



# What does an intermediate success rate mean? An analysis of a Piagetian liquid conservation task in the great apes

Chikako Suda\*, Josep Call

*Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig D-04103, Germany*

Received 28 January 2004; revised 27 October 2004; accepted 27 January 2005

---

## Abstract

The study investigates what an intermediate success rate means in bonobos, chimpanzees, and orangutans. Apes participated in liquid conservation experiments where they had to track the larger of two different quantities of juice after various kinds of transformations [Suda, C., & Call, J. (2004). Piagetian liquid conservation in the great apes (*Pan paniscus*, *Pan troglodytes*, and *Pongo pygmaeus*). *Journal of Comparative Psychology*, 118, 265–279]. When making a decision, apes sometimes demonstrated hesitant behavior, concurrently pointing to both alternatives or successively changing their choice. Moderately successful apes showed more hesitation than highly successful or unsuccessful apes. The results are consistent with the cognitive conflict model: The experiments created a higher degree of cognitive conflict on moderately successful apes than on very successful or unsuccessful apes. This indicates that an intermediate performance reflects the joint operation and potential conflict between two different cognitive strategies (identity and appearance) inherent to the Piagetian conservation task.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Hesitation; Cognitive conflict; Intermediate success

---

Animals often demonstrate intermediate success in a variety of cognitive tasks. For instance, when confronted with two alternatives in a choice situation (50% correct by chance) many studies often report a performance around 75% correct in things like the use of experimenter cues, analogical reasoning, or numerical competence

---

\* Corresponding author. Tel.: +49 341 3550 407; fax: +49 341 3550 444.

*E-mail address:* [suda@eva.mpg.de](mailto:suda@eva.mpg.de) (C. Suda).

(e.g. Beran, 2001; Call, Hare, & Tomasello, 1998; Vonk, 2003). Although subjects are clearly above chance in these studies, the interpretation of the results is less than straightforward because their performance is still far from perfect, and therefore it is unclear to what extent subjects have mastered the task. One possible interpretation of such results is that subjects are distracted or uninterested for a sizeable number of trials. Another possibility, certainly a more interesting one from a cognitive point of view, is that this intermediate performance reflects the joint operation and potential conflict between different cognitive strategies. For instance, in an object choice task where an experimenter demonstrates social cues, subjects have to not only associate the direction of a given cue with the reward but also overcome the spatial bias generated by finding food on a particular location in the previous trial.

One way to diagnose which of these two previous possibilities is more accurate consists of relating performance to behavioral measures of hesitation. This is based on the theory that cognitive conflict prompts hesitation, which is a manifestation of a secondary decision-making system as opposed to automatic responses (James, 1890/1981; Smith, Shields, & Washburn, 2003; Tolman, 1938, 1932/1967). According to this theory, metacognitive animals should show more hesitant behavior as their internal mental conflict intensifies. There is some data on nonhuman animals supporting this idea. Smith et al. (1995) used a threshold discrimination task and found that a dolphin and humans increased their hesitation as a function of task difficulty. Similarly, Scheumann and Call (2004) found that hesitation measures such as wavering or looking longer increased in fur seals as a function of the nature of the cue they received from an experimenter. Harder cues to follow produced more hesitation than more easily detected cues. In a similar vein, Leavens, Aureli, Hopkins, and Hyatt (2001) reported that chimpanzees increased the frequency of self-directed behavior such as scratching when task difficulty increased.

Although these studies support the notion that behavioral measures of hesitation can be used as indicators of internal conflict, it is unclear what type of conflict they are tapping into. For instance, hesitation and anxiety may arise in difficult trials because subjects have a low expectation of getting the reward. Thus, it is conceivable that a linear increase between success and hesitation is related to reward expectancy, rather than indicative of conflict generated by two competing cognitive strategies. One way to tease these two possibilities apart is to present a task that can be tackled by using two alternative strategies that produce different degrees of accuracy. Subjects that follow one or the other strategy should show low and comparable levels of hesitation whereas subjects that shift between strategies should show higher levels of hesitation. One task that affords two distinct strategies is liquid conservation.

Liquid conservation is defined as the ability to understand that liquid quantities remain constant even after changing their perceptual appearance (Piaget, 1941/1997). In a classic liquid conservation task, subjects are presented with a pair of identical transparent containers filled with the same amount of liquid. An experimenter then transfers one of the contents into another container so that the initial appearance of the liquid changes, and the subjects are asked to judge whether the quantities remain equal after the transformation. This procedure arguably induces potential conflict between a notion of identity (i.e. the liquid is the same and hence its quantity should remain unchanged) and

appearance (i.e. the liquid looks different after the transformation) (Cantor, 1983). A modified nonverbal version of liquid conservation task, “over-conservation”, has been used to test nonhuman great apes (Call & Rochat, 1996, 1997; Muncer, 1983; Suda & Call, 2004). The task begins by presenting apes with a pair of identical clear containers filled with different amounts of juice, the larger of which they likely select. After the apes’ first choice, an experimenter transfers the contents into a new pair of containers, changing the appearance of the liquid. A correct choice is to track the larger quantity, not being distracted by misleading perceptual information. Accordingly, there are two strategies available for the apes: a perceptual strategy and a content strategy (see also Suda & Call, 2004). The former corresponds to the animals’ vulnerability to deceptive appearances, while the latter is based on the notion of identity. In the perceptual strategy, subjects would select a certain shape of container over another or select a stimulus on a certain side over the other regardless of the contents. Subjects that follow the perceptual strategy would achieve a success rate of 50% because the combination of the quantities and containers, as well as which side the stimuli are on, are counterbalanced. In the content strategy, on the other hand, subjects would track the larger quantity based on the notion of identity, not being distracted by irrelevant perceptual information. Subjects that rely on the content strategy would achieve a perfect performance. Subjects whose scores fall between these two extremes may represent cases of cognitive conflict between the perceptual and content strategies.

To assess such a possibility, one can plot the percentage of successful trials (i.e. trials in which subjects obtained the larger quantity) against the percentage of trials in which subjects showed signs of hesitation. Based on our previous discussion we can present three models with their corresponding predictions regarding the relation between success and hesitation (see Fig. 1):

### *1.1. Cognitive conflict model*

Subjects will show more hesitation as the degree of cognitive conflict increases. Consequently, the model predicts that subjects using a single strategy will show less hesitation than subjects that are shifting between two strategies. The more heavily apes rely on a single strategy (and the closer the apes’ success rate nears either extreme), the less hesitant they will become. Therefore, this model predicts an inverted U-shaped curve between success and hesitation. A 75% success rate will produce the highest hesitation.

