International Journal of Comparative Psychology, 2009, 22, 43-60.

Assessing Generalization Within and Between Trap Tasks in the Great Apes

Gema Martin-Ordas and Josep Call
Max Planck Institute for Evolutionary Anthropology, Germany

Considerable research has been devoted to investigate the type of information that subjects use to solve tool-using tasks in which they have to avoid certain obstacles (e.g., traps) to retrieve a reward. Much of the debate has centered on whether subjects simply use certain stimulus features (e.g., the position of the trap) or instead use more functionally-relevant information regarding the effect that certain features may have on a moving reward. We tested eight apes (that in a previous study had succeeded in a trap-tube task) with one functional and two nonfunctional traps to investigate the features that they used to solve the task. Four of the eight subjects used functional features. Additionally, we presented 31 apes with a trap task that did not involve tools but required subjects to make an inference about the position of a hidden reward based on its displacement over a substrate with or without a trap. Subjects performed above chance levels (including from the first trial) in the experimental condition (unlike in the control conditions), suggesting that they took into account the effect that a trap may have on a reward. Third, we correlated the subjects performance in four trap tasks (3 involving tool-use and one without tool-use) and found positive correlations between some of the tasks. Our results suggest that apes possess some knowledge about the effects that traps have on slow moving unsupported objects. However, this knowledge was not robust enough to prevent the influence of certain practice and task effects. Moreover, subjects' knowledge may not have been abstract enough to allow them to establish broad analogies between tasks.

Tool-using tasks that require subjects to overcome obstacles to get a reward have been a major component of research devoted to investigate causal knowledge in primates and birds (e.g., Henrich, 2000; Köhler, 1925; Visalberghi & Limongelli, 1996). One task that has been used extensively in recent years is the trap-tube task (Visalberghi & Limongelli, 1994) in which subjects are presented with a stick, a clear tube with a trap on its bottom part and a reward placed inside the tube next to the trap and outside of the subject's direct reach. Subjects have to use the stick to get the reward out of the tube while avoiding the trap in which the reward may fall. Monkeys and apes find this task difficult (Limongelli, Boysen & Visalberghi, 1995; Povinelli, 2000; Visalberghi & Limongelli, 1994), although recent research has shown ways in which the subjects' performance can be improved. In particular, allowing subjects to choose whether they can rake or push the reward out of the tube (the original studies only allowed subjects to push the reward out of the tube) has substantially improved their performance (Martin-Ordas, Call & Colmenares, 2008; Mulcahy & Call, 2006; see also Seed, Tebbich, Emery, & Clayton, 2006).

Although solving the trap task, especially within a few trials, is consistent with the idea that subjects have knowledge about the causal relations between the elements of the problem, it does not necessarily prove it. There are other mechanisms such as innate predispositions to avoid traps or learned heuristics that can produce a high level of proficiency. For instance, subjects may solve the trap task not because they comprehend the effect that the tool and the trap may exert on

Correspondence concerning this article should be addressed to Gema Martin-Ordas, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany. (ordas@eva.mpg.de).

the reward's displacement but because they have learned to insert the tool in a particular location in relation to the trap position without understanding the effect that the trap has on the reward. Generalization tasks (often referred to as transfer tasks in that literature) have been used to determine whether subjects exploit specific stimulus features (e.g., the position, orientation, or the type of trap) or acquired habits to solve problems or, instead use more general principles about the function and the effect that those features may have on other task elements (e.g., the reward).

Traditionally, researchers have assessed generalization in trap tasks by administering the original task and then implementing some key modifications to this task. Visalberghi and Limongelli (1994) used the inverted tube condition in which the trap was rendered non-functional by flipping the tube 180^{o} such that the trap was above the tube and the bottom of the tube had no holes. Once the trap is non-functional, there is no need to avoid it. Results have been mixed. While in some studies subjects avoided the trap (Povinelli, 2000; Visalberghi & Limongelli, 1994), in others they did not (e.g., Mulcahy & Call, 2006; Tebbich & Bshary, 2004). Moreover, Silva, Page and Silva (2005) criticized this control procedure because the results derived from it are hard to interpret. Failing to avoid a non-functional trap is not diagnostic of lack of comprehension because subjects are not making any mistakes, whereas not avoiding the trap may result from subjects perceiving the tube with the inverted trap as a different problem. Note that inverting the trap not only involves a change in the functional properties of the trap but also a change in its configuration (i.e. the trap receptacle is flipped up). It is therefore hard to know whether subjects are responding to the change in function or configuration.

Seed et al. (2006) circumvented this problem by administering a two-trap tube task in which a functional and a non-functional trap were located on the bottom of the tube. Thus, the use of two traps created a "built-in" control condition. They presented subjects with two conditions: in one case, they had to push the reward into the non-functional trap to recover it from below whereas in the other condition, they had to avoid the functional trap and displace the reward over the non-functional trap in order to get the food. Seven rooks solved the two different trap problems. However, these tasks are still vulnerable to the criticism that subjects may use procedural rules, such as avoiding common features present in the tasks, (i.e. black disks) to solve the problem and not the functional properties of the traps. Therefore, Seed et al. (2006) conducted two additional transfer tests, in which the causal relations inherent in the task remained the same, but the appearance of the stimulus changed thus precluding the use of certain perceptual features. Only one of the seven rooks performed well in these transfer tasks. Although the Seed et al. (2006) trap setup represents an improvement over previous trap tasks and their associated control conditions, transfer tests involving the change of other elements such as the position of the trap in relation to the substrate or presence of new functional elements (i.e. bungs) can introduce a novelty factor that may negatively affect the subjects' performance. An alternative is to present a follow-up condition in which the only thing that changes is the presence of the trap hole while all other elements remain the same.

