Learning & Behavior
2005, 33 (1), 47-58

Methodological–conceptual problems in the study of chimpanzees' folk physics: How studies with adult humans can help

FRANCISCO J. SILVA, DANA M. PAGE, and KATHLEEN M. SILVA
University of Redlands, Redlands, California

In three experiments, we examined humans' folk physics (i.e., a naturally occurring and spontaneous understanding of the physical world), using variations of problems used to study chimpanzees' folk physics. Presented with trap-tube problems in two experiments, adult humans showed an unnecessary bias to insert a stick into the end of the tube farthest from the reward to push it out the other end. When presented with trap-table problems with ineffective trapping holes, people unnecessarily avoided the side with the hole. The similarity of humans' and chimpanzees' behavior on these tasks highlights methodological and conceptual problems in studies of chimpanzees' folk physics and suggests alternative explanations for their behavior.

Folk physics is a commonsense "understanding of the physical world that develops naturally and spontaneously" during the development of children and later forms adult humans' conception of how the world works (Povinelli, 2000, p. 2). Different conceptions of the world might be evident when, for example, someone chooses to pull a blanket with an object on it to retrieve the object, whereas someone else pulls a blanket with the object next to it. In the latter case, pulling the blanket does not allow one to retrieve the object.

It was 27 of these sorts of experiments that Povinelli (2000) and his collaborators conducted to elucidate not humans', but chimpanzees', folk physics. In the end, it seems that chimpanzees, unlike humans, know little if anything about phenomena that are unobservable (e.g., force or gravity) and, thus, cannot invoke abstract causal concepts to solve problems (Povinelli, 2000; Povinelli & Vonk, 2003).

Two sets of experiments that contributed to this conclusion involved the use of trap-tube and trap-table apparatuses. One type of trap-tube problem consists of a clear plastic tube with a hole (trap) at its center (see Limongelli, Boysen, & Visalberghi, 1995; Reaux & Povinelli, 2000; Visalberghi & Limongelli, 1994). A reward, such as a cookie, is placed inside the tube on either the right or the left side of the trap. To remove the reward, a chimpanzee inserts a stick tool into the end of the tube farthest from the reward and pushes it out the other end of the tube (see Figure 1A). If the chimpanzee inserts the stick into the end of the tube closest to the reward, the cookie will fall into the trap.

The problem is difficult for most chimpanzees. Combining the results obtained by Limongelli et al. (1995) and Reaux and Povinelli (2000), only 3 of 9 chimpanzees correctly performed at above-chance levels. To examine what their one successful ape (Megan) had learned, Reaux and Povinelli presented it with a test situation in which the trap was inverted (see Figure 1B). In this orientation, the trap had no effect on the reward. With the trap inverted (ineffective), Megan inserted the tool into the end of the tube farthest from the reward on 39/40 trials. Because the trap was not operational, the researchers concluded that there was no reason for Megan to avoid the trap and, thus, this animal and the other chimpanzees "did not understand how the trap functioned in the context of the causal interactions among the tool, the reward, and the trap itself" (Reaux & Povinelli, 2000, p. 131).

The trap-table problem consists of two rake tools that can be used to drag a reward along either a flat solid surface (successful retrieval) or a flat surface with a trapping hole cut into it (unsuccessful retrieval). Choosing to pull the rake on the side with the trap causes the reward to drop into the hole (see Figure 2A). In some variations of the trap-table problem, one side of the table has a hole in an ineffective position (see Figures 2B and 2C). Because the hole cannot trap the cookie, pulling either rake successfully retrieves the reward. Confronted with the trap-table problem shown in Figure 2A, only 1 of 6 chimpanzees performed at above-chance levels across all probe trials by choosing the correct rake; 3 of the 6 chimpanzees performed at above-chance levels across the last 10 trials (see Povinelli & Reaux, 2000). When presented with the problems illustrated in Figures 2B and 2C, the apes did not show a statistically significant preference for pulling one rake over the other. Although the inter-

---

We thank Armando Machado, Bill Timberlake, and two anonymous reviewers for their helpful comments and suggestions on earlier drafts of this article. Please address correspondence to F. J. Silva, Department of Psychology, University of Redlands, P.O. Box 3080, 1200 East Colton Avenue, Redlands, CA 92373-0999 (e-mail: francisco_silva@redlands.edu).

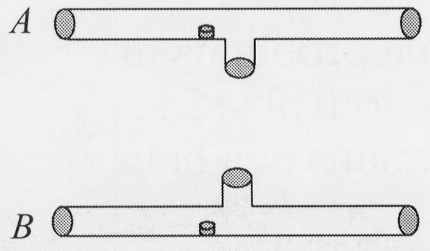

Figure 1. Trap-tube apparatus used to study folk physics in chimpanzees and monkeys in its standard configuration (A) and with the trap in an ineffective position (B). Inserting the stick tool into the end of the tube farthest from the reward successfully pushes the reward out the other end; inserting the tool into the end closest to the reward causes the reward to become trapped. Note that because the trap is equidistant from the ends of the tube, the reward is necessarily always closer to one end.

pretation of these results is not straightforward, Povinelli and Reaux's main conclusion was that because the chimpanzees did not readily and reliably solve these problems *without trial-and-error learning*, these animals appeared unable to invoke abstract concepts related to causality. However, through trial-and-error learning, some chimpanzees learned the perceptual (concrete) relations in the tasks.

As revealing as these experiments are, conclusions about chimpanzees' use of hypothetical constructs related to physical causality are premature. That chimpanzees did not readily solve some "physics problems" might mean that they did not invoke hypothetical constructs related to causality; however, their behavior could be due to countless other variables. For example, the apes' "misbehavior" during the trap-tube problem might have been the result of using a clear tube, which allowed the reward to be seen as it was being pushed (Reaux & Povinelli, 2000). For several species, a moving reward evokes behavior that interferes with performing a target response (see Breland & Breland, 1961). Also, continuous visual access to a reward might cause it to overshadow the characteristics of the trap, much as an unconditioned stimulus (US) presented simultaneously with a conditioned stimulus (CS) interferes with learning about the CS. Similarly, the ape that successfully solved the trap-tube problem but then kept inserting the stick into the end of the tube farthest from the reward when the trap was ineffective might have done so because this action resulted in the reward's traveling the least distance (see Limongelli et al., 1995). The immediacy and ease of retrieving the reward, rather than the trap's orientation, might have been the critical variable controlling Megan's behavior. Further experiments are needed to exclude these and other variables as reasons for the chimpanzees' difficulties solving the trap-tube problem (Anderson, 2001).

