Title:

In Search of Balance: A Review of Povinelli's *World without Weight*

Forthcoming in *Biology & Philosophy*

Cameron Buckner
Visiting Assistant Professor
University of Houston
513 Agnes Arnold Hall
Houston, TX 77204-3004
Phone: 713-743-3010
Fax: 713-743-5162

Humboldt Postdoctoral Fellow
Ruhr-University Bochum

cjbuckner@uh.edu

Review of Daniel J Povinelli, *World without Weight: Perspectives on an Alien Mind*. Oxford University Press, 2012, 353 pp.

Abstract: Povinelli and colleagues ask whether chimpanzees can understand the concept of weight, answering with a resounding "no". Their answer is justified by appeal to over thirty previously unpublished experiments. I here evaluate in detail Povinelli's arguments against his critical targets, questioning the assumption that such comparative questions will ever be resolved with a simple "yes" or "no".

The comparative study of animal cognition increasingly finds itself caught in persistent standoffs between proponents and skeptics who agree on all the data but disagree on their interpretation. Daniel Povinelli and his colleagues are well-known for their criticisms of evidence that chimpanzees can understand mental states or the causal properties of tools (Povinelli & Eddy 1996; Povinelli 1999); in *World without Weight*, they extend and systematize this critique to a new domain, arguing that chimpanzees cannot understand the concept of weight. The project is offered as the most compelling evidence yet for their Relational Reinterpretation Hypothesis, which holds that only humans can form theories involving higher-order relations that unite perceptually disparate situations under a common underlying explanation, for example in the way that redox chemistry explains burning, rusting, and metabolism as different forms of oxidation. Should our nearest living relative be unable to understand a force about which they have such ample opportunity to learn, the argument goes, it is unlikely that any other animal will be able to represent any other sort of theoretical relation either.

Such claims seem to contradict some of the most alluring recent findings in primatology. For example, wild and captive chimpanzees appear to select stone tools for nut-cracking on the basis of their weight, and even to flexibly calibrate their choices based on tradeoffs between how long they must carry a stone tool versus the durability of the nuts to be cracked (Sakura and Matsuzawa 1991; Schrauf and Call 2011). Povinelli dismisses these results on the grounds that sensory-motor representations of the effort required to lift the tools are sufficient to explain them, and no genuine theoretical representation of weight as an unobservable force is required. Additionally, Hanus & Call (2008) found that when they placed a food reward in one of two cups on a balance beam and left the other empty (the subjects could not see *which* was baited), their chimpanzees reliably chose, on the first trial, to explore the cup on the side of the beam that tilted downward. Hanus & Call claim this result provides evidence that chimpanzees choose by inferring “that the presence of the reward, and its weight in particular, caused the balance to tilt”. Povinelli holds this interpretation to be a direct challenge to his overarching hypothesis, but argues that these results can more parsimoniously be interpreted in terms of the chimpanzees having learned the obvious perceptual regularity that “things move downward when dropped”, and so followed a simple rule “always search for food below where it was released.” To test this interpretation of Hanus & Call’s results, Povinelli and colleagues conducted a similar experiment in which both a grape and a lead weight were placed on opposite sides of a balance beam, finding that their chimpanzees continued to prefer the cup that moves downward through sixteen test trials, even though the subjects repeatedly discovered in these trials that the bottom cup contained not the grape but rather the heavier lead weight.

Before delving further into this controversy, I should attempt to explain why work in this area is so fractious. I begin with some background: the field is currently characterized by two distinct approaches to answering comparative questions about cognitive capacities. On the first method, the researcher aims to find positive evidence that animals possess some sophisticated human capacity by attempting to devise a task that could only be solved by using the competence under investigation. On this methodology, the researcher tries to control for deflationary explanations of success, such as that the animal found some simple perceptual cue indicating the correct solution, or that they possess an inflexible reflex, heuristic, or simple form of associative learning that allows them to solve only a limited range of problems. If the animals pass the test under appropriate controls, the thought goes, then they must possess the capacity in question; and we can, as Povinelli puts it derisively, “log another similarity in the annals of comparative psychology” and call it a day.