### *1.2. Reward expectancy model*

Subjects will show more hesitation as the probability of getting the reward decreases. Thus, animals that are less rewarded will show more hesitation than animals that are more rewarded. Therefore, this model predicts an inverse linear relation between success and hesitation. A 75% success rate will produce an intermediate amount of hesitation that is higher than the one observed for a 100% success rate and smaller than that observed for a 50% success rate.

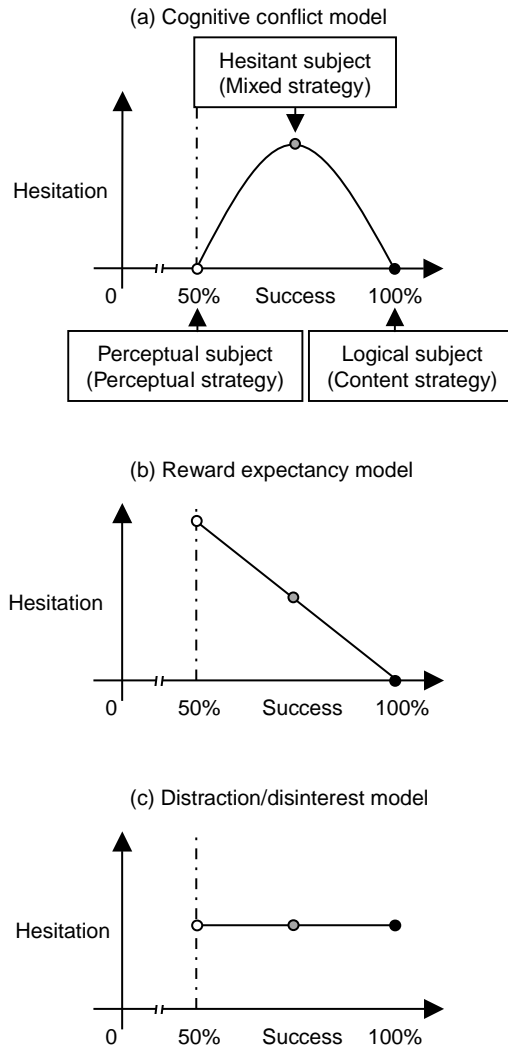


Fig. 1. Three hypothetical models regarding the relation between success and hesitation.

### 1.3. Distraction/disinterest model

Subjects' hesitation will be independent of success rate because subjects monitor neither cognitive conflict nor the amount of reward they get. According to this model, subjects will show the same amount of hesitation (including no hesitation at all) regardless of the extent of cognitive conflict. Therefore, this model predicts a constant relation between success and hesitation. A 75% success rate will produce the same amount of hesitation as a 100 or a 50% success rate.

In the current study, we tested the predictions derived from these models using data from a previous study on liquid conservation in apes (Suda & Call, 2004). We analyzed the distribution of apes' success, scored their hesitation responses, and related both measures to see which one of the three models presented is supported.

## 2. Methods

### 2.1. Subjects

Four bonobos, five chimpanzees, and five orangutans served as subjects in Experiment 1, while the same subjects except for one chimpanzee (Sandra), who refused to take part in the testing, participated in Experiments 2 and 3. Table 1 provides further descriptions of the subjects used in this study.

### 2.2. Task procedure

The task was originally designed for investigating the apes' understanding of Piagetian liquid conservation (Suda & Call, 2004). The apes were tested in indoor cages with a Plexiglas window that had a row of three identical holes. An experimenter sat in front of the window and placed a wooden platform so that it was about 6–8 cm away from the Plexiglas. She presented the apes with two unequal amounts of juice in a pair of identical clear cups at the center of the platform, and simultaneously moved each of the two stimuli in front of the far right and left holes. The experimenter then pushed the platform against the Plexiglas for the apes to indicate their choice by sticking out a finger through one of the holes. The apes generally (>90%) preferred the larger quantity. After the apes' first choice, the experimenter pulled back the platform and placed a new pair of empty

Table 1  
Age, sex, birthplace, and rearing histories of subjects included in the study

Subject	Species	Age (years)	Sex	Birthplace	Rearing history
Joey	<i>Pan paniscus</i>	19	M	Captivity	Hand-reared
Kuno	<i>Pan paniscus</i>	5	M	Captivity	Hand-reared
Limbuko	<i>Pan paniscus</i>	6	M	Captivity	Hand-reared
Ulindi	<i>Pan paniscus</i>	8	F	Captivity	Mother
Fifi	<i>Pan troglodytes</i>	8	F	Captivity	Mother
Fraukje	<i>Pan troglodytes</i>	25	F	Captivity	Hand-reared
Jahaga	<i>Pan troglodytes</i>	8	F	Captivity	Mother
Sandra	<i>Pan troglodytes</i>	8	F	Captivity	Mother
Ulla	<i>Pan troglodytes</i>	24	F	Captivity	Hand-reared
Bimbo	<i>Pongo pygmaeus</i>	22	M	Captivity	Mother
Dunja	<i>Pongo pygmaeus</i>	31	F	Captivity	Unknown
Pini	<i>Pongo pygmaeus</i>	14	F	Captivity	Mother
Toba	<i>Pongo pygmaeus</i>	8	F	Captivity	Mother
Walter	<i>Pongo pygmaeus</i>	13	M	Captivity	Mother

Note. M, male; F, female.

containers at the center of the platform. The experimenter transferred the liquid quantities into the new pair of containers, and thus the apes witnessed the stimuli undergoing transformation. The experimenter removed the empty containers immediately after the transfer. She then placed the new containers, which were now filled with the juice, in front of the far right and left holes of the Plexiglas, and pushed the platform against the Plexiglas. The apes were again allowed to indicate their choice by manual pointing after the transformation. A correct response was to track the larger quantity after the transformation. All trials were videotaped.

Table 2 summarizes the experimental conditions used in the current study, which is based on the first three experiments of our original liquid conservation study. There were four, four, and two testing conditions in Experiments 1, 2, and 3, respectively. As for the main analyses of the relationship between apes' hesitation and success (the model-fit analyses), we excluded three conditions of Experiment 2 (NTD, TS, & NTS conditions) because only one remaining condition (TD condition) was analyzable in this experiment.