In addition to more experiments, the study of chimpanzees' folk physics might benefit from studies with adult humans (Anderson, 2001). How? By helping researchers identify conceptual problems in studies of folk physics (see also Machado, Lourenço, & Silva, 2000; Whiten, 2001a) and by providing information about the range of people's responses when they are faced with problems similar to those presented to chimpanzees (Anderson, 2001; Horowitz, 2003). Consider the following three examples. First, if it is illogical to conclude that similar behavior between species implies similar cognitive processes (i.e., *argument by analogy*; Povinelli & Giambrone, 2000), is it not equally erroneous to conclude that different behavior implies a different underlying cognition (see also Uttal, 2004)? Yet for some laboratories studying chimpanzees' folk physics, different behavior between species implies different cognitions (e.g., Povinelli, 2000). Second, what would it mean if adult humans behaved like chimpanzees (Anderson, 2001; Whiten, 2001b)? Given the simple tasks used with chimpanzees, it would be difficult to conclude that people who behaved similarly on these tasks did not understand the relevant abstract concepts. Third, what would it mean if adult humans behaved differently from what would be predicted by some theories of folk physics? To clarify this last point, reconsider the trap-table problem in which a chimpanzee can pull a rake on the side of the table with no hole or a rake on the side of a table with a hole in an ineffective position (e.g., Figures 2B and 2C). Because the hole cannot trap the cookie, pulling either rake retrieves the reward, and hence, if chimpanzees can invoke the theoretical concept of *gravity*, they should be indifferent as to which rake they pull. If they do not have this high-level understanding, they will avoid the rake on the side of the table with the hole (Reaux & Povinelli, 2000).

On the surface, this is a sensible conceptualization: An animal that understands gravity seemingly has no need to avoid the side with the ineffective hole. A problem with this account, though, is that an animal might also avoid the hole *because* it understands gravity. Although the position of the hole relative to the rake and cookie makes it unlikely that the hole would interfere with the retrieval of the cookie, choosing the rake on the side of the table without the hole *guarantees* that it (or more specifically, gravity) will not interfere with retrieving the reward. Thus, an animal that avoided the rake on the side of the table with the hole might very well understand gravity. In problems of this sort, the task did not permit differentiation among alternative explanations, yet the researchers formed conclusions as if differentiation were possible (Machado & Silva, 2003; see also Machado et al., 2000). The choices made by adult humans could help us interpret chimpanzees' behavior, for it would be difficult to argue that humans who showed a bias to pull one of the rakes did not understand the hypothetical causal constructs in the task.

Just as the nature of chimpanzees' folk physics cannot be assumed (Povinelli, 2000), the nature of adult humans' folk physics should not be assumed. But by not studying adult humans' folk physics, researchers have implicitly assumed how these people view and would solve the tool use problems presented to chimpanzees (Anderson, 2001; see also Whiten, 2001b). To fill this void and illustrate how the study of folk physics can benefit from

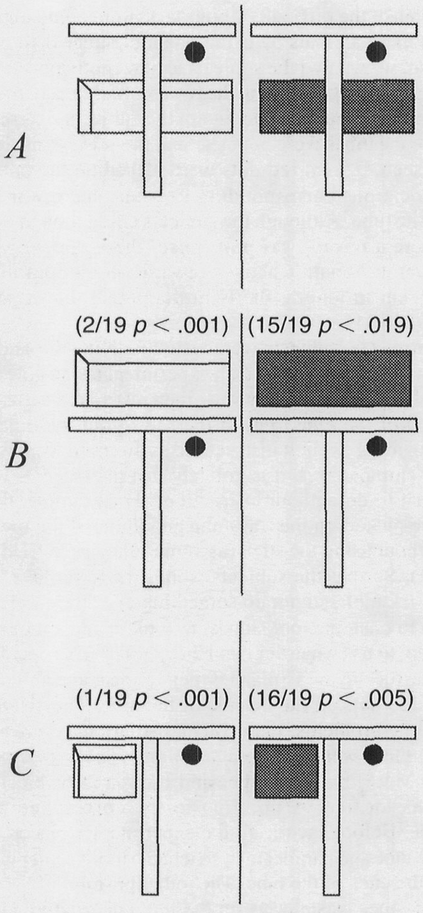

Figure 2. Various configurations of the trap-table apparatus used to study chimpanzees' folk physics. (A) A configuration in which pulling the rake on the left causes the reward to fall into the trapping hole, but pulling the rake on the right successfully retrieves the reward. In this configuration and those shown in panels B and C, the solid gray rectangle or square on the right side is simply a painted surface; it does not interfere with the retrieval of the reward. (B) A configuration in which pulling the rake on either side of the table should successfully retrieve the reward because the trapping hole on the left side is in an ineffective position above the rake and the reward. To minimize the number of figures, the data from Experiment 3 are presented above each rake, and this figure and these data are referred to in Experiment 3. The data show the number of people who chose to pull that rake and the results of a binomial test. (C) A configuration in which pulling the rake on either side of the table should successfully retrieve the reward because the trapping hole on the left side is in an ineffective position away from the path of the reward if the rake is pulled straight down. The data above each rake show the number of people who chose to pull that rake and the results of a binomial test. These data are discussed in Experiment 3. From *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works* (p. 140), by D. J. Povinelli, 2000, Oxford: Oxford University Press. Copyright 2000 by Daniel J. Povinelli. Adapted with permission.

investigations with adult humans, we presented adults with tool use problems similar to those presented to chimpanzees. In some cases, people were given actual trap-tubes, stick tools, and food rewards similar to those used with chimpanzees. In other cases, people were presented with diagrams of trap-tubes and trap-tables and were asked to make a choice that would solve the problem (i.e., obtain the reward). The general goal of these experiments was to use the behavior of adult humans to illustrate conceptual and methodological problems in studies of chimpanzees' folk physics and to show how studies with adult humans might help us identify and correct these problems.

EXPERIMENT 1

As has been summarized above, most chimpanzees cannot solve trap-tube problems. In Reaux and Povinelli's (2000) experiments, even the one ape that consistently avoided the trap later showed a strong unnecessary bias to insert the tool into the end of the tube farthest from the reward when the trap was ineffective. Reaux and Povinelli suggested two reasons for why this might be the case. First, most chimpanzees cannot keep enough causal principles in mind to solve the task. Second, the chimpanzee (Megan) that solved the task but did not seem to understand the causal principles underlying the trap could not inhibit the prepotent action of inserting the tool into the end of the tube farthest from the reward. Unfortunately, neither explanation lends itself to empirical refutation. The former is irrefutable because it is impossible to know how many causal principles a chimpanzee has in its mind at a particular moment; the latter is true by definition. A simpler testable explanation emerges from a closer examination of the trap-tube problem (see also Anderson, 2001).

In Reaux and Povinelli's (2000) experiment, the trap was equidistant from the ends of the tube, thereby ensuring that the reward was always closer to one end of the tube—even when the trap was inverted. In combination with other variables (e.g., visual access to a moving reward, overshadowing of the tube's features by the reward) or in isolation, the immediacy of the reward and/or the relative ease of moving it a shorter distance could interfere with learning about the trap. This variable seemed to control Megan's behavior. This same variable might also have controlled the behavior of the 2 successful chimpanzees in Limongelli et al. (1995). When presented with a tube where the trap was displaced from the center of the tube, 1 chimpanzee's first attempt to retrieve the reward was unsuccessful because it inserted the stick into the end of the tube farthest from the reward and pushed it into the trap. However, this was its only mistake. The second chimpanzee performed almost as well. It made only three errors in 30 trials, the earliest errors occurring on Trials 2 and 9. In another study, the strategy *insert the stick into the end of the tube farthest from the reward* was used successfully by the single capuchin monkey that reliably solved the trap-tube problem (see Visalberghi & Limongelli, 1994). When tested with a trap-tube with the trap displaced from the center, this monkey solved the problem on only 20% of the trials, because it kept inserting the stick into the end of the tube farthest from the reward.

