Learn Behav (2014) 42:337–347
DOI 10.3758/s13420-014-0151-0

Adult humans' understanding of support relations: an up-linkage replication

Francisco J. Silva · Merritt I. ten Hope · Ali L. Tucker

Published online: 23 August 2014

Abstract In an up-linkage replication, three experiments examined adult humans' folk physics, i.e., their naturally occurring and spontaneous understanding of the physical world, using a violation of expectation (VOE) task and stimuli similar to those used to study chimpanzees', monkeys', and rooks' folk physics. Unlike what has been reported with nonhuman primates, adult humans did not look longer at physically impossible than possible events, though they did rate the physically impossible events as more interesting and novel than the possible events. These results underscore that behavior during a VOE experiment has many possible causes, only one of which may be a subject's folk physics.

Keywords Up-linkage · Folk physics · Support relations · Looking time response · Humans

Introduction

When chimpanzees were shown objects in physically possible versus physically impossible support relations, they spent more time looking at the impossible relations (e.g., Cacchione & Krist, 2004). This was true when shown video clips of "contact/no-contact" (e.g., a banana on a box vs. a banana floating in the air near a box) and "amount of contact" (e.g., a banana that is entirely supported by a box vs. a banana that is only partially supported by a box) relations, but not "type of contact" (e.g., an apple supported by a box vs. an apple floating above a box but in contact with the side of another box) (see Figs. 2, 4, and 6 in Cacchione & Krist, 2004). These results led Cacchione and Krist to conclude that chimpanzees have some expectations about support relations that are "based on the perceived contact relation between two objects, both qualitatively (contact vs. no contact) and quantitatively (sufficient vs. insufficient amount of contact) but not on information about the objects' absolute orientation in space" (p. 147). A subsequent study with chimpanzees and Japanese monkeys corroborated Cacchione and Krist's results and their interpretations (Murai, Tanaka, & Sakagami, 2011). In contrast, studies with human infants suggest that they have an understanding of simple "contact/no-contact" support relations by 3 months of age, "type of contact" by about 4.5 to 5.5 months of age, and "amount of contact" by about 6.5 months of age (Baillargeon, 2002).

The dependent measure on which Cacchione and Krist (2004), Murai et al. (2011), and Baillargeon and her colleagues (e.g., Baillargeon & Hanko-Summers, 1990; Baillargeon, Needham, & DeVos, 1992; Hespos & Baillargeon, 2008; Needham & Baillargeon, 1990) based their conclusions was the subjects' looking time. This measure was obtained using a violation-of-expectation (VOE) task, which involves presenting subjects with events that violate everyday perceptions of physical laws versus events that conform to these laws. The former are referred to as "unexpected," "surprising," or "impossible" events; the latter are referred to as "expected," "unsurprising," or "possible" events. The rationale underlying this task is that people, chimpanzees, monkeys, and other animals have an intuitive understanding of their physical worlds, i.e., naïve or folk physics (Povinelli, 2000). When subjects detect an event that violates or contradicts their folk physics, they are interested in or surprised by this unexpected occurrence and thus look longer at it than at an event that does not violate their expectations (Murai et al., 2011; Wang, Baillargeon, & Brueckner, 2004). Increased looking times indicate that subjects possessed the expectation under investigation, detected the violation in the unexpected event, and were surprised by the violation, where the term

F. J. Silva (✉) · M. I. ten Hope · A. L. Tucker
Department of Psychology, University of Redlands, P.O. Box 3080, 1200 East Colton Avenue, Redlands, CA 92373, USA
e-mail: francisco_silva@redlands.edu

surprise is a label for "a state of heightened attention or interest caused by an expectation violation" (Wang et al., 2004, p. 168). The causal sequence is thus violation of expectation → increased interest → increased looking times.

The interpretation of differences in people's and other animals' looking times has been the subject of considerable debate, at the center of which is whether these differences are best explained by perceptual habituation/dishabituation or by violation of an expectation (e.g., Bogartz, Shinskey, & Speaker, 1997; Cashon & Cohen, 2000; Clearfield & Westfahl, 2006; Cohen & Marks, 2002; Schilling, 2000; Schöner & Thelen, 2006; but see Wang et al., 2004). In particular, the VOE task relies on presenting subjects with familiarization events (i.e., a "dress rehearsal" consisting of stimuli used in the experiment) and then presenting physically possible versus impossible events (or, for example, in other studies, mathematically correct vs. incorrect events; e.g., Wynn, 1992). Because the familiarization events are more similar to the physically possible than impossible events, a subject's looking response is dishabituated by the more dissimilar stimulus (i.e., the impossible event), resulting in longer looking times at this stimulus. In studies with nonhumans, familiarization trials are kept to a minimum to prevent subjects from becoming bored with the task and to reduce habituation to the familiarization events, which by contrast may create a bias to look longer at the perceptually novel impossible event (Murai et al., 2011). Whether differences in the looking response are the result of methodological (e.g., presence, type, and amount of familiarization trials) or subject variables (e.g., physical intuition) is an ongoing debate.

In the present study, we begin with an up-linkage replication (e.g., Hachiga, Silberberg, Parker, & Sakagami, 2009; Silberberg et al., 2013) of studies that examined what nonhuman animals understand about support relations (e.g., Bird & Emery, 2009; Cacchione & Krist, 2004; Murai et al., 2011). The basis for an up-linkage replication is as follows. A researcher attempts to demonstrate the degree of cognitive continuity between a nonhuman animal, often a monkey or an ape, and adult humans. But because monkeys, apes, and rooks cannot speak, researchers cannot use verbal reports to ascertain what these subjects know. As a result, what these subjects understand about a phenomenon is inferred from their behavior on a nonverbal task. Absent from this approach is data about how adult humans behave on a similar task. In lieu of data, researchers assume how adult humans would behave and then discuss the implications of the nonhuman animals' behavior as if actual data of people's performance existed (see e.g., Reaux & Povinelli, 2000). If an up-linkage replication with adult humans produces results that are inconsistent with interpretations of other animals' behavior on similar tasks, these interpretations are jeopardized (see Silva, Page, & Silva, 2005; Silva & Silva, 2006).

