannot implicate higher-order reasoning about the physics of tool use fared?

I begin with a not-so-surprising answer: I have yet to encounter a rigorous methodology for leveraging either the final or the developing forms of naturally-occurring tool-related behaviors in chimpanzees (or any other animals) in a way that can strongly implicate higher-order reasoning.³⁷ I do not claim such a theory is impossible (some may think they already have one, see Andrews, 2020), but I include myself among those who have no clear idea how one would use evidence about naturally developing behavioral forms to systematically avoid the *Asymmetric Dependency Problem* (see below). After all, if the *Asymmetric Dependency Problem* cripples our ability to make sound inferences about higher-order reasoning in the context of the current genre of experimental investigations, it applies with equal force to naturalistic observations. The problem casts a long shadow.

Despite this, I admit I am frequently tempted to think that better data from the natural ecologies of these organisms could (somehow) overcome this problem. Since the publication of *Folk Physics*, field researchers have documented not only both a greater diversity of tool-making and tool-using among wild species (e.g., Hicks et al., 2019; Sanz et al., 2009), but also, in some cases, have carefully mapped the developmental trajectories of these skills (Boesch et al., 2019). And quite understandable, as field researchers have acquired these hard-won data, they have offered general ideas about the kinds of causal understanding that may be involved in such tool use, as well as potential fitness advantages:

³⁶ This polemic is distinct from questions of whether Candy and her peers, raised in captivity, are different from their wild counterparts in ways that matter to the representational-level questions at hand (see below, "Claim 5: Chimpanzees Raised in Captivity"; see also, Povinelli, 2000, p. 15-19, 326-328; Povinelli & Vonk, 2004, p. 28; Povinelli, 2012, p. 62-65).

³⁷ Some systematic attempts have been made to use massive compendiums of anecdotes (solicited from lab and field workers alike) to implicate particular kinds of mental representations (e.g., O'Connell, 1995; Whiten & Byrne, 1988). Although scholars have moved away from the use of anecdotes in this manner (see Ramsay & Teichroeb, 2019), the anecdotally inclined reader will be heartened to discover that the journal *Behaviour* has recently implemented a formal mechanism for submitting and archiving new ones (Kret & Roth, 2020).

Notions of cause and effect could notably improve the performance of tool-use, particularly when selecting the raw material to make a tool. Appropriate choice of hammers allows a 30 to 43% energy gain in nut-cracking. Thus, understanding of cause and effect allows chimpanzees to mentally anticipate their actions and to choose tools adapted to specific purposes (Boesch & Boesch-Achermann, 2000, p. 233).

The brush-tipped fishing probe is not inherent in the structure of the herb stem, but rather arises from transformation of the raw material that shows evidence of being deliberate. In particular, the lengthwise pulling of the probe through partially closed teeth is a behaviour that was not observed in other contexts and was often repeated several times during brush manufacture. These transformations also increased the effectiveness of these tools. Our results indicate that chimpanzees have a mental template of the tool form, which is employed in crafting the tool prior to use and refining it during use... (Sanz et al., 2009, p. 295).

Data from all chimpanzee study sites suggest that a proportion of tool-using behaviors are regularly exhibited, while others are manifested only rarely and in response to specific stimuli. It is certain that there is a high degree of variation between sites. In contrast to the rigid patterns of tool use observed in other taxa, apes respond to particular situations with adaptive solutions, which may demonstrate an understanding of causality in external objects (Sanz & Morgan, 2007, p. 432).

The videos . . . suggest that adult Hawaiian crows clearly understand the problem (the food is in the hole) and the solution (use a stick to get the food out). However, the way Hawaiian crows treat their tools offers a possible clue to their working memory capacity . . . After each probe, they drop the tool before pecking at the food. In contrast, New Caledonian crows usually safeguard their tools in between probes by deliberately and carefully trapping them under a foot, on a branch, or in a hole. The difference could be related to working memory, motivation, or necessity In sum, even if a species understands the advantage of tool use, sufficient working memory and manipulation skill are probably important for long-term maintenance of the behavior (Uomini & Hunt, 2017, p. 206).

In this study, we systematically compared tool-transfer behavior between Goulougo and Gombe chimpanzees and found significant population differences in this form of scaffolding. ...Broader comparative studies will continue to inform us about the capacity for different types of scaffolding, including tool transfers, across species, while assessing multiple tool contexts within species will further illuminate how helping varies with task demands. Differentiating specific types of helping is also essential for elucidating the potential cognitive underpinnings of these behaviors (Musgrave et al., 2020, p. 974).

But can such promissory notes ever be cashed? In other words, can detailed descriptive data derived from observations of naturally-occurring tool use resolve the representational-level debates targeted by *Folk Physics* and its sequel, *Weight*? Data on how wild animals use and manufacture tools can undoubtedly implicate the kinds of information the animals *must somehow keep track of* as they interact with objects. But as we have seen with Candy and her coconut, many kinds of mental activity can be comfortably conflated under the catch-all terminology *causal understanding*, *working memory*, *mental templates*, *planning*, *inhibitory control* are uncontroversial—unless you take the *Conflation of Explanatory Levels* problem seriously (see especially, Penn et al., 2008a; Penn & Povinelli, 2009).

The recent interest in New Caledonian crows is a good illustration of how, in one way, most scholars are fully aware of these issues. Hunt's (1996) discovery that wild New Caledonian

crows were manufacturing and using hook-shaped tools in the wild excited (for good reason) great interest about the cognitive underpinnings of the behavior:

Crow tool manufacture had three features new to tool use in free-living nonhumans: a high degree of standardization, distinctly discrete tool type with definite imposition of form in tool shaping, and the use of hooks. The features only first appeared in the stone and bone tool-using cultures of early humans after the Lower Paleolithic, which indicates crows have achieved a considerable technical capability in their tool manufacture and use (Hunt, 1996, p. 249)

But rather than interpreting these natural history observations as settling representational-level claims, researchers turned to the laboratory to test various representational-level interpretations of the phenomenon (for a few examples among many, see Jelbert et al., 2019; Taylor et al., 2009a, b; von Bayern et al., 2009; Weir & Kacelnik, 2006; Wimpenny et al., 2009). And while my assessment (see below) is that none of these tests have implicated any of the higher-order mental operations at stake in *Folk Physics* and *Weight* (see also Bluff et al., 2007; Rutz et al., 2016), they have revealed possible (uncontroversial) mental operations involved in the tool-related skills of these birds, as well as possible ontogenetic constraints on their development. For now, I simply note that researchers never interpreted the spontaneously developing form of the behavior as adequate to distinguish between first- versus higher-order relational reasoning.³⁸

Thus, twenty years after *Folk Physics* was published, the question of how to use the natural trajectories of development (including critical weigh points along those trajectories) to implicate higher-order reasoning remains unsolved. The *Asymmetric Dependency Problem* implies that those natural developmental pathways can be fully explained by powerful cognitive engines related to first-order, perceptually-based relational reasoning. Again, this is distinct from *functional-level* debates about which species possess greater or lesser flexibility in their tool use and manufacture (Auersperg et al., 2011; Kabadayi & Osvath, 2017; Völter & Call, 2014).³⁹

I would be remiss if did not at least briefly mention one approach to thinking about the natural behavior of animals (including their tool-using behavior) that some might believe can overcome these problems: *interpretive anthropomorphism*. A lot has been written about anthropomorphism, and it is certainly not my intention here to directly engage with the multifarious debates over whether anthropomorphism is a fundamental fallacy (I don't think it is),⁴⁰ nor whether *being anthropomorphic* uniquely creates problematic science (I don't think that's true either), or whether intentionally preying upon the human capacity to think anthropomorphically can lead to widespread misunderstanding of the inferences warranted by a given empirical finding (I think it can and frequently does).⁴¹ Here, I simply wish to explore whether *the fact of anthropomorphism*

³⁸ Again, none of what I have said here should be construed as downplaying the importance of collecting such information from free-ranging populations of chimpanzees and other animals. Data from the wild is indispensable for building and testing evolutionary and socio-ecological theories of tool use (including the current controversy over whether the intra- and intergenerational transmission of tool use is driven by cultural or individual learning, see Bandini & Tennie, 2017; Tennie et al., 2009; see also Bernstein-Kurtycz et al., 2020). The issue at hand is not (and never was) about the relative importance of data derived from the field versus the laboratory, but about how such data sets match up with the *interpretative level* one seeks to address.

³⁹ Recent investigations of tool use in captive Goffin's cockatoos—a species not known to be natural tool-users—reveals a degree of flexibility comparable to that of chimpanzees, leading some to suggest that they may be on the brink of discovering tool use in the wild (see, e.g., Auersperg et al., 2012, 2016; Osuna-Mascaró & Auersperg, 2018).

⁴⁰ See Povinelli (1997) for a discussion of the possibility that animals might engage in their own species-typical "-morphisms" (e.g., *Panmorphisms* by chimpanzees).

⁴¹ Here is a gentle example. Alex Taylor and colleagues recently produced a report entitled, "New Caledonian crows behave optimistically after using tools" (McCoy et al., 2019). They quickly note, however, that "The

offers a good reason to use using species-typical behavioral forms to resolve debates over higher-order relational reasoning in animals, and what it would mean for the field of comparative psychology if it were the case.

A weak way of using anthropomorphism might begin with folk psychological theories (generated, of course, by humans) about the causal role that higher-order constructs play in *human* tool-use, and then use this as a source of generating hypotheses about functionally similar cases of animal tool use. One can (and I think, should) question the accuracy of such folk theories about the human case (see below, "Claim 3: Human Folk Psychology (About Human Folk Physics)"). Nonetheless, such *heuristic anthropomorphism* is not only widespread, but also unavoidable (Andrews, 2020; Asquith, 1984; Boesch, 2020; Povinelli, 1997).⁴² It certainly was the opening foil of *Folk Physics*:

...[I]t is possible simply to suppose from the outset that because chimpanzees are so closely related to us, and because they must confront more or less the same physical universe as we do, they understand the world in a very similar manner. We adopt a different approach. The purpose of the research . . . is to break down this supposition of similarity into a series of specific, testable hypotheses concerning chimpanzees' understanding of concepts such as gravity, force, mass, shape, and physical connection (to name just a few), and then subject these hypotheses to serious experimental scrutiny (Povinelli, 2000, p. viii).

But heuristic anthropomorphism, by definition points us back to the laboratory experiments and therefore does not attempt to use naturally developing behavioral forms to draw strong inferences about the presence of higher-order thinking.

However, there is a second way of using anthropomorphism to try to resolve debates over higher-order thinking animals—what Fisher (1996) called *interpretive anthropomorphism*. Summarized by Keeley (2004), interpretive anthropomorphism can be understood as "the explanatory gambit of interpreting an animal's traits as being caused by similar mechanisms or constituted in ways similar to human traits" (p. 529). For both Fisher and Keeley this again points back to standard scientific business-as-usual experimental practices to adjudicate. But for other scholars, interpretive anthropomorphism can be used, full stop, to generate strong inferences about higher-order processes. For lack of a better phrase, let us call this *strong explanatory anthropomorphism*.

Andrews (2020) examines and advocates this approach (among others) in her recent book, *How to Study Animal Minds*. As an example, she explores what she interprets as the spontaneous *pantomiming* communicative gestures of ex-captive juvenile orangutans under rehabilitation (see Russon & Andrews, 2011; Russon, 2018). First, the naturalistic descriptions and functional-level explanations:

...[we] noticed that the orangutans would sometimes act out what they wanted their human caregivers to do for them... We spent time with these orangutans and their human caregivers and learned what their typical practices and behaviors were. It was easy to see that the orangutans regularly gathered together in dusty areas, wrestling in the dirt, collecting handfuls of dust like children in a sandbox, and dumping it on their own heads.

terms optimism and pessimism in this context are short-hand labels for responses made to ambiguous cues from which, respectively, positively and negatively valenced affective states ("dimensional" emotions and/or moods) can be inferred without implying that these are consciously experienced" (p. 2737). Does this encourage the conflation of first- and higher-order representations? When all is said and done, ants behave optimistically, as well—or else they don't (see Czaczkes et al., 2018).

⁴²There are obvious and intricate connections here to Daniel Dennett's (1987) original formulation of the intentional stance and the voluminous literature that has blossomed from it.

We also soon came to expect that their human caregivers would clean the little orangutans after their play bouts, brushing the dirt from their heads with leaves. Given our observation of normal behavior...we were able to recognize a behavior as a request from an orangutan, Cecep, for Russon to clean him. Cecep approached and sat in front of Russon, picked up a leaf, and handed it to her. Russon used it to briefly clean Cecep's head, then dropped it on the ground. Cecep picked up and handed Russon another leaf, but this time she played dumb and just examined the leaf. After a few seconds Cecep took the leaf back from Russon, rubbed it on his own head while looking her in the eye, and then placed it on her notebook... (p. 43)

Next, the step involving *strong explanatory anthropomorphism*:

We interpreted this event as Cecep asking Russon to clean his head by handing her the leaf, and when she didn't respond as he expected, Cecep elaborated on his message by pantomiming—acting out—what he wanted Russon to do (p. 43)

The scholars then recovered many other incidents that Russon had previously recorded, ultimately amassing 62 incidents from her detailed field notes:

With this dataset, we were able to analyze the contexts in which the pantomimes were exhibited in order to determine the functions of these gestures. We found that in all but one case the orangutans used the gesture imperatively; they tended to use it to elaborate a prior failed message. In seven cases they used pantomime in a deceptive context, and in one case an orangutan pantomimed in a declarative context (p. 44).

Using these communicative data, Andrews (2020) then explicitly endorses becoming cultural and linguistic "anthropologists" among animals. While it is unclear (to me, at least) if she is specifically attempting to draw a strong inference about the presence of higher-order thinking (in the orangutans), her invocation of *the principle of charity*⁴³ suggests she might be. But whether she is or is not, I see no reason why this method could not be used to infer higher-order mental states such as <beliefs>, <desires>, <seeing>, etc., or in the case of tool use, <force>, <weight>, <gravity> (or for that matter, <time> and <space>, etc.).⁴⁴ Indeed, other researchers (while ignoring the issues raised by the *Unprincipled Titration Paradox*), pursue something in this ballpark when attempting to infer the cognitive states attending tool use in the wild (see below, "Do Data From the Wild Demonstrate...?"). Indeed, if one is inclined to open the interpretive floodgates

⁴³ For psychologists who may not appreciate the scope and importance of Andrews' (2020) invocation of the *Principle of Charity* in interpreting orangutan gestures, I suggest beginning with Feldman's (1998) overview of the issues involved when applying it to the utterances of other humans, which begins: "The principle of charity governs the interpretation of the beliefs and utterances of others. It urges charitable interpretation, meaning interpretation that maximizes the truth or rationality of what others think and say. Some formulations of the principle concern primarily rationality, recommending attributions of rational belief or assertion. Others concern primarily truth, recommending attributions of true belief or assertion. Versions of the principle differ in strength. The weakest urge charity as one consideration among many. The strongest hold that interpretation is impossible without the assumption of rationality or truth."

⁴⁴ Daniel Dennett has been making this point so clearly, and for so many years, that I can do no better than quote him directly: "Darwin's theory of evolution by natural selection unifies the world of physics with the world of meaning and purpose by proposing a deeply counterintuitive "inversion of reasoning"...to make a perfect and beautiful machine, it is not requisite to know how to make it"... Turing proposed a similar inversion: to be a perfect and beautiful computing machine, it is not requisite to know what arithmetic is. Together, these ideas help to explain how human intelligences came to be able to discern the reasons for all of the adaptations of life, including our own" (Dennett, 2009, p. 10061).

by using strong explanatory anthropomorphism, the possible attributions seem endless. This is especially true given that Andrews (2020) argues that in such cases, (1) no controlled laboratory tests are possible (or necessary) in order to generate strong inferences, and (2) no situational, contrastive observations are likely to add useful information (nor are they required to make the relevant inferences).

So, does strong explanatory anthropomorphism offer a way around *Folk Physics'* claim that naturally developing behavior cannot be used to implicate higher-order thinking? Many years ago, in the context of critiquing the argument by analogy for other minds, my colleagues and I spent a great deal of time examining the inferential flaws of such an approach (see, for example, Povinelli & Giambrone, 1999). I believe those same objections apply (in full force) here. Rather than reiterating them, however, let me offer a more general objection. Imagine one concedes that the fatal problems implicated by the *Unprincipled Titration Paradox* are real. If so, *the fact of anthropomorphism* (whether honed, morphed, elaborated, and/or constrained by living with orangutans or not) in no way undermines the fact that first-order explanations are both necessary and sufficient to explain the orangutans' behaviors—especially when there is good reason to believe such inferences are frequently incorrect in the human case (see below, "Claim 3: Human Folk Psychology of Human Folk Physics").

A tempting conclusion is that it may be best to embrace a pluralistic approach, letting all methods proceed, including experimental work, with the hope that convergent data will ultimately emerge. This certainly *sounds* right (and, if the past is any guide to the future, probably how things will play out anyhow). Alas, despite its ameliorative tone, I believe this is a fool's errand. Why? Because as I show next, the second opening argument of *Folk Physics* was only half correct: none of the experimental methods that were available then, nor any developed since, have the power to uniquely implicate the causal work performed by higher-order relational reasoning. If this is right, until the culture of this area of comparative psychology radically changes, there can be no converging evidence.

Claim 2: Tool-Using Experiments Can Assay Higher-Order Thinking

Background assumptions. While never stated explicitly, *Folk Physics* assumed (like many projects before and after it)⁴⁵ that standard psychological experiments (properly conceived) had the power to allow researchers to make strong inferences about either the presence *or* the absence of higher-order thinking:

...[I]t is possible simply to suppose from the outset that because chimpanzees are so closely related to us, and because they must confront more or less the same physical universe as we do, they understand the world in a very similar manner. We adopt a different approach. The purpose of the research . . . is to break down this supposition of similarity into a series of specific, testable hypotheses concerning chimpanzees' understanding of concepts such as gravity, force, mass, shape, and physical connection (to name just a few), and then subject these hypotheses to serious experimental scrutiny (Povinelli, 2000, p. viii).