The second approach is practiced almost exclusively by Daniel Povinelli and his colleagues, and involves the collection of large numbers of negative results by performing increasingly difficult variants of experimental tasks on a small number of captive animals. The goal of this approach is to find the limits of animal cognitive abilities—to demonstrate that there are classes of problems that they cannot solve, because they lack some general psychological ability. Since negative findings can have at least as many alternative explanations as positive findings—the animals might have misunderstood the task, lacked critical background learning required to recognize solutions, be distracted or otherwise unable to perceive important features of the task, or lack motivation—large numbers of results must be collected on a number of task variants to rule out alternative explanations for failure.

By now, criticism of Povinelli's approach is a genre in itself (e.g. Allen, 2002; Whiten, 2001), so I will briefly review, but not dwell on, the standard concerns about his previous book, *Folk Physics for Apes*. Many of the worries pertain to the purportedly atypical upbringing and daily life of his chimpanzees at the New Iberia Research Center: the "Project Megan" troop (see p 62-64) consisted of seven animals reared together as peers (i.e. no adults) in a captive environment. This upbringing raises concerns about how far one can generalize from their abilities to those of chimpanzees (or other animals) as a whole. For example, Megan's troop would have missed out on any age-critical instruction or enculturation involving adult teachers; their "concrete and chain-link" environment might have differed in variety and validity from their natural environments (Allen 2001, 697; Bulloch et al. 2008); and they were tested daily for decades on a variety of artificial tasks but at the end of the day were fed regardless of success. Povinelli dismisses these concerns about the troop as "more than a fair bit of silliness…[from] people who know little or nothing about them" (64). This response is not reassuring, given that chimpanzees from other groups have occasionally been reported to pass the same tests (Bulloch, Boysen, and Furlong 2008; Kaminski, Call, and Tomasello 2004) that Povinelli's animals failed (Povinelli and Eddy 1996; Reaux, Theall, and Povinelli 1999). Additionally, reviewers have worried about

Povinelli's approach to publishing—packaging large numbers of studies into monographs rather than subjecting each study to rigorous peer review. While conceding that the work did not go through the standard process of journal peer review, ten anonymous reviewers recommended *WwW* for publication; and three of those reviewers who saw a complete draft were distinguished full professors: a comparative psychologist, a developmental psychologist, and a comparative neuroscientist, all with significant backgrounds in primatology.[1]

In Povinelli's defense, he is forthright about his work's limitations; and given the headwinds facing primatology today, perhaps we must make do with what data can be gathered under imperfect conditions. The economic downturn has been especially hard on primatology, with several major labs—including both Boysen's lab at Ohio State and Povinelli's lab at New Iberia—having recently been closed. In the interim, however, questions about the psychological abilities of our nearest living relatives have only grown more pressing. This combination of high stakes and dwindling resources has produced an environment governed by a corollary of Kissinger's Principle: the interpretive controversies are so persistent precisely because the data are so thin.

Given these caveats, what else has changed in the intervening decade since *FP*? Povinelli now emphasizes a systematization of his ideas about the Reinterpretation Hypothesis into what he and Derek Penn call a "representational-level account", highlighted in *WwW* through a dramatized retelling of his initial conversation with Penn that convinced him of its truth. Whereas in earlier work Povinelli attempted to show that many positive results in animal cognition could more parsimoniously be explained in terms of simple classical or instrumental conditioning, he now explicitly acknowledges that animals are capable of remarkable cognitive flexibility. His emphasis has now shifted to describing the role-based analogy-making faculty that human minds purportedly possess but nonhuman minds lack. Such role-based representations supposedly enable humans to solve tasks in a domain-general fashion,

[1] Information obtained from private communication with Martin Baum, the book's editor at OUP.

transferring their knowledge from one task to others that are perceptually dissimilar but share the same underlying structure. The test for whether chimpanzees have such an understanding of weight—which Povinelli designates '*f(weight)*', to avoid confusion—is thus whether they can solve weight-related problems in a flexible, transferrable way that is not tied to idiosyncratic perceptual similarities or contingencies.

Utilizing this modus operandi, Povinelli and colleagues evaluate the understanding of their troop across a variety of tasks, demonstrating either that they fail at a task that *f(weight)* should allow them to solve, or solve the task only by using a perception-based strategy. In Experiments 1-8, the authors explore what chimpanzees *can* do: use sophisticated and flexible sensorimotor representations of weight to sort objects by the effort required to lift them, calibrate lift forces to object sizes before actually hefting them (thereby exhibiting the "positional overshoot effect", where an object is lifted with too much force when its weight is surreptitiously reduced from one trial to another), and associate color cues with differences in *effort-to-lift*. In Experiments 10-13, their chimpanzees' sorting ability apparently dropped to chance when required to make a choice about whether to place an object in the "light" or "heavy" bin before being allowed to heft them, though three of their subjects were eventually able to plan where to place novel heavy or light objects by color cues before hefting them.