Experiment 1

Chimpanzees and monkeys looked longer at videos that showed physically impossible support relations than at physically possible support relations when the examples consisted of "contact/no-contact" and "amount of contact," but not "type of contact" (e.g., Cacchione & Krist, 2004; Murai et al., 2011). When shown photographs depicting these same object relations, rooks spent more time looking at physically impossible than possible events (Bird & Emery, 2009). In Experiment 1, we used a VOE task with test trials only to measure how long adult humans looked at photographs of physically possible versus impossible events. Unlike in studies with human infants, chimpanzees, and monkeys (e.g., Baillargeon & Hanko-Summers, 1990; Cacchione & Krist, 2004; Murai et al., 2011; but see Wang et al., 2004), we did not use any familiarization trials and thus avoided the pitfalls of using these trials. Also, using adult human subjects allowed the subjects themselves to indicate how long they looked at an image instead of having observers and coders determine if and for how long a subject looked at a stimulus (which is what is done when infants and nonhumans are the subjects).

Method

Subjects Twenty-four students enrolled in Introduction to Psychology at a small liberal arts university in southern California volunteered for the study in partial fulfillment of the course's requirements.

Setting, apparatus, and stimuli The subjects were run individually in a laboratory space equipped with a desk and a computer. An iMac computer with a 21.5-in (51.6-cm) screen was used to present six images (see Fig. 1). The images in panels *a* and *b* show possible and impossible physical support relations of the type "contact/no-contact," panels *c* and *d* show possible and impossible relations of "type of contact," and panels *e* and *f* show possible and impossible relations of "amount of contact." Each image filled the screen and was presented once, randomly, within a block of the six images. Each block was presented three times; thus, the experiment consisted of 18 trials. Millisecond resolution of stimulus presentations and data recording were controlled by OpenSesame (Mathôt, Schreij, & Theeuwes, 2012).

Procedure and data analysis The experiment began with subjects sitting at the computer and reading the instructions on the screen: "You will be presented with several images; some may be shown more than once. Look at each image for as long as you wish. To look at another image, press the space bar twice. To load the first image, press any key." The duration that each image was presented was a subject's looking time for that presentation. A 1-second white screen was used as a

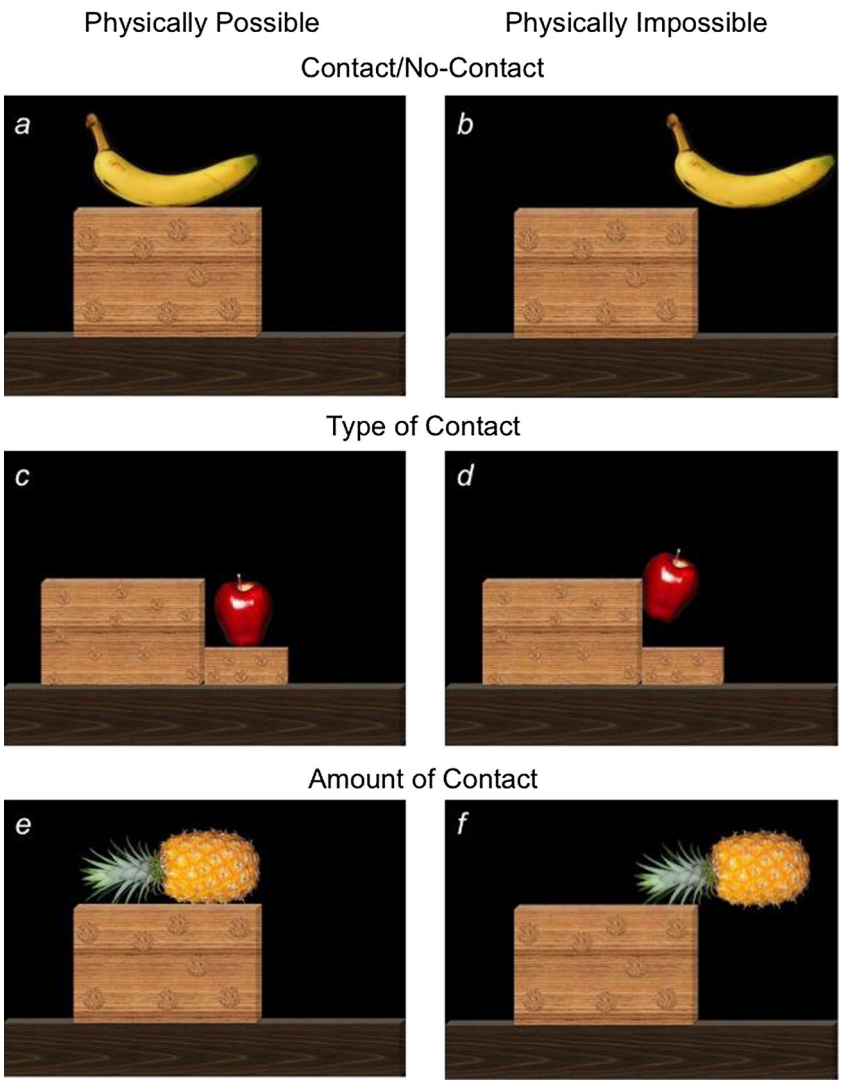

Fig. 1 Stimuli presented in Experiments 1, 2, and 3

distinct cue separating stimulus presentations. The duration that a subject looked at an image was normalized to the subject's longest looking time such that a relative looking time of 1.0 represented the longest duration that a subject looked at an image. Statistical significance was set to $\alpha = .05$. Bonferroni corrections were used for planned comparisons.

Results and discussion

Figure 2 shows the subjects' normalized looking times to each stimulus across blocks of trials. An analysis of variance (ANOVA) with three within-subjects factors (*type of event*: physically possible or impossible; *type of object relation*: contact/no-contact, type of contact, amount of contact; and *block*) showed main effects of object relation [$F(2, 46) = 5.27$, $p = .009$, $\eta_p^2 = .19$] and block [$F(2, 46) = 51.86$, $p < .001$, $\eta_p^2 = .69$], but no main effect of event-type [$F(1, 23) = 3.65$, $p = .07$]. There were no statistically significant interactions. Tukey's Honestly Significant Difference (HSD) test