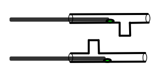
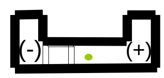
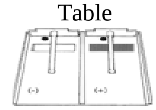
A complementary approach to the within-task generalization mentioned above is the between-task generalization approach. Here subjects perform the original task and they are subsequently assessed on a second task that differs both in the problem layout and the response mode, but it is still functionally equivalent to the first task. There are two other tasks that are functionally equivalent to the trap-tube: the trap-table (Povinelli, 2000) and the trap-platform (Martin-Ordas, et al., 2008). The table task consists of a flat surface divided into two sections. One section has a hole cut in it creating an effective trap so that a reward displaced over it is lost; the other section has a fake trap painted on it with dimensions and position identical to the effective trap. One reward is placed behind each trap and subjects are given a choice of pulling one of two rakes. The heads of the rakes are placed on the table behind the rewards. Less than 10% of the chimpanzees tested in this setup are able to select the correct alternative (i.e., the tool positioned by the fake trap) above chance levels within 20 trials (Girndt, Meier, & Call, 2008; Povinelli, 2000). However, Girndt, Meier, and Call (2008) found that 74% of the subjects succeeded within 10 trials when subjects did not have to select one of the two pre-positioned rakes but could decide where to insert the tool to rake one of the two rewards.

The platform-trap task is a mixture of the tube and the table task (see Table 1 for similarities and differences among these tasks). The task consists of an inverted U-shaped platform (the subject is facing this inverted U-shaped platform) with a hole at the top of one of the two platform arms. The reward is placed on the center of the platform and the trap is located either to the right or to the left side of the reward. In order to solve the task, subjects have to displace the reward away from the trap and bring it within reach down the arm of the U-shaped platform (see Table 1). Martin-Ordas et al. (2008) also tested a trap-tube task in which subjects could either rake or push the reward and the platform task. Although around 50% of subjects performed above chance in each task after only 36 trials, there was no correlation between tasks, that is, no evidence of transfer across tasks. However, one potential confound is that each task required different motor responses. It has been shown that specific features unrelated to the comprehension of the task such as the type of action (i.e. push the reward) and the location of the tools (i.e. tools already positioned on the apparatus) can seriously affect the subjects' performance in trap tasks (Girndt, Meier, & Call, 2008; Mulcahy & Call, 2006). In other words, requiring the use of certain responses may mask the knowledge that subjects may possess about the critical features of the task. One solution to this problem consists of presenting a task in which the motor response required is minimized.

We had three goals in the current study. First, we investigated the question of within-task generalization by administering a modified version of the trap-tube task to eight great apes that had succeeded in a previous study with the trap-tasks. We focused on whether subjects were able to solve the trap tube task by using the presence of causal features (i.e. the presence of the hole) or non-causal features (e.g., the position, color, and shape of the trap receptacle). We used a procedure that minimized the methodological problems associated with changing both causal and non-causal elements of the task and reducing drastically the introduction of novel features.

Second, we addressed the question of between-task generalization by administering a trap task that required no tools or complex manipulations. One advantage of this task is that it minimized the potential problems arising from testing subjects on different trap tasks that require different motor responses. In particular, this new task, called the gap task, simply required subjects to take into account the position of the trap to infer the location of a hidden reward under one of two upside down cups after one of the cups had been displaced over a trap and nothing had fallen from it. In addition, the inclusion of control conditions allowed us to investigate whether subjects showed an intrinsic avoidance for those cups in close proximity to the trap. A positive correlation between the tube/table task and the gap task would support the idea that subjects have some knowledge about the causal relations between the elements in those tasks.

Table 1
Similarities and differences among the original trap tube, the trap table and the trap platform tasks.

	Tube 	Platform 	Table 
Physical Support	Tube	Platform	Platform
Tool insertion	Yes	Yes	No
Reward movement	Left Right	Left Right Straight down	Straight down
Action	Rake Push	Rake	Rake

[1] "Tool insertion" indicates whether subjects had to insert the tool in the apparatus or whether the tools were already pre-positioned.
[2] "Reward movement" indicates in which direction subjects could displace the reward with the tool in the apparatus.
[3] "Action" refers to the technique used to get the reward out of the apparatus.

Third, we collated the data from the various traps tasks (including the new task from Experiment 2) that had been conducted in our laboratory with the same subjects and calculated the correlation between tasks. We included the following four tasks: The gap task from Experiment 2, the trap-platform task (Martin-Ordas et al., 2008), and the modified versions of the trap-tube (Martin-Ordas et al., 2008) and the trap-table (Girndt, Meier, & Call, 2008). This is a particularly important exercise given that assessing what level of causal knowledge subjects possess requires convergent evidence across multiple tasks and to date only one study, to the best of our knowledge, had used this correlational approach in the context of trap tasks.