As I explain below, I now see this argument as only half correct: While our experiments may have the power to implicate the *absence* of higher-thinking about objects, they do not have the power to implicate its *presence*.

⁴⁵ For putative (and mostly rhetorical) defenses of the argument that standard "behaviorist" experimental designs could be used to diagnosis higher-order thinking in animals (e.g., theory of mind), see especially, Premack and Woodruff (1978), Premack (1988) and (alas) Povinelli and Eddy (1996a).

The experimental logic of Folk Physics (and Weight). Similar to our investigations of social cognition, the strategy of *Folk Physics* and *Weight* was to present Candy and her companions with tests designed to distinguish between competing representational-level accounts of the kinds of tool-using behaviors that are slowly mastered in free-ranging populations of chimpanzees (typically between 3-8 years of age; see for example, Boesch & Boesch-Achermann, 2000; Boesch et al., 2019; Inoue-Nakamura & Matsuzawa, 1997; Sousa et al., 2009). To do so, we once again followed the method of presenting them with stimuli and conditions that we hoped could pry apart variables nature typically conjoins, not only during the natural development of tool use in the wild, but also during the many spontaneous bouts of tool use had observed in Candy and her peers outside of testing (see below, "Candy's Family").⁴⁶

The broad idea was to micro-genetically track the development of our apes' physical reasoning skills as they interacted with more and more tools in our experimental settings, *even as they were developing many forms of tool use outside of testing.*⁴⁷ Although there were notable exceptions, in both *Folk Physics* and *Weight* we rarely attempted to train our apes exhaustively on such tests. Instead, we typically assessed their incremental performances in small groups of trials (usually four).⁴⁸ Why? Because for the most part we knew they could learn just about any tool-using task we gave them. Our goal, on the contrary, was to examine their transfer within and across types of tool use and manufacture. As we made clear at the time, *the tests were never an attempt to determine what our chimpanzees could or could not learn to do.* Here's an interim assessment, about half-way through the project:

...a fairly clear pattern can be seen in the results of the experiments presented thus far. Initially, our apes performed as if they had no understanding of the relevant folk physics of the problem at hand. However, with additional opportunities for learning, their performances improved, and indeed, in some cases there was evidence that the apes detected and used the same relevant perceptual features of the task as humans. However, in each case...this knowledge did not transfer easily to perceptually novel [sic], but conceptually similar tasks (Povinelli, 2000, p. 205).⁴⁹

⁴⁶ Some reviewers of *Folk Physics* worried that our chimpanzees (who began the studies at about 6 years of age and completed them around 11 years of age) might have been too young to display their full tool-using abilities (e.g., Allen, 2002; Hauser, 2001). Although understandable, the kinds of tool use that are frequently offered as evidence for higher-order understanding (e.g., nut-cracking), develop between 2.5 to 5 years of age in chimpanzees, with additional mastery up to 8-9 years of ages (see Boesch & Boesch-Achermann, 2000). Nonetheless, we were quite cognizant that our project was chasing both our apes' age and their experience (see Povinelli & Eddy, 1996a, Chapter 6; Povinelli, 2000). This is one (but not the only) reason we deployed longitudinal methodologies on several key tasks (e.g., Reaux et al., 1999). If nothing else, each time point clearly showed how easily our thinking about their social and cognitive skills could be fooled (and double fooled) by our personal folk psychology (see above, "Fooled" & "Double Fooled"). For what it is worth, the published results of *Seeing*, *Folk Physics* and *Weight* all followed our chimpanzees as they matured from juveniles into adolescence, and for *Weight*, into full adulthood. This latter point has been lost on some who have confused task experience with age.

⁴⁷ Later in this essay, I offer a glimpse into the spontaneous development use of tools that our chimpanzees displayed in their everyday lives as the 59 experiments reported in *Folk Physics* and *Weight* were being conducted (see below, "Candy's Family"; see also, Table 1).

⁴⁸ As wild chimpanzees learn to crack nuts, for example, they have far more than four, eight, sixteen, thirty-two, etc. trials to learn how to bring their sensory-motor skills and cognitive capacities into alignment for mastery of the skill. Because of our audacity in attempting to ask about the precise nature of those sensory-motor and cognitive skills, great violence has been perpetrated on Candy and her companions by researchers who continue to conflate their experience on a given task with their cognitive abilities (see below, "Claim 5: Chimpanzees Raised in Captivity").

⁴⁹ We stressed this simple fact throughout *Folk Physics*: "We close this chapter by emphasizing we have little doubt that, with considerably more experience on their part, and considerably more patience on ours, [more]

A typical example was their training and testing with rakes and hook tools. After they became experts on one version of the hook task, we developed and evaluated various models that made different predictions about expected *patterns of transfer* depending on which features of the tasks were influencing their behavior (see also Povinelli & Frey, 2016).⁵⁰ Whether they could *learn through experience* was never the aim of our project. Chimpanzees are biologically pre-prepared to engage in object related actions such as pulling, pounding, shaking, dropping, etc. (see especially, Boesch & Boesch-Achermann, 2000, p. 201-224; see also, Boesch et al., 2019). Thus, demonstrating that they could engage in such activities in captivity was of no particular interest to us.

Return of the *Asymmetric Dependency Problem*. But if Jennifer and I were correct that the *Asymmetric Dependency Problem* crippled any available (or easily conceivable) experimental inferences about higher order social cognition in animals (i.e., theory of mind), and this applied to all other domains of higher-order cognition (see Vonk & Povinelli, 2006), an urgent question arose for our laboratory. If our measurement tools could not implicate the *presence* of higher-order thinking, could we continue to use *specific patterns of transfer* to test for its absence? This was an especially important concern as we were already several years into our investigations of whether chimpanzees possess a higher-order understanding of *<weight>*. Getting the right answer to this question remains as vital today as it was then.

So, can narrowly circumscribed patterns of transfer between tests yield strong inference regarding the *absence* of higher-order thinking? Although I continue to believe they can (see below), in *Weight* my colleagues and I tried to emphasize the opposite problem: namely, the inferential weaknesses of the high-level models. Indeed, we bent over backwards to clarify the ultimate source of these models: our intuitive (folk) ideas about how *<weight>* influences our own behavior. For example:

Our *folk model* of this [box-pulling] task predicted that, if our chimpanzees were able to represent [*<weight>*], they would interpret their caretakers' distinctively different behaviors with the boxes as evidence that one was "heavy" and the other was "light," thus leading them to pull the light box. On the other hand, if [*<weight>*]-related representations were unavailable to our apes, we would expect their choices to be equally distributed between the two boxes. Of course, one can immediately question whether this folk model is an accurate description of the cognitive operations a human might use to solve such a task (Povinelli, 2012, p. 147).

Our point here was that we were generally flying blind: assuming we knew how our own higher-order representations articulate with our behavior in such situations (see below, "Claim 3: Human Folk Psychology (of Human Folk Physics) is an Unstable Foundation").

Setting aside the question of whether there are any patterns of results that could have supported the inference that our apes were engaged in higher-order thinking, what results actually

of our apes ... could have learned to solve the basic trap-tube problem (p. 131). "Reflecting on the results of both the trap-tube and trap-table tasks, it seems clear that chimpanzees will uncover the regularities inherent in such simplistic problems... [T]he learning curves displayed by all of the apes across the trap table experiments ...highlight the central role...that direct feedback plays in their acquisition of such competences" (p. 147). In *Weight*, we spent even more time giving our apes full mastery of tasks in order to conduct what we had hoped would be more informative transfer tests.

⁵⁰ Many of these models were not even specifically about higher-order constructs (see Povinelli, 2000, p. 267). In retrospect, I wonder if this was because we were already realizing that the only thing we directly measure was which first-order perceptually-based relations they learned most readily.

emerged in our *Weight* experiments? After many transfer tests—tests that confronted our apes with relations ranging from those likely to be highly developmentally canalized, to those that were functional analogues of using weight as a tool in the wild, to those that were completely arbitrary (e.g., sorting objects based on weight)—a consistent pattern emerged. The predictions of the higher-order folk models were rarely if ever supported, whereas our chimpanzees displayed immediate (from the standpoint of our experimental tests) understanding of many first-order relations involving *effort-while-lifting* and/or *effort-to-lift* (see Povinelli, 2012, Chapters 3 & 7). We concluded that this overall pattern in the transfer tests implicated the absence of reasoning about *<weight>*.

This may feel troubling to some. If first-order, perceptually-based relational reasoning involving representations such as *effort-to-lift* or *effort-while-lifting* are both necessary and sufficient to account for the development of skills such as nut-cracking with stones, then why didn't our chimpanzees evince immediate evidence of understanding of all of the causal regularities in our weight tests? The answer is the same as why young chimpanzees in the wild take years to become effective at cracking nuts: experience is required (see Boesch & Boesch-Achermann, 2000; Estienne et al, 2019). More to the point, as Osiurak et al. (2010) have noted, despite their mastery of hammer stones, those wild chimpanzees do not use weight *for any other purpose*. It is a truism of cognitive science that given an organism's starting resources, some first-order relations will be far easier to learn than others (see Clark & Thornton, 1997; Povinelli & Penn, 2011; Povinelli, 2012).⁵¹

But the more pressing question remains: If the *Asymmetric Dependency Problem* is correct, and an entire genre of experimental tests cannot implicate the presence of higher-order thinking, then how can we claim that the very same tests can implicate its absence? Although a lot can be said here, the most straightforward response is that the two claims are logically consistent. As Jennifer and I noted many years ago in the theory of mind debate: "We are simply proposing that a pattern of results of type 'x' could be produced by either [first- or higher-order social cognition], but that a pattern of results of type 'y' would be expected for [a first-order system] but not [a higher-order one]" (Povinelli & Vonk, 2004, p. 20).

Thus, the causal asymmetry highlighted by the *Asymmetric Dependency Problem* applies to the general relationship between first- and higher-order mental representations. To wit, showing that Candy can learn to deploy a hook-shaped tool in order to create a mechanical force vector sufficient to retrieve a banana, does not suffice to conclude she needs *<force>* representations to do so. On the other hand, a consistent pattern of transfer failure among conceptually, but not perceptually, related tasks might be able to implicate its absence. Whether this can be achieved in a sufficiently precise manner to produce a strong argument is open to debate (see Povinelli, 2012, Chapters 11 & 12).⁵²

Current status of animal folk physics: burying the target rather than striking it. In the twenty years since *Folk Physics*, comparative psychologists have been busy. They have secured an experimental cache of results that would make scrub jays jealous. In the early days, much of this

⁵¹ To elaborate, our argument is that first-order, perceptually-based representations like *effort-while-lifting* and *effort-to-lift* allow chimpanzees (and other species) to enter certain tests and immediately deploy robust first-order relational reasoning, whereas those same representations do not easily prepare them for tasks unconnected to their evolutionary ecology (e.g., spatial sorting of objects based on *effort-while-lifting*). I invite the interested reader to compare and contrast the experimental results reported in Chapters 3 versus 4 of *Weight* (see also Schrauf & Call, 2009).

⁵² I should note that the fate of the *Asymmetric Dependency Problem* is independent of this claim. If we cannot develop a convincing strategy to use transfer failures to build a case for the absence of higher-order thinking, this simply makes the problem we face viz. higher-order relational reasoning in animals even more intractable.

activity swirled around a few tasks that examined (for example) whether slight alterations in existing experimental designs or species could speed up rates of learning (see Weir & Kacelnik, 2006). Some research provided startling convergence with the findings in *Folk Physics* (e.g., Herrmann et al., 2008). Others' research was interpreted as contradicting out results, despite the fact that it did not (e.g., see the detailed analysis of Manrique et al. (2010) by Povinelli & Vonk, 2012). Other research showed that chimpanzees could learn to avoid a trap faster when they were allowed to use their fingers instead of a tool (e.g., Seed et al., 2009). But beyond these narrow confines, an armada of new procedures arose, involving a variety of new species. A wide range of claims attended them. For just one example, consider the explosion of studies that followed Bird and Emery's (2009a) invention of the Aesop's fable paradigm (for a discussion of the folk scientific framing this paradigm, see Barker and Povinelli [2019b]).

From my perspective as an anthropologist trained in evolutionary biology, *there are many important reasons to investigate these phenomena that have nothing to do with whether animals possess higher-order thinking*. For example, one might be interested in testing functional-levels of explanations for why tool use is so rare in nature (e.g., Hunt et al. 2013), despite how easily it can sometimes be drawn out in captivity (see, e.g., Beck, 1980; and footnote 36). Or, one might be interested in testing functional-level models about the relationships among brain size, tool use, and social organization (e.g., Emery & Clayton, 2004)—or between tool use and animal cultural evolution (Laland, & Janik, 2006; Schuppli & van Schaik, 2019; Whiten, 2017). Indeed, much of the work in the comparative psychology of tool use these days seems at least as much oriented toward these vital, functional-level questions as the ones that motivated *Folk Physics*.

However, as the theory of mind debate illustrates, the enterprise can easily go awry when terms like *causal understanding*, *complex cognition*, *tool functionality*, *flexible tool selection* (to name a few) begin to conflate levels of analyses. The conflation of functional- and representational-level claims inevitably leads to a downstream muddling of the distinction between specific first-versus higher-order representational-level hypotheses (see Penn & Povinelli, 2009). False contrasts begin to obscure the core questions that motivated projects like *Seeing*, *Folk Physics* and *Weight* in the first place: Do animals possess the equivalent of the human capacity to reason in a systematic, structural, role-based manner over disparate perceptual relations or not (see Penn et al. 2008a)? In what follows, I conclude that, with respect to *this question*, an entire generation of research has, in effect, buried the target rather than striking it.

To begin, several illustrative (and, I claim, typical) examples can be found in the contributions to the August 2020 special issue of *Animal Behavior and Cognition*. Consider Jordan et al.'s (2020) introduction and summary of their intricate studies of monkeys with cups and puzzle boxes: "Despite studies showing that some nonhuman primates can discriminate between functional and non-functional tools, whether they achieve this by recognizing an object's physical properties or via associative learning of perceptual cues remains contested" (p. 365) and "...this group of experiments adds to the current literature suggesting that capuchins—but not squirrel monkeys—are sensitive to the functional properties of objects, and specifically to solidity" (p. 388). From the point of view of the questions that gave birth to *Seeing*, *Folk Physics*, and *Weight*, the distinction between *recognizing an object's physical properties* and *associative learning of perceptual cues* is incoherent: the relevant physical properties of an object *just are* the first-order perceptually-based information that an animals must possess for higher-order thinking.⁵³ And the *Asymmetric*

⁵³ Examples abound of how this kind of conflation can be promoted by subtle linguistic slippage. Seed and Byrne (2010): "Behaviour like this raises the intriguing possibility that the animals represent the physical properties and forces involved in the tool-using event in an abstract, conceptual way: in terms of properties such as rigidity, continuity, and connectedness. The simpler alternative is that the animals' thinking is grounded in perceptual features of the objects (their shape, feel, or spatial orientation). Psychological experiments have often capitalized on tool-using (or proto-tool-using) behaviour to try to tease these alternative explanations apart. For many years, laboratory studies gave results supporting the simpler

Dependency Problem shows that such first-order representations are necessary and sufficient to explain the relevant behaviors in any given task (Penn et al., 2008a; Povinelli & Vonk, 2004). As we explained in *Folk Physics*:

...the chimpanzee will rarely be fooled by superficial alterations of the task. Changes in the color, size, or even general perceptual form of the tool or platform will rarely befuddle the experienced chimpanzee. ... This is because the experienced chimpanzee will have already 'seen through' this level of surface features, and has located the perceptual features and spatial arrangements of the objects that yield the outcome desired. ... [O]f equal significance ... is that the chimpanzee will freely substitute any tool that will generate contact, and indeed, may even avoid the hook tool [for example] if some other tool will make the requisite contact more effectively (Povinelli, 2000, p. 307).

Based on existing data from the empirical literature, our projects presupposed that chimpanzees possess stable, atomistic perceptual symbols that correspond to the set of abstracted physical/functional properties/features of the tools they use in any situated task. Whether these symbols correspond to what our folk psychology of *their folk physics* assumed beforehand is largely beside the point.

At a structurally equivalent level, Amodio et al. (2020) wield the constructs of *complex physical cognition*, *tool selectivity*, and *contingency learning*. Although these phrases may or may not pick out something useful for building functional-level explanations, whatever they do pick out is orthogonal to the representational-level debate highlighted in *Folk Physics*. In other words, the phenomena isolated by such phrases are not logically inconsistent with, nor are they competing explanations for, the representational-level distinctions at play. Instead, they are labels that capture various aspects of first-order, perceptually-based relational reasoning (see Penn et al., 2008a; Povinelli & Penn, 2011). Indeed, to the extent these labels thrive inside discussions of higher-order thinking is, I submit, because of their tight connection with the prevailing scientific folklore about the dangers of invoking explanations based on 'associative learning'—a catchall term that has nothing to do with the representational-level questions raised by *Folk Physics* and *Weight* (see above, "Experimentalists React").⁵⁴

From a different point of view, consider an example discussed by Alex Taylor (2020) in his thoughtful reflection piece in the special issue. He considers a study by von Bayern et al. (2009) in which four New Caledonian crows were initially shown how to use their beak to push down (collapse) a platform to get a worm. When the platform could no longer be directly accessed, two of the four birds dropped stones into a tube above it. Taylor notes that two hypotheses can be offered to account for this, one involving the first-order relation *contact-leads-to-collapse-and-access-to-worms*, the other involving a higher-order relation *<force>-<causes>-collapse-and-access-to-worms*. Using the Seven-Step Program, it becomes immediately obvious that any *<force>* description depends upon the bird also possessing a *contact* description (see Povinelli, 2000, Chapter 12). And, because *contact* is necessary and sufficient to explain the results, there is no role for the higher-order (*<force>*) representation to play in explaining the results. It is only *our* folk psychology of the folk physics of the apparatus that adds the notion of *<force>*.⁵⁵ The same

explanation. *Even chimpanzees seemed to be using perceptually-based information rather than an abstract notion of object properties*" (italics added, p. R1035). Until young scholars are trained to detect the flawed contrasts made in such passages, there will be little hope for making progress in this field.