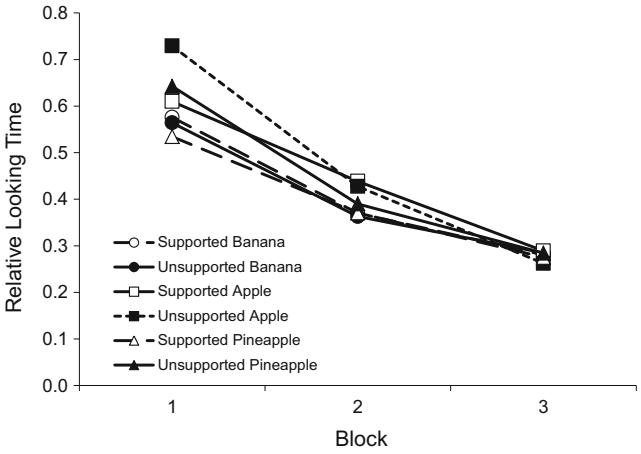

Fig. 2 Mean relative looking times for each stimulus across blocks of trials. As shown in Fig. 1, bananas were used in "contact/no-contact" object relations, apples were used in "type of contact" relations, and pineapples were used in "amount of contact" relations

confirmed that the subjects spent more time looking at images of "type of contact" (M = 0.46, SD = 0.28) than "contact/no-contact" (M = 0.41, SD = 0.25) (p = .01) and "amount of contact" (M = 0.42, SD = 0.26) (p = .04) and that looking times decreased across blocks of trials (ps < .004).

To determine if there were differences in how long subjects looked at the physically possible versus impossible events within an object relation category (i.e., contact/no-contact, type of contact, amount of contact), we conducted planned comparisons on the pairs of images within each category for Block 1, the block where the looking response could not be influenced by prior presentations of the same image. None of the comparisons were statistically significant [ts(23) < 1.62, ps > .11]. Figure 3 shows a sample of individual subjects' data. The most notable feature is a decrease in looking times with repeated presentations of the same image.

The purpose of Experiment 1 was to conduct an up-linkage replication (e.g., Hachiga et al., 2009; Silberberg et al., 2013) to see whether adult humans behaved similarly to chimpanzees, monkeys, and rooks. Based on monkeys' and

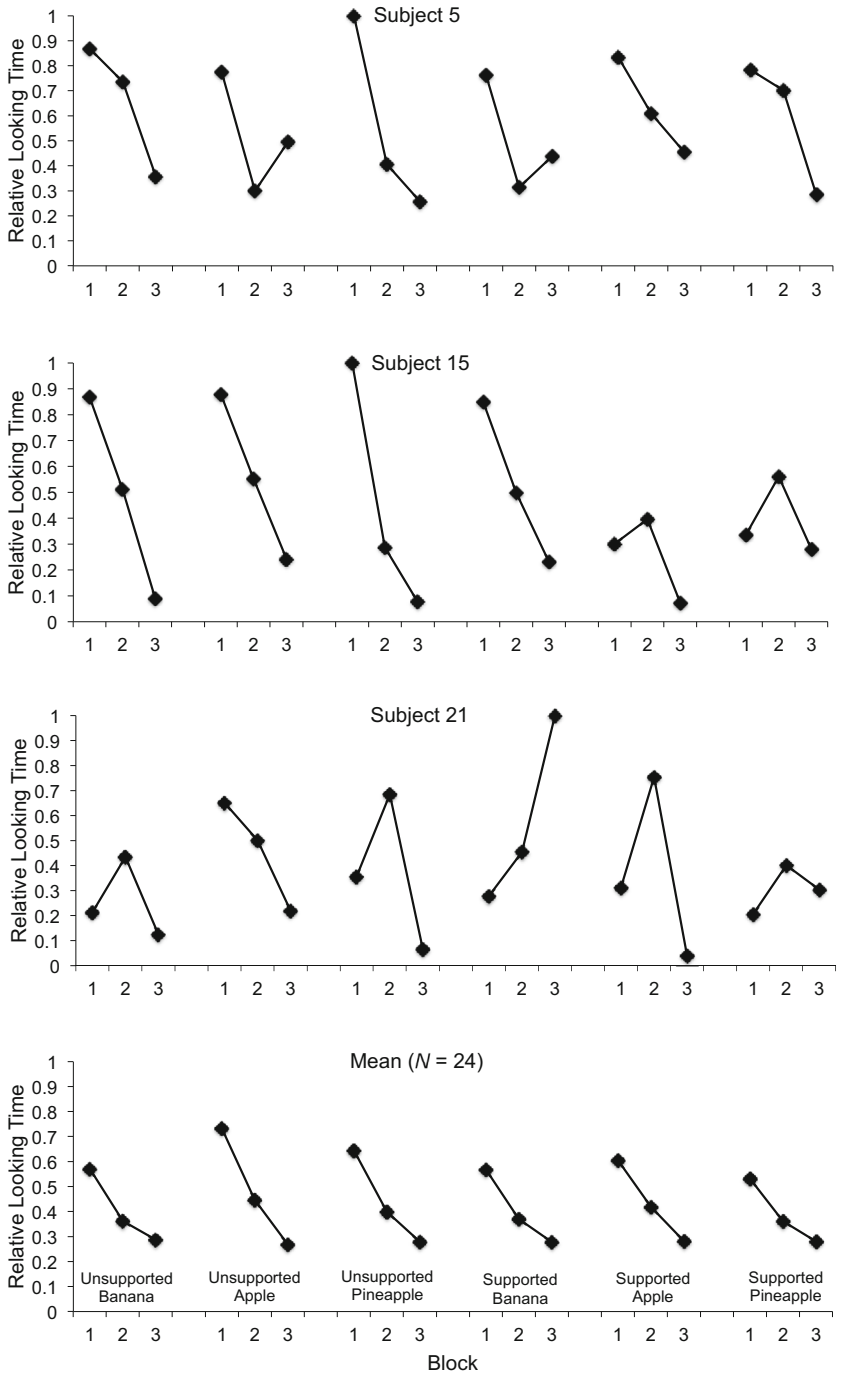

Fig. 3 Sample of individual subjects' data and the group data similarly formatted

chimpanzees' looking times during VOE tasks, researchers concluded that these animals possess intuitions about the support relations "contact/no-contact" and "amount of contact," but not "type of contact" (e.g., Cacchione & Krist, 2004; Murai et al., 2011). In contrast, because the participants in Experiment 1 did not look longer at the physically impossible than possible events, it seems that adult humans do not have an intuitive understanding of the aforementioned support relations. In this case, absence of evidence is not evidence of absence, for it is unlikely that the adult humans in Experiment 1 did not understand that unsupported objects would fall. Even human infants seem to understand these support relations (see Baillargeon, 2002). This means that the results of Experiment 1 were due to something other than a lack of understanding related to support relations.