Table 2  
Descriptions of testing conditions used in the analyses

Condition <sup>a</sup>	Description
<i>Experiment 1</i>	
Clear-same (CS)	Unequal quantities of juice (20 and 50 ml) transferred from identical clear cups to identical clear cups
Clear-different (CD)	Unequal quantities of juice (20 and 50 ml) transferred from identical clear cups to clear cup and clear tube. Larger quantity transferred in cup in half of trials (LC trials). Larger quantity transferred in tube in the other half (LT trials)
Opaque-same (OS)	Unequal quantities of juice (20 and 50 ml) transferred from identical clear cups to identical opaque cups
Opaque-different (OD)	Unequal quantities of juice (20 and 50 ml) transferred from identical clear cups to opaque cup and opaque tube. Larger quantity transferred in cup in half of trials (LC trials). Larger quantity transferred in tube in the other half (LT trials)
<i>Experiment 2</i>	
Transfer-different (TD)	Different quantities of juice (20 and 50 ml) transferred from identical clear cups to clear dish and clear tube. Larger quantity transferred in dish in half of trials (LD trials). Larger quantity transferred in tube in the other half (LT trials)
No-transfer-different (NTD)	Different quantities of juice (20 and 50 ml) presented in clear dish and clear tube. Larger quantity presented in dish in half of trials (LD trials). Larger quantity presented in tube in the other half (LT trials)
Transfer-same (TS)	Same quantities of juice (50 ml) transferred from identical clear cups to clear dish and clear tube
No-transfer-same (NTS)	Same quantities of juice (50 ml) presented in clear dish and clear tube
<i>Experiment 3</i>	
4-Cup	Different quantities of juice (14 and 28 ml) transferred from identical clear cups to single clear cup and set of 4 clear cups. Larger quantity transferred in single cup in half of trials (LS trials). Larger quantity transferred in multiple cups in the other half (LM trials)
8-Cup	Different quantities of juice (14 and 28 ml) transferred from identical clear cups to single clear cup and set of 8 clear cups. Larger quantity transferred in single cup in half of trials (LS trials). Larger quantity transferred in multiple cups in the other half (LM trials)

<sup>a</sup> Each condition had 24 trials. The placement of unequal quantities left or right, as well as those of different containers, was counterbalanced.

The excluded conditions either involved equal amounts of juice as test stimuli (i.e. there were no correct and wrong choices) or did not allow the apes to witness the transformation of stimuli (i.e. the apes could not use the content strategy). The hesitation data of the three excluded conditions were used in separate analyses, together with the other one condition of Experiment 2, to investigate the effect of liquid transfer on the apes' hesitation.

### 2.3. Coding

We coded the apes' hesitant behavior shown after the transformation of stimuli from the videotapes. Hesitant behavior was defined as making multiple choices within a single trial by either simultaneously pointing at both alternatives or by successively changing a choice. These two types of hesitant behavior were defined as follows (see Fig. 2):

*Pointing at both alternatives (PB)*. A subject simultaneously sticks its fingers through both the right and left holes in the Plexiglas.

*Changing a choice (CC)*. A subject sticks its finger(s) through one of the holes in the Plexiglas, and changes to the other hole.

The apes' hesitant behavior was coded after the transformation of the stimuli. More precisely, coding of the hesitant behavior started when the experimenter began to move the transformed liquid quantities from the center of the platform toward the right and left holes in the Plexiglas. The coding ended when 1 s had passed after the experimenter placed the platform against the Plexiglas for the subjects to indicate a choice.

We coded the apes' choices made after the transformation of stimuli based on the following criteria. If the apes' fingers went through only one of the holes in the Plexiglas when the platform touched against the Plexiglas (after the transformation of stimuli), the apes were coded as having chosen a stimulus on that side. If the apes stuck out their fingers through both of the right and left holes at the moment of the platform touching the Plexiglas, the side of which the apes first touched the platform became their choice. If the apes simultaneously touched both sides of the platform, the side where the apes' fingers last remained became their choice. If this judgment was impossible, the side of the hole where their fingers remained was considered as their choice. If the apes concurrently pulled back their fingers from the holes, the side on which the apes first inserted their fingers again through either right or left holes in the Plexiglas became their choice.

The experimenter watched all of the tapes and recoded the subjects' hesitant behavior and choice. A second observer, who was blind to the purpose of this study, coded 20% of testing trials. Interobserver reliability with the main coder was excellent for both the subjects' hesitant behavior ( $\kappa = .87$ ) and choice ( $\kappa = .97$ ).

### 2.4. Data analysis

In the model-fit analyses, we calculated the percentage of trials in which a subject demonstrated hesitant behavior (either of PB and CC) for each testing condition, and obtained each subject's mean percentage of hesitation within experiment (hereafter "hesitation score"). For instance, if a subject showed hesitant behavior in 15, 10, 30, and 15% of trials in the four conditions of Experiment 1, her hesitation score in this experiment would be  $(15 + 10 + 30 + 15)/4 = 17.5\%$ . Likewise, we calculated each subject's mean