⁵⁴ I have grown convinced that standard verbal descriptions of the theoretical variables in the experimental literature are too imprecise to allow progress in these debates. They must be replaced by a set of shared, formal notations (see also, footnotes 23 & 31). I hope the Seven-Steps program is a foot in this direction.

⁵⁵ Furthermore, New Caledonian crows in general (and perhaps even these specific birds) were already known to use sticks and other materials to make contact with out-of-reach objects. Thus, although the

analysis can be applied, *mutatis mutandis*, to Barrett and Benson-Amram (2020) and DeLong and Burnett's (2020) investigations of the spitting behavior of elephants and orangutans.⁵⁶

In direct contrast to the preceding examples, consider the work presented in the special issue by Kersken et al. (2020) on object individuation by capuchin monkeys. Stripped of unnecessary theoretical baggage, these investigators show how great strides can be made toward imagining (and testing) less folk psychologically tainted ways of thinking about *thinking-about-the-physical-world*. Yet even here, despite all their excellent work, they balk at the full implication of the *Asymmetric Dependency Problem*:

...the type of sensitivity to property and kind information uncovered by the paradigm we used is compatible both with the presence and with the absence of sortal concepts and essentialist beliefs. We therefore conclude that, despite the considerable empirical evidence accumulated in recent years, Tinklepaugh's ...view that there is no "true evidence" of the "representative factors" underlying object individuation remains correct: many theories remain compatible with the evidence, and only further experimental and theoretical work can provide a fuller picture (p. 362).

While it is true that many theories remain *logically* compatible with essentialist beliefs, the first-order account is both necessary and sufficient to explain the results. Therefore, the higher-order descriptions are unnecessary. Furthermore, the *Asymmetric Dependency Problem* provides a principled explanation of why future studies of this kind will never be able to implicate the higher-order ones.

I close this section with the claim that all studies I have examined in the burgeoning literature on physical cognition—more importantly, all studies that fall within the identified genre—succumb under the scrutiny of the Seven-Step Program. I know this will be dissatisfying to many (perhaps most) experimentalists, especially those trained to think we can experiment our way out of such boxes with the right control conditions.⁵⁷ The best I can do is invite the skeptical reader

representational-level debate does not hinge upon the answer to this question, the specific goal-directed description of their intention in dropping the stones (i.e., *make the platform collapse*) may or may not be accurate (see Povinelli & Henley, 2020, Appendix). It is possible they are simply attempting to make contact with the worm in any manner possible, and the collapse of the platform initially occurs as an experimentally constrained consequence. Some support for this idea comes from a subsequent test with these same birds in which they were given a choice between sticks and stones. All birds chose sticks all of the time. To be clear, future tests might be able to differentiate between when pokes with sticks are intended to spear a worm versus when they are intended to collapse a platform (see von Bayern et al., 2009, Table 1). My point here is how easy it is for our folk psychological interpretations to be wide of the mark even for largely inconsequential claims about generally uncontroversial issues (i.e., that an animal's behavior was goal-directed/intentional in the specific manner we assume), and how our assumptions can, downstream, color our inferences about higher-order thinking. Indeed, from one perspective, *Folk Physics* is just one big case study of this *human* phenomenon.

⁵⁶ To be fair, both of these teams emphasize that the exact cognitive mechanisms underwriting the spitting behavior remain unclear. I would urge a gentle rephrasing of this conclusion: Because these studies fall under the umbrella of the *Asymmetric Dependency Problem*, no variants of them will ever implicate higher-order thinking (see the formal analysis in Povinelli & Henley, 2020, Appendix). As an aside, both Ghirlanda and Lind (2017) and Hennefield et al. (2018) demonstrate that in related tasks involving water displacement, regardless of whatever else is involved, incremental trial and error learning characterizes the results (see also the contributions in Barker & Povinelli, 2019b).

⁵⁷ In particular, some scholars may attempt to mitigate this conclusion as follows: *I knew right away those studies had local methodological flaw x, y, and z. But our studies have controls a, b, c...* Given such deeply canalized commitments to producing more and more experiments within the same genre, it may be as impossible to correct this misunderstanding as it would be to read one's way out of Borges' *Library of Babel* one book at a time.

to seriously engage with the method, applying it to whichever experimental investigations one finds most compelling. Ty Henley's masterful deconstruction of the recent study by Jelbert et al. (2019)—which claimed that New Caledonian crows have a higher-order concept of *<weight>*—may be a helpful tool in this regard (see Perspective Piece 3, this issue; see also, the formal analysis of the spitting-water-in-a-tube-task provided in Povinelli & Henley, 2020, Appendix).⁵⁸

In the meantime, I turn to another issue raised by *Folk Physics* that has so far appeared only informally in this essay: How seriously can we trust the folk psychologically-derived descriptions of even our own folk physics?

Claim 3: Human Folk Reasoning (about Human Folk Physics) is an Unstable Foundation for Tests of Chimpanzee Folk Physics

Folk physics stressed that while it is manifestly obvious that humans possess a higher-order folk physics, we may nonetheless be dramatically mistaken about *when* and *how often* our object-oriented behavior is modulated by higher-order thinking. We urged a reexamination of how higher-order thinking causally connects to everyday tool use in humans:

...we must again be careful not to mislead our reader. ...[W]e do not suppose that humans use high-level judgements related to the folk physics of transfer of force each time we use a stick to retrieve and out-of-reach objects. On the contrary, we supposed that on many occasions...folk physics are not involved. However, this should not obscure the fact that humans *can* invoke such concepts when circumstances demand it (Povinelli, 2000, p. 161-162).

This argument mirrored the claim we had made repeatedly in the context of social cognition: The human ability to represent *<mental states>* as such may not play as large a role in generating and/or attending our behavior as we think (see, e.g., Povinelli & Giambrone, 1999). Penn and Povinelli (2013) offered the most general version of the claim:

Notwithstanding the monumental impact our uniquely human system for reasoning about higher-order relations and analogical inferences has had on human cognition, we suspect that humans nevertheless overestimate the importance and cognitive efficacy of our symbolic-relational abilities. As Povinelli's original Reinterpretation hypothesis first suggested, the vast majority of humans' everyday social interactions do not engage our uniquely human ToM system. The role of explicit mentalistic theorizing in human affairs is more post-hoc than we folk would like to admit—and often misguided to boot. Indeed, our species' cognitive system for reasoning about higher-order symbolic relations does not merely subserve our unique linguistic, logical, causal reasoning and mentalistic abilities. It also subserves our inveterate predilection to reinterpret the behavior of heterospecifics in mentalistic terms...and many other uniquely human delusions (p. 77).

Other scholars have made similar points (see Bermudez, 2003).⁵⁹

⁵⁸ Again, the point may be as narrow as it is important: *Investigations of higher-order thinking that involve physical objects should be understood as being limited by the same logical roadblock (i.e., the Asymmetric Dependency Problem) as studies of theory of mind.*

⁵⁹ Two seemingly unrelated research programs can be considered in light of this problem. First, Francesco Silva and colleagues have shown that the performance of adult humans on the tasks we presented to our chimpanzees in *Folk Physics* is strongly influenced by local perceptual factors that are at odds with higher-order folk physical constructs (Silva & Silva, 2006; Silva et al., 2005, 2008, 2014). In many cases, this leads

All of this lies at the heart of our *reinterpretation hypothesis*: the claim that higher-order systems are (a) uniquely human and (b) were grafted into ancestral, first-order cognitive developmental pathways. On this hypothesis, humans did not replace our ancestral first-order relational reasoning abilities; rather, we reinterpret their outputs within a higher-order framework. Initially, this hypothesis was offered as a way of understanding why the claim that chimpanzees and other species engage in complex social behaviors is not inconsistent with the claim that theory of mind is uniquely human (for early defenses of this view, see Povinelli, 1996; Povinelli & Giambrone, 1999). Jennifer Vonk and I (2006) later expanded the reinterpretation hypothesis to several other domains, and Penn et al. (2008a) formalized the representational-level claims of the hypothesis, generalizing it to all domains of cognition.

The reinterpretation hypothesis thus offers not only an evolutionary framing of so-called dual systems approaches to understanding human social and physical cognition (in my view, better described as a *myriad* systems approach),⁶⁰ but also, because of the manner in which the two systems were purportedly interconnected in the evolutionary history of humans, it provides a jumping off point for exploring why humans may often be quite mistaken about the causal linkages between first- versus higher-order thinking.⁶¹ For example, we have speculated that the widespread idea that higher-order social cognition (i.e., theory of mind) evolved to cope with the dynamics of concurrent social interactions might be false (e.g., Povinelli et al., 2000). While such higher-order-thinking may or may not occur in such contexts, it is at least as plausible that those dynamics are frequently handled by first-order representations, occasionally modulated by higher-order representations built during offline reflection. Clearly, the same reasoning can be applied to tool use. Indeed, the work by Francesco Silva and colleagues on the dissociation between human folk physics and tool use can be interpreted as evidence to support this claim. Their work suggests that our commonsense descriptions of the processes underlying our tool use may not match up with the actual processes involved (see details in footnote 59).

In the past few years, my colleague, John Pruett, and I have worked with our research team to empirically explore this idea in the context of human social cognition. For example, in one study, we attempted to determine the contexts in which adults report making theory of mind attributions (Bryant et al., 2013). After some training on the procedure, we gave adults electronic devices to wear that interrupted their ongoing daily activity 30 times during a day. During these interruptions,

them to make the same "errors" as our chimpanzees. Second, Lucy Cheke and her colleagues (2012) found that human children must await their eighth birthday before they perform better than crows on certain tool-using tests involving water displacement. Do these findings imply that crows possess the capacity for relevant higher-order folk physical reasoning, whereas adult humans and children younger than eight do not? Consistent with the argument advanced here, an alternative interpretation of both sets of findings is that, in humans, higher-order notions of <force>, <mass>, <gravity>, <intrinsic connection>, etc. do not play the direct role in mediating our simple tool-using behaviors that we frequently assume them to play. To wit, the framework laid out by the *Asymmetric Dependency Problem* is consistent with two interpretations of these kinds of data: (a) young and adult humans may possess elements of a higher-order folk physics but simply do not (immediately) recruit such knowledge in these cases, and (b) solving such tasks does not require a higher-order folk physics. Note that this in no way implies that human higher-concepts never causally interact with our behavior.

⁶⁰Several recent, detailed theoretical analyses have reached nearly identical conclusions. For example, in a masterful deconstruction of temporal cognition, Hoerl and McCormack (2019) cut through a jungle of experiments with apes and birds and show how first-order "temporal updating" systems are necessary and sufficient to explain the animal results, leaving (by default) the higher-order, "temporal reasoning" system uniquely human.

⁶¹There are obvious connections here to philosophical debates about the role of higher-order thought in both theories of human intentional action and consciousness. Here, I simply note that these debates need not be settled in order to hold open the possibility that we may often be mistaken about the role that specific mental states play in modulating specific behaviors.

the participants were prompted to quickly categorize their current thought as being about an action, a mental state, or miscellaneous.⁶² They also recorded whether the thought was about themselves or others, and whether they were alone or with other people. Finally, they provided a short free-form description of the thought. The data suggested that these adults (1) spent more time thinking about actions than mental states, (2) exhibited more self- than other-directed thought when alone, and importantly, (3) made mental state attributions more frequently when they were not interacting with others than when they were doing so. And, as predicted by our hypothesis about the primary function of theory of mind attributions, action (but not mental state) thoughts about others occurred more frequently when participants were interacting with other people. (There was also an increase in the frequency of both action and mental state attributions about the self when participants were alone as opposed to socializing.)

In another study, we narrowed our focus to nonverbal episodes of joint attention (see Shaw et al., 2017). We recruited pairs of adult participants to cooperate to find specific kinds of images scattered about the walls of a room. Our question was whether we would detect a systematic co-occurrence between mental state thoughts and joint attention episodes. We monitored the participants' behavior on live video and interrupted them during episodes when joint attention (behaviorally defined) was either occurring or not occurring. The subjects instantaneously recorded their thoughts using procedures similar to those described above. The task was highly effective in eliciting many spontaneous episodes of joint attention. However, when we examined participants' reported thought contents, the results showed that *joint attention and thoughts about mental states did not systematically co-occur*.

While in no way definitive, these results offer at least face value support for the claim made in *Folk Physics* (and elsewhere) that the causal linkages between human higher-order thinking and ongoing dynamical actions, such as interacting with others or using tools, are not particularly transparent to our folk psychology. In the study just described, for example, even the cases in which joint attention and thoughts about mental states did co-occur, there is no evidence that the reported higher-order state *caused* the behavior in question. To raise just one possibility, the higher-order accounts could be rapid confabulations. Thus, the claim made in *Folk Physics*, and the challenge it poses, remains as foundationally important today as it did twenty years ago: *If we rely upon our folk psychology to imagine how our folk physics is causally linked to our interactions with tools (for example), and then (implicitly or explicitly) use this as the basis for the design of our experiments, we may be standing on a very unsteady foundation.*⁶³

All of this raises the inevitable question: *If humans are not particularly great at recovering the detailed connections between our higher-order thinking and our behavior, then why did higher-order thinking evolve in the first place—what is its function?*

As discussed in *Folk Physics* and elsewhere, the reinterpretation hypothesis offers three kinds of (positive) reasons for why higher-order thinking may have evolved, even if human folk psychological beliefs about the connection between such representations and our behavior is frequently mistaken. Crucially, none of these arguments require a flawless ability (in hindsight or foresight) to get the causal connections between our folk psychology/physics and our behavior correct:

The Good Enough is Better than Nothing Function. It may well be that the original evolutionary function of higher-order thinking was not to endow humans with whole new

⁶² Before the study began, participants were instructed on how to categorize their thoughts according to strict definitions and examples.

⁶³ I have made this (seemingly) uncontroversial claim to many experimentalists. The typical reaction? *I agree, and that's really bad when that happens. But we don't base our experiments on our folk psychology.* This common retort is contradicted by the informal nature of the hypotheses offered in most research reports. The work of Francesco Silva and colleagues (e.g., 2014) described in footnote 59, bears directly on this problem.

"forms" of behavior, but to deliver additional traction on existing ones—that is, more skillfully deploying ones that were in full operation (via first-order reasoning) long before our species appeared on the scene. To pick a random example, even minor increases in successful acts of deception might yield important fitness advantages—and the same can be said of tool use (see, e.g., Povinelli, 2000, p. 68). Indeed, given that the most general treatment of the reinterpretation hypothesis applies to all domain of cognition (Penn et al., 2008), these incremental fitness advantages can be seen as playing out across all of the daily activities of early humans, and culturally evolving across time.⁶⁴

The Predictive Function. Even if the utility of higher-order thinking is less important in our moment-to-moment predictions about the behavior of people and objects than we think, this does not entail that we never use such descriptions for predictive functions. Off-line, effortful, explicit considerations of the <beliefs>, <desires> or <emotions> of ourselves and others might allow humans to make better *Good Enough* predictions, ones superior to an individual without a theory of mind. The same can be said about higher-order thinking related to folk physics. Off-line effortful higher-order thinking may uncover causal relationships that might take much longer to discover without it. Eventually, through cultural ratchet effects, these forms of tool-use could become so dependent upon cultural transmitted forms of higher-order thinking that no system restricted to first-order thinking could ever achieve them (the network of gravity wave detectors built to confirm Einstein's theory of general relativity comes to mind). This idea seems uncontroversial and consistent with a variety of (competing) ideas about *human* social and technological cultural evolution (cf. Henrich, 2015; Heyes, 2018; Tomasello, 2014). Furthermore, to the extent that claims of higher-order thinking are judiciously excised, this idea is broadly compatible with current discussions about animal culture, as well (cf. Henrich & Tennie, 2018; Schuppli & van Schaik, 2019; Whiten, 2019; see especially Osiurak & Heinke, 2018).

The Explanatory Function. Much higher-order thinking undoubtedly serves to assist humans in building narratives (stories) about why the world works the way it does (cf. Carr, 2008; Velleman, 2003). These explanations are traded in natural language, and are used to build alliances and create cultural traditions—all of which generate explicit (if fuzzy) reasons to justify what we do (Povinelli & Prince, 1998, pp. 90-92). Language not only transmits information, it is inherently connected to calling forth and carrying out action (Austin, 1962). The imperative communicative gestures of chimpanzees likewise call forth action, but because human language is also inherently higher-order (and thus riddled with folk psychological and folk physical ideas steeped in metaphor and analogy), higher-order relational thinking is available to augment ancestral imperative and blind informational functions of communication by generating explanatory narratives. These narratives not only build ever-richer descriptions of people and objects, but also create shared reasons for behaving in particular ways. And because they are explicit, these reasons can be rapidly manipulated, and can thus provide input to higher-order forms of planning explicitly connected to our higher-order conceptions of <past>, <present> and <future> (see Hoerl & McCormack, 2019, for a robust deconstruction of the evidence that animals explicitly reason about time).⁶⁵

⁶⁴ To be clear, these folk psychological and folk physical descriptions could be *false* and still confer fitness advantages, a point stressed in the introductions of both *Folk Physics* and *Weight* alike.

⁶⁵ Communication systems that transmit information from sender to receiver are widespread in the animal kingdom. On the standard account of the famous waggle-dance of bees, for example, the behavior of forager bees communicates the distance, direction and quality of a food source (Wenner et al., 1967). This in no way nakedly implies that bees possess the higher concepts of <distance>, <direction> or <quality>, let alone

I do not offer these three possible functions of higher-order thinking to make any bold claims about their phylogenetic origins. I simply present them as reasons for believing that there is no contradiction among the claims embedded in this claim: The capacity for folk psychological and physical thinking may be (a) uniquely human, (b) evolutionarily useful, and (c) not a particularly great guide to the (actual) causal interactions between higher-order thinking and our behaviors.