The remainder of the experiments in *WwW* are offered as evidence for what chimpanzees *cannot* do. Experiment 14 tested whether their sorting ability was based on a mere association between color cues and sorting locations by familiarizing their chimpanzees with object weight during free play rather than by training them on the sorting task. In this case, the chimpanzees did not exhibit evidence of transfer, leading Povinelli and colleagues to conclude that the subjects solved the initial task merely by associating color cues with sorting locations.[2] Experiments 15-23 explored whether chimpanzees could use *f(weight)* to reason about causal influence of weight as an impediment, such as the tendency

[2] A response is that the color cues were adequate for predicting success on the earlier task, so there would have been no reason for the chimpanzees to adopt a strategy that sorted by weight rather than by color.

of heavy items to prevent chimpanzees from pulling food rewards to them with strings, and of heavy projectiles to dislodge food rewards in a series of ball drop tasks. While in most cases the chimpanzees could eventually learn the solution to the problem—e.g., to launch a heavy ball down a ramp to dislodge an apple, rather than another ball that is too light—in each case they acquired the solution too slowly, or failed to transfer to novel conditions (e.g. when the actual collision was hidden by an occluder) relevant to meeting Povinelli's standards for *f(weight)*. Finally, Experiments 25-29 showed that while chimpanzees are competent with the "close imposter" concepts *balance*, *support*, and *deformation*, these representations again fall short of a domain-general *f(weight)*; and Experiments 30-32 reviewed results suggesting that human children succeed on several of the same tasks that chimpanzees fail, and thus, at least by the age of 5, possess a robust representation of *f(weight)*.

The primary problem with Povinelli and colleagues' case against their opponents is that in offering their revised "representation-level" account, they did not include the semantics for those representations. In other words, in attempting to move the debate from "cognition vs. association" to "relational reinterpretation vs. everything else" they have provided no argument as to why genuine competence with the concept *weight* (or any other theoretical concept) must be interpreted as the fully abstract, non-perceptual, domain-general *f(weight)*, as opposed to something more graded and moderate.

There are a number of reasons to eschew such strictures. For one, Povinelli's targets never invoke fully domain-general ability in their positive claims. Indeed, in the debate over social cognition proponents have explicitly disavowed Povinelli's similarly "black and white" interpretation of "theory of mind", claiming instead to favor a "graded notion" in which animals can display some sensitivity to the mental states of others in some contexts but not in others (Bugnyar 2007; Call & Tomasello 2008). Moreover, it is unclear that even humans possess such a domain-general, abstract competence with any of the theoretical concepts with which they are judged competent. For example, Landy & Goldstone

(2007) found that the validity judgments of even high-scoring undergraduates on simple algebra problems were sensitive to formally irrelevant manipulations of the equations, such as changes of spacing amongst the symbols and the names chosen for variables. Under Povinelli's criteria for competence, these undergraduates must apparently be deemed incompetent with simple mathematical concepts like *validity* and *equivalence*. As Penn, Holyoak, & Povinelli (2008) eventually concede, even adult humans only approximate the relevant computational architecture; and that humans approximate this ideal to a greater degree than animals should come as no great surprise.

This concession undermines Povinelli's case against opponents such as Hanus & Call (2008). Anticipating that their balance experiment could be explained in terms of a simple bias for the salient motion of the downward trajectory independently of its cause, Hanus & Call performed a "non-causal" control condition in which the balance was pushed down by the experimenter rather than by the weight of the food. In this case, the chimpanzees showed no preference for the lower cup. Povinelli dismisses the adequacy of this control on the grounds that since the experimenters waited four seconds before pushing down the cup, this fell outside the "Michottean window" of 250 milliseconds required for causes to be perceived as connected to their effects (Michotte 1963). A problem with this worry is that the chimpanzees did not actually observe the contact between the food and the cup that moved the scale, whereas visible contact between objects is typically required for Michotte-style causal perception.[3] Furthermore, the outcome of the "lead weight and grape" trials offered by Povinelli as support for his interpretation of the results is ambiguous; the chimpanzees were given only 16 trials in which both the grape and the lead weight were placed in the cups, but they had never before been exposed to a trial in which two objects were placed in the cups simultaneously, nor were they able to observe, during the critical test blocks, that both cups simultaneously contained objects. The worry is that the chimpanzees might have induced the rule from familiarization and prior trials that only one cup

[3] I am grateful to Josep Call, who originally raised these worries in private communication.

will contain an object at a time; and if following that rule, one would predict that they would continue to choose the lower cup even if they understood something robust about *weight*. Crucially, overturning the lower cup and finding the lead weight would not be evidence against this rule, and sixteen trials may not have been sufficient for chimpanzees to realize that the problem type had been changed to this more complex format.