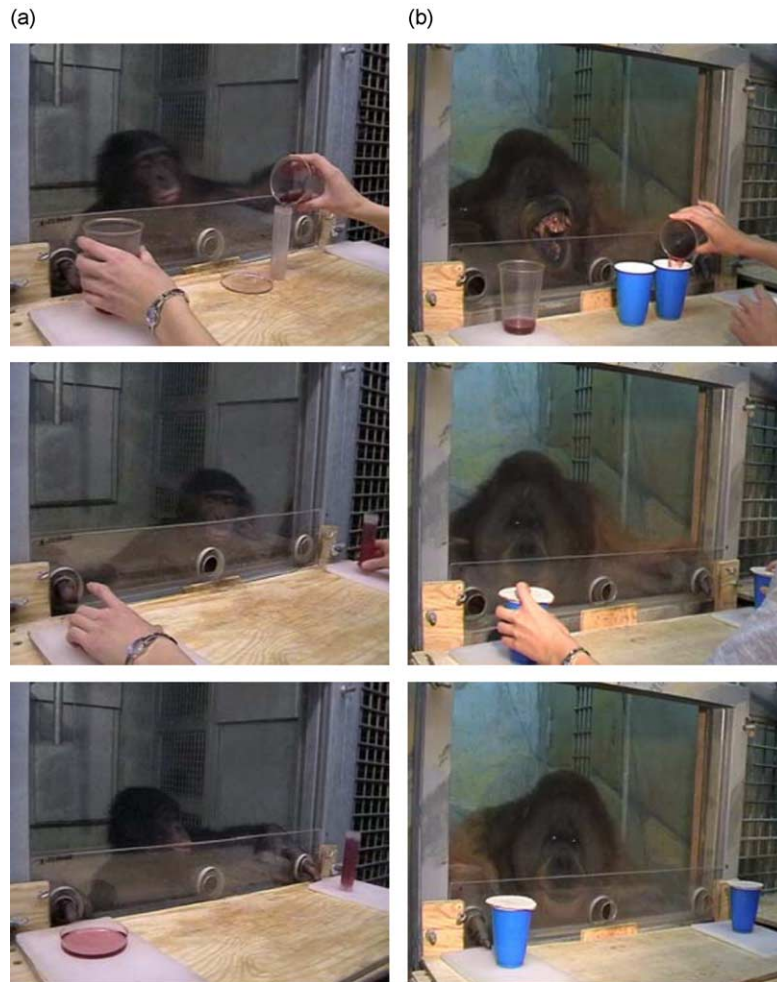


Fig. 2. Two types of hesitant behavior demonstrated by apes. (a) Pointing at both alternatives (PB). (b) Changing a choice (CC).

percentage of correct responses within experiment (hereafter “success score”). For example, if a subject were correct in 60, 70, 75, and 65% of trials in the four conditions of Experiment 1, her success score in this experiment would be  $(60 + 70 + 75 + 65) / 4 = 67.5\%$ . As for Experiment 2, there was only one condition included in the model-fit analyses, and hence subjects’ percentages of hesitation and success in this condition were simply used as their hesitation scores and success scores. We analyzed the relationship between the subjects’ hesitation and success within experiment and investigated which model best fit the data. More specifically, we plotted the subjects’ data points with their success score and hesitation score on the  $x$ - and  $y$ -axis, respectively, and evaluated the distribution of the data points. We also divided the apes into three groups:



Low (50–66.6%), Moderate (66.7–83.3%), and High (83.4–100%), depending on their success score in each experiment, and compared the groups' means of hesitation scores within experiment.

In addition to the model-fit analyses, we investigated whether the transfer of liquid had any effect on the apes' hesitation. We calculated a percentage of trials in which a subject demonstrated hesitant behavior (either of PB and CC) for each testing condition of Experiment 2. In this experiment, there were four testing conditions, two of which had unequal quantities as test stimuli and the other two had equal quantities as stimuli. When unequal quantities were used, the apes could witness the transformation of liquid in one condition, while they could see only the end result of the transformation in the other condition. Likewise, with the same quantities being presented, the apes could witness the transformation in one condition, whereas the transformation was hidden in the other. An ANOVA was conducted on the apes' percentage of hesitation with two factors: whether or not apes witnessed the liquid transformation, and whether test stimuli were equal or unequal quantities. In all of the analyses, an alpha level was set to .05.

### 3. Results

There were no significant species differences in the apes' hesitation score across the experiments. An ANOVA on the apes' hesitation score with experiment as a within-subjects factor and species as a between-subjects factor revealed no effect of experiment,  $F(2,20)=0.03$ ,  $P=.970$ , species,  $F(2,10)=1.33$ ,  $P=.307$ , or an interaction between the two factors,  $F(4,20)=0.09$ ,  $P=.985$ . The distribution of the subjects' data points on the y-axis could not be explained by species differences. Therefore, we analyzed the data points of the three species altogether.

Fig. 3 presents the apes' hesitation score as a function of their success score in Experiments 1–3. Plotting the apes' data points produced an inverted U-shaped curve. A quadratic curve significantly fit the distribution of data points in Experiment 1,  $\text{adj. } r^2 = .31$ ,  $P=.018$ , whereas a linear correlation hardly predicted the data points,  $\text{adj. } r^2 = -.08$ ,  $P=.779$ . Similarly, a quadratic curve significantly fit the data points in Experiment 2,  $\text{adj. } r^2 = .53$ ,  $P=.006$ , whereas the distribution was far from matching a linear trend,  $\text{adj. } r^2 = .05$ ,  $P=.230$ . The results of Experiment 3 pointed the same direction as in Experiments 1 and 2. A quadratic fit failed to reach a significant level,  $\text{adj. } r^2 = .16$ ,  $P=.069$ , but it still predicted the data points better than a linear trend did,  $\text{adj. } r^2 = -.08$ ,  $P=.787$ . The distribution of hesitation and task performance was consistent with the cognitive conflict model: apes with intermediate success scores were more hesitant than highly successful or unsuccessful apes.

In order to analyze the data from a slightly different perspective, we divided the apes into three groups depending on their success score: Low (50–66.6% success), Moderate (66.7–83.3% success), and High (83.4–100% success) groups, and compared the groups' mean hesitation scores. Fig. 4 presents mean hesitation scores as a function of success group. An ANOVA on the apes' hesitation score with success group as a factor revealed a significant effect of group in Experiment 1,  $F(2,11)=4.11$ ,  $P=.047$ . There was a significant quadratic trend in hesitation across the three groups,  $F(1,11)=8.12$ ,  $P=.016$ ,