In reflecting on these issues at the end of *Folk Physics*, I offered some speculations about how the evolutionary emergence of the narrative function of higher-order thinking may be tangled up with the human capacity for generating explanations in the first place. This, of course, raised the immediate question of whether chimpanzees (or other animals) do the same. What particularly interested us was the question of whether they ask, *Why?*

Claim 4: Do Chimpanzees Ask, *Why?*

Candy and her coconut offer an excellent case study to consider this question. When she first approaches her coconut and sniffs it, is her exploratory activity ushered on by *reason-seeking*? If the conclusions of *Folk Physics* were correct—that Candy does not reason about higher-order phenomena—then, by definition, her exploratory behaviors would not be generated by explicit reasons (although, to be fair, her behaviors could still be said to be *directed by reasons* that are not explicitly represented as such).⁶⁶ To put it colloquially, chimpanzees and other animals could possess powerful *What?-systems* (systems designed by evolution to recover the information that builds first-order causal representations), but no *Why?-systems* (systems designs to build explicit reasons related to causal factors). The poignant example at the end of *Folk Physics* remains thought-provoking. If Candy were to observe a chicken crossing the road, and we could somehow ask her why it did so, without any higher-order thinking she would simply reply, *Yes*. Candy might even be better than us at recalling details of the crossing event. Nonetheless, she would never appeal to *<wants>* or *<desires>* or *<emotions>*.

At the time, we had already mapped out (and partially undertaken) a lengthy program of experiments to try to ask Candy and her companions this question directly. Again, using our own folk psychology as a guide, we designed numerous experiments we hoped could reveal if they possessed the ability to report more than 'yes' to the chicken-crossing-the-road question. We arranged a number of scenarios in which we believed a behavioral dependent measure could tell us whether they were engaging in either retrospective or prospective causal diagnostic reasoning. In short, whether they were asking *why* things happened, or would happen in the future.

Although only two of these studies were ever published (Povinelli & Dunphy-Lelii, 2001; Povinelli & Frey, 2016), they remain particularly interesting to me. Part of the reason why is that although there are numerous explanations for why any particular behavior exhibited by the chimpanzees on these tests should or should not be taken as strong evidence of asking *why* (i.e., reason-seeking), I am not fully convinced these tests are part of the genre of experiments ensnared

<communication>. The particular claims in this paragraph, in contrast, hinge on the explicitness of the information contained in human languages (see also footnote 44).

⁶⁶ Almost ten years ago now, my colleague (and now dear friend), Caroline Arruda, engaged me in a discussion as to whether chimpanzees are intentional agents. My naïve reaction was, *Of course chimpanzees are agents!* While sympathetic, Caroline patiently helped me see that because of the role that *having reasons* plays in many conceptual analyses of agency (among other debates), settling the question of chimpanzee agency was not trivial. Our work together has generated what we both now believe is a solid foundation for the claim that regardless of whether chimpanzees possess the higher-order conceptual resources to endorse their reasons as their own, they do stand in a *directed relationship* to reasons. This directed relationship highlights that chimpanzees possess belief- and desire-like states that mediate their goal-directed actions, regardless of whether they represent those states as such (for more details, see Arruda & Povinelli, 2016, 2018). Thus, they can be said to be agents (secret or otherwise).

by the *Asymmetric Dependency Problem*. Thus, I was heartened to read the essay by Cristoph Völter and colleagues (2020) in the August 2020 special issue in which they outline new strategies for tackling this important, but difficult question. After the work reported by Povinelli and Dunphy-Lelii (2001), our laboratory group worked hard to develop tighter contrasts to help us distinguish between the operation of the purported *What?*- and *Why?*-systems, but in the end I was never convinced we succeeded. Thus, as our unpublished studies in this area collect dust in our archives, Völter et al.'s (2020) ideas stand as a possible roadmap forward.

Claim 5: Chimpanzees Raised in Captivity Can Shed Light on *Chimpanzee* Cognition

Two persistent concerns among scholars interested in higher-order reasoning in animals are, first, whether results obtained in captivity can be generalized to free-ranging populations, and second, the role these respective data ought to play in generating explanations of their behavior (for a recent proposal, see Andrews, 2020). As a physical anthropologist and zoologist who has conducted my share of fieldwork studying monkeys and apes in the forests of Central and South America and Indonesia, I have grappled with these issues for more than forty years. Predictably, both of these concerns were raised by reviewers of *Folk Physics* (e.g., Allen, 2002; Kahn Jr., 2003; Hauser, 2001; Whiten 2001). Boesch's (2020) essay in the August 2020 special issue drives home the point that this contentious issue remains alive and well.⁶⁷

I admit at the outset that at times these questions feel quite personal. How could they not? After all, much of what has been written about our chimpanzees is so vitriolic and dark—not to mention unconnected to reality—that it is often hard to know where to begin (see Wise, 2000). But, after taking a deep breath, I realized that addressing these arguments directly might help to expose how a confusion between the *representational* and *functional-level* claims made in *Seeing, Folk Physics* and *Weight* colored their reception. In turn, this might help isolate the core reason why such dismissive arguments were primarily (though not exclusively) directed against our captive chimpanzees, not captive chimpanzees in general: because our *interpretations* of the data differed from other scholars, not because the data itself differed (see above, "Double Fooled"). Finally, it might illuminate how these confusions continue to haunt both experimental comparative psychologists and field researchers alike.⁶⁸

Are captive chimpanzees abnormal and therefore unable to speak for wild chimpanzees?

At the outset: there can be no debate over whether the experiences of captive and free-ranging chimpanzees are wildly different, the latter inhabiting incomparably larger ranges and complexities of environments. Nor, in my mind, is there any debate that all chimpanzees should be provided with the most ethologically appropriate environments possible. Of course, the word *possible* immediately raises the complex intersection of devastating habitat loss, poaching, the ethics of captivity, and defining and balancing moral rights.⁶⁹ For what it's worth, I believe making progress on these questions is far more important for chimpanzees than determining the precise nature of their cognitive skills. Nonetheless, the ethical questions are distinct from questions of whether Candy and her companions grew up in a social and physical environment that drew out the kinds of cognitive abilities found in chimpanzees in the wild. While the answers to the latter may inform answers to the former, the reverse is unlikely to be the case.

⁶⁷ Most of the conceptual issues I have raised in this essay come together here. To be absolutely clear, I believe the debates over using data from the field versus captivity most closely track the *interpretation of results*, not the results themselves (or, for that matter, the actual lives of any particular group of animals).

⁶⁸ It also provides an opportunity to explore one of the most interesting questions in the debates over animal intelligence: the *reasons* that folks (including myself) give for believing what they do (see Barker & Povinelli, 2019b).

⁶⁹ For part of my perspective on these matters, see Povinelli and Preuss (2012).

In his contribution to the special issue dedicated to *Folk Physics*, Boesch (2020) takes direct aim at these issues. Based on a confusing intersection of claims, he paints a fanciful picture of cognitive deficits of Candy and her companions: Megan, Apollo, Jadine, Brandy, Kara, Mindy, and their offspring, Lance, Keagan, Brayden. Some of the brushstrokes of his painting include (see Boesch, 2020, Table 1):

- Small declines in time spent swimming through Morris water mazes in fragile strains of Wistar rats separated from their mother for 12-14 hours;
- The impairments exhibited by the severely maltreated monkeys of Harry Harlow;
- Higher rates of agnostic behavior in the play behavior of young orphaned chimpanzees;
- Retrospective MRI data on possible differences in brain development in peer-versus mother-reared captive chimpanzees (with unspecified life histories) *not linked to any functional-level differences in behavior or cognition and not compared to wild-living animals* (cf. below, "Candy's Family").

Given that there is no way of directly combating the unsubstantiated application of these data to our animals, I will leave it to the reader to decide their relevance to the lives of our chimpanzees (see below "Group Megan: Social Complexity and Spontaneous Tool Use"). Instead, through the lens of my decades of experience as a *physical-anthropologist/zoologist-using-the-methodological-tools-of-the-psychologist*, I will approach the question of the external validity of our investigations of the family of precious chimpanzees that we were privileged to study for twenty years by examining their real lives. Before that, however, I begin with a realistic look at the stresses and the traumas (or, if one wishes, the <suﬀering>) of wild chimpanzees.

Romanticism: Ignoring the cries of wild chimpanzees. First, for those who have not conducted fieldwork with nonhuman primates (as both Boesch and I have), romanticizing their lives is tempting. There is something inescapably powerful about seeing animals in the ecologies in which they evolved, deploying the functional abilities they have evolved. But building an argument that well-treated and well cared for animals in captivity are abnormal in the sense that they cannot shed important light on the cognitive abilities their species, requires more than a belief in evolution. It requires a sober look at the stresses and traumas that wild chimpanzees experience on a daily basis. Indeed, the heretofore unacknowledged suffering of wild animals has led to growing calls to intervene into nature and alleviate it.⁷⁰

To begin, there is a large, broad and growing literature documenting the pervasive stress experienced by populations of wild animals, in taxa ranging from fish, to birds to mammals (including primates). Stressors can begin *in utero* and continue throughout development, triggered by nutritional stress, predation avoidance, illness, social interactions, and infection (Almasi et al., 2012; Benowitz-Fredericks et al., 2008; Blas et al., 2007; Boonstra, 2005, 2013; Boonstra et al., 1998; Breuner & Hahn, 2003; Chapman, 2007; Clinchy et al., 2004; Creel et al., 2013; du Dot et al., 2009; Giesing et al., 2011; Hawlena & Schmitz, 2010; Hayward & Wingfield, 2004; Landys et al., 2011; Love & Williams, 2008; Love et al., 2013; Sheriff et al., 2012). In free-ranging chimpanzees, Goodall (1986) has linked the death of several chimpanzees at Gombe to nutritional stress during the dry season, and van de Rut-Plooij & Plooij (1988) have linked infant illness to social stress. The long-term stress effects on mothers and juvenile chimpanzees due to infanticide (and cannibalism) among wild chimpanzees is unknown but concerning (see Arcadi & Wrangham,

⁷⁰ For a discussion about the ethics of wild animal suffering and the moral imperative to do something about it, see Cowen (2003), Everett (2001), Faria and Paez (2015), Fink (2005), Johannsen (2020, in press), McMahan (2015), Moen (2016), Ng (1995) and Sapontzis (1987).

1999; Watts & Mitani, 2000; Wilson et al., 2004). It is also widely known that free-ranging populations of chimpanzees suffer from chronic and cyclical parasitic and viral infections (Bakuza & Nkwengulila, 2009; Huffman et al., 1997; Wallis & Lee, 1999; Woodford et al., 2002), including chimpanzees at Kibale National Park in Uganda, Tai National Park in Cote d'Ivoire, and Gombe National Park, Tanzania. These infections result in chronic, well described, and painful lesions, and have been linked with morbidity and mortality (e.g., pathologic lesions associated with *Oesophagostomum* sp., Terio et al. 2018; see Krief et al., 2010; Terio et al., 2011). In one study examining the known causes in over 130 chimpanzee deaths, 58% were from illness, including respiratory ailments, polio, mange, and wasting (Williams et al., 2008). Some of these diseases (including polio) have been transmitted to chimpanzees (and other wild animal populations) by field researchers themselves (Dunay et al., 2018; Kaur et al., 2008; Köndgen et al., 2008). Other significant fractions of deaths may be due to interspecific aggression, especially in individuals under 20 years of age (Williams et al., 2008). Other stressors exist as well. Skeletal collections reveal that upwards of one-third of chimpanzees and other great apes suffer long bone fractures (likely from falls), in some cases presumably debilitating ones, as well as skull puncture wounds (Jurmain, 1997). Thus, stressors among wild chimpanzees are early, often in some cases chronic, and their effects long lasting. Many are likely still poorly characterized.

Despite all of this, in comparing wild chimpanzees to captive chimpanzees, Boesch (2020) indulges his reader to imagine that wild chimpanzees live in an idyllic natural state designed by evolution to maximize their mental and emotional wellbeing. He offers no mention of the *chronic stress and trauma they experience*. I could go on, regaling the readers with the *folk psychologically horrifying* things I have witnessed in free-ranging primates in Southeast Asia and Central and South America, including particularly disturbing images of forced copulations between sub-adult male and adult female orangutans. The main point is this: evolution is a fitness-maximizing equation wherein the physical and emotional wellbeing of animals matters only to the extent it aids them in the game of foraging and surviving to reproduction.

Thus, there is every reason to believe that relieved from the stresses of nature, raised and cared for in challenging social and physical environments, provided with stable nutrition, dental and general health care, our chimpanzees, at least, may have been less stressed, more free to elaborate upon their natural cognitive skills in ways they might not otherwise be possible. To those who find such an idea implausible, consider the following. By my estimates, each of our chimpanzees engaged in over 15,000 testing sessions, each session comprised of 10-15 of what we conveniently call trials.⁷¹ These testing sessions confronted our chimpanzees not just with the same classes of functional problems chimpanzee encounter in the wild, but in many cases considerably more challenging variants of them.⁷² And lest we forget, all of this testing was stacked upon the rich experiences of their everyday lives together (see below, "Candy's Family").

I end this section by inviting the reader to turn to Sarah Dunphy-Lelii's poignant narrative about the stresses experienced by wild chimpanzees (see Perspective Pieces 4, this issue). After reading it, I invite readers to ask themselves: If a wild chimpanzee's newborn infant was about to

⁷¹ See Povinelli and Henley (2020) for a more detailed discussion of the problematic notion of a *trial* in comparative psychology.

⁷² It is worth noting that some folk-psychologically assumed effects of environment on tool-using abilities disappear under close scrutiny. For example, Boesch (e.g., 1991; Boesch & Boesch-Achermann, 2000) argued that active teaching by mothers played a strong role in prompting acquisition of nut cracking among infant chimpanzees. Detailed recent analyses by his team reveal these effects are all but nonexistent (Estienne et al., 2019; see also footnote 82). A similar picture may be emerging for behaviors once thought to be 'culturally' transmitted (see Bandi et al., 2020; Bernstein-Kurtycz et al., 2020; Fiore et al., 2020). This does not mean maternal inputs are unimportant to development, just that our folk psychology is probably not a good guide to uncovering the causal pathways involved. (Although somewhat tangential, Boesch and Boesch-Achermann's (2000) apparent observations of a secret symbolic code in chimpanzee drumming behavior quickly disappeared, as well.)

fall victim to infanticide, but instead were somehow rescued by morally-minded human bystanders (see footnote 70), and then raised in a materially-enriched environment with other chimpanzees and humans (comparable to what Candy and her peers experienced), would we expect that infant's life-time stress and trauma to be greater or less than if she had survived in wild? Neither the answers, nor the implication of the answers, are obvious. But they do help to raise pertinent questions.

Candy's family: Social complexity and spontaneous tool use. Against Boesch's strangely Disney-like portrayal of wild chimpanzees (having spent his life studying them, he knows better), let me now sketch a realistic portrait of the lives of Candy and her companions (see Figure 2; see also Povinelli, 2012, p. 62-65). They were raised together from birth, except for Megan and Apollo who joined the peer group when they were about a year old. Initially, they were primarily cared for by a loving caretaker. They were bottle-fed, held and rocked for hours a day. At just a few weeks of age, they were already interacting with each other. Over the first year of their lives, they gradually transitioned to another (equally patient and loving) caretaker who remained their primary human attachment (and trainer) for the next 18 years. They lived in a spacious compound that included multiple interconnected indoor-outdoor environments where they could climb, swing, explore, break up in to small groups, hide from each other, play, and forage.⁷³

Figure 2

Candy and Her Companions, Growing Up Ape, in a Human World



⁷³ The schematic drawing of their living quarters that Boesch (2020) reproduced from *Folk Physics* was just that, a schematic. It did not include the material enrichment of the environment, nor their indoor living quarters, nor were the animals drawn precisely to scale.

From the beginning, their environment was filled with objects that they used in innumerable ways, including as spontaneous tools (see below). These included toys, blankets, buckets, balls, burlap sacks, clothing, novel foodstuff barrels, balls, hay, tires, fire hoses, rope swings. They were given a wide variety of foodstuffs including seeds, nuts and fruits and vegetables, the latter including kiwis, pumpkins, coconuts, watermelons, broccoli, celery, onions, string beans, broccoli, potatoes, peanuts, pecans, fresh and frozen juices—with additional periodic opportunities of extractive foraging involving elaborate puzzle feeders. Their trainers and caretakers (especially our student volunteers) constantly challenged Candy and her companions with new objects and experiences. Their compound was frequently filled with hay, creating opportunities for foraging for hidden food, as well as additional opportunities for deception. The coming and goings of birds, cats, raccoons, monkeys in the distance, overhead airplanes, automobiles etc., all elicited their attention and vigilance, triggering, alarm calls, food calls, and reunion vocalizations.

I could go on, but given that (a) the spatial scale of these activities does not begin to approximate that experienced by wild chimpanzees, and (b) the social and physical interactions experienced by our chimpanzees differ in important ways from those of wild chimpanzees (although they involve extraordinary complexities of its own), we can proceed directly to the central question: what are the abstract set of experiences necessary to draw out the range of mental abilities of chimpanzees found in wild chimpanzees? And more to the point, how would we know?

I can think of two data sets that could bear on this question.

The first (and most obvious) data set is the naturalistic observations of the everyday behavior of our family of chimpanzees (after all, naturalistic observations of wild chimpanzees are the relevant contrast case for this discussion). Here we can say a lot. To the naked eye, Candy, Megan, and the rest of the gang, *developed the full range and complexity of chimpanzee social behavior* (briefly: deception, gaze-following, dominance-submission displays, ally recruitment, reconciliation after fights, sexual behavior, play signals and behavior, begging gestures, boundary patrols, alarm calls, food calls, reunion calls, food discovery calls, an enormous range of spontaneous tool invention and use, embracing after separations, social grooming, play behavior, comfort-giving, aggression, simple bedding behaviors, communicative gestures and, as they grew older and started having babies of their own, maternal and allo-parenting behavior). From time to time we even witnessed episodes of spontaneous (and arguably "cooperative")⁷⁴ hunting, usually involving small birds, lizards, and rodents. They even negotiated how to share the spoils afterwards. The list goes on and on.