As with any null effects, there are many possible reasons why the participants did not look longer at the physically impossible than possible events. Perhaps the lack of familiarization trials contributed to the results of Experiment 1, though exactly how is unclear because the rationale for familiarization trials in a VOE task is ambiguous. For example, Murai et al. (2011) stated that familiarization trials serve to "acquaint the participants with all the materials used for the following test events" and to "inform them that the object was obviously detached from the platform" (pp. 218-219). Unfortunately, the phrase "acquaint the participants" is vague; it neither specifies the processes (e.g., habituation, extinction) involved or the responses (e.g., orienting, looking) affected. Equally vague is what it means for human researchers to inform monkeys and chimpanzees and why it is necessary to inform them that an object is obviously detached from a platform before testing if they will look longer at objects detached from a platform. Moreover, a study that did not use familiarization trials showed that rooks spent more time looking at physically impossible than possible events (see Bird & Emery, 2009). Although it is possible that the lack of familiarization trials in Experiment 1 contributed to the difference in the results observed with adult humans versus monkeys and chimpanzees, it should be noted that each stimulus was presented three times. Thus, familiarization trials of a sort (e.g., Block 1) were a part of Experiment 1.

A second explanation concerns the stimuli used in Experiment 1 versus those used in similar studies with monkeys and chimpanzees. The stimuli in Experiment 1 consisted of static photographs. In studies with nonhuman primates, the stimuli consisted of dynamic video clips (e.g., Cacchione & Krist, 2004) where an agent (e.g., an experimenter's hand) interacts with an object (e.g., a banana) to produce a physically possible (e.g., placing a banana on a box) or impossible outcome (e.g., letting go of a banana and then having it float in space). Perhaps the use of a series of photographs or a video might make the impossible outcome more surprising in the sense that a common action produced an uncommon outcome. Or perhaps people might look longer at a video of a physically impossible event in an attempt to see how the impossible event was accomplished. The likelihood of these explanations being correct, though, is lessened by the observation that rooks spent more time looking at static images of physically impossible than possible events (see Bird & Emery, 2009). Ultimately, the effect of static photographic versus dynamic video presentations of possible and impossible events on people and other animals' looking times can only be resolved empirically.

A third explanation is that looking is not controlled by a violation of an expectation that evokes interest, but rather that looking times are affected by different aspects of a stimulus such as its novelty, complexity, abstractness, size, color, or many other dimensions (Bogartz et al., 1997). Thus, when nonhuman primates look longer at a physically impossible event, it could be because these events are more novel than the physically possible events. Monkeys and chimpanzees, even laboratory ones, have seen objects resting on other objects many times, but have not seen physically impossible events such as a banana floating above a table. Rather than these animals' folk physics, monkeys' and chimpanzees' looking times during a VOE task could be a function of the novelty of the stimuli. Novel stimuli elicit more orienting (looking) than more familiar stimuli (Pavlov, 1927). Stating that a novel stimulus that depicts a physically impossible event elicits longer looking times because it elicited a violation of an expectation does not add much to the observation that novel stimuli elicit more orienting. In Experiment 1, perhaps the images of impossible events were not novel enough in comparison to their possible counterparts. We examine this possibility, along with the possibility that the stimuli were not differentially interesting to produce different looking times, in the next two experiments.

Experiment 2

Whether because of the stimuli's characteristics (e.g., novelty, complexity) or because a violation of the expectation that unsupported objects will fall did not occur, the images in Experiment 1 did not appear to elicit different amounts of interest, which may be why looking times at the physically possible and impossible events were similar. Regardless of the exact reasons for heightened interest, e.g., violation of an expectation, novelty or complexity of a stimulus, the VOE task assumes that differences in looking times reflect differences in subjects' interest. In Experiment 2, we asked adult humans to rate the six images used in Experiment 1 in terms of how interesting they were. We chose to focus on the images' interest because researchers have hypothesized that interest is the critical factor in generating the looking response during

VOE tasks (see Murai et al., 2011; Wang et al., 2004). If the results of Experiment 1 occurred because the stimuli did not elicit different amounts of interest, then people's ratings of how interesting they found the stimuli should produce a pattern of results similar to those of Experiment 1.

Method

Subjects, setting, apparatus, and procedure Twenty students enrolled at a small liberal arts university in southern California volunteered for the study. They were shown the images in Fig. 1 and asked to rate each image in terms of how interesting it was on a scale from 0 (*least interesting*) to 10 (*most interesting*). To promote comparison among stimuli, all six images were projected simultaneously (as shown in Fig. 1 but without the categorical labels) on a screen at the front of a classroom. Also, presenting all images simultaneously prevented the rating of any image from being influenced by a prior image (i.e., sequence effects of the ratings).

Results and discussion

Figure 4 shows how interesting the subjects found the six images. An ANOVA with two within-subjects factors (*type of event*: physically possible or impossible; *type of object relation*: contact/no-contact, type of contact, amount of contact) showed that subjects rated the impossible events as more interesting than the possible events [$F(1, 19) = 17.17$, $p < .001$, $\eta_p^2 = .48$]. Tukey's HSD test showed that this was true for each pair of stimuli within each type of event ($ps < .001$). The ANOVA also showed that subjects rated the object relations differently [$F(2, 38) = 3.82$, $p = .03$, $\eta_p^2 = .17$]. Tukey's HSD test revealed that subjects rated the "amount of contact" images ($M = 5.18$, $SD = 2.55$) as more interesting than the "contact/no-contact" images ($M = 4.38$, $SD = 2.70$) ($p = .02$) but not the "type of contact" images ($M = 4.85$, $SD = 2.39$) ($p = .51$). These latter two events did not differ ($p = .24$). There was no interaction between event-type and object relation [$F(2, 38) = 0.45$, $p = .64$].