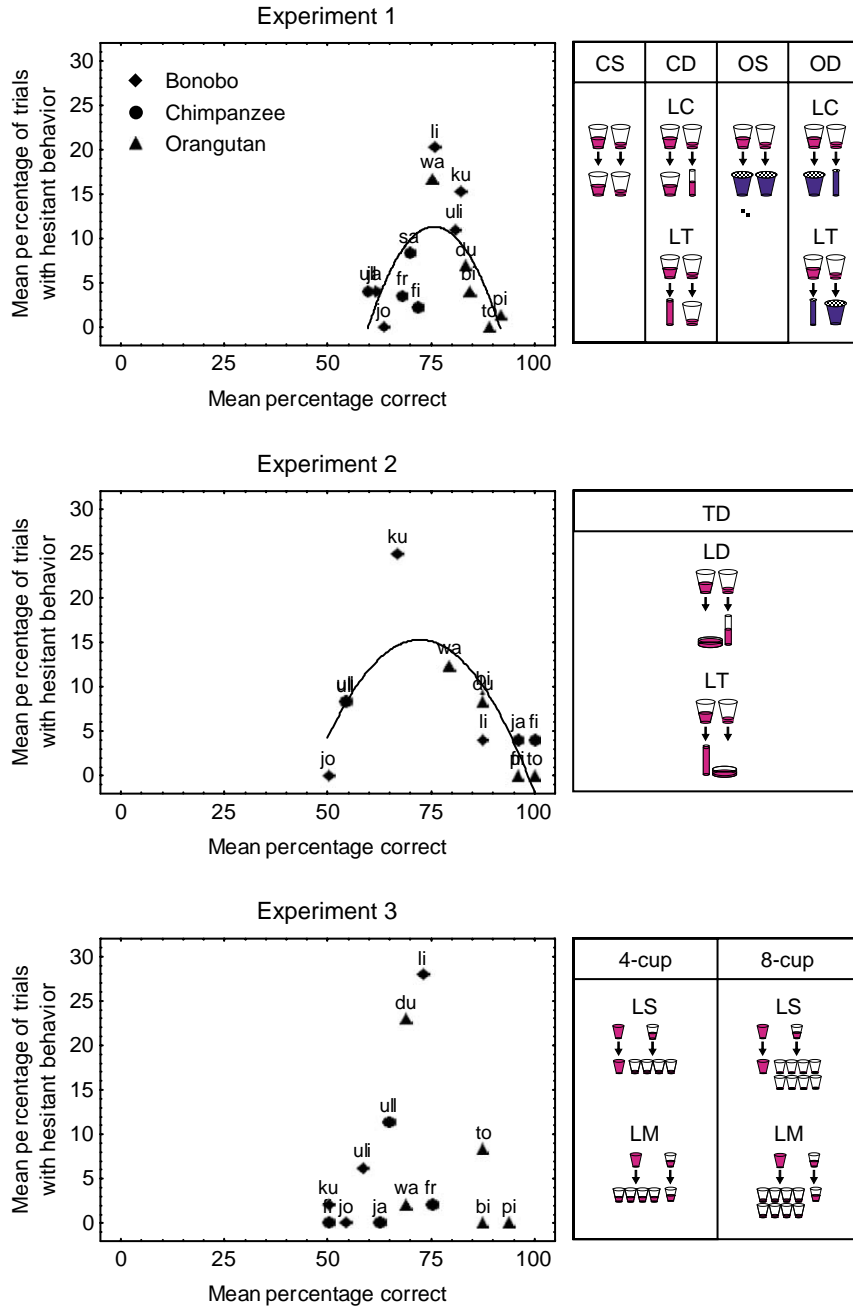


Fig. 3. Distribution of apes' hesitation scores and success scores in Experiments 1–3, along with used testing conditions. Letters above data points indicate subjects' names.

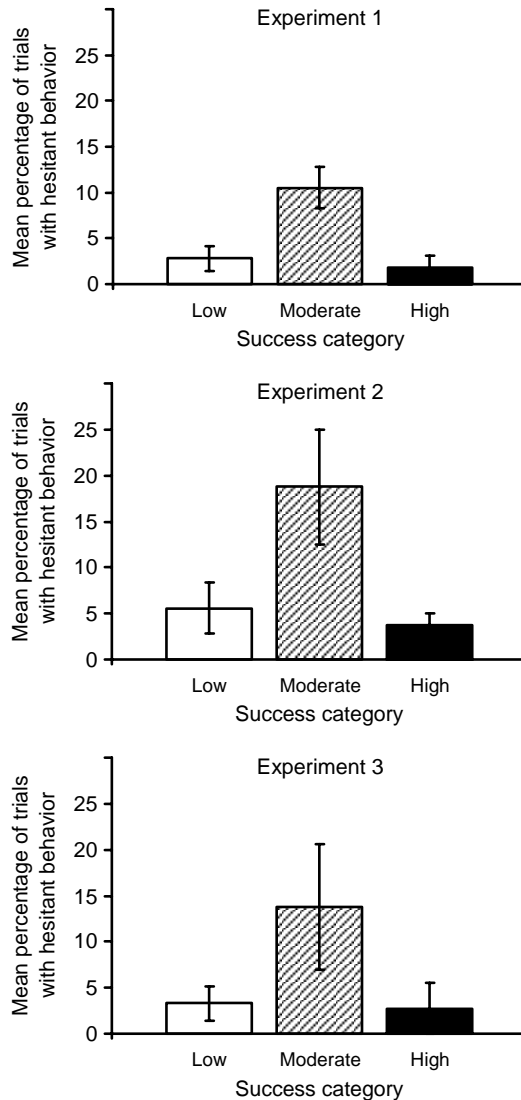


Fig. 4. Group mean of hesitation scores ( $\pm$ SEM) as a function of success group category.

but no linear trend was found,  $F(1,11)=0.05$ ,  $P=.834$ . The Moderate group demonstrated more hesitant behavior than the Low and High groups did, Low:  $M=2.8$ ,  $SD=2.4$ ; Moderate:  $M=10.5$ ,  $SD=6.4$ ; High:  $M=1.9$ ,  $SD=2.1$ . Likewise, a significant effect of group on hesitation was found in Experiment 2,  $F(2,10)=8.33$ ,  $P=.007$ . There was a significant quadratic trend in hesitation across the three groups,  $F(1,10)=14.83$ ,  $P=.003$ , but no linear trend,  $F(1,10)=0.33$ ,  $P=.580$ . The Moderate group demonstrated a higher mean of hesitation scores than the Low and High groups did, Low:  $M=5.6$ ,  $SD=4.8$ ;

Middle:  $M=18.8$ ,  $SD=8.8$ ; High:  $M=3.8$ ,  $SD=3.6$ . In Experiment 3, the results followed the same pattern as in Experiments 1 and 2. In particular, the Moderate group had a higher mean of hesitation scores than the Low and High groups did, Low:  $M=3.3$ ,  $SD=4.7$ ; Middle:  $M=13.8$ ,  $SD=13.6$ ; High:  $M=2.8$ ,  $SD=4.8$ , although an ANOVA on hesitation yielded no significant effect of group,  $F(2,10)=2.21$ ,  $P=.161$ . Thus, apes with an intermediate success were more hesitant than very successful and unsuccessful apes. These findings further supported the cognitive conflict model.