Collectively, these results suggest that chimpanzees and capuchins that solve the trap-tube problem do so by inserting the tool into the end of the tube farthest from the reward. The strategy is unnecessary, but not detrimental, when the trap is inverted (e.g., Reaux & Povinelli, 2000). However, this strategy is unsuccessful when the trap is displaced from the center of the tube and the reward is placed on the side of the trap nearest to the center of the tube. In this configuration, pushing the reward by inserting the stick into the end of the tube farthest from the reward will cause it to fall into the trap. In the case of a capuchin, this behavior occurred repeatedly (Visalberghi & Limongelli, 1994); for 2 chimpanzees, this behavior occurred very infrequently, but it did occur early in testing (Limongelli et al., 1995). Moreover, just because these 2 chimpanzees did not use the strategy *insert the stick into the end of the tube farthest from the reward* when the trap was off-center does not mean that they did not use this strategy when the trap was centered in the tube. They could have used this strategy when the trap was centered and then, after an error or two when the trap was off-center, quickly learned another strategy.

In Experiment 1, adult humans were presented with trap-tube problems that controlled for the position of a reward relative to the ends of tubes with no traps, with traps, and with ineffective traps. According to one conceptualization of this problem (e.g., Reaux & Povinelli, 2000), there is no reason for people who understand the causal features of the task to systematically insert the tool into a particular end of a tube when a trap is ineffective. However, according to the alternative interpretation outlined above, a person who understands the causal features of the task may nevertheless insert the tool into the end of the tube that minimizes the distance necessary to push the reward out the other end. The general question this experiment sought to answer was the following. Which variables control people's behavior of inserting the tool to remove the reward: (1) the presence of a trap or traps, (2) the presence of an ineffective trap or traps, and/or (3) the distance of the reward to an end of the tube?

Method

Subjects. Ten undergraduate students of traditional college age volunteered for the study.

Setting and Apparatus. Three different tube apparatuses were used. All were constructed of 1.9-cm-diameter (0.75-in.) PVC tubing that was 51 cm long. One tube contained no traps (see Figure 3A); a second tube contained a single 1.9-cm-diameter trap 6 cm from one of the ends of the tube (see Figure 3B); and a third tube contained two 1.9-cm-diameter traps, each of which was 6 cm from an end of the tube (see Figure 3C). With the exception of the location of the trap, the overall dimensions of the apparatuses were similar to that used with chimpanzees (e.g., Reaux & Povinelli, 2000). The tubes were individually mounted onto a 5 × 11 cm foam brick partially hollowed to support the tube.

Although clear tubes are normally used with chimpanzees, we chose to use white (opaque) PVC because this would increase the difficulty of the task and because we planned to use transparent "tubes" in Experiment 2. More specifically, because the trap-tube problem is difficult for chimpanzees and capuchins to learn, we sought to make the difficulty of our task more comparable to those used with these animals by using opaque, rather than transparent, tubes. Also, we wanted the subjects to stay motivated across all trials. We thought that occluding the reward might make the task more challenging and, thus, help maintain the subjects' interest.

Because the tubes were not clear and the reward inside could not easily be seen, 0.5-cm red dots were placed on the outside of the tubes at locations corresponding to where the reward might be placed in the tube. Although the subjects could look inside the tube to see where a reward was positioned, these dots provided additional visual information about a reward's initial location. A wood dowel (55 cm in length, 0.79-cm diameter) served as the tool. Chocolate M&Ms served as the rewards.

Procedure. The subjects were tested individually and were told that they were participating in an experiment to see how people use tools to solve problems. They were then instructed to insert the tool into an end of the tube to push the M&M out the other end. Although providing explicit instructions is different from what is possible with chimpanzees, it is unlikely that the subjects would have known what to do without instructions. For example, the subjects might have picked up the trap-tube and dumped the reward out of an end without using the stick (assuming they even tried to remove the M&M). So that the subjects would try to remove the reward with the stick tool and not do something else, they were explicitly instructed to push the tool slowly, to look inside a tube as often as they wished, to use whatever hand they preferred to operate the tool, to rotate the tube in the *xy* plane if they wished, and about the meaning of the red dots on the outside of the tube. They were also told that for some problems, there was no solution, whereas for other problems, there were multiple solutions. The experimenter then placed the M&M inside the tube and showed the subject where the reward was located by pointing to the corresponding red dot on the tube. Before each trial, the experimenter placed the tool in front of the tube and parallel to it, so that the tool's center was equidistant from the ends of the tube. The initial position of the trap in the single-trap tubes was always on the left side relative to a subject.

On separate trials, an M&M was placed at three positions in the tube without any traps, at three positions in each of the single-trap and double-trap tubes when the traps were operational, and at these same positions when the traps were inverted (ineffective). Figure 4 shows the configurations that were used, grouped by common causal characteristics of the tasks. Group 1 consisted of configurations in which the reward was closer to one end of the tube and inserting the tool into either end successfully pushed out the reward. Group 2 consisted of configurations in which the reward was equidistant from both ends of the tube and inserting the tool into either end successfully pushed out the reward. Group 3 consisted of tubes with effective traps, where the reward was closer to one end

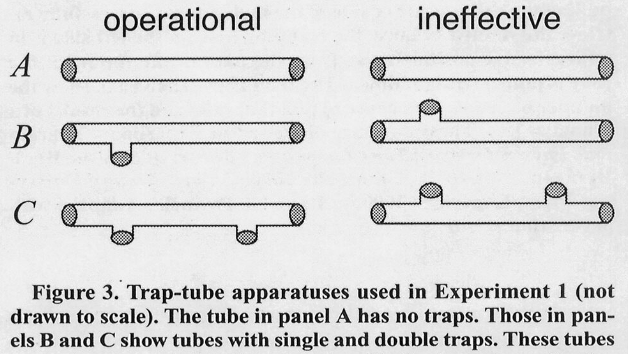

Figure 3. Trap-tube apparatuses used in Experiment 1 (not drawn to scale). The tube in panel A has no traps. Those in panels B and C show tubes with single and double traps. These tubes were also presented to the subjects with the traps in ineffective positions.