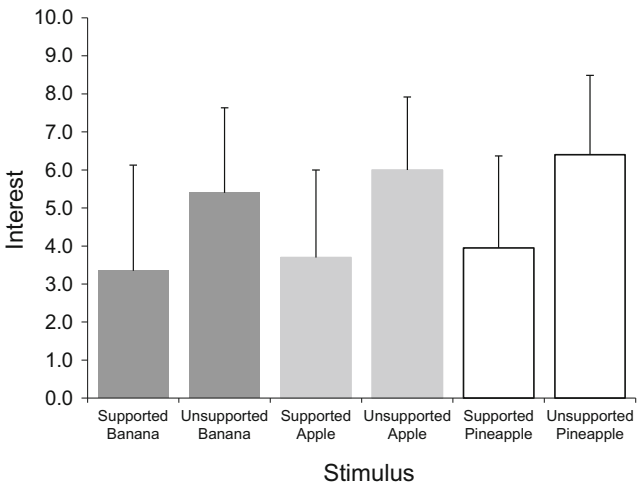

Fig. 4 Mean interest rating for each image. Error bars show standard deviations. Bananas were used in "contact/no-contact" object relations, apples were used in "type of contact" relations, and pineapples were used in "amount of contact" relations

If the interest an image elicits is the main reason for looking times during a VOE task (see Murai et al., 2011; Wang et al., 2004), then the results of Experiment 2 should have been similar to those of Experiment 1. But the results of the two experiments differed. Although subjects in Experiment 2 rated the impossible events as more interesting than the possible events, there were no differences in the looking times during these two types of events in Experiment 1. Also, although subjects in Experiment 2 rated the "amount of contact" images as more interesting than the "contact/no-contact" images but not the "type of contact" images, the subjects in Experiment 1 spent more time looking at images of "type of contact" than "contact/no-contact" and "amount of contact."

It is important that the subjects in Experiment 2 discriminated between the impossible and possible images even though these consisted of static photographs. If static photographs were sufficient to produce discrimination between physically possible and impossible events in Experiment 2, then these same images should have been sufficient to produce differences in looking times in Experiment 1 if the conceptual foundations of the VOE task are correct. Of course, procedural differences between the two experiments may explain the different results. The images in Experiment 1 were presented randomly and serially, and required subjects to produce the next image by pressing the space bar. The images in Experiment 2 were presented simultaneously in a single configuration by a researcher. Focusing on procedural differences, though, raises the question: What are the rules for translating a violation of an expectation into an overt response (e.g., looking at a stimulus)?

There is little in the rationale of the VOE task that answers this question. If a subject possesses the expectation under investigation (e.g., that an unsuspended object should fall or otherwise not hang mysteriously in the air), then the presentation of a physically impossible event will violate the expectation if the subject detects the violation. If each of the preceding occurs, the violation of the expectation elicits surprise, which produces increased looking at the impossible event relative to a comparable event that did not violate a subject's expectation. The strength of the violation is measured by the looking time such that the more that a subject is surprised, the longer it will look at the event. Because surprise is a label for heightened attention or interest (Wang et al., 2004), surprise and heightened interest are manifestly the same things.

Given the conceptual foundation of the VOE task, we could assume that the subjects in Experiment 1 did not possess the expectation under investigation, did not notice anything to elicit a violation of an expectation, or were not surprised by

the violation. The first reason is unlikely, for it is hard to imagine that adult human subjects did not possess an expectation that unsupported objects fall. The second reason states what must necessarily be true: A violation of an expectation can only happen when a subject notices something to cause a violation of an expectation. The third reason, that a subject will look longer at an event that violated an expectation only if it is surprised by the violation, is untestable in nonverbal subjects. The only evidence for surprise is an increase in looking times, which is the same evidence used to determine that a violation of an expectation occurred (Mix, 2002). Without an increase in looking times, there is no evidence of an expectation, a violation of an expectation, or surprise. Whether separately or in combination, these three reasons directly related to the conceptual foundation of the VOE task do not clarify the results of Experiment 1 and thus do not clarify why Experiments 1 and 2 produced different patterns of results.

Experiment 3

Experiment 3 examined whether the results of Experiment 1 reflect differences in the novelty of the images. More novel stimuli elicit more stimulus-directed responding (orienting and looking). The effects of stimulus novelty were evident in Experiment 1 in that the subjects spent less time looking at the stimuli across blocks. Might the results of Experiment 1 have occurred because the stimuli were equally novel? If so, then people's ratings of how novel they found the stimuli should produce a pattern of results consistent with those of Experiment 1.

Method

Subjects, setting, apparatus, and procedure Fourteen students enrolled at a small liberal arts university in southern California volunteered for the study. They were shown the images in Fig. 1 and asked to rate each image in terms of how novel it was on a scale from 0 (*least novel*) to 10 (*most novel*). To promote comparison among stimuli, all six images were projected simultaneously (as shown in Fig. 1 but without the categorical labels) on a screen at the front of a classroom.

Results and discussion

Figure 5 shows how novel the subjects found the six images. An ANOVA with two within-subjects factors (event-type and object relation) showed that subjects rated the impossible events as more novel than the possible events [$F(1, 13) = 181.75$, $p < .001$, $\eta_p^2 = .93$]. Tukey's HSD test showed that this was true for each pair of stimuli within each object relation ($ps < .001$). The ANOVA also showed that the subjects rated the different object relations similarly in terms of their novelty [$F(2, 26) = 1.87$, $p = .17$] and that there was no interaction between event-type and type of object relation [$F(2, 26) = 0.13$, $p = .88$].

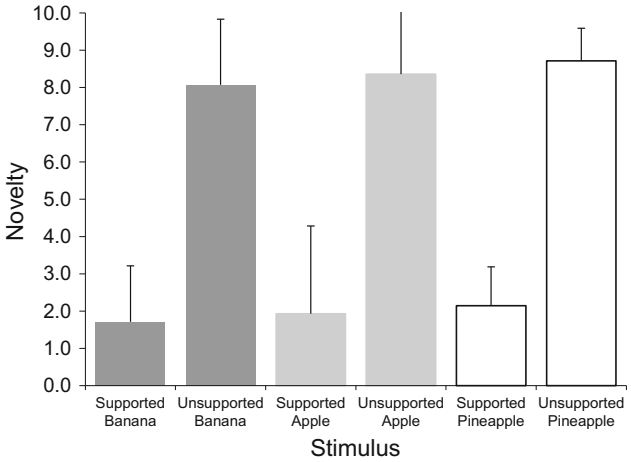

Fig. 5 Mean novelty rating for each stimulus. Details are the same as for Fig. 4

If the novelty of an image is the key reason for looking times during a VOE task, then the results of Experiment 3 should have been similar to those of Experiment 1. But the results of these two experiments differed. Although subjects in Experiment 3 rated the impossible events as more novel than the possible events, there were no differences in the looking times during these two types of events in Experiment 1. Also, although all of the object relations in Experiment 3 were rated as equally novel, subjects in Experiment 1 spent more time looking at images of "type of contact" than "contact/no-contact" and "amount of contact." Other than the decrease in looking times across blocks in Experiment 1, it does not appear that a stimulus' novelty contributed significantly to the results of that experiment.