Experiment 1: Within-task Generalization: The Holeless Trap

We administered three conditions. First, all subjects received the functional trap condition that consisted of a tube with a functional trap. Second, subjects received the fake trap condition, which consisted of the same setup as the original task condition except that there was no hole in the tube through which the reward could fall into. Those subjects that failed the fake trap condition (i.e. continued to avoid a non-functional trap) received the painted trap condition that consisted of a tube mounted inside a plastic box that had a trap receptacle painted on one side thus, simulating the appearance of the trap in the original task condition. Thus, the first condition presented a functional trap whereas the last two conditions presented non-functional traps. Avoiding the trap in the original condition but not in the fake trap condition would suggest that subjects recognized the presence of the hole as an essential feature of the problem. In contrast, failing to avoid the fake trap would suggest that subjects did not focus on the functional properties of the trap. Data on the painted trap condition would help us refine the nature of the non-functional features that subjects used to solve the problem. In particular, this procedure would allow us to distinguish between an explanation based on avoiding the position of the trap and a simpler explanation based on avoiding a black painted area.

Methods

Subjects

Eight apes (3 chimpanzees, 2 bonobos, 2 orangutans and 1 gorilla) housed at the Wolfgang Köhler Primate Research Center in the Leipzig Zoo participated in this experiment (see Table 2). All had participated and succeeded in a previous trap tube experiment (Martin-Ordas et al., 2008). There was 1 male and 7 females with an age range of 7 to 18 years.

Apparatus

There were three apparatuses: the functional trap tube, the fake trap tube and the painted trap tube corresponding to each of three conditions. The functional and the fake trap tubes closely followed the design by Martin-Ordas et al. (2008). For both conditions, we used two tubes (see Figure 1). Each consisted of a 40 cm long x 5 cm diameter Plexiglas tube with an off-center opaque trap (8 cm wide x 15 cm long) located in its bottom part. One tube had the trap on the right side (8 cm from the right end of the tube) and the other had the trap on the left side (8 cm from the left end of the tube). The functional and the fake trap tubes looked exactly the same except for the trap: in the functional trap condition the trap was attached to the tube and connected by a hole, whereas in the fake condition the trap was attached to the tube but no hole connected the tube with the trap. Even though, the painted trap task look different than the fake trap and the functional trap tasks, no new element was incorporated to the task [e.g. bungs (Seed et al., 2006)] and the position of the painted trap was the same as in the functional trap and fake trap conditions.

Table 2
Name, gender, age, rearing history, study participation and previous experience on tool-use tasks.

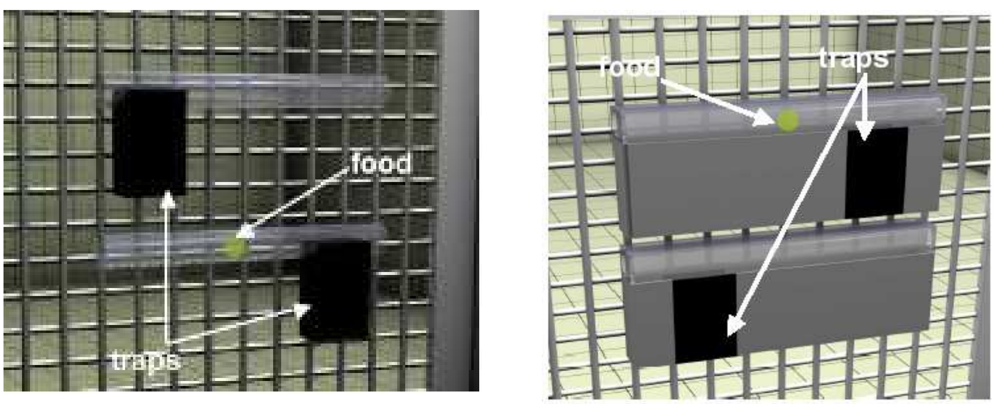
Subject	Gender	Age (years)	Rearing History	Study Participation	Previous Experience on Tool-use Tasks
Chimpanzee					
Robert	M	34	Nursery raised	2	c
Frodo	M	14	Mother raised	2	c
Unyoro	M	10	Nursery raised	2	c
Patrick	M	10	Mother raised	2	c
Alex	M	6	Nursery raised	2	c
Fraukje	F	31	Nursery raised	2	a,c
Dorien	F	27	Nursery raised	2	c
Trudi	F	14	Mother raised		a,c
Sandra	F	14	Mother raised	1,2	a,
Jahaga	F	14	Mother raised	2	a,c
Fifi	F	14	Mother raised	1,2	a[2],b,c
Pia	F	8	Mother raised	1,2	a[2],c
Alexandra	F	8	Nursery raised	2	c
Annett	F	8	Nursery raised	2	c
Bonobo					
Joey	M	25	Nursery raised	2	a,b,c
Limbuko	M	12	Nursery raised	2	a,c
Kuno	M	11	Nursery raised	1,2	a[2],c
Ulindi	F	14	Mother raised	1,2	a[1],b,c
Yasa	F	10	Unknown	2	a,c
Orangutan					
Bimbo	M	27	Nursery raised	2	a,b,c
Walter	M	18	Mother raised	2	b,c
Dunja	F	34	Mother raised	2	a,c
Pini	F	19	Mother raised	1,2	a[2],b,c
Dokana	F	18	Mother raised	1,2	a[2],b,c
Toba	F	13	Mother raised	2	b,c
Padana	F	10	Mother raised	2	a,c
Gorilla					
Gorgo	M	26	Nursery raised	2	a,c
Ndiki	F	28	Mother raised	2	a,c
Bebe	F	28	Mother raised	2	a,c
Viringika	F	12	Mother raised	1,2	a[1],b,c
Ruby	F	10	Nursery raised	2	c