One item deserves special attention in the context of *Folk Physics*. As we saw with Candy and her coconut, our chimpanzees spontaneously (*read*: outside of testing) exhibited innumerable instances and types of tool use. From 1991 to 2009, our laboratory kept a series of hardbound notebooks for our staff to record interesting observations that were not part of our testing sessions (which constituted only a small fraction of the day for any individual chimpanzee). Every member of our team was introduced to these notebooks and encouraged to write in them regularly. Collectively, they became known as the *Log Book of Interesting Occurrences*. As part of my work on this essay, I quickly combed through the first ten years of entries, noting the ones that dealt with the spontaneous use of tools. I then sorted each entry into the functional categories depicted in Table 1.⁷⁵

⁷⁴ I am not implying that our chimpanzees conceptualized the higher-order understanding of the roles they played as they trapped and ate these small animals. Indeed, despite the great fanfare about the uniquely challenging (and higher-order) basis of group hunting in chimpanzees (see Boesch, 2005; Boesch & Boesch-Achermann, 2000), it is worth noting that an analysis by Gilby and Connor (2010) concluded that, compared to other social predators, there is nothing special about chimpanzee cooperative hunting.

⁷⁵ All by itself, the category of play-start (see Goodall, 1986)—using an object as tool to engage a companion in play—yielded 43 instances, involving 14 discrete kinds of objects (balls, PVC tube, burlap sack, cardboard

Table 1

Categories of Spontaneous Tool-Using (Illustrated with One Example Each) Observed During Informal Observations of Candy and Her Companions

Functional Category ^a	Example ^b
hooking	Brandy hooks an angled PVC pipe under a perch and swings
hanging	Megan threads burlap sack over a perch, grabs both ends and hangs
raking	Megan uses a PVC stick to drag in potato
squashing	Candy uses a coconut to crush a mango
hammering	Brandy uses a wooden box to hammer open a pecan
throwing (display)	Apollo throws balls while displaying
throwing (aimed)	Kara threatens student caretaker, then throws Kong toy at her
probing (body)	Megan picks into Jadine's nose with a piece of hay and extracts mucus
probing (insects)	Candy and Apollo use straw to poke at termites in rotting wood planter
probing (dead animals)	Jadine probes a dead bird with a stick
probing (with modification)	Kara chews a plastic straw repeatedly until it fits into puzzle feeder
transport	Brandy collects toys inside a plastic bin, carries them indoors, then dumps them out and plays with them
pushing	Mindy use a rubber wash tub to push balls indoors
containing	Brandy fills plastic hard hat at spigot and adds monkey chow and onions
cleaning	Kara wipes Jadine's bottom with towel then sniffs it
play-start	<i>See footnote 75</i>
fishing	Kara inserts cardboard strip toward out of reach ants and eats them
missile	Jadine flings ball at snake
levering	Apollo uses screwdriver to pry apart metal covered on food bin
clubbing	Kara hits Brandy with a hook tool
dragging (display)	Jadine drags burlap sack and metal drum while displaying to passing strangers
spinning	Megan turns hard hat upside down, sits in it, and spins in circles
dipping	Candy repeatedly dips toy into cup and sucks off clinging juice
bridging	Kara uses her foot to slide a plastic barrel into the empty space between two perches and uses it to cross between them
sliding	Mindy, Megan, Apollo use a plastic lid to slide across a wet floor
dropping	Apollo drops beam on Brandy while she sleeps, then jumps on her
self-tickle	Brandy embraces plastic block and wrestles with it until she laughs
wipe	Megan steps in feces and limps to cloth sack and wipes her foot on it
hiding	Mindy repeatedly crawls inside burlap sack and encloses herself
mopping	Kara, Mindy, and Apollo use cotton sheets to wipe water off floor
stirring	Kara fills bowl at spigot, adds monkey chow, then inserts toy and stirs
sponging	Kara chews up paper towels and uses the wad to soak up juice, then squeezes into her mouth in bottom of a barrel

Note. I derived this list by browsing through ten (of eighteen) years of anecdotal entries in a series of informal log books (collectively, *The Log Book of Interesting Occurrences*) in which our research group haphazardly recorded things we found personally intriguing about the natural daily behavior of Candy and her family (see the main text for more detail). I stress that these records are in no way systematic and were not guided by any instructions. ^aThese categories are not intended to be exclusive or exhaustive and could be expanded or condensed to fit existing categories of tool-use in wild chimpanzees. ^bMost categories contained three or more examples. Excluding play-start (see main text), the categories with the highest instances were transport (14), probing (12), containing (9), probing (9), missile (7), hiding (5).

box, unwanted food shell, plastic carts, toy, hard hats, cups, bucket, rope, brick, dead bird, block of Styrofoam, and not specified). Furthermore, I did not individuate, for example, balls of different sizes and colors, so this should be considered a conservative estimate. Thus, in this one category alone (amassed from a haphazard set of observations spanning only half of the time we were privileged to study these precious animals), I documented great diversity in tool types. The same can be said for most of the other categories in Table 1. However, I quickly add the caveat—cribbed directly from Boesch and Boesch-Achermann's (2000) ethnography of the Tai chimpanzees—that the data reported above and in Table 1 "is only a small fraction of the intelligent behaviors we observed, and contains none of those we saw but did not understand" (p. 229).

Thus, at a purely functional-level of description, our chimpanzees spontaneously developed tool use for a wide range of purposes. Indeed, a number of our tests were explicitly designed to probe their understanding of tool using behaviors we saw them carrying out spontaneously.⁷⁶ Were these identical to those developed by wild chimpanzees? Certainly not, but every time we offered them the opportunities to spontaneously discover functionally equivalent forms of tool use as their free-ranging counterparts, they did so (for example, enrichment devices requiring them to use stick-like probes to dip for honey and other foodstuffs).⁷⁷

None of what I have just said is intended to prove that Candy, Megan and their companions and offspring developed precisely the same cognitive skills as wild chimpanzees. But I do believe it supports the following claim: *If we knew nothing about wild chimpanzees, and the only information we had about chimpanzees was from detailed observations of the daily behavior of Candy and her peers, then whatever conclusions one is tempted to make from the (actual) range and complexity of wild chimpanzees, one could have made from observations of our chimpanzees* (see below, "Do data from the wild demonstrate?").⁷⁸ To use an animal idiom, sauce for the goose is good for the gander.

Gedankenexperiments: Testing wild chimpanzees. The second data set that could bear on the question of whether the kinds of experiences received by Candy and her peers is sufficient to draw out the relevant range of chimpanzee cognition is, unfortunately, imaginary. Nonetheless, it is worth considering. Imagine a data set consisting of the results of our tests administered to wild chimpanzees. Two variants of this thought experiment directly speak to the underlying issues driving Boesch's (2020) polemics:

Gedankenexperiment No. 1: Present the tasks of *Folk Physics* and *Weight* to young chimpanzees growing up in the wild. Would they display the same pattern of results we obtained? Even setting aside the complex experimental cleaving of the variables our chimpanzees encountered, as we have seen, data from the wild suggests a long slow process of skill acquisition even for the tool-using abilities they do develop.

Gedankenexperiment No. 2: Present our tasks to wild *adult* chimpanzees who have already acquired the full range their population's tool-using skills. Which tasks would prove trivial, which more difficult? Would their skill acquisition and pattern of transfer look similar to our chimpanzees? If *Folk Physics* and *Weight* taught me anything, it is that my folk intuitions based on the final form of our chimpanzees' skilled behaviors was a very bad guide to performance on conceptually-related transfer tests (see "Featured anecdote: Scientists are human, too," Povinelli, 2012, p. 298-301). Would the same hold true for assumptions about wild chimpanzees?

⁷⁶ An example may suffice. After seeing Jadine poke a dead bird with a stick, we experimentally showed that our chimpanzees spontaneously used a stick when retrieving a banana *that was within their reach but placed alongside an alarming object* (see Povinelli et al., 2010). They also used the stick to poke at the alarming objects. Chimpanzees in the wild engage in similar behaviors, although it has been glossed as spear use (see Pruetz & Bertolani, 2007). Another example we have widely discussed involved their use of tools to obstruct their own vision (e.g., placing an opaque bucket over their heads). These behaviors inspired the design of some of the tests we used to determine if they understood <seeing> (see above, "Folk Psychology meets *Folk Physics*").

⁷⁷ Martin-Ordas et al. (2012) report that captive chimpanzees, bonobos, and orangutans are able to use up to five stick tools in sequence in a problem-solving test.

⁷⁸ Another regret about my myopia. Because I always took this fact to be self-evident, I failed to use the observational tools I had honed as both an anthropologist and psychologist to systematically collect and report these observations alongside our experimental data.

Some scholars prefer to address this question flipped on its head, arguing that there is a *prima facie* contradiction between the everyday behavior of free-ranging chimpanzees and the findings from our experiments. Although I have already dispensed with this objection above (our chimpanzees exhibited the very behaviors that prompt the high-level attributions to wild chimpanzees), this idea is so persistent, so folk psychologically tempting, that it is worth taking a deeper dive into some specific case studies.

Each of the case studies drives home the challenge not just of using naturalistic data to draw inferences about higher-order thinking, but, as we have seen, the comparable use of experimental data, as well.⁷⁹

Do Data from the Wild Demonstrate that Data from Captivity are Invalid?

The question of <intrinsic connection>. Many who have not thought deeply about the contrast between the functional acquisition of skills, versus representational-level questions about how those skills are acquired, found the results of *Folk Physics* shocking. Boesch's (2020) musing about our studies of <*intrinsic connection*> is an excellent example. He writes:

Take the notion of connectivity or contact studied by Povinelli and his team. In a famous series of experiments, he placed chimpanzee subjects in front of a food they could acquire only by pulling at a handle or rope either placed in proximity to the food, or connected with it. His peer-group chimpanzees performed often at chance-level selecting equally the connected or not-connected food. Imagine these same chimpanzees in a tree 40 meters above the ground! What would happen if they jumped on a dead branch not connected to the tree trunk or on a far too thin branch given their weight? How could such a chimpanzee try to capture a monkey that runs and jumps full-speed between trees to escape? (p. 484)

Here, Boesch displays a deep misunderstanding of the representational-level questions at stake. Indeed, this set of experiments was designed *with these seemingly contradictory issues in mind*.⁸⁰ We discussed the problem at length in *Folk Physics*:

⁷⁹ The challenges of the *Asymmetric Dependency Problem* apply equally to both naturalistic and current experimental data sets, even though this is rarely acknowledged. Indeed, the *Experimental Necessity Dilemma* illustrates how controversies over these two sources of data intersect to create circular reasoning on the part of experimentalists when they assert that lab tests are necessary to disambiguate the representational basis of observations from the field. Working through examples can train one to see why this is a *general* worry, not an *ad hoc* concern over particular sets of results (see above, "A Seven-Step Program to Recovery"). Povinelli and Vonk (2012), for example, offer a detailed case study of such circular reasoning in the context of studies by Manrique et al. (2010) attempting to argue higher-order representations of <*rigidity*> and <*floppiness*> in great apes. To be fair, similar circular reasoning can probably be found in the background assumptions of much of our own work (e.g., Povinelli and Frey (2016), although, in this latter case, the documented *failures of transfer* may ultimately be more informative (see above, "Return of the *Asymmetric Dependency Problem*").

⁸⁰ My background in physical anthropology and zoology led me to use the well-understood behavior of free-ranging chimpanzees to design the tests we used in *Seeing, Folk Physics* and *Weight* around. Each design was inspired by behavioral observations of wild chimpanzees. This served as a gut check on the question of the external validity of the tasks. Even tasks that no wild living chimpanzee would ever encounter were grounded in what chimpanzees actually do in the wild. This illustrates the error of Boesch's (2020) complaint that, in our systemic review of the best evidence in the entire field of comparative psychology (i.e., Penn et al., 2008a), we do not cite a single publication on wild chimpanzees. The reason we did not do so is because while the anthropologist in me knows that naturalistic data can provide indispensable information for functional-level theories about the proximate and ultimate utility of these behaviors, it cannot address the representational-level questions. For the record, we did analyze *experimental research* on wild baboons.

The data provide strong support for at least some aspects of the Köhlerian view that chimpanzees do not have a notion of connection deeper than mere contact. If this is so . . . how does the ape come to explain why in some cases their actions on an intermediary object (a tool) yield co-varied movement in a goal object, but in other cases they do not? Perhaps this question can be posed most succinctly by considering the apes' actions on objects in their world. When an ape grasps a hammer stone sitting upon a pile of rocks, surely the ape does not expect the other rocks to rise along with the hammer stone... On the surface, this would seem to raise trouble for the Köhlerian position. ...However, such events are only problematic if one assumes that chimpanzees seek coherent explanations among separate events in the first place. If, on the other hand, chimpanzees act upon objects in the world, detect specific regularities, and use them as default assumption about how the world is likely to behave, then the kinds of effects we have reported in this chapter can rest quite comfortably alongside the sorts of actions chimpanzees perform all the time. If the ape receives considerable experience (both through its spontaneous play and our experimental settings) that the post remains attached to the platform, but possesses not underlying explanation or account of why this is the case, then merely seeing the post set upon the platform may not initially offer any good reason (from their perspective) for believing that the resulting perceptual form has dramatically different affordances than the similar forms with which the ape is familiar (Povinelli, 2000, p. 252-253).

We went on to explain that animals in general (including our chimpanzees) have innumerable experiences in which roughly similar objects, when acted upon, sometimes result in the movement of other (in reality, physically connected) objects, whereas sometimes they do not:

Armed with a theory of causal mechanisms, the human easily explains the difference within a coherent framework. Armed only with the perceptual evidence, however, the ape may merely assume thatin situations in which a goal object is out of reach, but. ...some intermediary object...is contacting the goal, the intermediate object can be acted upon to move [it]. ...Thus, in summary we envision chimpanzees as possessing excellent perceptual discrimination abilities and thus able to make roughly the same 'contact' versus 'no contact' judgments as humans. ...with experience, their judgments about contact (or imminent contact) are used to generate robust expectations about the contingencies between their actions on an intermediary object and the movements of a goal object ...However, our results show that such judgmentsneed not be attended by parallel interpretations of underlying physical connection (Povinelli, 2000, p. 253).

Thus, Boesch (2020) falls prey to the *Conflation of Explanatory Levels* problem:

Functional level: Chimpanzees learned to build reliable expectations about the co-varied movement (or lack thereof) among objects in their everyday lives that support skilled behaviors.

Representational-level debate question under examination: Are first-order, perceptually-based representations both necessary and sufficient to account for the ability of chimpanzees to master relevant cases of higher-order representations of <connection> to explain co-variation of movement (and the lack thereof) among various cases of *objects-in-contact*?

We further expanded upon how such a theory could explain the natural behavior of both chimpanzees *and* humans (see above, "Human Folk Psychology (of Human Folk Physics)");

Povinelli, 2000, p. 297-328). Thus, to elaborate on Boesch's example, one need not see a chimpanzee holding a monkey by the tail (while eating it alive) to realize that his characterization of the issues at stake is far off base.

The question of <weight>. Boesch (2020) makes structurally identical errors in his discussion of *Weight*. He summarizes it as a book "about [Povinelli's] peer-group⁸¹ chimpanzees' notion of weight and their limitations in generalizing and understanding it" (p. 484). This summary misses the entire point of the decade of research that went into the book. As we have seen, *Weight* was an attempt to set aside functional-level explanations invoking "weight", and investigate the representational-level distinction between first-order, perceptually-based constructs of weight (e.g., *effort-to-lift*, *effort-while-lifting*, etc.), and higher-order, structural, role-based representations of <weight> (see the distinctions laid out in Povinelli, 2012, Table 1.1.). Having once again revealed a deep misunderstanding of the irreducible point of our project, Boesch (2020) asks:

How would such a chimpanzee be able to crack nuts in the African forest where so many different potential "hammers" are found, of which only a small minority are functional? ... Not only can Tai chimpanzees appreciate the need to adapt the weight of the hammer to the hardness of the nuts to crack, but they do so by selecting a hammer purely by looking at it and were not observed to manipulate it before use ... Thereby, they demonstrated an uncanny ability to evaluate the unseen properties of tools. (p. 484).⁸²

⁸¹ While it is true that Candy and her companions were formed as a peer group, with the addition of their offspring (Lance, Kegan and Brayden) they quickly became a *family group*.

⁸² In an attempt to widen the mental distance between the two populations of chimpanzees, Boesch (2020) argues: "Furthermore, [the Tai chimpanzees] are not only able to select weight correctly, but at the same time select the size and hardness of the hammer conditional upon the distance they will need to transport it to the anvil where they intend to crack the nuts... Thereby, they demonstrated an uncanny ability to evaluate the unseen properties of tools. Sure, Tai chimpanzees grew up in an environment where processing nuts to eat them is essential for many months of the year, and they saw their mothers do so each year. That is a different world from the one that shaped the cognition of Povinelli's chimpanzees" (p. 484). Fallacies abound. First, he conflates *experience* with particular relations, with the underlying cognitive capacities that do or do not support them. Second, he ignores the fact that, in many cases, the same skills were learned by both groups of chimpanzees. Third, he does not mention the fact that our chimpanzees learned relations never mastered by wild chimpanzees. Fourth, he conflates "unseen" properties of objects with higher-order representations of objects. Fifth, he invokes maternal scaffolding effects despite his own team's recent publication of a detailed, elegant, and powerful longitudinal analysis, which reveals that the suspected forms of maternal input do not, in fact, influence the acquisition of nut-cracking in the Tai chimpanzees:

Against our predictions, we found that neither the general tendency of producing different forms of practicing opportunities provided by mothers ... nor nut-sharing ... promoted immatures' skills, measured as achievement of task understanding, probability that they successfully cracked a nut (efficacy), and number of nuts cracked per minute of nut-cracking (efficiency). However, using hammers that were just previously used by their mothers had a strong positive effect on immatures' efficiency and, seemingly, efficacy. In addition, immatures' efficiency positively correlated with maternal efficiency (Estienne et al., 2019, p. 10-11).