Such ambiguity is typical of Povinelli's critical experiments, which often confound increasingly diagnostic assessment of the competence under dispute with increasing inferential complexity of the task at hand. Domain-general conceptual competence, it seems, also requires open-ended inferential sophistication. If we reject this assumption and its corollary that *weight* must be interpreted as *f(weight)*, the data leave us with a complex and multifaceted picture of chimpanzee's understanding of *weight*: they can at least sort objects by *effort-to-lift*, program lift forces before lifting, learn to sort objects by size and color cues correlated with weight, reason about many concomitants of weight such as *deformation*, *support*, and *balance*, and distinguish a case where an unobserved object presses down upon a balance beam vs. the case where the beam moves down due to at least some other reasons. Chimpanzees should be expected to fail on many other weight-relevant tasks, but it is of much greater theoretical interest to explain these errors in their own right rather than to chalk them up to the monolithic absence of a faculty that not even humans fully possess (Seed et al. 2012).

The most lasting scholarly contribution of *WwW* is likely to be Povinelli and colleagues' meticulous trial-by-trial description of the patterns of learning and error observed in troop Megan across dozens of experiments. In many cases, the authors are left puzzled and intrigued by these patterns. For example, Povinelli reports extreme variability in the number of trials required to reach criterion in one of the weight sorting tasks, between 390 trials and 1562 (one animal did not even reach criterion by 2422 trials). Why do some chimpanzees master the sorting task so much faster than others? Moreover, on several transfer tasks, especially one involving a transfer of sorting ability from jugs to balls, the data demonstrated a peculiar “U-shaped” pattern, with initial success, followed by a decline in performance, and then a gradual increase back up to criterion. Povinelli conjectures that this pattern is due to lower levels of difficulty in associating sorting locations with color and shape cues, compared to cues based on their weight, with initial transfer due to transfer of the color cue strategy and later performance diminished due to slower re-learning of the weight-based strategy. Another example involved Experiment 15, in which Povinelli and colleagues attempted to determine whether apes could use observational cues to infer an object’s weight by allowing subjects to watch three investigators pushing and pulling a heavy box across the floor with great effort. The chimpanzees are reported as going “crazy, hooting and hollering for the full three to five minutes of the demonstration” (151). Did they perceive the box as a disputed resource, and did their interpretation of the demonstration explain why all but one of the chimpanzees selected the heavy box—the “incorrect” choice, because it was too heavy for them to move alone, and thereby obtain the food reward resting on top of it—on their first test trial?

Near the end of the book, Povinelli calls for the comparative exploration of a “‘causal workspace’ within which various weight-related representations can do their work” (292). This workspace would be fleshed out from a combination of more precise models of psychophysics, memory, motivation, learning, and inference, and held responsible to fine-grained patterns of learning rate, error, and individual differences. Such a theory would attempt to explain why, rather than just report that, particular species and individuals find certain problem types easier or more difficult than others. Through developing such models, we can ask about the biological cost/benefit tradeoffs offered by various representational strategies in different environments in a more rigorous and responsible way. In service of this goal, a comparison of failures would be at least as interesting as a comparison of successes (e.g. assessing whether animals make the same errors that humans do regarding “impetus principles” in folk physics—Kozhevnikov & Hegarty, 2001), rather than just, upon finding a task that humans can pass but animals cannot, "logging another difference in the annals of *contrastive* psychology" and calling it a day.