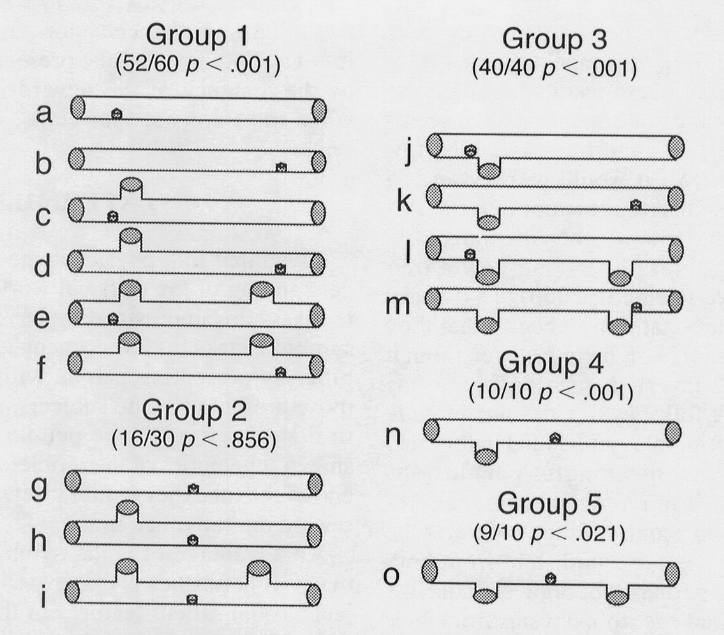

Figure 4. All configurations of the trap-tube apparatus (not drawn to scale) used in Experiment 1, grouped according to their common causal features. The data show the number of insertions of the tool into the end of the tube farthest from the reward for Groups 1 and 3, into the right end of the tubes for Groups 2 and 5, and into the end of the tube closest to the trap for Group 4. The *p* values are from a binomial test.

of the tube and inserting the tool in the end of the tube farthest from the reward successfully pushed it out. Group 4 consisted of a single-trap tube with the reward centered within the tube; inserting the tool into the end of the tube closest to the trap was necessary to push out the reward. Group 5 consisted of a double-trap tube with the reward centered within the tube; inserting the tool into either end of the tube did not push out the reward. Because the subjects could rotate the tube in the *xy* plane, the final location of the single trap was not necessarily on a subject's left side, as depicted in the figure. With reference to the configurations shown in Figure 4, 5 of the subjects were presented the configurations in the order *a*, *g*, *b*, *j*, *n*, *k*, *c*, *h*, *d*, *l*, *o*, *m*, *e*, *i*, and *f*; the other 5 subjects were presented these configurations in the reverse order. Unless otherwise noted, two-tailed binomial tests were used to analyze the results.

Results and Discussion

All 10 subjects attempted to remove the reward from all 15 positions in the three tubes. Figure 4 shows the configurations that were used, grouped by the common causal characteristics of the tasks described above, and the number of insertions of the tool into the end of the tube farthest from the reward for Groups 1 and 3, into the right end of the tubes for Groups 2 and 5, and into the end of the tube closest to the trap for Group 4. A binomial test for Group 1 shows that the subjects had a strong bias to insert the tool into the end of the tube farthest from the reward when there were no traps and ineffective traps (52/60; $p < .001$).[1] In contrast, when the reward was equidistant from the ends of tubes without traps and ineffective traps (Group 2), the subjects inserted the tool about the same number of times into each end of the tubes (16/30 chose the right end; $p < .856$). When the tool had to be inserted into the end of a tube farthest from the reward (Group 3) or nearest to the trap (Group 4), every subject made the correct response (40/40 and 10/10, respectively; *p*s $< .001$, one-tailed). When the reward was placed in the center of a double-trap tube (Group 5), the subjects showed a bias to insert the tool into the right end of the tube (9/10; $p < .021$). We will discuss reasons for this result below and will examine this effect further in Experiment 2.

Because of the theoretical significance of ineffective traps in studies of folk physics, the data for configurations *c*, *d*, *e*, and *f* were analyzed separately. On 36 out of 40 trials ($p < .001$), the tool was inserted into the end of the tube farthest from the reward. For configurations *h* and *i*, the tool was inserted equally into both ends of the tube on 10 out of 20 trials ($p = 1$).

The general question Experiment 1 was designed to answer was, Which variables control people's behavior of inserting a tool into tubes with and without traps to push out a reward? The results indicated that, unlike chimpanzees, adult humans readily insert the tool into the end of the tube that permits the reward to be retrieved without its falling into a trap. The results also indicated that the distance of a reward to the end of a tube determines into which end of a tube people insert a tool to push the reward out the other end. Specifically, the subjects inserted the tool into the end of the tube that minimized the distance that the reward traveled. This response occurred when there were tubes with no traps and ineffective traps. There was no bias to insert the tool into a particular end of a tube when the reward was equidistant from both ends. These results indicate that the distance of a reward from the end of a tube is a determinant of adult humans' behavior during trap-tube problems.

In relation to folk physics, these results highlight the difficulties of forming conclusions on the basis of seemingly critical experiments or conditions (see Shettleworth, 1998; Whiten, 2001a). It would be awkward to suggest that adult humans inserted the tool into the end of the tube farthest from the reward when the traps were inverted because they did not understand how a trap functioned (e.g., Reaux & Povinelli, 2000). That every person avoided every functional trap indicates that they understood the critical features of the traps, even though their behavior during the inverted trap configurations made it appear that their folk physics did not include knowledge of the causal principles embedded in the task. Given that we cannot draw meaningful conclusions about the nature of people's folk physics from their behavior during ineffective trap configurations of trap-tube problems when the reward is not equidistant from both ends of a tube, we should perhaps not draw conclusions about chimpanzees' or monkeys' folk physics from their performance during similar situations (e.g., Reaux & Povinelli, 2000; Visalberghi & Limongelli, 1994). A variety of configurations and test trials—ideally, with different sizes, lengths, and types of stimuli—are necessary to avoid misinterpreting people's and animals' folk physics (Visalberghi & Tomasello, 1997). Similar experiments in which the hedonic value of a reward is varied might also be informative. In studies with chimpanzees, the animals are probably well motivated to retrieve the food reward, and this motivation, in conjunction with other features of the task, might interfere with their learning the causal structure of the problem. It is unlikely that the M&Ms used in Experiment 1 were equally motivating to the subjects. The subjects' histories of solving puzzles and responding to instructions were probably the primary reasons why they attempted to solve the task. To evaluate the importance of the motivational characteristics of the reward on chimpanzees' folk physics, future experiments with these animals could, for example, compare their success in solving the trap-tube problem when foods of different palatability are used (e.g., Visalberghi & Addessi, 2000).

One unusual result concerned the subjects' bias to insert the tool into the right end of the double-trap tube when the traps were operational. In this configuration, inserting the tool into either end of the tube resulted in the trapping of the reward. Why, then, did most of the subjects insert the tool into the right end of the tube? Perhaps this result is a sample error or reflects the fact that 8 out of 10 of the subjects were right-handed and tried to solve this unsolvable problem with their preferred hand.

The outcomes of Experiment 1 show that humans, on one hand, and chimpanzees and monkeys, on the other hand, differ in their abilities to solve trap-tube problems. Exactly why these species differ is an unresolved issue (see Limongelli et al., 1995; Povinelli, 2000; Reaux & Povinelli, 2000; Shettleworth, 1998) that was not addressed in this experiment. What Experiment 1 did show was that adult humans' behavior during trap-tube problems is controlled, first, by the presence of traps and, second, by the distance of the reward from the end of a tube when the traps are ineffective and when there are no traps.