a = trap tube and trap platform tasks [a[1]=subjects who only solved the tube task; a[2]=subjects who solved the trap tube and trap platform task (Martin-Ordas et al., 2008)]
b= trap tube task (Mulcahy & Call,2006)
c= trap table task (Girndt, Meier, & Call, 2008)

We also used two Plexiglas tubes for the painted trap condition (see Figure 1), each 40 cm long with a diameter of 5 cm. Each tube was placed inside a Plexiglas box (40 cm long x 5 cm deep x 20 cm high). The boxes were opaque except for the upper part (40 cm long x 5 cm high), where the tubes were attached. A black rectangle (8 cm wide x 15 cm long) simulating the trap was painted on the front wall of each box. One box had the black rectangle on the right side (8 cm from the right end of the tube) and the other on the left side (8 cm from the left end of the tube). The distance between the reward and the trap was identical in the three conditions: 8 cm. All tubes had a small hole (1 cm diameter) drilled in their back wall to allow the experimenter to bait them. The tool consisted of a straight wooden dowel (0.5 cm) with a length of 40 cm. We used grapes halves as rewards.

Procedure

Subjects were tested individually in their sleeping cages after they were separated from their groupmates. Young infants stayed with their mothers while the test took place. There were three conditions: the functional trap tube, the fake trap tube and the painted trap tube. In the three conditions, the experimenter fastened both trap-tubes to the mesh inside the testing cage. The tubes were placed one above the other with their respective traps facing opposite sides (Figure 1). The location of the trap (left-right) for each tube changed across sessions, but remained the same during a session. In the fake trap tube and painted trap tube conditions and before the test started, the experimenter placed four rewards inside the two openings of each tube. Subjects were allowed to inspect the tubes without holes by dragging the rewards out of them without any tool (with their fingers). Then, the experimenter began the test. The procedure was the same for the three conditions. The experimenter placed the reward inside one of the tubes through the baiting hole and gave the tool to the subject through the mesh right above the apparatus. In order to get the reward, subjects had to insert the stick through one of the sides of the tube and rake or push the reward out of the tube. In the functional trap tube condition, subjects could get the reward only from one side of the tube (side with no trap), while in the fake trap tube and the painted trap tube conditions, they could get the reward from both sides of the tube.



(a) functional trap-tube and fake trap-tube (b) painted trap-tube

Figure 1. Experimental setup for the functional trap tube (a), fake trap tube (a) and painted trap tube (b) conditions

Subjects received the conditions in the following order: first, subjects were given the functional trap tube; second, they were presented with the fake trap tube. Those subjects who did not avoid the fake trap received again the functional trap tube condition. In contrast, those subjects who avoided the fake trap at above chance levels received the painted trap tube condition, followed again by the functional trap tube condition. We administered a total of four 12-trial sessions in the functional trap tube condition: two sessions at the beginning of the experiment and two at the end.

Subjects received three 12-trial sessions in each of the other two conditions. The position of the trap and which tube was baited were counterbalanced across trials within a session so that the trap appeared the same number of times to the left and to the right of the subject. Each session lasted approximately 20 to 30 minutes.

Scoring and Analyses

We videotaped all trials. In the functional trap tube condition, we scored whether subjects succeeded in retrieving the reward; in the fake trap tube and painted trap tube conditions, we scored from which side of the tube they obtained the reward: either from the side of the trap or the side without the trap. We calculated the percentage of trials in which subjects got the reward (functional trap tube condition) and the percentage of trials in which subjects obtained the reward from one side or the other (fake trap tube and painted trap tube conditions). We used an analysis of variance (ANOVA) to investigate the effect of condition of the percent of correct trials. We also assessed whether subjects performed above chance levels in each condition using a one-sample t-test (with 50% as the chance expected value). We used the binomial test to assess the subjects' performance in the first trial and also to analyze whether certain individuals differed from chance. All statistical tests were two-tailed.

Results

Table 3 presents the percentage of trials in which subjects avoided the trap in the functional (first and second tests) and fake trap conditions. An ANOVA with condition as within-subject factor indicated that there were significant differences between conditions ($F_{2,14} = 6.01$, $p = 0.013$). Post-hoc LSD analyses indicated that subjects avoided the trap significantly less in the fake compared to the first functional trap condition ($p = 0.030$). In contrast, there were no significant differences between the first and the second functional trap conditions (p=0.080) or the fake and the second functional trap conditions ($p = 0.072$). Overall, subjects avoided the trap significantly above chance in all conditions ($t_7 > 4.48$, $p < 0.005$, in all conditions).