The cognitive conflict model best predicted the relationship between the subjects' hesitation and their success rate. However, the distribution of the data points could be an artifact of some apes' ambiguous manual pointing. Perhaps, all of the subjects relied on a single strategy, but some might have ended up with achieving an intermediate success because the experimenter had trouble interpreting their choices correctly. To rule out this possibility, we excluded the trials in which the subjects demonstrated hesitant behavior and calculated each subject's success score from the remaining trials (hereafter "unambiguous success score"). We compared the apes' hesitation score and unambiguous success score just as in the previous model-fit analyses. If the apes' ambiguous choices were responsible for the relationship between hesitation and task performance, the data points would no longer produce an inverted U-shaped curve. (The data points should get diverted to either 50 or 100% success.) Alternatively, if an inverted U-shaped curve would remain after the above adjustments, an inadvertent effect of apes' ambiguous choices could not explain the hesitation–success relationship. The results were consistent with the latter. Plotting the data points continued to produce an inverted U-shaped curve. Fig. 5 presents the apes' hesitation score as a function of their unambiguous success score in Experiments 1–3. In Experiment 1, a quadratic curve remained to fit the data points at a significant level,  $\text{adj.}r^2=.31$ ,  $P=.019$ , and it predicted the data points much better than a linear correlation did,  $\text{adj.}r^2=-.06$ ,  $P=.641$ . Likewise, in Experiment 2, a quadratic curve remained to fit the data points at a significant level,  $\text{adj.}r^2=.59$ ,  $P=.002$ , while a linear correlation poorly predicted the data points,  $\text{adj.}r^2=-.06$ ,  $P=.562$ . In Experiment 3, neither a linear correlation nor a quadratic curve fit the data points at a significant level. Therefore, the relation between hesitation and success was not a byproduct of apes' ambiguous choices, at least for Experiments 1 and 2.

One could argue that the hesitation score is an individual characteristic of the subject (i.e. some animals always hesitate while others rarely do so) rather than something that depends on the particular cognitive conflict induced by each of the tasks. Table 3 summarizes each subject's hesitation score and success score in the three experiments. For comparison, the table also includes each subject's scores of hesitation and success in the first choice of Experiment 1 (in which the apes were presented with a pair of identical clear cups containing different quantities of juice prior to the transformation of the liquid). Success scores were categorized into three groups, Low (50.0–66.6%), Moderate (66.7–83.3%), and High (83.4–100%). As for hesitation scores, we divided a maximum hesitation score (27.9%, which was demonstrated by a bonobo subject, Kuno in Experiment 3) by three to make three categories of hesitation, Low (0.0–9.3%), Moderate (9.4–18.6%), and High (18.7–27.9%). All of the apes except for two subjects (Limbuko & Ulla) rarely hesitated prior to the liquid transformation in Experiment 1. Thirteen animals completed all of the three experiments, while the remaining one subject participated in only Experiment 1. Six subjects out of

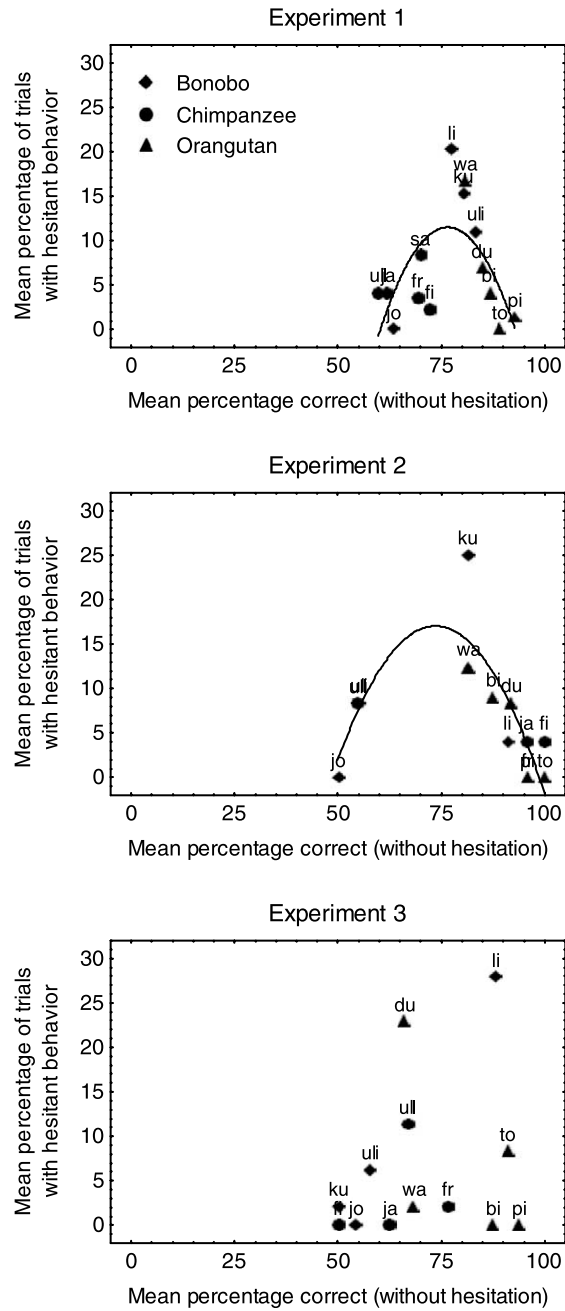


Fig. 5. Distribution of apes' hesitation scores and unambiguous success scores in Experiments 1–3. Letters above data points indicate subjects' names.

Table 3  
Patterns of subjects' hesitation and success across the experiments

Subject	1st choice		Experiment 1		Experiment 2		Experiment 3		Pattern of hesitation
	Suc	Hes	Suc	Hes	Suc	Hes	Suc	Hes	
<i>Bonobos</i>									
Joey	H	L	L	L	L	L	L	L	Consistent
Kuno	H	L	M	M	M	H	L	L	Inconsistent
Limbuko	H	M	M	H	H	L	M	H	Inconsistent
Ulindi	H	L	M	M	L	L	L	L	Inconsistent
<i>Chimpanzees</i>									
Fifi	H	L	M	L	H	L	L	L	Consistent
Fraukje	H	L	M	L	H	L	M	L	Consistent
Jahaga	H	L	L	L	H	L	L	L	Consistent
Sandra	H	L	M	L	–	–	–	–	–
Ulla	H	H	L	L	L	L	L	M	Inconsistent
<i>Orangutans</i>									
Bimbo	H	L	H	L	H	L	H	L	Consistent
Dunja	H	L	M	L	H	L	M	H	Inconsistent
Pini	H	L	H	L	H	L	H	L	Consistent
Toba	H	L	H	L	H	L	H	L	Consistent
Walter	H	L	M	M	M	M	M	L	Inconsistent

Note. Suc, success; L, low success rate (50.0–66.6%); M, moderate success rate (66.7–83.3%); H, high success rate (83.4–100%); Hes, hesitation; L, low hesitation rate (0–9.3%); M, moderate hesitation rate (9.4–18.6%); H, high hesitation rate (18.7–27.9%); 1st choice, 1st choice in Experiment 1.

the 13 clearly showed different patterns of hesitation and success across the experiments, which means that these animals occupied different regions in a hesitation–success distribution depending on the experiment. The other seven subjects demonstrated a low level of hesitation rather consistently across the experiments.