EXPERIMENT 2

Given that folk physics is the naturally occurring understanding of the physical world, it should be possible to assess this understanding in verbal creatures by using symbolic tasks. In Experiment 2, we assessed this possibility by presenting people with the same problems as those presented to the subjects in Experiment 1. Unlike in that experiment, the people in Experiment 2 were shown schematics of the problems and were asked which end of the tube they would insert the tool into to push the reward out the other end.

An advantage of using symbolic tasks to assess humans' folk physics is that experimenters can direct subjects to the salient features of the task without concern for such things as the relative importance of the diameters of the tubes and tools, whether the tube is clear or opaque, whether the reward is food or a coin, and the like. In the symbolic realm, subjects are simply asked which end of a tube they would insert a stick into to push a reward out the other end.

Experiment 2 assessed whether people would behave on a symbolic folk physics task similarly to the way they did on the real task in Experiment 1. With one exception, there were no probable reasons for why the adult humans in Experiment 2 would not behave exactly as the people did in Experiment 1. The exception concerned the subjects' behavior in Experiment 1 of inserting the tool into the right end of a double-trap tube when the reward was between the two traps. If this was a sample error or the result of most of the subjects being right-handed and trying to solve a difficult (indeed, unsolvable) problem with their preferred hand, we would not expect this result with a symbolic task in which no actual manipulation of tools and rewards are required and larger numbers of subjects are used. Experiment 2 should thus provide some information about whether handedness is expressed in the symbolic realm or whether its expression requires actual experiences with tools and trap-tube apparatuses.

Method

Subjects. Twenty-four undergraduate students of traditional college age volunteered for the study. The students were given extra credit for their participation.

Setting, Apparatus, and Procedure. The subjects were given a 15-item answer sheet with *left* or *right* next to each number. The stimuli shown in Figure 5 were displayed individually on a screen at the front of a room from an overhead projector. With reference to the letters next to each configuration in Figure 5, the stimuli were presented in the following order: *g*, *a*, *b*, *o*, *l*, *m*, *n*, *j*, *k*, *i*, *e*, *f*, *h*, *c*, and *d*.

The subjects were tested as a group and were informed that they were participating in an experiment to see how people use tools to solve problems. They were then asked to indicate into which end of

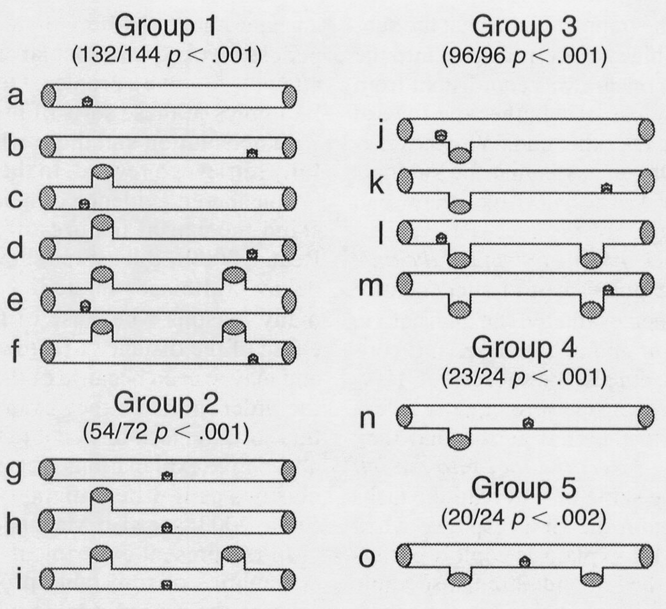

Figure 5. All configurations of the trap-tube apparatus (not drawn to scale) used in Experiment 2, grouped according to their common causal features. The data show the number of insertions of the tool into the end of the tube farthest from the reward for Groups 1 and 3, into the left end of the tubes for Groups 2 and 5, and into the end of the tube closest to the trap for Group 4. The p values are from a binomial test.

a tube, left or right, they would insert a stick to push a reward out the other end. The subjects were told that some tubes contained one or two traps and how to identify an operational versus an ineffective trap. The subjects also were told that it was physically impossible to push a reward across a trap; thus, any such attempts would cause the loss of the reward. Finally, the subjects were told that there was no solution for some problems and multiple solutions for other problems. Unless otherwise noted, two-tailed binomial tests were used to analyze the results.

Results and Discussion

Figure 5 shows the configurations that were used, grouped by the common causal characteristics of the tasks described above, and the number of insertions of the tool into the end of the tube farthest from the reward for Groups 1 and 3, into the right end of the tube for Groups 2 and 5, and into the end of the tube nearest the trap for Group 4. A binomial test for Group 1 indicated that the subjects showed a strong bias to insert the tool into the end of the tube farthest from the reward when there were no traps or there were ineffective traps (132/144; $p < .001$). In contrast, when the reward was equidistant from the ends of the tubes without a trap and with ineffective traps (Group 2), the subjects showed a bias to insert the tool into the left end of the tube (54/72; $p < .001$). When the tool had to be inserted into the end of a tube farthest from the reward (Group 3) or nearest to the trap (Group 4), in all but one instance every subject chose correctly (96/96 and 23/24, respectively; ps $<$ .001, one-tailed). When the reward was placed in the center of a double-trap tube (Group 5), the subjects generally inserted the tool into the left end of the tube (20/24; $p < .002$).

When the traps were ineffective and the reward was closer to one end of the tube (configurations *c*, *d*, *e*, and *f*), the tool was inserted into the end of the tube farthest from the reward on 88 out of 96 trials ($p < .001$). When the traps were ineffective but the reward was placed equidistant from the ends of the tube (configurations *h* and *i*), the subjects indicated that they would insert the stick into the left end on 35 out of 48 trials ($p < .003$).

Although the subjects responded to symbolic versions of the same tasks as those presented to people in Experiment 1, most of the results of Experiment 1 were replicated in Experiment 2. When asked to choose which end of a tube they would insert a stick into to push a reward out the other end, the subjects in Experiment 2 chose the end that permitted the reward to be pushed out without its falling into the trap. Also, they indicated that they would insert the tool into the end of the tube that minimized the distance that the reward traveled. This response occurred when there were no traps in a tube and when there were ineffective traps.

There were two differences between the results of Experiments 1 and 2. A significant number of the subjects in Experiment 2 indicated that they would insert the tool into the left end of tubes with no traps and those with ineffective traps when the reward was centered in the tube. A significant number of the people in Experiment 2 also indicated that they would insert the tool into the left, not the right, end of the double-trap tube to try to remove the reward positioned between the traps. It seems that the subjects in Experiment 2 had a bias to insert the tool into the left end of the tube when the reward was equidistant from the ends of the tube, regardless of whether the reward could actually be pushed out the other end. Why these results occurred is unclear. One guess is that the subjects evaluated the options from left to right like they read words on a page. They might, for instance, have evaluated the option *insert the stick into the left end of the tube* and concluded that this response would successfully push out the reward. They then evaluated the alternative, *insert the stick into the right end of the tube*, and concluded that this response would also be successful. Having determined that both responses were equally effective, the subjects stuck with the first action that they concluded would work (i.e., *insert the tool into the left end of the tube*). Perhaps the same bias to read the tasks from left to right and to retain the first response when both are equally unsuccessful explains people's preference to insert the stick into the left end of the unsolvable trap-tube. Thus, the unifying explanation for the otherwise unusual biases shown when the reward was placed equidistant from the ends of tubes with ineffective traps or no traps (Experiment 2) or equidistant from the ends of an unsolvable trap-tube (Experiments 1 and 2) is that people may be exercising their biases to read information from left to right or to use their preferred hand in attempts to solve a difficult problem. Although handedness was not expressed during the symbolic task, a serial order of evaluation bias may have occurred (cf. Guyla & Colombo, 2004). Expressing handedness seems to require an actual trap-tube apparatus and a stick tool.