However, there were important individual differences. Table 3 shows the subjects' individual performance in each condition. Whereas four subjects stopped avoiding the trap in the fake condition, four others continued to avoid it. Interestingly, subjects who had previously succeeded on both the trap tube and trap platform tasks in the Martin-Ordas et al. (2008) study (Table 2) also avoided the fake trap less often than those who had only solved the trap tube task ($t_6 = -3.10$, $p < 0.05$). Nevertheless, the four subjects that failed to avoid the trap in the fake condition did not avoid the trap in the painted trap condition ($t_3 = 2.332$, $p = 0.102$). All subjects responded correctly in the first trial of the first functional trap condition (Binomial test: $p = 0.008$), but only 3 out of 8 (38%) did so in the second functional trap condition. Subjects were not above chance in the first trial of the fake trap (7 out of 8; Binomial test: $p = 0.070$) or painted trap conditions (Binomial test: $p = 1$).

Table 3
Percentage of correct responses in the three trap conditions and their associated p values (binomial test).

	functional (first exposure)		**fake**		**painted**		**functional (second exposure)**	
Subject	%correct	p	%correct	p	%correct	p	%correct	p
Fifi	91.66	<0.001	66.66	0.065	-	-	83.33	0.002
Sandra	83.33	0.002	86.11	<0.001	50	1	83.33	0.002
Pia	87.5	<0.001	77.77	0.001	58.33	0.405	71	0.064
Kuno	83.33	0.002	58.33	0.405	-	-	66.66	0.152
Ulindi	75	0.023	77.77	0.001	52.77	.868	79.16	0.007
Pini	87.5	<0.001	66.66	0.065	-	-	83.33	0.002
Dokana	91.66	<0.001	50	1	-	-	79.16	0.007
Viringika	95.83	<0.001	91.66	<0.001	58.33	0.405	100	<0.001

We found no evidence that subjects increased or decreased their performance across sessions in any of the four conditions (first functional trap: $t_7 = 0.813$, $p = 0.443$; second functional trap: $t_7 = -1.158$, $p = 0.285$; fake trap: $F_{2,14} = 0.401$, $p = 0.667$; painted trap: $F_{2,6} = 0.179$, $p = 0.840$).

Discussion

Our results showed a mixed picture. Although overall subjects avoided the functional trap more often than the non-functional traps, subjects still avoided the non-functional fake trap, but not the painted trap, above chance levels. Moreover, there were important individual differences in trap avoidance. Whereas four subjects only avoided the functional trap, four others also avoided the fake non-functional trap. None avoided the painted non-functional trap.

The evidence reported here indicates that at least some subjects distinguished between those features that could cause the reward to be lost from those features that could not. This result is hard to explain if subjects are merely learning an instrumental response (insert the stick in a particular location depending on the trap position) without any knowledge of the features of the task because the safest approach would be to continue using the same response. The fact that those same subjects were also the ones that had performed better in a previous study (Martin-Ordas et al., 2008) lends credence to the hypothesis that they may have possessed more general knowledge about what makes a trap effective. One remaining possibility is that subjects did not avoid the fake trap because they had a general predisposition against displacing food over holes but without attributing any causal properties (i.e. they can make the food fall) to them. This would also explain why they performed well not just on the trap-tube task but also the platform-trap task in the Martin-Ordas et al. (2008) study. We will address this issue in the next Experiment.

Our control condition of rendering the trap ineffective could be criticized on the same grounds as the inverted trap tube; that is, given that there is no cost in continuing to avoid the ineffective trap, it is not suitable to conclude that subjects lack a causal understanding of the task (Seed et al., 2006; Silva et al., 2005; Tebbich, Seed, Emery, & Clayton, 2007). Indeed, the data for those four subjects that avoided the non-functional fake trap are hard to interpret. However, the results of the painted condition demonstrate that they were not using the position, color and shape of the trap as a discriminative cue to decide where to insert the tool. There is another aspect that complicates the interpretation of this experiment. Subjects avoided the trap significantly more often in the first functional condition than in the fake condition, but they only showed a tendency in the same direction between the fake trap condition and the second functional trap condition. Indeed, there was a tendency to reduce trap avoidance when comparing the first and the second functional trap conditions. This result is hard to explain if subjects were solely attending to the causal features of the task.

The use of between-task generalization is one way to investigate whether subjects use specific features to solve problems or more general processes. If the latter is true, one would expect subjects to be able to generalize their skill to new functionally equivalent problems, despite the differences between the new task and the original task in terms of the setup and its response requirements. In the next experiment we changed the setup substantially and investigated whether apes solve a new trap problem and whether their performance correlated with that of other trap tasks. This new setup also allowed us to test the possibility that subjects had succeeded in trap task because they had a predisposition for refraining from displacing rewards over holes.

Experiment 2: Between-task Generalization: The Gap Task

In this experiment subjects did not have to use a tool to get a reward while avoiding a trap. Instead they had to infer the location of the reward (placed under an upside down cup) based on the cups' trajectory on a platform in the presence of a trap.