General discussion

The purpose of this study was to examine whether adult humans, like chimpanzees and monkeys in particular but also rooks and infants, would look longer at images of physically impossible than possible events. Unlike many studies that use VOE tasks, we did not use familiarization trials that complicate the interpretation of subsequent looking responses (Clearfield & Westfahl, 2006; Schilling, 2000). In contrast to nonhuman primates, adult humans did not spend more time looking at physically impossible than possible events during a VOE task. However, adult humans did rate the impossible events as more interesting (Experiment 2) and novel (Experiment 3) than the possible events. This was true regardless of the type of object relation. Discrepancies between how

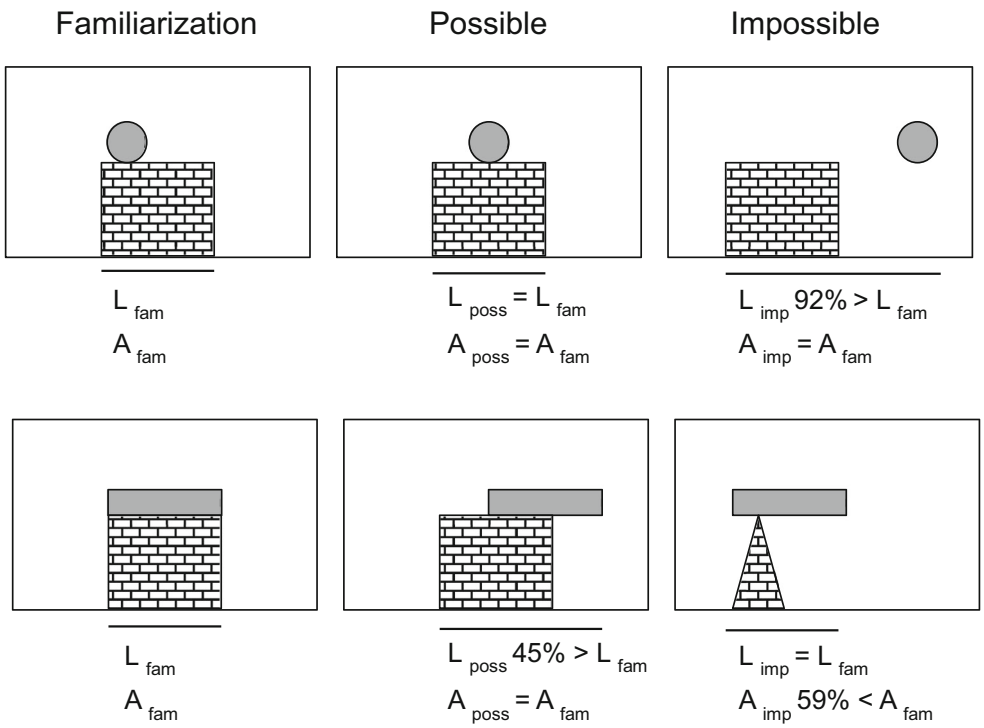

Fig. 6 Schematic illustrating the lengths (L) and areas (A) of the stimulus configurations in different events: familiarization (fam), physically possible (poss), and physically impossible (imp). The horizontal line below each panel shows how the lengths of the configurations were measured from the leftmost object to the rightmost point of the rightmost object. The total area occupied by a configuration was obtained by summing the areas of all objects in an event

interesting and novel subjects rated the stimuli (Experiments 2 and 3) and how long subjects looked at the images (Experiment 1) suggest that the key factors controlling looking times were not interest or novelty.

The results of the three experiments reported here do not prove that monkeys' and chimpanzees' folk physics do not consist of an understanding of the support relations "contact/no-contact" and "amount of contact" (see Cacchione & Krist, 2004; Murai et al., 2011). What the present study does is question whether these animals' ability to discriminate between some types of object relations warrants a rich interpretation of their folk physics (Charles & Rivera, 2009; Haith, 1998; Mix, 2002). The soundness of using VOE tasks to infer what nonverbal subjects understand about physical causality is not improved by stating that other researchers make equivalent interpretations (e.g., Cacchione & Krist, 2004). Paraphrasing Mix (2002), despite common interpretations of behavior during VOE tasks, the conceptual foundations of these tasks are built on shifting sand.

The conclusion that apes and monkeys possess similar intuitions about support that differ in part from those of humans (Cacchione & Krist, 2004; Murai et al., 2011) is not based on comparisons of these animals' behavior to adult humans' behavior on VOE tasks. The conclusion is based on nonhuman primates' behavior (looking) during VOE tasks in comparison to the researchers' views that different looking times between physically impossible and possible events mean that apes' and monkeys' folk physics is similar to humans'. Ideally, nonhuman primates' behavior should be compared to adult humans' behavior on a VOE task. Before researchers studying nonhuman animals can make strong claims about evolutionary continuity, they need to know how adult humans behave on a similar task. Just as chimpanzees' folk physics should not be assumed (Povinelli, 2000), people's folk physics should not be assumed. Adult humans' behavior is surprisingly chimpanzee-like on tasks used to study these animals' folk physics (Silva et al., 2005; Silva & Silva, 2006). Failure to produce in an up-linkage study the result that was up to then assumed to occur in adult humans raises questions about the validity of the interpretation of the nonhuman animal's behavior (e.g., Silberberg et al., 2013).

Using adult human subjects and a VOE task, we did not obtain results similar to those reported with nonhuman primates – despite evidence that adult humans viewed the physically impossible events as more interesting and novel than the physically possible events. The problem is that although studies with nonhuman primates show that subjects discriminated between stimuli, the reasons for the discrimination are unknown. Stimulus discrimination may be caused by something

Table 1 Difference Between the Lengths of the Stimulus Configurations in the Possible versus Impossible Events in Cacchione and Krist (2004)

Experiment	Object Relation	% Difference
1	Contact/No-Contact	36.5
2	Type of Contact	0
3	Amount of Contact	27.8

Note. Positive values indicate that the stimulus configuration for the impossible event was longer than the configuration for the possible event.