These unexplained outcomes notwithstanding, Experiment 2 showed that symbolic tasks can be used to assess people's understanding of the causal features of trap-tube problems when there is an obvious right or wrong response and when the reward is closer to one end of the tube. Adult humans' behavior during trap-tube problems is controlled, first, by the presence of traps and, second, by the distance of the reward from the end of a tube when there are no traps and when traps are ineffective. Moreover, Experiment 2 shows again the difficulty of forming conclusions about a person's (or an ape's) folk physics on the basis of behavior on seemingly critical experiments or conditions. On the basis of the subjects' responses to Groups 3 and 4, we conclude that they understood the concept of gravity, but their biases in Groups 1, 2, and 5 suggest that they did not understand how a trap functions (see Reaux & Povinelli, 2000). For the reasons stated earlier, we should not accept this last conclusion.

EXPERIMENT 3

In Experiment 3, we explored another challenge in studies of folk physics—namely, the a priori inability of an experiment to differentiate whether apes, monkeys, or people invoked a particular theoretical concept in their attempts to solve a problem (see also Shettleworth, 1998). Examples of these sorts of problems include trap-tables with no solution and those where any response successfully retrieves a reward. In these circumstances, any bias is considered evidence of a subject's inability to understand the causal features of the task (see Povinelli & Reaux, 2000). But as was shown in Experiments 1 and 2, behavioral biases in studies of folk physics occur for many reasons—because of the presence of traps, because of the distance of a reward from an end of a tube, and maybe even because of the handedness of subjects or the order in which they evaluate a task. In problems of this sort, the task does not permit differentiation among alternative explanations, yet the researchers formed conclusions as if differentiation was possible (Machado & Silva, 2003; see also Machado et al., 2000).

In the present experiment, we presented people with schematics of trap-table problems in which pulling a rake on the side of a table with a hole in an ineffective position retrieved the reward; pulling the other rake also retrieved the reward, because there was no hole at all on that side of the table (see Figures 2B and 2C). Presented with this and related trap-table problems, chimpanzees learn to be indifferent as to which rake they pull (Povinelli & Reaux, 2000). This suggested to Povinelli and Reaux that chimpanzees can learn the perceptual (concrete) relations of the task.

It is also potentially true that a bias to pull the rake on the side of the table without the hole could be construed as confirmatory evidence of an ape's or a person's understanding of the abstract causal features of the task. Although the position of the hole relative to the rake and the reward makes it unlikely that the hole would interfere with the retrieval of the reward, choosing the rake on the side of the table without the hole *guarantees* that gravity will not interfere with retrieving the reward. Viewed in this manner, the strongest evidence for an understanding of how the trap functioned might result from a bias to pull the rake on the side of the table without a trap, because this action guarantees that the reward will not be trapped. In Experiment 3, we sought to assess which rake people would pull when given a choice to pull a rake on the side of the table with no hole or a rake on the side with the hole in a seemingly ineffective position (see Figures 2B and 2C).

Method

Subjects. Nineteen undergraduates of traditional college age participated in this experiment. The students were given extra credit for their participation.

Setting, Apparatus, and Procedure. The subjects were tested individually. Everyone was shown the schematics illustrated in Figures 2B and 2C, was told how to recognize the holes and how to recognize the painted shape on the tables, was asked to write which rake, *left* or *right*, of each pair he or she would pull in an attempt to retrieve the reward, and was asked to provide in writing a single-sentence explanation for his or her choice. Two-tailed binomial tests were used to analyze the results.

Results and Discussion

For the tasks in Figures 2B and 2C, 15 out of 19 and 16 out of 19 ($ps < .02$ and .005) people, respectively, indicated that they would pull the rake on the side of the table with no hole; 2 out of 19 indicated they would pull either rake. An analysis of the reasons the subjects provided for their choices showed that 15 out of 19 (Figure 2B) and 16 out of 19 (Figure 2C) indicated that pulling the rake on the side without a hole increased the likelihood that the reward would not be lost.

These results show that a bias to pull a rake on the side of a trap-table without a hole should not be considered evidence that a subject did not understand the causal features of the task (e.g., Povinelli & Reaux, 2000). To do so would necessitate arguing that the people in this experiment did not understand the physics embedded in the task. Such an argument would be inappropriate, especially given the subjects' self-reported reasons for their choices, which indicate clearly that they understood the critical features of the task. Thus, the use of humans in this experiment helped identify a conceptual–methodological difficulty in similar studies of chimpanzees' folk physics and suggested an alternative explanation for a bias to avoid the rake on the side of the table with the ineffective trap—namely, that humans want to avoid the danger of a reward's falling into a trap.

GENERAL DISCUSSION

The constellation of results from the three experiments presented here show that adult humans will display behavioral biases in tool use tasks similar in causal structure to those presented to chimpanzees. Because many critical tests in studies of chimpanzees' folk physics predict indifference, a behavioral bias is considered evidence of an animal's inability to understand the causal features of a task. But it would require a degree of incredulity to suggest that the people in the experiments presented here could not accurately represent the critical causal features of the tasks they were asked to solve (see also Shettleworth, 1998). Given that the subjects easily avoided pushing a reward into a trap and justified their choices to avoid rakes on the side of a table with a hole by invoking a physical property of a hole (i.e., things can fall into it), we can conclude that the adults in all three experiments understood the causal principles embedded in the tasks. Their choices during configurations with ineffective traps only made it appear that they did not understand the principles. Thus, this study highlights a pitfall of forming conclusions about chimpanzees' folk physics on the basis of similar biases: Behavioral biases during tool use experiments have causes that may reveal little about the nature of a subject's folk physics.

In addition, the present study illustrates the need for careful conceptual analyses in studies of chimpanzees' folk physics. A closer examination of the premises and conceptual scaffolding of some well-known studies of chimpanzees' folk physics reveals a variety of errors (see Machado et al., 2000): experiments that could not distinguish a priori between a subject who understood versus one who did not understand hypothetical causal features (e.g., a chimpanzee might avoid pulling a rake on the side of a table with a hole in an effective position because it understands gravity, not because it does not), explanations that violated Morgan's canon (e.g., behavioral biases occur because "chimpanzees do not invoke a priori theoretical concepts [such as gravity] to mediate their use of tools" [Reaux & Povinelli, 2000, p. 131], rather than because of the distance of the reward from an end of the tube), irrefutable explanations (e.g., the number of causal principles in a chimpanzee's mind), explanations that are true by definition (e.g., a repetitive action occurred because a chimpanzee could not inhibit this prepotent action), and logical inconsistencies (e.g., arguing that similar behavior among species does not imply a similar underlying cognition, but concluding that dissimilar behavior implies a different underlying cognition), to name a few.