In particular, we presented apes with a platform with square holes 10 cm from the center of the platform. A pair of opaque cups was positioned upside down next to each trap. One of the traps was covered with a solid plastic cover. We hid a reward under the cup closest to the covered hole and displaced both cups in succession over the holes and towards the ends of the platform. Subjects did not have to manipulate tools to avoid a trap but simply had to select the correct cup to get the reward. One advantage of this procedure is that it minimized the potential problems associated with motor constraints or biases that may appear when manipulating a tool. Another advantage is that the procedure did not involve the subjects displacing objects over traps and it also allowed us to assess whether subjects showed a general avoidance for objects located near traps.

Methods

Subjects

Thirty-one apes (14 chimpanzees, 5 bonobos, 7 orangutans, and 5 gorillas) were tested in this experiment. All had participated in the previous experiments (see Table 2). There were 11 males and 20 females with an age range of 6 years to 34 years.

Apparatus

We used a grey plastic platform with two holes (80 cm wide x 25 cm long) located 10 cm from the center of the platform to present the task (see Figure 2). The platform was mounted on an L-shaped metal support frame perpendicular to a solid Plexiglas panel with three holes cut on its bottom part equidistant from each other (two on opposite corners and another one in the center). Two grey opaque plastic pieces (17 cm wide x 15 cm long) and a clear Plexiglas piece (17 cm wide x 15 cm long) were used to cover the platform holes in different conditions. Two opaque blue plastic cups (7 cm diameter x 12.5 cm high) were used to cover and displace the reward on the platform and an opaque screen was used to bait one of the cups outside of the subject's view. We used grapes, banana pieces or monkey chow pellets as rewards.

Figure 2. Experimental setup for the gap task.

Procedure

The experimenter sat behind the platform facing the subject who was located behind the Mesh panel. The platform rested on a metal support that allowed the experimenter to pull it away to prevent subjects from touching the test materials during their presentation. He placed the two cups upside down side by side on the center and forward position of the retracted platform. Then the experimenter raised the screen, showed the reward to the subject and placed it between the two cups and alternatively lifted and dragged both cups over the reward's position capturing it with a predetermined cup. At this point subjects still did not know under which of the two cups the reward had been hidden. Then, the experimenter removed the screen and dragged the left cup towards the left side crossing over the left hole and leaving the cup in front of the left hole in the panel on the far left side of the platform. Next, the experimenter executed the same action on the cup on the right side of the platform. Upon completing both cup displacements, the experimenter pushed the platform forward and allowed the subject to select one of the cups by inserting a finger in one of the two extreme holes and touching the cup. We administered four different conditions:

Clear-Open condition. The experimenter covered one hole with the transparent Plexiglas piece and left the other hole uncovered. The reward was always on the side where the Plexiglas piece had been placed.

Opaque-Opaque condition (control)**.** The experimenter covered each of the holes with one of the opaque plastic pieces.

Post Clear-Open condition (control). The experimenter followed the same procedure as in the previous condition. However, once he dragged both cups towards the left side and right side of the platform, respectively, he removed both opaque plastic pieces and covered the hole next to the baited cup with the Plexiglas piece and left the other hole uncovered.

Clear-Clear condition (control). This condition was identical to the Opaque-Opaque condition except that Plexiglas pieces rather than opaque plastic pieces covered the holes. This condition allowed us to assess whether subjects could detect the food location when the holes were covered by Plexiglas pieces**.**

Each subject received two 12-trial sessions composed of the Clear-Open and Opaque-Opaque conditions (6 trials per condition per session). We counterbalanced the order of these conditions within subjects. Subsequently, subjects received one session (12 trials) in the Post Clear-Open condition and one session (12 trials) in the Clear-Clear condition in this order. We counterbalanced the location of the reward in each condition so that it appeared the same number of times on each of the sides with the stipulation that it could not appear more than three consecutive trials on the same side.

Data Scoring and Analyses

We videotaped all trials. We scored the first cup touched by the subject as her choice. Our dependent variable was the percent of correct trials. To analyze the results of the gap task, we used nonparametric statistics because the Opaque-Opaque control condition violated the assumption of homogeneity of variance. Therefore we used the Kruskal-Wallis test to assess species differences and the Wilcoxon test to investigate the effect of condition of the percent of correct trials. We also used the Wilcoxon test to assess whether subjects performed above chance levels in each condition (with 50% as the chance expected value). We used the binomial test to assess the subjects' performance in the first trial. All statistical tests were two-tailed. To analyze the generalization across tasks, we used Pearson's correlation. Additionally, we conducted a factor analysis that included the scores of the four tasks to detect potential clusters of tasks. We plotted the distribution of the four tasks based on the two main factor components.