Finally, the effect of liquid transformation on the apes' hesitation was investigated. Experiment 2 had four testing conditions, which were generated by varying two factors: whether or not the apes witnessed the transformation of liquid, and whether testing stimuli

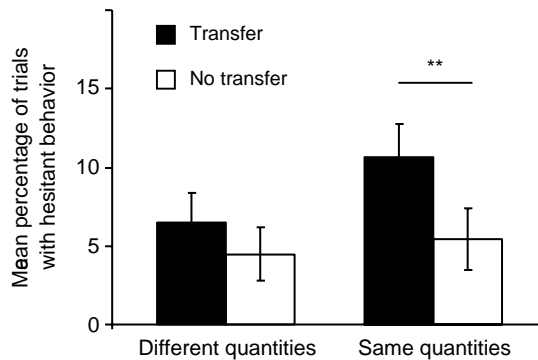


Fig. 6. Mean percentage of trials in which subjects showed hesitant behavior ( $\pm$ SEM) as a function of liquid transfer and used quantities.  $**P < 0.01$ .

were equal or unequal quantities. Fig. 6 presents the apes' percentage of hesitation as a function of condition. An ANOVA on the apes' hesitation with liquid transformation and liquid quantity as within-subjects factors revealed a nearly significant effect of transformation,  $F(1,12)=4.59$ ,  $P=.053$ , and a significant interaction between transformation and quantity,  $F(1,12)=5.58$ ,  $P=.036$ , but no effect of quantity,  $F(1,12)=1.73$ ,  $P=.213$ . Post hoc pairwise comparisons using the Bonferroni–Holm procedure revealed that the apes hesitated more when seeing liquid transformation than when not seeing it, and that this pattern reached a significant level only when liquid quantities were equal,  $P=.005$ . Thus, witnessing liquid transformation induced hesitation in the apes, which suggests that the apes in fact experienced conflict between the notion of identity and appearance.

#### 4. Discussion

The cognitive conflict model best fit the relationship between the apes' hesitation and their success rate. Moderately successful apes tended to show more hesitant behavior than very successful or unsuccessful apes in our series of liquid conservation experiments. Note that the other two models that we considered failed to explain our data satisfactorily. The simple monitoring of reward gain predicts that apes will hesitate more as the amount of their reward gain decreases, which we did not find in the current study. Similarly, the distraction/disinterest model also failed to explain the distribution of hesitation and success observed in this study because this model predicts that hesitation will not be related to success.

The key point for interpreting this result is that our experiments likely created more intense cognitive conflict in moderately successful apes than in very successful or unsuccessful apes. A liquid conservation experiment, by its nature, induces potential conflict between the notion of identity and appearance. The existence of this conflict was confirmed by the finding that liquid transformation, particularly when pairs of equal quantities were involved, increased the apes' hesitation. As long as subjects relied on a single strategy (either the content or perceptual strategy), they would not face any cognitive conflict. A single strategy led subjects to respond to test stimuli in a consistent way. The content strategy, without exception, directed subjects to track the larger quantity, and the perceptual strategy always directed subjects to select the container that appeared to have a larger quantity. In contrast, subjects would face cognitive conflict if they entertained the two strategies. Consistent with this idea, we found that apes with a moderate success were more hesitant than very successful or unsuccessful apes.

There are three lines of evidence that further supports the cognitive conflict account. First, unsuccessful apes with a chance performance actually deployed a highly consistent strategy of selecting a certain container over another (e.g. a dish over a tube), rather than making a random choice (Suda & Call, 2004). Second, all of the apes showed a container preference (i.e. the perceptual strategy), albeit to different degrees, and the apes as a group showed some ability to track the larger quantity when the contents were occluded (i.e. the content strategy) (Suda & Call, 2004). Third, our previous studies showed that alternative strategies such as response perseveration were not viable mechanisms for the apes' performance (Call & Rochat, 1997; Suda & Call, 2004). It is, therefore, plausible that

the apes in fact used these two distinctive strategies to solve our conservation task, and that a 50% correct performance corresponded to the perceptual strategy while a 100% correct performance corresponded to the content strategy. Accordingly, the apes' hesitant behavior, which was examined in the current study, seemed to be a reflection of their internal cognitive conflict between the two competing strategies.

To what extent the apes understood their own cognitive conflict (that is, not only they experienced it but they were also aware of it) is an important issue for animal metacognition, but it is beyond the scope of the current study. From the present findings alone, we cannot draw strong conclusions on whether the apes were explicitly aware of their own internal conflict or the conflict "automatically" triggered hesitant behavior in the apes. One such non-metacognitive explanation is that hesitation might result from a cognitive load rather than metacognition. Subjects with a chance or perfect performance might have recruited fewer cognitive resources by activating only a single strategy than did moderately successful subjects, who might have activated both of the two competing strategies. The introduction of an explicit behavioral measure (such as an escape response to be made before an actual presentation of a task) is vital for assessing a further question on metacognition (see Hampton, 2001).

Despite this limitation, our study offers at least two interesting features compared to previous studies examining animal metacognition. First, the apes' hesitant behavior in the current study (i.e. pointing at both alternatives, or changing a choice successively) was not reinforced in any way. The apes gained little benefit from showing hesitant behavior in terms of increasing their reward gain, and thus their hesitant behavior was spontaneous. Consequently, the distribution of this behavioral measure cannot be the result of differential reinforcement. In contrast, some of the previous studies used a behavioral response, with which animals could potentially increase their reward gain (*escape response*: Hampton, 2001; Shields, Smith, & Washburn, 1997; Smith et al., 1995; Smith, Shields, Allendoerfer, & Washburn, 1998; Smith, Shields, Schull, & Washburn, 1997; *looking response*: Call & Carpenter, 2001; Hampton, Zivin, & Murray, 2004). Note also that animals were given hundreds of trials until they reached a mature level of performance (with an exception of Call and Carpenter, 2001), which was not the case in the current study. Other studies used a non-rewarding behavior (e.g. scratching) as an indicator of cognitive conflict, but they failed to rule out the possibility that the behavior reflected perceived reward contingency (Leavens et al., 2001; Scheumann & Call, 2004). Therefore, only the current study clearly teased animals' hesitation from their reward expectation, without imposing extensive training on them.