The conceptual analyses presented in this article and the results of the experiments that emerged from these analyses show how studies of adult humans' folk physics can help identify conceptual and methodological problems in studies of chimpanzees' folk physics. Despite this contribution, a determined critic may argue that comparisons between the behavioral biases shown by the people in this study and chimpanzees in other studies of folk physics are inappropriate. First, the tasks used by the people in the present study and those normally used with chimpanzees are different. For example, chimpanzees are usually asked to push out a reward from a transparent, not an opaque, tube. Perhaps people would have behaved differently than the chimpanzees had they been asked to solve a trap-tube problem with a clear tube. Also, perhaps the instructions given to the subjects affected their behavior. Finally, the use of symbolic tasks might be relatively uninformative, because similar tasks cannot be used with chimpanzees and monkeys.

For a variety of reasons, these and similar criticisms do not undermine the fundamental contributions of the present study. First, it is hard to conceive of a plausible explanation for why someone's behavior during the trap-tube problem would differ if the tube was transparent. Besides, the subjects could and did routinely look inside the tubes while pushing out the rewards. Second, the schematics of the tubes used in Experiment 2 were neither explicitly opaque nor clear. If anything, the tubes were seemingly transparent. Third, investigators of chimpanzees' folk physics have themselves argued that invoking the clarity of the tube as a significant causal variable is an example of "hyper-naturalism" and "beside the main point" (Reaux & Povinelli, 2000, p. 131) of comparative studies of folk physics. Fourth, the combined results of Experiments 1 and 2 showed that it is appropriate to assess humans' folk physics with symbolic tasks, because the subjects in these experiments solved the trap-tube problems similarly. Although it is possible that the subjects would have behaved differently on actual trap-table problems, instead of the symbolic ones used in Experiment 3, it is hard to imagine why they would, given their justifications for their choices. Their main justification—that avoiding the rake on the side with the hole guaranteed that the reward would be successfully retrieved—would almost certainly be invoked to solve actual trap-table problems. Finally, with regard to the role of instruction, there is no doubt that the subjects' behavior was partially controlled by the experimenter's instructions, for they would not know what to do otherwise. Critically, though, the instructions did not suggest *how* they should solve a problem.

Another reason for questioning the appropriateness of comparing the behavior of people in the present study with chimpanzees' behavior relates to the single-trial method used here, rather than the multiple-trial method used in studies of apes' and monkeys' folk physics (e.g., Hauser, Pearson, & Seelig, 2002; Povinelli, 2000). In these studies, an animal's behavior is recorded across many trials. It is possible, then, that a person who is allowed multiple opportunities to solve a problem might do so in different ways and converge on a response bias different from his or her initial response. Although this is possible, it is also relatively unimportant because, in the study of folk physics, an animal's first response during test conditions is the critical response (see Limongelli et al., 1995; Povinelli, 2000). Ideally, there would not even be any pretraining, because such training undermines the study of a subject's *naturally occurring* understanding of the physical world (Povinelli, 2000). Thus, a person's or a chimpanzee's first response during a physics problem, preferably without any pretraining, is the gold standard for assessing folk physics (Chappell & Kacelnik, 2002; Povinelli, 2000).

Finally, in a related point, primatologists may object that the present study shows only how adult humans solve some tool use problems, but not how chimpanzees or monkeys solve similar problems. For example, perhaps the distance of a reward from the end of a tube is not the critical determinant of chimpanzees' behavior during ineffective trap conditions. Although this is possible, the analyses of Limongelli et al.'s (1995), Reaux and Povinelli's (2000), and Visalberghi and Limongelli's (1994) studies presented above are highly suggestive that this is an important variable. More significant, perhaps, is that the present experiments and the analyses suggest specific experiments that can be conducted with nonhuman primates to investigate the generality of the analyses and effects reported here. Most important, the present study shows that the more serious problems in studies of chimpanzees' folk physics are conceptual and relate to the meaning of different outcomes. More experiments alone will not solve these problems (Machado et al., 2000; see also Povinelli & Vonk, 2003).

It is important to emphasize that the outcomes of the experiments presented here should not lead one to conclude that humans' folk physics is different from or similar to chimpanzees' folk physics. Our goal was simply to draw attention to the methodological and conceptual problems in studies of chimpanzees' folk physics and to show how studying adult humans' folk physics may, in some instances, elucidate conceptual–methodological infelicities, help interpret results, and suggest alternative explanations. Just as the nature of chimpanzees' folk physics cannot be assumed (Povinelli, 2000), the nature of adult humans' folk physics should not be assumed. The study of folk physics can benefit from information about how people view and solve the tool use problems presented to chimpanzees (Anderson, 2001). For similar reasons, the study of primates' folk physics could also benefit from investigations of other animals' folk physics. An ongoing study in our lab of the behavior of domestic cats during a trap-table problem reveals a cost of studying primates exclusively. Presented with a modified trap-table problem similar to the one in Figure 2A, 1 cat never responded correctly above chance levels after more than 200 trials. This cat was insensitive to the presence of an effective trapping hole on one side of the trap-table. A 2nd cat, though, chose to retrieve the food on the side of the table without the hole on almost 100% of the trials after the first 10 trials, results that are much better than those obtained with chimpanzees. Later, when given test trials in which the food was placed in front of the hole (i.e., the trap was ineffective; see Figure 2B), this cat distributed its behavior evenly across both sides of the table. This cat's behavior suggests that it understood the abstract causal principles embedded in the trap-table problem (see Povinelli & Reaux, 2000). Additional experiments showed that covering the food with a piece of cloth increased the 1st cat's correct responses.

Following the reasoning advanced in some studies of chimpanzees' folk physics, these preliminary results with cats suggest that they might have a better understanding of the causal features of a physics problem than chimpanzees do. This seems odd. It is more likely then that there is something wrong with the premises that led to this conclusion. In this regard, research carried out to investigate nonprimates' folk physics may also contribute to the study of chimpanzees' folk physics (e.g., Chappell & Kacelnik, 2002; Tebbich & Bshary, 2004; Weir, Chappell, & Kacelnik, 2002).