Table 2 Differences Between the Lengths of the Stimulus Configurations in the Possible versus Impossible Events and Familiarization versus Test Trials in Murai et al. (2011)

Experiment & Stimulus	Object Relation	% Difference					
		Length			Area		
		Possible vs. Impossible	Familiarization vs. Possible	Familiarization vs. Impossible	Possible vs. Impossible	Familiarization vs. Possible	Familiarization vs. Impossible
Exp 1, Set 1 vs. Set 3	Contact/No-Contact	40.5	-2.6	36.8	0	0	0
Exp 1, Set 2 vs. Set 4	Contact/No-Contact	37.9	-5.1	37.9	0	0	0
Exp 2, Set 1 vs. Set 3	Type of Contact	0	-24.8	-23.5	-9.7	0	0
Exp 2, Set 2 vs. Set 4	Type of Contact	0	-30.0	-29.1	-9.5	0	0
Exp 3, Set 1 vs. Set 5	Amount of Contact	-8.5	-9.6	-18.9	-16.2	-28.3	-39.9
Exp 3, Set 2 vs. Set 6	Amount of Contact	3.2	-11.3	-9.3	-20.4	-33.2	-46.9
Exp 3, Set 3 vs. Set 7	Amount of Contact	-6.4	-16.1	-20.0	-31.2	-19.4	-44.9
Exp 3, Set 4 vs. Set 8	Amount of Contact	-10.3	-11.8	-23.0	-39.8	-23.3	-53.8

Note. Positive values indicate that the stimulus configuration for the impossible event was longer or occupied more area than the configuration for the possible event, or that the configuration during a test trial was longer or occupied more area than the configuration during the familiarization trial. The dashed lines enclose the largest differences between the "Familiarization vs. Possible" and the "Familiarization vs. Impossible" configurations.

other than a subject's folk physics. For example, the length of the stimulus configuration (as measured along a horizontal line from the leftmost point to the rightmost point; see examples in Fig. 6) in the test events in Figs. 2, 4, and 6 of Cacchione and Krist (2004) shows that the impossible test events were about 36.5 % and 27.8 % longer than the possible test events for the "contact/no-contact" and "amount of contact" object relations, respectively, but not the "type of contact" relation (see Table 1).[1] Recall that Cacchione and Krist's chimpanzees looked longer at the physically impossible than possible events of the "contact/no-contact" and "amount of contact" object relations, but not the "type of contact" relation. The differences in looking times between possible and impossible events may have been due to differences in the lengths of the stimulus configurations that comprise these two events. Or, could the key be that the impossible events are more dissimilar to the familiarization events than the possible events are to the familiarization events? This question cannot be answered from Cacchione and Krist's article because they did not provide a photograph of the stimuli used in the familiarization trials; however, Murai et al. (2011) did provide photographs of all familiarization events and physically possible and impossible events used in their study.

Table 2 shows the differences between the lengths of the stimulus configurations in the physically possible versus impossible events and between the familiarization and test trials in Murai et al. (2011). Here, too, the impossible test events were substantially longer (40.5 % and 37.9 %) than the possible test events for the "contact/no-contact" (Experiment 1), but not "type of contact" object relation (Experiment 2) as shown in the last frame of video – the one presented for half of the duration of a trial – in Murai et al.'s Figs. 1, 4, and 5. More importantly, a comparison of the stimulus configurations in the familiarization trials with their corresponding test trials shows that the lengths of the configurations in the possible test events were more similar to those in the familiarization trials than the impossible test events were to the familiarization trials for the "contact/no-contact" object relations and "amount of contact" relations in all stimulus sets except Set 2 versus Set 6 and Set 3 versus Set 7 (Experiment 3). That is, the percent-difference between familiarization and possible events (2.6 % and 5.1 %) was less than the difference between familiarization and impossible events (36.8 % and 37.9 %) for the "contact/no-contact" object relation. The same pattern was true for the "amount of contact" relation, except for Set 2 versus Set 6 and Set 3 versus Set 7, the differences of which are within 5 % of each other. For the "type of contact" object relation, the differences between the lengths of the stimulus configurations in the familiarization events versus physically possible events were similar to the differences between the lengths of the configurations in the familiarization events versus the physically impossible events (24.8 % vs. 23.5 % and 30 % vs. 29.1 %). In no instance was the familiarization event more similar to the physically impossible than to the possible event in terms of the length of the stimulus configurations.

[1] The lengths of the stimulus configurations were obtained by enlarging images of Figs. 2, 4, and 6 from Cacchione and Krist (2004) and then measuring the lengths of the configurations as illustrated in Fig. 6 of the present article. These lengths did not include the image of the experimenter's arm because an arm was not present in the last (2-second) frame of the video clips used by Murai et al. (2011), which we analyzed to produce the results in Table 2. More specifically, we analyzed the lengths of the stimulus configurations in the last frame of Figs. 1, 4, and 5 from Murai et al. and we measured the area occupied by the objects in these frames. For irregularly shaped objects such as pears and bell peppers, we estimated their areas using a rectangle. Because events containing pears were never compared to events containing bell peppers, for example, any error resulting from this estimation was irrelevant to the analyses of the similarities and differences between physically possible and impossible events and between familiarization trials and test trials. Erring on the side of caution, we considered percentages within 5 % to be equivalent.

A similar analysis of the area occupied by the objects in the last frame of the video clips showed that the impossible test events consisted of stimulus configurations that occupied substantially less area than the configurations in the possible events for the "amount of contact" object relation. More importantly, relative to the area occupied by the objects in the familiarization trials, the area occupied by the stimuli in the corresponding impossible events varied from 39.9 % to 53.8 % less – differences that are substantially larger than those between the familiarization trials and the physically possible test events (19.4 % to 33.2 %).

The preceding analysis suggests that the pattern of stimulus discrimination shown by chimpanzees and monkeys in two studies that used VOE tasks to study these animals' folk physics could have occurred because the impossible events were more dissimilar (in terms of the lengths of the stimulus configurations in these events or the area occupied by the configurations) than the possible events were to the familiarization events. When there were no substantial differences in the lengths or areas of the stimulus configurations between the familiarization trials and the physically possible and impossible events, looking times at these events were undifferentiated. Rather than concluding that monkeys' and chimpanzees' folk physics lack an understanding of "type of contact" (see Cacchione & Krist, 2004; Murai et al., 2011), these animals' failure to discriminate between impossible and possible events may have occurred because the stimulus configurations in these events were similar in two salient dimensions (i.e., length and area) to the familiarization events.