This final point should be stressed and translated: *infants who were fortunate enough to be around stones that were more efficient, were more efficient*. Notwithstanding these important findings, Boesch and colleagues elect to continue to describe the naturalistic observations as *teaching* (see Boesch et al., 2019). As a physical anthropologist, I want to be clear that the problem arises not from the functional label of *teaching*, but from the apparent confusion it causes him (and perhaps others) in understanding what is (and it is not) at stake in the representational-level debates (for other important functional-level data on teaching in chimpanzees, see Musgrave et al., 2020).

So, how *could* wild chimpanzees do all these things without representing weight? The answer is simple: *they could not*. What is more, our chimpanzees could not have displayed all the immediate and learned competences they displayed in *Weight* without "weight." What Boesch misses is the distinction between the functional- and representational-level questions:

Functional level: After much experience, chimpanzees learn to pick the best stones, and after considerably more experience, learn to distinguish between hammer stones and different nuts to maximize their foraging efficiency

Representational-level question: Which of the necessary (and well-understood) perceptually-based representations of weight (e.g., *effort-to-lift*, *effort-while-lifting*, *size*, *texture*) do chimpanzees require in order to learn the relations involved in their skills with hammer stones, and are they sufficient to account for their behavior? Is there any explanatory work to be done by positing the (additional) presence a higher-order, structural, role-based representations of <weight>?

This main point should be hammered home: Boesch's (2020) argument fails to engage with the distinction between perceptual-based relational reasoning and higher-order relational thinking (see Penn et al., 2008a). The *Asymmetric Dependency Problem* illustrates why the failure to make this distinction leaves Boesch's (2020) arguments inert.⁸³ Of course chimpanzees represent weight. The open question is whether they represent <weight>.

The question of <seeing>. Boesch (2020) offers a final case study, this one involving social cognition, and in particular, the debate I discussed at the outset of this essay: the question of whether chimpanzees possess the higher-order concept of <seeing>:

Povinelli's chimpanzees had difficulties distinguishing between a human with a bucket on the head, from one with his visible eyes. His chimpanzees were seen to beg for food equally frequently to both of them! On the other side, wild chimpanzees seem especially sensitive to the gaze of social group members, although this has not been studied systematically in nature. For example, a low-ranking chimpanzee will greet a higher-ranking individual as long as the dominant acknowledges his submission with a peaceful gaze (and not with a head movement). If, however, the dominant refuses to look at him, the low-ranking individual will invariably start to scream ... How would one of Povinelli's chimpanzees fare in such a social environment? (p. 484)

So, how would our chimpanzees fare in that kind of environment? *They lived in it and danced through it with aplomb*. For anyone who finds Boesch's (2020) dialectic compelling, let me be clear: On whatever timescale one chooses to analyze our chimpanzees' social interactions (daily, hourly, by the minute or second), they skillfully deployed the subtlest of these communicative gestures as they navigated their way through the social dynamics he mentions, both in their social interactions with each other, as well as with us (see above, "Candy's Family"). What is more, even our *experimental work* revealed their strong sensitivities to the visual behavior of others, including the movement of just the eyes alone (see above, "Folk Psychology Meets Folk Physics"). Once again:

⁸³ As I have shown throughout this essay, this problem applies with equal force to the entire genre of experimental investigations targeting higher-order thinking in animals (for examples involving weight, see Hanus & Call, 2008; Schrauf & Call, 2009, 2011; see also, Povinelli & Henley, 2020).

Functional level question: Do chimpanzees respond to eyes, face, and bodily orientations to build and navigate through complex social environments? Answer: *Yes, chimpanzees everywhere do this.*

Representational-level question: Are the observed behaviors generated solely by diverse, rich, robust, elaborate, sophisticated, first-order, perceptually-based representations (which the humans and chimpanzees alike both possess), or must they also possess a higher-order, structural, role-based interpretation of the eyes and faces of others (e.g., *<seeing>* or *<attention>*)?

At the risk of being repetitive, Boesch (2020) misses the (only) question at stake in our work. In contrast, what seems to be at stake in Boesch's (2020) essay is his faith in his intuitions about how wild chimpanzees would perform when the natural joints of their free-flowing behavior are pulled apart. Once again, however, given that our chimpanzees displayed the same free-flowing complexities of behavior, what's sauce for the goose is good for the gander.

Stressing Data

To be sure, there are legitimate *a priori* reasons to wonder if the investigation of one group of captive chimpanzees can capture the full scope of the cognitive abilities to all members of the species. But also, *to be sure*, there are strong reasons to doubt Boesch's blunt deprivation arguments as directed against the chimpanzees I was privileged to work with for twenty years. To summarize why:

- Candy and her family developed the full range of chimpanzee social behavior, as well as an extremely large range of spontaneous tool using behavior. Thus, it seems reasonable to conclude that the enriched environment we provided (while not comparable to the wild) contained the abstract social and physical experiences necessary to draw out their species-typical behavior and cognitive skills.
- When viewed across their lifespan, our apes achieved the same functional-level skills that wild chimpanzees master. Indeed, in many cases, they developed far more elaborated skills on such tasks than wild chimpanzees.
- Our chimpanzees confronted (and solved) problems of greater complexity than any ever reported by wild chimpanzees.
- Our chimpanzees had experiences with tool use that cleaved apart the causal features of tools in a way that that likely created far more complex problems than those experienced by most (all?) populations of wild chimpanzees—and all of that was *stacked on top of the time they spent using tools in their spontaneous daily behaviors*.
- Both captive and wild chimpanzees experience stressors. These overlap but each set has many unique elements. It is not at all obvious how to compare the severity of the stressors experienced by our chimpanzees to those experienced by free-ranging populations.
- Finally, even if one decided to ignore all experimental research on captive chimpanzees, the *Asymmetric Dependency Problem* would block any strong inferences about higher-order cognition in wild chimpanzees.

Finally, I believe this exploration of captive and wild chimpanzees reveals something else: the incredulity generated in our minds when we witness how skills which we think ought to hang together, do not, in fact, appear to hang together. Boesch's (2020) essay thus neatly illustrates the naked force of folk psychology turned on itself:

BOESCH: How can the results of *Seeing*, *Folk Physics* and *Weight* be meaningful given the complex behaviors I have witnessed *with my own eyes* in wild chimpanzees!

POVINELLI: In the same way that the complex spontaneous behaviors that we have witnessed in our chimpanzees can be meaningful—right alongside our experimental results.

According to the hypotheses my colleagues and I have been developing over the past two-and-a-half decades, the contradiction resides in our (folk) thinking, not in the observations themselves. Fooled. Then double fooled. And undoubtedly triple fooled back again. The reconciliation begins with an honest reflection on the *morphizational* power of human folk psychology. This requires quelling the internal shouts of our folk psychology. Only after that can one check oneself into the Seven-Step Program to Recovery.

Candy's Coconut Redux

I end by inviting readers to consider how the issues I have explored in this essay are deeply interconnected.

- First, the *Asymmetric Dependency Problem* prevents us from using *either* naturalist observations or experiments to implicate the presence of higher-order thinking in animals (although it may allow stronger inferences about its absence).⁸⁴
- Second, when experimentalists ignore this problem, they confront the same inferential limitations as field researchers, hence the *Experimental Necessity Dilemma*. Although their experiments allow them to say with far more precision *which* first-order representations their animals are responsive to, their conclusions about the presence of higher-order states are no more founded than those based upon naturalistic observations.
- Third, the illusion that this is not so—that we *are* making steady progress toward answering this question—is perpetuated by folk scientific ideas that herd researchers into the *False Models Contrast Problem*. But because the models they are attempting to adjudicate between are not incompatible, researchers inadvertently back into the *Unprincipled Titration Paradox*, which leaves them scrambling to conjure *novel stimuli* to present to their animals. But because these stimuli cannot, in fact, be distinctly novel given the representational commitments that give rise to their entire research operation, *any given experiment ultimately collapses under the weight of its own theoretical incoherence*. As we have seen, any stimuli to which an animal can respond coherently is not, by definition, *novel*—regardless of whether a given animal has or has not previously encountered the precise exemplar of that stimuli created by the experimentalist (see footnote 22).

⁸⁴ *Nota bene*: The *Conflation of Explanatory Levels* alone would likely, by itself, block strong inference derived solely from naturalistic data collected in the wild (see above, "Claim 1: Natural Tool Use Cannot").

None of this implies that the ongoing studies of physical and social cognition in animals are unimportant. On the contrary, I am delighted to see how they continue to address vital debates about the proximate and ultimate functions of the behaviors under scrutiny.

Still, the aim of *Folk Physics* and *Weight* was never about functional-level claims about *complex cognition, causal reasoning* or *intelligent tool use*. These descriptions were assumed, not *to-be-tested-for*. Our work was no more and no less than one attempt to address the long-standing question of whether chimpanzees, like us, think about things as abstract as <gods>, <ghosts> and <gravity>, as well as things equally abstract, but which, for some reason, seem less so—things like <weight>, <shape>, <force> and <connection>. Regardless of what one concludes about the overall pattern of results detected by our projects, the work of Candy and her companions have inarguably revealed a set of core conceptual-methodological questions that block any effort to investigate higher-order thinking in animals.

On a personal note, I understand that much of what I have said herein is hard to accept. We humans want *reasons* for why things happen—in particular, reasons that extend beyond the realm of the perceptual world. Chimpanzees may or may not seek such reasons or wield explanations based upon them. But one thing seems certain: If, in our ongoing drive to determine whether other animals possess higher-order mental states, we continue to inadvertently import those constructions into the premises of our arguments (*read*: experimental designs), they will always appear in the conclusions (*read*: results). This is the very definition of vicious circularity (see Perspective Pieces 1-3). Unless or until we discover news ways of approaching these impasses (cracking the toughest of all the tough nuts that the field faces)—or decide to set the problem aside altogether—our musings about Candy's cogitations as she hammers away with her coconut will remain as ambiguous as ever.

Acknowledgments

The arguments I make in this essay have been formulated through innumerable discussions with many close colleagues over the past 35 years. I have highlighted the especially important contributions of several of these individuals throughout this essay. To this list, I would like to mention two others. First, thanks to the folklorist and friend, Brandon Barker. Our work together had deepened my understanding of the connection between my own would-be investigations of animal folk psychology (and folk physics) and animal (and scientific) folklore. Second, a very special debt is owed to the philosopher (and also dear friend), Caroline Arruda, who commented on successive drafts of this manuscript. I am deeply grateful for the amount of time she devoted in helping me make the arguments contained herein both more precise and more broadly accessible.

References

- Allen, C. (2002). A skeptic's progress. *Biology and Philosophy*, 17(5), 695-702.
- Almasi, B., Rettenbacher, S., Müller, C., Brill, S., Wagner, H., & Jenni, L. (2012). Maternal corticosterone is transferred into the egg yolk. *General and Comparative Endocrinology*, 178, 139–144.
- Amici, F., Aureli, F., Visalberghi, E., & Call, J. (2009). Spider monkeys (*Ateles geoffroyi*) and capuchin monkeys (*Cebus apella*) follow gaze around barriers: Evidence for perspective taking? *Journal of Comparative Psychology*, 123(4), 368-274.
- Amodio, P., Boeckle, M., Jelbert, S. A., Ostojić, L., & Clayton, N. S. (2020). How flexible is tool use in Eurasian jays (*Garrulus glandarius*)? *Animal Behavior and Cognition*, 7(3), 270-287.
- Anderson, J. R. (2001). Review of folk physics for apes: The chimpanzee's theory of how the world works, by Daniel Povinelli. *Animal Behaviour*, 61, 1042-1043.
- Andrews, K. (2005). Chimpanzee theory of mind: Looking in all the wrong places? *Mind & Language*, 20(5), 521-536.
- Andrews, K. (2016). Snipping or editing? Parsimony in the chimpanzee mind reading debate. *Metascience*, 25(3), 377-386.

- Andrews, K. (2020). *The animal mind: An introduction to the philosophy of animal cognition*. Routledge.
- Arcadi, A. C., & Wrangham, R. W. (1999). Infanticide in chimpanzees: Review of cases and a new within-group observation from the Kanyawara study group in Kibale National Park. *Primates*, *40*(2), 337-351.
- Arruda, C. T., & Povinelli, D. J. (2016). Chimps as secret agents. *Synthese*, *193*(7), 2129-2158.
- Arruda, C. T., & Povinelli, D. J. (2018). Two ways of relating to (and acting for) reasons. *Mind & Language*, *33*, 441-459.
- Asquith, P. J. (1984). The inevitability and utility of anthropomorphism in description of primate behaviour. In R. Harré & V. Reynolds (Eds.), *The Meaning of Primate Signals*, (pp. 138–176). Cambridge University Press.
- Auersperg, A. M., von Bayern, A. M., Gajdon, G. K., Huber, L., & Kacelnik, A. (2011). Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. *PLoS One*, *6*(6), e20231.
- Auersperg, A. M., Szabo, B., von Bayern, A. M., & Kacelnik, A. (2012). Spontaneous innovation in tool manufacture and use in a Goffin's cockatoo. *Current Biology*, *22*, R903-R904.
- Auersperg, A. M., Borasinski, S., Laumer, I., & Kacelnik, A. (2016). Goffin's cockatoos make the same tool type from different materials. *Biology Letters*, *12*(11), 20160689.
- Austin, J. L. (1962). *How to do things with words*. Oxford University Press
- Bakuza, J. S., & Nkwengulila, G. (2009). Variation over time in parasite prevalence among free-ranging chimpanzees at Gombe National Park, Tanzania. *International Journal of Primatology*, *30*(1), 43-53.
- Bandini, E., & Tennie, C. (2017). Spontaneous reoccurrence of “scooping”, a wild tool-use behaviour, in naïve chimpanzees. *PeerJ*, *5*, e3814.
- Barker, K. B., & Povinelli, D. J. (2019a). Anthropomorphomania and the rise of the animal mind: A conversation. *Journal of Folklore Research*, *56*(2-3), 71-90.
- Barker, K. B., & Povinelli, D. J. (Eds.). (2019b). *Folklore as an (unlikely) frame for science, The Aesop's fable paradigm* [Special issue]. *Journal of Folklore Research*, *56*(2-3), 1-291.
- Barrett, L. P., & Benson-Amram, S. (2020). Can Asian elephants use water as a tool in the floating object task? *Animal Behavior and Cognition*, *7*(3), 310-326.
- Barth, J., Reaux, J. E., & Povinelli, D. J. (2005). Chimpanzees' (*Pan troglodytes*) use of gaze cues in object-choice tasks: Different methods yield different results. *Animal Cognition*, *8*(2), 84-92.
- Beck, B. B. (1980). *Animal tool behaviour: The use and manufacture of tools by animals*. Garland STPM Publishing.
- Benowitz-Fredericks, Z. M., Shultz, M. T., & Kitaysky, A. S. (2008). Stress hormones suggest opposite trends of food availability for planktivorous and piscivorous seabirds in 2 years. *Deep Sea Research II*, *55*, 1868– 1876.
- Bermudez, J. L. (2003). The domain of folk psychology. In A. O'Hear (Ed.), *Mind and persons* (pp. 25-48). Cambridge University Press.
- Bernstein-Kurtycz, L. M., Hopper, L. M., Ross, S. R., & Tennie, C. (2020). Zoo-housed chimpanzees can spontaneously use tool sets but persevere on previously successful tool-use methods. *Animal Behavior and Cognition*, *7*(3), 288-309.
- Bird, C. D., & Emery, N. J. (2009a). Rooks use stones to raise the water level to reach a floating worm. *Current Biology*, *19*(16), 1410-1414.
- Blas, J., Bortolotti, G. R., Tella, J. L., Baos, R., & Marchant, T. A. (2007). Stress response during development predicts fitness in a wild, long lived vertebrate. *Proceedings of the National Academy of Sciences*, *104*, 8880– 8884.
- Bluff, L. A., Weir, A. A. S., Rutz, C., Wimpenny, J. H., & Kacelnik, A. (2007). Tool-related cognition in New Caledonian crows. *Comparative Cognition & Behavior Reviews*, *2*(1), 1–25.
- Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour*, *41*(3), 530–532.
- Boesch, C. (2005). Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously. *Behavioral and Brain Sciences*, *28*(5), 692-692.
- Boesch, C. (2007). What makes us human (*Homo sapiens*)? The challenge of cognitive cross-species comparison. *Journal of Comparative Psychology*, *121*(3), 227-240.
- Boesch, C. (2020). Mothers, environment, and ontogeny affect cognition. *Animal Behavior and Cognition*, *7*(3), 474-489.
- Boesch, C., & Boesch-Achermann, H. (2000). *The chimpanzees of the Tai Forest*. Oxford University Press.