Second, our liquid conservation experiments entailed two distinctive strategies (i.e. the perceptual strategy and the content strategy), which inevitably instructed the apes to produce contradictory responses. A 50% success corresponded to the perceptual strategy, whereas a 100% success corresponded to the content strategy. Accordingly, task performance was determined by the degree of the apes' dependence on these strategies. In contrast, the difficulty of the previous tasks relied on either animals' perceptual constraints (e.g. auditory or visual discrimination) or memory limits. Therefore, the current study seems to provide the first evidence that the use of two different cognitive strategies results in cognitive conflict in apes.



The current study also revealed two important aspects of apes' hesitation and cognitive conflict. First, the apes' hesitation seemed to be somewhat case-dependent, as it is in humans (summarized in [Smith et al., 2003](#)). Several subjects clearly varied a level of hesitation according to a task in the current study. An individual, who was highly hesitant in one experiment, might hesitate rarely in another. That the apes' hesitation was to some degree flexible and was not totally constrained by their individual characters is comparable to the nature of human hesitation. The flexibility of an individual's hesitation is also consistent with the theory that hesitation emerges from certain circumstances where cognitive conflict is especially intense.

Second, the apes' hesitation might illustrate successive stages of their conceptual understanding. In order to achieve a true concept of liquid conservation, subjects must recognize their internal conflict between the notion of identity and appearance, and overcome this conflict. Hence, the distribution of the apes' hesitation and success might have demonstrated three stages of conceptual acquisition: (1) no recognition of the conflict and no understanding of the concept; (2) recognition of the conflict and an incomplete understanding of the concept; and (3) conquest of the conflict and a mature understanding of the concept. One could not obtain such rich information if he simply observed animals' percentage of correct responses. The conflict-inducing paradigm has proven to be a useful method for exploring what animals understand about underlying principles of cognitive tasks.

Moreover, the current study complements other recent studies that have used explicit measures of metacognition such as escaping in difficult trials ([Hampton, 2001](#); [Shields et al., 1997](#); [Smith et al., 1995, 1998, 1997](#)) or seeking additional information upon encountering incomplete data to solve a task ([Call, 2005](#); [Call & Carpenter, 2001](#)). The study of apes' metacognition had been neglected until recently, although this ability appears to have great importance for current hotly debated issues such as theory of mind, self-awareness, or even consciousness ([Call, 2003](#); [Hampton, 2005](#); [Smith et al., 2003](#)).

In conclusion, the current study suggests that the intermediate performance of three species of nonhuman great apes in a Piagetian conservation task is due to the cognitive conflict generated by two competing strategies: identity and appearance. Our data failed to support the idea that an intermediate performance is due to inattentiveness or a reflection of perceived reward gain. The current findings also suggest that combining hesitation and performance data can be a very fruitful approach to study the strategies that subjects use to solve a variety of problems and the level of comprehension of those strategies. Finally, these results are consistent—albeit they do not conclusively prove—the hypothesis that apes may be aware of their own cognitive conflict. Further research on comparative metacognition using explicit measures can shed some light onto the evolution and development of animal cognition.

## **Acknowledgements**

Applicable animal care and use rules and guidelines were followed in the conduct of this research. We thank the staff of the Wolfgang Köhler Primate Research

Center for their support. We also thank Daniel Stahl for statistical advice and Friederike Schmitz for helping with reliability assessment. C.S. would like to thank Laurence King for his useful comments on the original version of this manuscript, and Keith Jensen for grammatical correction. A part of this study was presented at the COE 21 International Symposium “Diversity of Cognition: Evolution, Development, Domestication, and Pathology” in Kyoto, Japan in September, 2003, and at the XX Congress of the International Primatological Society (IPS) in Torino, Italy in August, 2004.

## References

- Beran, M. J. (2001). Summation and numerosness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *115*, 181–191.
- Call, J. (2003). On linking comparative metacognition and theory of mind. *Behavioral and Brain Sciences*, *26*, 341–342.
- Call, J. (2005). The self and the other: A missing link in comparative social cognition. In H. Terrace, J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 321–341). New York: Oxford University Press.
- Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, *4*, 207–220.
- Call, J., Hare, B. A., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition*, *1*, 89–99.
- Call, J., & Rochat, P. (1996). Liquid conservation in orangutans (*Pongo pygmaeus*) and humans (*Homo sapiens*): Individual differences and perceptual strategies. *Journal of Comparative Psychology*, *110*, 219–232.
- Call, J., & Rochat, P. (1997). Perceptual strategies in the estimation of physical quantities by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, *111*, 315–329.
- Cantor, G. N. (1983). Conflict, learning, and piaget: Comments on Zimmerman and Blom’s “Toward an empirical test of the role of cognitive conflict in learning”. *Developmental Review*, *3*, 39–53.
- Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, *98*, 5359–5362.
- Hampton, R. R. (2005). Can rhesus monkeys discriminate between remembering and forgetting? In H. Terrace, J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 272–295). New York: Oxford University Press.
- Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition*, *7*, 239–246.
- James, W. (1981). *The principles of psychology*, vol. 1. Cambridge: Harvard University Press (Original work published 1890).
- Leavens, D. A., Aureli, F., Hopkins, W. D., & Hyatt, C. W. (2001). Effects of cognitive challenge on self-directed behaviors by chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, *55*, 1–14.
- Muncer, S. J. (1983). “Conversations” with a chimpanzee. *Developmental Psychology*, *16*, 1–11.
- Piaget, J. (1997). *The child’s conception of number*. London: Routledge (Original work published 1941).
- Scheumann, M., & Call, J. (2004). The use of experimenter-given cues by South African fur seals (*Arctocephalus pusillus*). *Animal Cognition*, *7*, 224–230.
- Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General*, *126*, 147–164.
- Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: General*, *124*, 391–408.
- Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, D. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General*, *127*, 227–250.

- Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition*, *62*, 75–97.
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, *26*, 317–373.
- Suda, C., & Call, J. (2004). Piagetian liquid conservation in the great apes (*Pan paniscus*, *Pan troglodytes*, and *Pongo pygmaeus*). *Journal of Comparative Psychology*, *118*, 265–279.
- Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review*, *45*, 1–41.
- Tolman, E. C. (1967). *Purposive behavior in animals and men*. New York: The Century (Original work published 1932).
- Vonk, J. (2003). Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first- and second-order relations. *Animal Cognition*, *6*, 77–86.