Results

Figure 3 presents the mean percent of correct trials as a function of condition. There were no significant differences between species in the Clear-Open experimental (Kruskal-Wallis test: $\chi^2 = 4.74$, df = 3, $p = 0.19$) and Opaque-Opaque control conditions (Kruskal-Wallis test: $\chi^2 = 5.43$, df = 3, $p = 0.14$). Therefore we pooled together the data from the various species. Subjects performed significantly better in the experimental than the Opaque-Opaque control condition (Wilcoxon test: $z = 3.30$, $p = 0.001$, N = 26). Moreover, subjects were above chance in the Clear-Open experimental (Wilcoxon test: $z = 4.36$, $p < 0.001$, N = 26), but not in the Opaque-Opaque control condition (Wilcoxon test: $z = 0.65$, $p = 0.51$, N = 27). This difference was already evident in the first trial, Sign test: $p = 0.019$. Twenty-two out of 31 subjects (71%) responded correctly in the first trial of the Clear-Open experimental condition (Binomial test: $p = 0.029$), but only 11 out of 31 subjects (35%) did so in the Opaque-Opaque control condition (Binomial test: $p = 0.15$). Subjects also failed to perform above chance levels in both the Post Clear-Open (Wilcoxon test: $z = 1.70$, $p = 0.09$, N = 17), and Clear-Clear control conditions, (Wilcoxon test: $z = 1.37$, $p = 0.17$, N = 14).

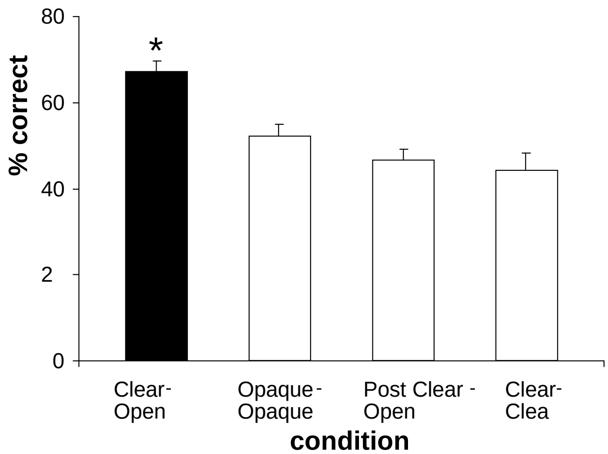


Figure 3. Mean percentage of correct trials in the experimental condition (clear-opem) and in the control conditions (opaque-opaque, post clear-open and clear-clear).

Table 4 presents the correlations between the four trap tasks (gap task, trap-platform task, trap table and modified version of the trap tubes). There was a significant correlation between the table and the gap tasks and between the table and the tube conditions. A factor analysis on the four tasks indicated that the two main components depicted in Figure 4 accounted for 65% of the variance. All tasks except for the platform task tended to cluster together (Figure 4). Indeed the platform task showed both the lowest loadings on the first component and the highest ones on the second component.

Table 4
Correlations between the four trap tasks.

	gap	**table**	**tube**	**platform**
gap	-	r = 0.49** n = 24	r = 0.04 n = 17	r = 0.01 n = 17
table	-	-	r = 0.421* n = 17	r = 0.09 n = 17
tube	-	-	-	r = -0.10 n = 19

*$p < 0.05$
**$p < 0.01$

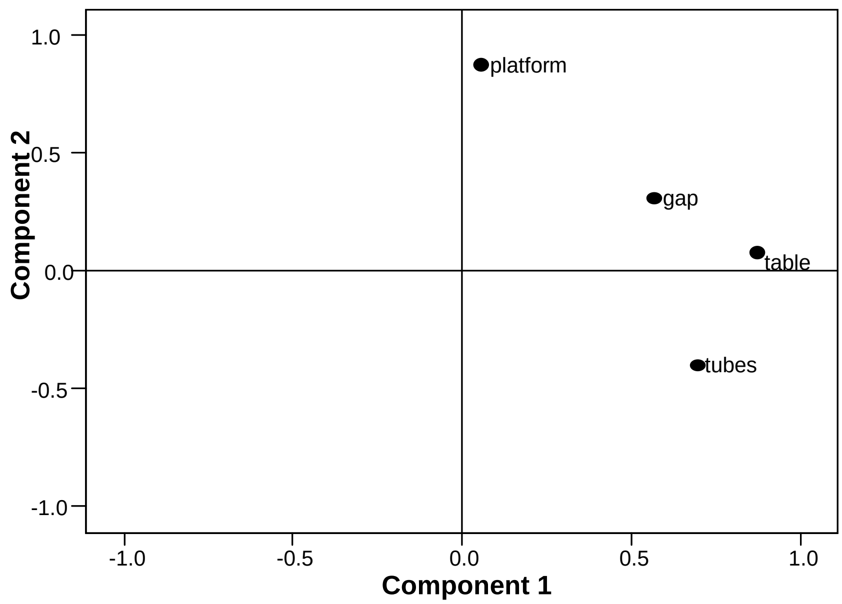


Figure 4. Factor analysis of the four trap tasks based on two factor components.

Discussion

Overall, subjects found the reward above chance levels in the Clear-Open experimental condition but not in any of the three control conditions. This difference was already evident in the first trial. Additionally, the Post Clear-Open control condition showed that they did not have any predisposition for avoiding objects located next to holes. We found no evidence of inter-specific differences. When we analyzed the level of subjects' performance across the different tasks (gap task, trap-platform task, trap table and modified version of the trap tubes), we found positive correlations between the table task and the gap and tube tasks. However, the performance in the trap-platform did not correlate with any of the other tasks.