- Boesch, C., Bombjaková, D., Meier, A., & Mundry, R. (2019). Learning curves and teaching when acquiring nut-cracking in humans and chimpanzees. *Scientific Reports*, 9(1), 1-14.
- Boonstra, R. (2005) Equipped for life: The adaptive role of the stress axis in male mammals. *Journal of Mammalogy*, 86, 236–247.
- Boonstra, R. (2013). Reality as the leading cause of stress: Rethinking the impact of chronic stress in nature. *Functional Ecology*, 27(1), 11-23.
- Boonstra, R., Hik, D., Singleton, G. R., & Tinnikov, A. (1998). The impact of predator-induced stress on the snowshoe hare cycle. *Ecological Monographs*, 68, 371– 394.
- Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, 119(2), 145-154.
- Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition*, 10(4), 439-448.
- Breuner, C. W., & Hahn, T. P. (2003) Integrating stress physiology, environmental change, and behavior in free-living sparrows. *Hormones and Behavior*, 43, 115– 123.
- Bugnyar, T., Stöwe, M., & Heinrich, B. (2004). Ravens, *Corvus corax*, follow gaze direction of humans around obstacles. *Proceedings of the Royal Society of London B*, 271(1546), 1331-1336.
- Bryant, L., Coffey, A., Povinelli, D. J., & Pruett Jr., J. R. (2013). Theory of mind experience sampling in typical adults. *Consciousness and Cognition*, 22, 697-707.
- Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). ‘Unwilling’ versus ‘unable’: Chimpanzees’ understanding of human intentional action. *Developmental Science*, 7(4), 488-498.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187-192.
- Carr, D. (2008). 1. Narrative explanation and its malcontents. *History and Theory*, 47(1), 19-30.
- Chapman, C. A., Saj, T. L. & Snaith, T. V. (2007) Temporal dynamics of nutrition, parasitism, and stress in colobus monkeys: Implications for population regulation and conservation. *American Journal of Physical Anthropology*, 134, 240– 250.
- Cheke, L. G., Loissel, E., & Clayton, N. S. (2012). How do children solve Aesop's Fable? *PLoS One*, 7(7), e40574.
- Christiansen, S. L. (2018). Mediating potency and fear: Action movies’ affect. *Cultural Studies*, 32(1), 43-62.
- Civelek, Z., Call, J., & Se-d, A. M. (2020). Inferring unseen causes: Developmental and evolutionary origins. *Frontiers in Psychology*, 11, 1-17.
- Clark, A., & Thornton, C. (1997). Trading spaces: Computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*, 20(1), 57-66.
- Clinchy, M., Zanette, L., Boonstra, R., Wingfield, J. C., & Smith, J. N. (2004). Balancing food and predator pressure induces chronic stress in songbirds. *Proceedings of the Royal Society of London B*, 271(1556), 2473-2479.
- Cook, R. G., & Smith, J. D. (2006). Stages of abstraction and exemplar memorization in pigeons’ category learning. *Psychological Science*, 17, 1059–1067.
- Cowen, T. (2003). Policing nature. *Environmental Ethics*, 25, 169-182.
- Creel, S., Dantzer, B., Goymann, W., & Rubenstein, D. R. (2013). The ecology of stress: Effects of the social environment. *Functional Ecology*, 27, 66– 80.
- Czaczkes, T. J., Brandstetter, B., di Stefano, I., & Heinze, J. (2018). Greater effort increases perceived value in an invertebrate. *Journal of Comparative Psychology*, 132(2), 200–209.
- Dacey, M. (2016). Rethinking associations in psychology. *Synthese*, 193(12), 3763-3786.
- DeLong, C. M., & Burnett, C. (2020). Bornean orangutans (*Pongo pygmaeus pygmaeus*) use water as a tool in the floating object task. *Animal Behavior and Cognition*, 7(3), 327-342.
- Dennett, D. (1987) *The intentional stance*. MIT Press.
- Dennett, D. (2009). Darwin's “strange inversion of reasoning”. *Proceedings of the National Academy of Sciences*, 106(Supplement 1), 10061-10065.
- de Waal, F. B. (1986). The integration of dominance and social bonding in primates. *The Quarterly Review of Biology*, 61, 459–479.
- du Dot, J. T., Rosen, D. A. S., Richmond, J. P., Kitaysky, A. S., Zinn, S. A., & Trites, A. W. (2009). Changes in glucocorticoids, IGF-I and thyroid hormones as indicators of nutritional stress and subsequent refeeding in Steller sea lions (*Eumetopias jubatus*). *Comparative Biochemistry and Physiology A*, 152, 524–534.

- Dunay, E., Apakupakul, K., Leard, S., Palmer, J. L., & Deem, S. L. (2018). Pathogen transmission from humans to great apes is a growing threat to primate conservation. *EcoHealth*, 15(1), 148-162.
- Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, 306(5703), 1903-1907.
- Estienne, V., Cohen, H., Wittig, R. M., & Boesch, C. (2019). Maternal influence on the development of nut-cracking skills in the chimpanzees of the Tai forest, Côte d'Ivoire (*Pan troglodytes verus*). *American Journal of Primatology*, 81(7), e23022.
- Everett, J. (2001). Environmental ethics, animal welfarism, and the problem of predation: A Bambi lover's respect for nature. *Ethics and the Environment*, 6, 42-67.
- Farrar, B. G., Altschul, D. M., Fischer, J., van der Mescht, J., Placì, S., Troisi, C. A., Vernouillet, A., Clayton, N. S., & Ostojic, L. (2020). Trialling meta-research in comparative cognition: Claims and statistical inference in animal physical cognition. *Animal Behavior and Cognition*, 7(3), 419-444.
- Faria, C., & Paez, E. (2015). Animals in need: The problem of wild animal suffering and intervention in nature. *Relations*, 3, 7-13.
- Feldman, R. (1998). Charity, principle of. In *The Routledge Encyclopedia of Philosophy*. Taylor and Francis. Retrieved 11 Oct. 2020, from <https://www.rep.routledge.com/articles/thematic/charity-principle-of/v-1>.
- Fink, C. K. (2005). The predation argument. *Between the Species*, 13(5), 1-16.
- Fisher, J. A. (1996). The myth of anthropomorphism. In M. Bekoff & D. Jamieson (Eds.). *Readings in animal cognition* (pp. 3-16). MIT Press.
- Fiore, A. M., Cronin, K. A., Ross, S. R., & Hopper, L. M. (2020). Food cleaning by Japanese macaques: Innate, innovative or cultural? *Folia Primatologica*, 91, 433-444
- Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15(5), 447-452.
- Gallagher, S., & Povinelli, D. J. (2012). Enactive and behavioral abstraction accounts of social understanding in chimpanzees, infants, and adults. *Review of Philosophy and Psychology*, 3(1), 145-169.
- Ghirlanda, S., & Lind, J. (2017). 'Aesop's fable' experiments demonstrate trial-and-error learning in birds, but no causal understanding. *Animal Behaviour*, 123, 239-247.
- Giesing, E. R., Suski, C. D., Warner, R. E., & Bell, A. M. (2011). Female sticklebacks transfer information via eggs: Effects of maternal experience with predators on offspring. *Proceedings of the Royal Society B*, 278, 1753-1759.
- Gilby, I. C., & Connor, R. C. (2010). The role of intelligence in group hunting: Are chimpanzees different from other social predators. In E. Lonsdorf, S. R. Ross, T. Matsuzawa, & J. Goodall (Eds.), *The mind of the chimpanzee: Ecological and experimental perspectives* (pp. 220-233). University of Chicago Press.
- Glorioso, G. C., Kuznar, S. L., Pavlic, M., & Povinelli, D. J. (in press). Still no solution to non-verbal measures of analogical reasoning: A reply to Walker & Gopnik (2017). *Cognition*. doi.org/10.1016/j.cognition.2020.104288
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Harvard University Press
- Grueneisen, S., Duguid, S., Saur, H., & Tomasello, M. (2017). Children, chimpanzees, and bonobos adjust the visibility of their actions for cooperators and competitors. *Scientific Reports*, 7(1), 1-10.
- Halina, M. (2015). There is no special problem of mindreading in nonhuman animals. *Philosophy of Science*, 82(3), 473-490.
- Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, 18(9), R370-R372.
- Hanus, D., & Call, J. (2011). Chimpanzee problem-solving: Contrasting the use of causal and arbitrary cues. *Animal Cognition*, 14(6), 871-878.
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59(4), 771-785.
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1), 139-151.
- Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, 101(3), 495-514.
- Hauser, M. D. (2001). Elementary, my dear chimpanzee. *Science*, 291(5503), 440.
- Hawlana, D., & Schmitz, O. J. (2010). Physiological stress as a fundamental mechanism linking predation to ecosystem functioning. *American Naturalist*, 175, 537-556.

- Hayashi, M., & Matsuzawa, T. (2003). Cognitive development in object manipulation by infant chimpanzees. *Animal Cognition*, 6(4), 225-233.
- Hayward, L. S., & Wingfield, J. C. (2004). Maternal corticosterone is transferred to avian yolk and may alter offspring growth and adult phenotype. *General and Comparative Endocrinology*, 135, 365–371.
- Healy, S. D. (2012). “Heavy words are so lightly thrown” (from what difference does it make? The smiths) A review of Daniel J. Povinelli (2012) World without weight: Perspectives on an alien mind. *Journal of Evolutionary Psychology*, 10(3), 159-161.
- Hennefield, L., Hwang, H. G., Weston, S. J., & Povinelli, D. J. (2018). Meta-analytic techniques reveal that corvid causal reasoning in the Aesop’s Fable paradigm is driven by trial-and-error learning. *Animal Cognition*, 21(6), 735-748.
- Henrich J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species and making us smarter*. Princeton University Press
- Henrich, J., & Tennie, C. (2018). Cultural evolution in chimpanzees and humans. In M. N. Muller, R. W. Wrangham, & D. R. Pilbeam (Eds.), *Chimpanzees and human evolution* (pp. 645-702). Harvard University Press.
- Herrmann, E., Wobber, V., & Call, J. (2008). Great apes' (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, *Pongo pygmaeus*) understanding of tool functional properties after limited experience. *Journal of Comparative Psychology*, 122(2), 220–230.
- Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, 21(1), 101-114.
- Heyes, C (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.
- Hicks, T. C., Kühl, H. S., Boesch, C., Dieguez, P., Ayimisin, A. E., Fernandez, R. M., Zungawa, D. B., Kambere, M., Swinkels, J., Menken, S. B. J., Hart, J., Mundry, R., & Roessingh, P. (2019). Bili-Uéré: A chimpanzee behavioural realm in Northern Democratic Republic of Congo. *Folia Primatologica*, 90(1), 3-64.
- Hobson, R. P. (1996). On not understanding minds. In D. J. Povinelli & T. J. Eddy (Eds.), What young chimpanzees know about seeing [Monograph]. *Monographs of the Society for Research in Child Development*, 61 (2, Serial No. 247), pp. 153-160.
- Hoerl, C., & McCormack, T. (2019). Thinking in and about time: A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, 42, e244.
- Hostetter, A. B., Cantero, M., & Hopkins, W. D. (2001). Differential use of vocal and gestural communication by chimpanzees (*Pan troglodytes*) in response to the attentional status of a human (*Homo sapiens*). *Journal of Comparative Psychology*, 115(4), 337–343.
- Hostetter, A. B., Russell, J. L., Freeman, H., & Hopkins, W. D. (2007). Now you see me, now you don’t: Evidence that chimpanzees understand the role of the eyes in attention. *Animal Cognition*, 10(1), 55-62.
- Huffman, M. A., Gotoh, S., Turner, L. A., Hamai, M., & Yoshida, K. (1997). Seasonal trends in intestinal nematode infection and medicinal plant use among chimpanzees in the Mahale Mountains, Tanzania. *Primates*, 38, 111–125.
- Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature*, 379(6562), 249-251.
- Hunt G. R., Gray R. D., & Taylor A. H. (2013). Why is tool use rare in animals? In C. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 89–118). Cambridge University Press.
- Inoue-Nakamura, N., & Matsuzawa, T. (1997). Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 111(2), 159–173.
- Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., Taylor, A. H., & Clayton, N. S. (2019). New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proceedings of the Royal Society B*, 286(1894), 20182332.
- Jo, Y. S., Heymann, G., & Zweifel, L. S. (2018). Dopamine neurons reflect the uncertainty in fear generalization. *Neuron*, 100(4), 916-925.
- Johannsen, K. (in press). *Wild animal ethics. The moral and political problem of wild animal suffering*. Routledge.
- Johannsen, K. (2020). To assist or not to assist? Assessing the potential moral costs of humanitarian intervention in Nature. *Environmental Values*, 29(1), 29-45.

- Jordan, E. J., Townrow, L. A. J., Wright, C. I., & Seed, A. M. (2020). Understanding solidity: Investigating knowledge of a functional object property in brown capuchin monkeys (*Sapajus apella*) and common squirrel monkeys (*Saimiri sciureus*). *Animal Behavior and Cognition*, 7(3), 365-391.
- Jung, W. P., Kahrs, B. A., & Lockman, J. J. (2015). Manual action, fitting, and spatial planning: Relating objects by young children. *Cognition*, 134, 128-139.
- Jurmain, R. (1997). Skeletal evidence of trauma in African apes, with special reference to the Gombe chimpanzees. *Primates*, 38(1), 1-14.
- Kabadayi, C., & Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science*, 357(6347), 202-204.
- Kahn Jr., H. (2003). Ape cognition and why it matters for the field of psychology: Essay review of *Folk physics for apes: The chimpanzee's theory of how the world works* by Daniel J. Povinelli. *Human Development*, 46, 161-168.
- Kahrs, B. A., Jung, W. P., & Lockman, J. J. (2014). When does tool use become distinctively human? Hammering in young children. *Child Development*, 85(3), 1050-1061.
- Kaminski, J., Call, J., & Tomasello, M. (2004). Body orientation and face orientation: Two factors controlling apes' begging behavior from humans. *Animal Cognition*, 7(4), 216-223.
- Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), 224-234.
- Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., & Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proceedings of the National Academy of Sciences*, 116(42), 20904-20909.
- Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: Can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, 105, 211-221.
- Karin-D'Arcy, R., & Povinelli, D. J. (2002). Do chimpanzees know what each other see? A closer look. *International Journal of Comparative Psychology*, 15, 21-54.
- Kaur, T., Singh, J., Tong, S., Humphrey, C., Clevenger, D., Tan, W., Szekely, B., Wang, Y., Li, Y., Muse, E. A., Kiyono, M., Hanamura, S., Inoue, E., Nakamura, M., Huffman, M. A., Jiang, B., & Nishida, T. (2008). Descriptive epidemiology of fatal respiratory outbreaks and detection of a human-related metapneumovirus in wild chimpanzees (*Pan troglodytes*) at Mahale Mountains National Park, Western Tanzania. *American Journal of Primatology*, 70(8), 755-765.
- Keeley, B. L. (2004). Anthropomorphism, primatomorphism, mammalomorphism: Understanding cross-species comparisons. *Biology and Philosophy*, 19(4), 521-540.
- Kenward, B., Rutz, C., Weir, A. A., & Kacelnik, A. (2006). Development of tool use in New Caledonian crows: inherited action patterns and social influences. *Animal Behaviour*, 72(6), 1329-1343.
- Kersken, V., Zhang, D., Gomez, J-C., Seed, A., & Ball, D. (2020). Capuchin monkeys individuate objects based on spatio-temporal and property/kind information: Evidence from looking and reaching measures. *Animal Behavior and Cognition*, 7(3), 343-364.
- Köndgen, S., Kühl, H., N'Goran, P. K., Walsh, P. D., Schenk, S., Ernst, N., Biek, R., Formenty, P., Mätz-Rensing, K., Schweiger, B., Junglen, S., Ellerbrok, H., Nitsche, A., Briese, T., Lipkin, W. I., Pauli, G., Boesch, C., Leendertz, F. H. (2008). Pandemic human viruses cause decline of endangered great apes. *Current Biology*, 18(4), 260-264.
- Kret, M. E., & Roth, T. S. (2020). Anecdotes in animal behaviour. *Behaviour*, 157(5), 385-386.
- Krief, S., Vermeulen, B., Lafosse, S., Kasenene, J. M., Nieguitsila, A., Berthelemy, M., L'Hostis, M., Bain, O., & Guillot, J. (2010). Nodular worm infection in wild chimpanzees in Western Uganda: A risk for human health? *PLoS Neglected Tropical Diseases*, 4(3), e630.
- Laland, K. N., & Janik, V. M. (2006). The animal cultures debate. *Trends in Ecology & Evolution*, 21(10), 542-547.
- Landys, M., Goymann, W. & Slagsvold, T. (2011). Rearing conditions have long-term consequences for stress responsiveness in free-living great tits. *General and Comparative Endocrinology*, 174, 219-224.
- Loretto, M. C., Schloegl, C., & Bugnyar, T. (2010). Northern bald ibises follow others' gaze into distant space but not behind barriers. *Biology Letters*, 6(1), 14-17.
- Love, O. P., McGowan, P. O., & Sheriff, M. J. (2013). Maternal adversity and ecological stressors in natural populations: the role of stress axis programming in individuals, with implications for populations and communities. *Functional Ecology*, 27(1), 81-92.