In short, *WwW* offers us two distinct Povinellis: the irascible "sheriff of comparative psychology" (296) and the inspirational, data-driven visionary. The sheriff dismisses his critics as having watched too much *Finding Nemo* and *The Lion King*, and quixotically hopes to prove, against an allegedly powerful and intransigent consensus, that animals lack a capacity that admittedly even humans only imperfectly approximate and rarely deploy. On the other hand, Povinelli the visionary aims to reinvigorate comparative psychology by ushering it away from pass-fail tests generated by armchair intuitions and towards specific processing models of animal learning and inference held responsible to the finer points of the data. The link between the two Povinellis is tenuous; rewarmed Newell & Simon is not the shot in the arm that 21st-Century comparative theory so badly needs, for the critical narrative derived from it tells us nothing about how animals actually *do* solve tasks, and is beholden to the same inflated, folk-psychological notions of *weight* and *theory of mind* that Povinelli campaigns to eradicate.

What barriers prevented Povinelli and colleagues from leaving behind the controversy and simply reforming by example? It is worth pondering one of Derek Penn's favorite quips, that human cognition "sticks out on this planet like an elephant's trunk" (ix). Notably, comparative morphologists have not hesitated to investigate "trunk-like" appendages in a number of other species, such as tapirs, saiga, and elephant shrews. Rather than getting embroiled in a contentious dispute between "trunk proponents" and "trunk skeptics" who disagree on criteria for "genuine" trunkhood, morphologists have assumed from the beginning that other appendages will be trunk-like in some respects and not others, concentrating their efforts on the development of a conceptual taxonomy adequate to the full range of similarities and differences. In this case, theorists have found it useful to describe a category of structures called "muscular hydrostats" which share a distinctive cluster of morphological and biomechanical properties (Witmer, Sampson, and Solounis 1999; Smith & Kier 1989). Why must the dialectic rather divide between two extreme positions when comparisons involve human psychological traits? As evidence piles up on both sides of comparative cognition's many interpretive standoffs, it becomes less likely that questions such as "can chimpanzees understand weight?" will be resolved with a simple affirmative or negative so much as dissolved by immersion in a much larger body of empirical data and comparative theory.

Literature Cited

Allen, C. 2002. "A skeptic's progress." *Biology and Philosophy* 17: 695–702.

Bulloch, MJ, ST Boysen, and EE Furlong. 2008. "Visual attention and its relation to knowledge states in chimpanzees, Pan troglodytes." *Animal Behaviour* 76(4): 1147–1155.

Hanus, D, and J Call. 2008. "Chimpanzees infer the location of a reward on the basis of the effect of its weight." *Current Biology* 18(8): R1–R2.

Kaminski, J, J Call, and M Tomasello. 2004. "Body orientation and face orientation: two factors controlling apes' begging behavior from humans." *Animal Cognition* 7: 216–223.

Kozhevnikov, Maria, and Mary Hegarty. 2001. "Impetus beliefs as default heuristics : Dissociation between explicit and implicit knowledge about motion." *Psychonomic Bulletin & Review* 8(3): 439–453.

Landy, D, and RL Goldstone. 2007. "How abstract is symbolic thought?" *Journal of Experimental Psychology: Leaning, Memory, and Cognition* 33(4): 720–733.

Michotte, A. 1963. *The perception of causality.* New York: Basic Books.

Povinelli, Daniel, and Timothy Eddy. 1996. "What young chimpanzees know about seeing." *Monographs of the Society for Research in Child Development* 61(3): 1–189.

Reaux, JE, LA Theall, and DJ Povinelli. 1999. "A longitudinal investigation of chimpanzees' understanding of visual perception." *Child Development* 70(2): 275–290.

Sakura, O, and T Matsuzawa. 1991. "Flexibility of Wild Chimpanzee Nut-cracking Behavior Using Stone Hammers and Anvils: an Experimental Analysis." *Ethology* 87: 237–238.

Schrauf, C, and J Call. 2011. "Great apes use weight as a cue to find hidden food." *American Journal of Primatology* 73: 323–334.

Seed, A, E Seddon, B. Greene, and J. Call. 2012. "Chimpanzee 'folk physics': bringing failures into focus." *Philosophical Transactions of the Royal Society B.* 367: 2743–2752.

Smith, K. K. & Kier, W. M. (1989). "Trunks, tongues, and tentacles: moving with skeletons of muscle." American Scientist 77: 29-35.

Whiten, A. 2001. “Tool tests challenge chimpanzees.” *Nature* 409: 133.

Witmer, L. M., S. D. Sampson, and N. Solounias. 1999. "The proboscis of tapirs (Mammalia: Perissodactyla): a case study in novel narial anatomy." Journal of Zoology, London 249:249-267.