A conceptual analysis (see Machado, Lourenço, & Silva, 2000; Machado & Silva, 2007) of the VOE task and researchers' interpretations of subjects' looking times is important because studies such as Cacchione and Krist's (2004) are the basis for subsequent studies of what animals understand about support relations (e.g., Murai et al., 2011). One such study concluded that rooks have a better understanding of support relations than chimpanzees because these birds spent more time looking at physically impossible than possible support relations of "type of contact" (Bird & Emery, 2009). But consider that the two images of physically impossible "type of contact" events used in Bird and Emery's study were an average of 20.5 % longer than the images that showed physically possible events (see Bird & Emery, 2009, Fig. 1, Experiment 2 stimuli). Could it be that the rooks spent more time looking at the "type of contact" physically impossible stimulus configurations because these were longer than the physically possible configurations instead of because the birds understood the difference between physically impossible versus possible support relations? Even the control experiment (Experiment 4) consisted of impossible configurations that were an average of 14 % longer than the possible configurations (see Bird & Emery, 2009, Fig. 1, Experiment 4 stimuli).

Behavior during studies of folk physics occurs for many reasons, only one of which may be a subject's folk physics (Silva et al., 2005; Silva & Silva, 2006). Whatever the reasons for looking times during a VOE task, the results of the present study and the analyses above suggest that researchers should be cautious about ascribing sophisticated cognitive understanding to nonhuman primates in the absence of studies of the rules underlying looking (Haith, 1998; Mix, 2002), studies of the habituation of the looking response (e.g., Bogartz et al., 1997; Kaplan & Wener, 1986; Schöner & Thelen, 2006), and studies that use VOE tasks with adult humans.

Author Note We thank Armando Machado and Kathleen Silva for their feedback and assistance, and two anonymous reviewers for their helpful suggestions, particularly with regard to up-linkage replications.

References

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 47–83). Oxford, England: Blackwell.

Baillargeon, R., & Hanko-Summers, S. (1990). Is the top object adequately supported by the bottom object? Young infants' understanding of support relations. *Cognitive Development, 5,* 29–53. doi:10.1016/0885-2014(90)90011-H

Baillargeon, R., Needham, A., & DeVos, J. (1992). The development of young infants' intuitions about support. *Early Development and Parenting, 2,* 69–78. doi:10.1002/edp.2430010203

Bird, C. D., & Emery, N. J. (2009). Rooks perceive support relations similar to six-month-old babies. *Proceedings of the Royal Society B: Biological Sciences, 277,* 147–151. doi:10.1098/rspb.2009.1456

Bogartz, R. S., Shinskey, J. L., & Speaker, C. J. (1997). Interpreting infant looking: The event set × event set design. *Developmental Psychology, 33,* 408–422. doi:10.1037/0012-1649.33.3.408

Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: Intuitions about support in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 118,* 140–148. doi:10.1037/0735-7036.118.2.140

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy, 1,* 403–428. doi:10.1207/S15327078IN0104_4

Charles, E. P., & Rivera, S. M. (2009). Object permanence and method of disappearance: Looking measures further contradict reaching measures. *Developmental Science, 12,* 991–1006. doi:10.1111/j.1467-7687.2009.00844.x

Clearfield, M., & Westfahl, S. M.-C. (2006). Familiarization in infants' perception of addition problems. *Journal of Cognition and Development, 7,* 27–43. doi:10.1207/s15327647jcd0701_2

Cohen, L. B., & Marks, K. S. (2002). How infants process addition and subtraction events. *Developmental Science, 5,* 186–212. doi:10.1111/1467-7687.00220

Hachiga, Y., Silberberg, A., Parker, S., & Sakagami, T. (2009). Humans (*Homo sapiens*) fail to show an inequity effect in an "up-linkage" analog of the monkey inequity test. *Animal Cognition, 12,* 359–367. doi:10.1007/s10071-008-0195-7

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior & Development, 21,* 167–179. doi:10.1016/S0163-6383(98)90001-7

Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition, 107,* 304–316. doi:10.1016/j.cognition.2007.07.009

Kaplan, P. S., & Werner, J. S. (1986). Habituation, response to novelty, and dishabituation in human infants: Tests of a dual-process theory of visual attention. *Journal of Experimental Child Psychology, 42,* 199–217. doi:10.1016/0022-0965(86)90023-8

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

Machado, A., & Silva, F. J. (2007). Toward a richer view of the scientific method: The role of conceptual analysis. *American Psychologist, 62,* 671–681. doi:10.1037/0003-066X.62.7.671

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44,* 314–324. doi:10.3758/s13428-011-0168-7

Mix, K. (2002). Trying to build on shifting sand: Commentary on Cohen and Marks. *Developmental Science, 5,* 205–206. doi:10.1111/1467-7687.00221_2

Murai, C., Tanaka, M., & Sakagami, M. (2011). Physical intuitions about support relations in monkeys (*Macaca fuscata*) and apes (*Pan troglodytes*). *Journal of Comparative Psychology, 125,* 216–226. doi:10.1037/a0022099

Needham, A., & Baillargeon, R. (1990). Intuitions about support in 4.5-month-old infants. *Cognition, 47,* 121–148. doi:10.1016/0010-0277(93)90002-D

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Translated and edited by G. V. Anrep. London: Oxford University Press.

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

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). New York, USA: Oxford University Press.

Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy, 1,* 389–402. doi:10.1207/S15327078IN0104_2

Schöner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. *Psychological Review, 113,* 273–299. doi:10.1037/0033-295x.113.2.273

Silberberg, A., Parker, S., Allouch, C., Fabos, M., Holberman, H., McDonald, L., … Wyatt, L. (2013). Human risky choice in a repeated-gambles procedure: An up-linkage replication of Lakshminarayanan, Chen and Santos (2011). *Animal Cognition, 16,* 907–914. doi: 10.1007/s10071-013-0623-1

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. *Learning & Behavior, 33,* 47–58. doi:10.3758/bf03196049

Silva, F. J., & Silva, K. M. (2006). Humans' folk physics is not enough to explain variations in their tool-using behavior. *Psychonomic Bulletin & Review, 13,* 689–693. doi:10.3758/bf03193982

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition, 93,* 167–198. doi:10.1016/j.cognition.2003.09.012

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749–750. doi:10.1038/358749a0