- Love, O. P., & Williams, T. D. (2008). The adaptive value of stress-induced phenotypes: Effects of maternally derived corticosterone on sex-biased investment, cost of reproduction, and maternal fitness. *American Naturalist*, *172*, E135–E149.
- Lurz, R. W. (2009). If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem. *Philosophical Psychology*, *22*(3), 305-328.
- Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. MIT press.
- Lurz, R. W., Kanet, S., & Krachun, C. (2014). Animal mindreading: A defense of optimistic agnosticism. *Mind & Language*, *29*(4), 428-454.
- Lurz, R., & Krachun, C. (2011). How could we know whether nonhuman primates understand others' internal goals and intentions? Solving Povinelli's problem. *Review of Philosophy and Psychology*, *2*(3), 449.
- Lurz, R., Krachun, C., Mahovetz, L., Wilson, M. J., & Hopkins, W. (2018). Chimpanzees gesture to humans in mirrors: Using reflection to dissociate seeing from line of gaze. *Animal Behaviour*, *135*, 239-249.
- Manrique, H. M., Gross, A. N.-M., & Call, J. (2010). Great apes select tools on the basis of their rigidity. *Journal of Experimental Psychology*, *36*(4), 409–422.
- Martin-Ordas, G., Schumacher, L., & Call, J. (2012). Sequential tool use in great apes. *PLoS One*, *7*(12), e52074.
- McCoy, D. E., Schiestl, M., Neilands, P., Hassall, R., Gray, R. D., & Taylor, A. H. (2019). New Caledonian crows behave optimistically after using tools. *Current Biology*, *29*(16), 2737-2742.
- McMahan, J. (2015). The moral problem of predation. In A. Chignell, T. Cuneo, & M. Halteman (Eds.). *Philosophy comes to dinner: Arguments on the ethics of eating* (pp. 268-294). Routledge.
- Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (*Pan troglodytes*) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154-162.
- Met, A., Miklósi, Á., & Lakatos, G. (2014). Gaze-following behind barriers in domestic dogs. *Animal Cognition*, *17*(6), 1401-1405.
- Moen, O. M. (2016). The ethics of wild animal suffering. *Nordic Journal of Applied Ethics*, *1*, 91-104.
- Musgrave, S., Lonsdorf, E., Morgan, D., Prestipino, M., Bernstein-Kurtycz, L., Mundry, R., & Sanz, C. (2020). Teaching varies with task complexity in wild chimpanzees. *Proceedings of the National Academy of Sciences*, *117*(2), 969-976.
- Ng, Y.-K. (1995) Towards welfare biology: Evolutionary economics of animal consciousness and suffering, *Biology and Philosophy*, *10*, 255-285.
- O'Connell, S. M. (1995). Empathy in chimpanzees: Evidence for theory of mind? *Primates*, *36*(3), 397-410.
- Osiurak, F., & Heinke, D. (2018). Looking for intoelligence: A unified framework for the cognitive study of human tool use and technology. *American Psychologist*, *73*(2), 169–185.
- Osiurak, F., Jarry, C., & Le Gall, D. (2010). Grasping the affordances, understanding the reasoning: Toward a dialectical theory of human tool use. *Psychological Review*, *117*(2), 517–540.
- Osuna-Mascaró, A., & Auersperg, A. (2018). On the brink of tool use? Could object combinations during foraging in a feral Goffin's cockatoo (*Cacatua goffiniana*) result in tool innovations. *Animal Behavior and Cognition*, *5*, 229-234.
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008a). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(2), 109-130.
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008b). Darwin's triumph: Explaining the uniqueness of the human mind without a deus ex machina. *Behavioral and Brain Sciences*, *31*(2), 153-178.
- Penn, D. C., & Povinelli, D. J. (2009). On becoming approximately rational: The relational reinterpretation hypothesis. In S. Watanabe, L. Huber, A. Blaisdel, & A. Young (Eds.), *Rational animals, irrational humans* (pp. 23-43). Keio University Press.
- Penn, D. C., & Povinelli, D. J. (2013). The comparative delusion: Beyond behavioristic and mentalistic explanations for nonhuman social cognition. In H. S. Terrace & J. Metcalfe, (Eds.), *Agency and joint attention* (pp. 62-81). Oxford University Press.
- Povinelli, D. J. (1988). Intention and attribution in primates: A new quest for mind. *Yale Graduate Journal of Anthropology*, *1*, 37-48.
- Povinelli, D. J. (1991). *Social intelligence in monkeys and apes*. [Unpublished doctoral dissertation]. Yale University.
- Povinelli, D. J. (1996). Growing up ape. In D. J. Povinelli & T. J. Eddy (Eds.). What young chimpanzees know about seeing [Monograph]. *Monographs of the Society for Research in Child Development*, *61*, (2 Serial No. 247), pp. 174-189.

- Povinelli, D. J. (1997). *Panmorphism*. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, anecdotes and animals* (pp. 92-103). State University of New York Press.
- Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford University Press. [Reprinted, 2003 with revisions and a new Forward by Nicholas Humphrey]
- Povinelli, D. J. (2012). *World without weight: Perspectives on an alien mind*. Oxford University Press.
- Povinelli, D. J. & Barker, K. B. (2019). Introduction: The perplexities of water. *Journal of Folklore Research*, 56(1-2), 1-13.
- Povinelli, D. J., Bering, J., & Giambrone, S. (2000). Toward a science of other minds: Escaping the argument by analogy. *Cognitive Science*, 24, 509-541.
- Povinelli, D. J., Bierschwale, D. T., & Cech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, 17, 37-60.
- Povinelli, D. J. & Dunphy-Lelii, S. (2001). Do chimpanzees seek explanations? Preliminary comparative investigations. *Canadian Journal of Experimental Psychology*, 55, 93-101.
- Povinelli, D. J., Dunphy-Lelii, S., Reaux, J. E., & Mazza, M. P. (2002). Psychological diversity in chimpanzees and humans: New longitudinal assessments of chimpanzees' understanding of attention. *Brain, Behavior and Evolution*, 59, 33-53.
- Povinelli, D. J. & Eddy, T. J. (1996a). What young chimpanzees know about seeing [Monograph]. *Monographs of the Society for Research in Child Development*, 61 (2, Serial No. 247).
- Povinelli, D. J., & Eddy, T. J. (1996b). Chimpanzees: Joint visual attention. *Psychological Science*, 7, 129-135.
- Povinelli, D. J., & Eddy, T. J. (1996c). Factors influencing young chimpanzees' (*Pan troglodytes*) recognition of attention. *Journal of Comparative Psychology*, 110, 336-345.
- Povinelli, D. J., & Eddy, T. J. (1997). Specificity of gaze-following in young chimpanzees. *British Journal of Developmental Psychology*, 15, 213-222.
- Povinelli, D. J., & Frey, S. (2016). Constraints on the exploitation of the functional properties of objects in expert tool-using chimpanzees (*Pan troglodytes*). *Cortex*, 82, 11-23.
- Povinelli, D. J., & Giambrone, S. (1999). Inferring other minds: Failure of the argument by analogy. *Philosophical Topics*, 27, 167-201.
- Povinelli, D. J., & Henley, T. (2020). More rope tricks reveal why more task variants will never lead to strong inferences about higher-order causal reasoning in chimpanzees. *Animal Behavior and Cognition*, 7(3), 392-418.
- Povinelli, D. J., Theall, L. A., Reaux, J. E., & Dunphy-Lelii, S. (2003). Chimpanzees spontaneously modify the direction of their gestural signals to match the attentional orientation of others. *Animal Behaviour*, 65, 71-79.
- Povinelli, D. J., Reaux, J. E., & Frey, S. H. (2010). Chimpanzees' tool use provides evidence for separable representations of hand and tool even during active use within peripersonal space. *Neuropsychologia*, 48, 243-247.
- Povinelli, D. J., & Penn, D. C. (2011). Through a floppy tool darkly: Toward a conceptual overthrow of animal alchemy. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 69-88). Oxford University Press.
- Povinelli, D. J. (2012). Afterword: World without apes? In D. J. Povinelli (Ed.), *World without weight* (pp. 330-332). Oxford University Press.
- Povinelli, D. J., & Prince, C. G. (1998). When self met other. In M. Ferrari, & R. J. Sternberg (Eds.), *Self-awareness: Its nature and development* (pp. 37-107). Guilford Press.
- Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Science*, 7, 157-160.
- Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language*, 19, 1-28.
- Povinelli, D. J., & Vonk, J. (2012). Case study: Levels of understanding rigid and floppy tools. In D. J. Povinelli (Ed.), *World without weight* (pp. 309-318). Oxford University Press.
- Premack, D. (1988). "Does the chimpanzee have a theory of mind" revisited. In R. Byrne, & A. Whiten (Eds.), *Machiavellian intelligence* (pp. 160-179). Oxford University Press.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515-526.

- Pruetz, J. D., & Bertolani, P. (2007). Savanna chimpanzees, *Pan troglodytes verus*, hunt with tools. *Current Biology*, 17(5), 412-417.
- Ramsay, M. S., & Teichroeb, J. A. (2019). Anecdotes in primatology: Temporal trends, anthropocentrism, and hierarchies of knowledge. *American Anthropologist*, 121(3), 680-693.
- Reaux, J. E., Theall, L. A., & Povinelli, D. J. (1999). A longitudinal investigation of chimpanzees' understanding of visual perception. *Child Development*, 70, 275-290.
- Roche, M. (2013). Povinelli's problem and introspection. *Review of Philosophy and Psychology*, 4(4), 559-576.
- Russon, A. E. (2018). Pantomime and imitation in great apes: Implications for reconstructing the evolution of language. *Interaction Studies*, 19(1-2), 200-215.
- Russon, A., & Andrews, K. (2011). Orangutan pantomime: Elaborating the message. *Biology Letters*, 7(4), 627-630.
- Rutz, C., Sugasawa, S., Van der Wal, J. E., Klump, B. C., & St Clair, J. J. (2016). Tool bending in New Caledonian crows. *Royal Society Open Science*, 3(8), 160439.
- Sanz, C., Call, J., & Morgan, D. (2009). Design complexity in termite-fishing tools of chimpanzees (*Pan troglodytes*). *Biology Letters*, 5(3), 293-296.
- Sanz, C. M., & Morgan, D. B. (2007). Chimpanzee tool technology in the Goualougo Triangle, Republic of Congo. *Journal of Human Evolution*, 52(4), 420-433.
- Sapontzis, S. F. (1987) *Morals, reason and animals*. Temple University Press.
- Schrauf, C., & Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. *Animal Cognition*, 12(4), 567-574.
- Schrauf, C., & Call, J. (2011). Great apes use weight as a cue to find hidden food. *American Journal of Primatology*, 73(4), 323-334.
- Schuppli, C., & van Schaik, C. P. (2019). Animal cultures: How we've only seen the tip of the iceberg. *Evolutionary Human Sciences*, 1, e2, 1-13.
- Seed, A., & Byrne, R. (2010). Animal tool-use. *Current Biology*, 20(23), R1032-R1039.
- Seed, A., & Call, J. (2009). Causal knowledge for events and objects in animals. In S. Watanabe, L. Huber, A. Blaisdell, & A. Young (Eds.), *Rational animals, irrational humans* (pp. 173-188). Keio University.
- Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes*, 35(1), 23-34.
- Shaw, J. A., Bryant, L., Malle, F. M., Povinelli, D. J., & Pruett, J. R. (2017). The relationship between joint attention and theory of mind in neurotypical adults. *Consciousness and Cognition*, 51, 268-278.
- Sheriff, M. J., Wheeler, H., Donker, S. A., Krebs, C. J., Palme, R., Hik, D., & Boonstra, R. (2012). Mountain-top and valley-bottom experiences: The stress axis as an integrator of environmental variability in arctic ground squirrel populations. *Journal of Zoology*, 287, 65-75.
- Silva, F. J., Page, D. M., & Silva, K. M. (2005). Methodological-conceptual problems in the study of chimpanzees' folk physics: how studies with adult humans can help. *Animal Learning and Behavior*, 33(1), 47-58.
- Silva, F. J., & Silva, K. M. (2006). Humans' folk physics is not enough to explain variations in their tool-using behavior. *Psychonomic Bulletin and Review*, 13(4), 689-693.
- Silva, F. J., Silva, K. M., Cover, K. R., Leslie, A. L., & Rubalcaba, M. A. (2008). Humans' folk physics is sensitive to physical connection and contact between a tool and reward. *Behavioural Processes*, 77(3), 327-333.
- Silva, F. J., ten Hope, M. I., & Tucker, A. L. (2014). Adult humans' understanding of support relations: An up-linkage replication. *Learning and Behavior*, 42, 337-347.
- Sober, E. (2016). Replies to Kristin Andrews's, Gordon Belot's, and Patrick Forber's reviews. *Metascience*, 25, 393-403.
- Sousa, C., Biro, D., & Matsuzawa, T. (2009). Leaf-tool use for drinking water by wild chimpanzees (*Pan troglodytes*): Acquisition patterns and handedness. *Animal Cognition*, 12(1), 115-125.
- Taylor, A. H. (2020). Folk physics for crows? *Animal Behavior and Cognition*, 7(3), 452-456.
- Taylor, A. H., Hunt, G. R., Medina, F. S., & Gray, R. D. (2009a). Do New Caledonian crows solve physical problems through causal reasoning? *Proceedings of the Royal Society B*, 276(1655), 247-254.
- Taylor, A., Roberts, R., Hunt, G., & Gray, R. (2009b). Causal reasoning in New Caledonian crows: Ruling out spatial analogies and sampling error. *Communicative and Integrative Biology*, 2(4), 311-312.

- Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B*, 364(1528), 2405-2415.
- Terio, K. A., Lonsdorf, E. V., Kinsel, M. J., Raphael, J., Lipende, I., Collins, A., Li, Y., Hahn, B. H., Travis, D. A., & Gillespie, T. R., (2018). Oesophagostomiasis in non-human primates of Gombe National Park, Tanzania. *American Journal of Primatology*, 80(1), e22572.
- Terio, K. A., Kinsel, M. J., Raphael, J., Mlengeya, T., Lipende, I., Kirchhoff, C. A., Gilagiza, B., Wilson, M. L., Kamenya, S., Estes, J. D., Keele, B. F., Rudicell, R. S., Liu, W., Patton, S., Collins, A., Hahn, B. H., Travis, D. A., & Lonsdorf, E. V. (2011). Pathologic lesions in chimpanzees (*Pan troglodytes schweinfurthii*) from Gombe National Park, Tanzania, 2004–2010. *Journal of Zoo and Wildlife Medicine*, 42(4), 597-607.
- Theall, L. A., & Povinelli, D. J. (1999). Do chimpanzees tailor their attention-getting behaviors to fit the attentional states of others? *Animal Cognition*, 2, 207-214.
- The Harvard Crimson (1990, September 10). Renowned Psychologist B. F. Skinner Dies at 86. Retrieved from <https://www.thecrimson.com/article/1990/9/10/renowned-psychologist-bf-skinner-dies-at/> on July 24, 2020.
- Tinbergen, N., & Perdeck, A. C. (1951). On the stimulus situation releasing the begging response in the newly hatched Herring Gull chick (*Larus argentatus argentatus* Pont.). *Behaviour*, 3(1), 1-39.
- Tomasello, M. (1996). Chimpanzee social cognition. In D. J. Povinelli & T. J. Eddy (Eds.), What young chimpanzees know about seeing [Monograph]. *Monographs for Society for Research in Child Development*, 61 (2, Serial No. 247), pp. 161-173.
- Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
- Tomasello, M., & Call, J. (2006). Do chimpanzees know what others see—or only what they are looking at? In S. Watanabe, L. Huber, A. Blaisdel, & A. Young (Eds.), *Rational animals, irrational humans* (pp. 371-384). Keio University Press.
- Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7(4), 153-156.
- Tomasello, M., & Call, J. (2008). Assessing the validity of ape-human comparisons: A reply to Boesch (2007). *Journal of Comparative Psychology*, 122(4), 449–452.
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35(2), 137-154.
- Uomini, N., & Hunt, G. (2017). A new tool-using bird to crow about. *Learning & Behavior*, 45(3), 205-206.
- van de Rut-Plooi, H. H., & Plooi, F. X. (1988). Mother-infant relations, conflict, stress and illness among free-ranging chimpanzees. *Developmental Medicine & Child Neurology*, 30(3), 306-315.
- Velleman, J. D. (2003). Narrative explanation. *The Philosophical Review*, 112(1), 1-25.
- Völter, C. J., & Call, J. (2014). The cognitive underpinnings of flexible tool use in great apes. *Journal of Experimental Psychology*, 40(3), 287–302.
- Völter, C. J., Lambert, M. L., & Huber, L. (2020). Do nonhumans seek explanations? *Animal Behavior and Cognition*, 7(3), 445-451.
- von Bayern, A. M., Heathcote, R. J., Rutz, C., & Kacelnik, A. (2009). The role of experience in problem solving and innovative tool use in crows. *Current Biology*, 19(22), 1965-1968.
- Vonk, J., & Povinelli, D. J. (2006). Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. In E. Wasserman, & T. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 363-387). Oxford University Press.
- Wallis, J., & Lee, R. D. (1999). Primate conservation. The prevention of disease transmission. *International Journal of Primatology*, 20, 803–826.
- Watts, D. P., & Mitani, J. C. (2000). Infanticide and cannibalism by male chimpanzees at Ngogo, Kibale National Park, Uganda. *Primates*, 41(4), 357-365.
- Weir, A. A., & Kacelnik, A. (2006). A New Caledonian crow (*Corvus moneduloides*) creatively re-designs tools by bending or unbending aluminium strips. *Animal Cognition*, 9(4), 317-334.
- Wenner, A. M., Wells, P. H., & Rohlf, F. J. (1967). An analysis of the waggle dance and recruitment in honey bees. *Physiological Zoology*, 40(4), 317-344.
- Whiten, A. (2001). Tool tests challenge chimpanzees. *Nature*, 409(6817), 133-133.
- Whiten, A. (2017). Culture extends the scope of evolutionary biology in the great apes. *Proceedings of the National Academy of Sciences*, 114(30), 7790-7797.
- Whiten, A. (2019). Cultural evolution in animals. *Annual Review of Ecology, Evolution, and Systematics*, 50, 27-48.

- Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*, *11*(2), 233-244.
- Williams, J. M., Lonsdorf, E. V., Wilson, M. L., Schumacher-Stankey, J., Goodall, J., & Pusey, A. E. (2008). Causes of death in the Kasekela chimpanzees of Gombe National Park, Tanzania. *American Journal of Primatology*, *70*(8), 766-777.
- Wilson, M. L., Wallauer, W. R., & Pusey, A. E. (2004). New cases of intergroup violence among chimpanzees in Gombe National Park, Tanzania. *International Journal of Primatology*, *25*(3), 523-549.
- Wimpenny, J. H., Weir, A. A., Clayton, L., Rutz, C., & Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in New Caledonian crows. *PLoS One*, *4*(8), e6471.
- Wise, S. (2000). *Rattling the cage: Toward legal rights for animals*. Da Capo Press.
- Woodford, M. H., Butynski, M. T., & Karesh, W. B. (2002). Habituating the great apes: The disease risks. *Oryx*, *36*, 153-160.
- Zentall, T. R., Wasserman, E. A., Lazareva, O. F., Thompson, R. K. R., & Rattermann, M. J. (2008). Concept learning in animals. *Comparative Cognition & Behavior Reviews*, *3*, 13-45.