The recurring message of this article is that studying folk physics in adult humans contributes to research into and the theory of chimpanzees' folk physics. In addition, the present study highlights two absent but important elements in studies of folk physics: (1) an awareness that what an experimenter determines to be the causal features of the task may be very different from what a subject perceives to be the causal features of the task and (2) a need for experiments in which one type of causal understanding predicts a specific response and another type of understanding (or lack thereof) predicts another specific response. With regard to the first deficiency, an awareness of a discrepancy between an animal-centered and an experimenter-centered view of the causal structure of the task is of potential importance for the conclusions drawn (Horowitz, 2003; Shettleworth, 1998). The results of all three experiments showed that the subjects were sensitive to more than an experimenter-centered view of the causal structure of the tasks. The smallest change in a task can produce a large change in what are considered by a subject to be the critical causal features of a problem (Horowitz, 2003). That is, a small change in a task might make an animal think like a human or vice versa. Viewed this way, null effects in studies of folk physics are relatively uninformative, because the absence of an effect may be due to many variables, only one of which relates to a subject's naturally occurring understanding of causality (Machado & Silva, 2003). The issue cannot be dismissed by claiming that changing a task to make a chimpanzee behave like a human is beside the point in studies of folk physics and then asserting that the point *is* that chimpanzees do not behave like humans on a particular task (e.g., Povinelli, 2000). This is wrong, for effects or their absence do not exist independently of the means used to study them. The task itself is not trivial or theoretically unimportant (Timberlake, 1990, 2001, 2002). What would it mean, for example, if chimpanzees readily behaved like adult humans and solved the trap-tube problem if an opaque rather than a transparent tube was used, if a shorter rather than a longer tube was used, if a token rather than a primary reward was used, and so on? What should we conclude then about the nature of chimpanzees' and humans' folk physics?

With regard to the second deficiency, a great deal of theoretical significance is attached to null effects. Typically, in studies of folk physics, one explanation is pitted against another, and these are supposed to manifest themselves behaviorally as a specific effect for one type of understanding and as a null effect (indifference) for another type of understanding. In most cases, the experiments are construed in such a way that an understanding of hypothetical causal variables should lead to indifference during a particular test and an inability to understand these variables should produce a particular behavioral bias. But as philosophers of science and statisticians remind us, predicting null effects is a weaker test of a hypothesis than is predicting a specific behavioral outcome. The results of Experiment 3 further remind us that, contrary to some initial conceptualizations, null effects may sometimes tell us little in studies of folk physics. In Experiment 3, a bias to pull the rake on the side of the table without any hole—not indifference to which rake was pulled—was the best indicator that the subjects understood the critical causal features of the task. Until such tests are common, the value of additional experiments is debatable.

REFERENCES

Anderson, J. R. (2001). [Review of the book *Folk physics for apes: The chimpanzee's theory of how the world works*]. *Animal Behaviour*, **61**, 1042-1043.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, **16**, 681-684.

Chappell, J., & Kacelnik, A. (2002). Tool selectivity in a non-primate, the New Caledonian crow (*Corvus moneduloides*). *Animal Cognition*, **5**, 71-78.

Guyla, M., & Colombo, M. (2004). The ontogeny of serial-order behavior in humans (*Homo sapiens*): Representation of a list. *Journal of Comparative Psychology*, **118**, 71-81.

Hauser, M., Pearson, H., & Seelig, D. (2002). Ontogeny of tool use in cottontop tamarins, *Saguinus oedipus*: Innate recognition of functionally relevant features. *Animal Behaviour*, **64**, 299-311.

Horowitz, A. C. (2003). Do humans ape? Or do apes human? Imitation and intention in humans (*Homo sapiens*) and other animals. *Journal of Comparative Psychology*, **117**, 325-336.

Limongelli, L., Boysen, S. T., & Visalberghi, E. (1995). Comprehension of cause–effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, **109**, 18-26.

Machado, A., Lourenço, O., & Silva, F. J. (2000). Facts, concepts, and theories: The shape of psychology's epistemic triangle. *Behavior & Philosophy*, **28**, 1-40.

Machado, A., & Silva, F. J. (2003). You can lead an ape to a tool, but . . . : A review of Povinelli's *Folk physics for apes: The chimpanzee's theory of how the world works*. *Journal of the Experimental Analysis of Behavior*, **79**, 267-286.

Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: Oxford University Press.

Povinelli, D. J., & Giambrone, S. (2000). Escaping the argument by analogy. In D. J. Povinelli, *Folk physics for apes: The chimpanzee's theory of how the world works* (pp. 9-72). Oxford: Oxford University Press.

Povinelli, D. J., & Reaux, J. E. (2000). The table-trap problem. In D. J. Povinelli, *Folk physics for apes: The chimpanzee's theory of how the world works* (pp. 132-148). Oxford: Oxford University Press.

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, **7**, 157-160.

Reaux, J. E., & Povinelli, D. J. (2000). The tube-trap problem. In D. J. Povinelli, *Folk physics for apes: The chimpanzee's theory of how the world works* (pp. 108-131). Oxford: Oxford University Press.

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. New York: Oxford University Press.

Tebbich, S., & Bshary, R. (2004). Cognitive abilities related to tool use in the woodpecker finch, *Cactospiza pallida*. *Animal Behaviour*, **67**, 689-697.

Timberlake, W. (1990). Natural learning in laboratory paradigms. In D. A. Dewsbury (Ed.), *Contemporary issues in comparative psychology* (pp. 31-54). Sunderland, MA: Sinauer.

Timberlake, W. (2001). Integrating niche-related and general process approaches in the study of learning. *Behavioural Processes*, **54**, 79-94.

Timberlake, W. (2002). Niche-related learning in laboratory paradigms: The case of maze behavior in Norway rats. *Behavioural Brain Research*, **134**, 355-374.

Uttal, W. R. (2004). Hypothetical high-level cognitive functions cannot be localized in the brain: Another argument for a revitalized behaviorism. *Behavior Analyst*, **27**, 1-6.

Visalberghi, E., & Addessi, E. (2000). Response to changes in food palatability in tufted capuchin monkeys, *Cebus apella*. *Animal Behaviour*, **59**, 231-238.

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause–effect relations in tool-using capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, **108**, 15-22.

Visalberghi, E., & Tomasello, M. (1997). Primate causal understanding in the physical and social domains. *Behavioural Processes*, **42**, 189-203.

Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science*, **297**, 981.

WHITEN, A. (2001a). Theory of mind in non-verbal apes: Conceptual issues and the critical experiments. In D. M. Walsh (Ed.), *Naturalism, evolution, and mind* (Royal Institute of Philosophy Supplement, No. 49, pp. 199-233). Cambridge: Cambridge University Press.

WHITEN, A. (2001b). Tool tests challenge chimpanzees: Do apes fail to understand physics, or do we fail to understand them? *Nature*, **409**, 133.

NOTE

1. Because these analyses consisted of data contributed by subjects from multiple trials, someone's choice on trial $x + 1$ might not be independent of his or her choice on trial x. However, the same could be true of any person or machine that flips a coin. Flip $x + 1$ may not be independent of flip x. Although there is no certainty that the agent that makes a choice or flips a coin is unbiased from trial to trial, binomial tests are routinely used to analyze choices made by the same subject across several trials in studies of tool use (e.g., Povinelli, 2000; Tebbich & Bshary, 2004), and this test is appropriate when there are only two possible outcomes and we can reasonably assume that trial $x + 1$ is independent of trial x. Alternatively, a one-sample t test could also have been used.

(Manuscript received June 16, 2004;
revision accepted for publication September 13, 2004.)