These results cannot be reduced to an explanation such as using the presence of a hole as a discriminative cue to find the food because subjects did not succeed in the condition in which the hole appeared after displacing the cups. Likewise, the use of inadvertent cues, such as the baiting procedure or reflection of the food on the Plexiglas pieces, cannot account for these results because subjects failed to choose the correct cup when there was no hole on the platform. Additionally, given that 71% of the subjects responded correctly in the first trial of the experimental condition, learning how to respond during the test cannot explain our findings. Therefore, our results support the idea that subjects may be using inferential reasoning to solve the gap task – a skill that has been documented in several previous studies (Call, 2004, 2006; Call & Carpenter, 2001; Premack &

Premack, 1994). However, unlike some of these studies, in which subjects could choose on the basis of witnessing an empty container (Call, 2006; Call & Carpenter, 2001) or the experimenter consume the contents of one of the containers (Premack & Premack, 1994), our subjects had to choose on basis of the presence or absence of the trap (and the effects that it may have had on the reward).

The positive correlations found between tasks that differed both in their setup and response requirements suggested that subjects were able to generalize their knowledge about particular tasks to other functionally equivalent trap tasks. However, this generalization was not extensive given that we found no significant correlations between most of the tasks. Most striking was the absence of any correlation between the platform task and the other tasks.

General Discussion

We found some evidence suggesting that apes have some knowledge about the relation between elements in trap tasks. On the one hand, the within-task trap-tube generalization test indicated that some subjects attributed a special importance to the presence of the hole, and even those that did not, they did not merely use the position of a black patch (painted trap) to solve the task. On the other hand, the between-task generalization revealed positive evidence of generalization between some functionally equivalent trap tasks that differed in their task elements and response modes. Most telling was the success in the gap task (which required inference by exclusion) and its positive relation with some of the tool-using tasks. We also found a positive correlation between the trap tube and the table tasks.

The combination of generalization data from within and between tasks leads us to postulate that subjects possess some general knowledge (as opposed to task-specific) about the effects that traps can cause on rewards moving over them. Recall that each task presented a different type of trap that required a different response to obtain the reward. Moreover, subjects often succeeded on the first trial in which they were exposed to the task. These results cannot be explained by invoking associations between familiar situations because they involved tasks not encountered before. Indeed, they suggest that subjects could be extracting the relationship between the elements of the different tasks. However, one could argue that subjects had a predisposition (learned or not) to simply avoid displacing the reward over traps, a feature that all tasks shared, but without understanding the functional properties of traps. Yet, this explanation cannot account for the results of the gap task in which subjects did not have to displace the reward at all but simply observe the displacements of cups and touch one cup. Another alternative is that subjects may have had a generalized aversion to objects being displaced over traps. However, this explanation again cannot account for some of the available data. Seed, Call, Emery, and Clayton (2008) found that chimpanzees can either avoid displacing the rewards over traps or push them into traps depending on whether the reward gets trapped or released from the apparatus, respectively. This means that there is no generic predisposition for avoiding traps. Instead, apes treated traps as a part of the problem or the solution depending on the situation.

The positive evidence for generalization cannot obscure the fact that there was also a lack of generalization at the individual and task level for both within and between tasks. At the individual level, some subjects failed the non-functional (fake trap) condition whereas other subjects that passed this condition nevertheless showed a subsequent performance decrease in the functional task. We do not think that such a decrease was due to a lack of motivation because subjects participated willingly throughout the experiment. Intead, we favour the idea that such reduction may constitute a predisposition to solve the problem in a fixed and specific manner even though there are more appropriate methods of solving it; that is, subjects could have learned to not avoid the fake trap in the fake trap condition such that this predisposition might bias them to not avoid the functional trap. This phenomenon has been called the *Einstellung* effect and it has been previously described for human adults in problem solving situations (Luchins & Luchins, 1959). At the task level, the generalization did not occur across all tasks. One could argue that if causal knowledge was implicated in the solution of these tasks, there should have been positive correlations among all tasks. However, there are some factors that may help explain the absence of a perfect generalization among tasks. From a practical point of view, it is conceivable that the different motor and perceptual requirements to solve each task introduced some noise that contributed to reduce the correlation across tasks. All other things being equal, those tasks requiring motor responses that were easier to execute would automatically lead to higher scores and consequently reduce the correlation between tasks. Similarly, the perceptual components of the task need to be carefully considered. Two spatial arrangements that the experimenter judged as functionally equivalent may not be perceived in the same way by the apes. Moreover, it is likely that the different spatial arrangement of each task generated some novelty effect that may have caused a dissimilar appearance of the problems. Brown (1990) showed that children also rely on surface features when they are faced with a spontaneous generalization tasks involving tool use. Even though, the domain-specific or task-specific knowledge might be fundamental to build up conceptual knowledge (Brown, 1990; Mandler, 2000), a dependence on stimulus-specific attributes in a transfer task may diminish the flexible use of the mechanisms involved in the solution of the problems (e.g., Oden, Thompson, & Premack, 1990).

From a theoretical point of view, Martin-Ordas et al. (2008) argued that failing to find evidence of positive transfer between tasks does not necessarily mean that subjects do not perceive the causal relations between the elements within each task. It may be that subjects are not comparing those relations across tasks. In other words, analogical rather than causal knowledge may be the reason for the lack of generalization across tasks. Obviously, subjects that can establish analogical relations between tasks would also be able to recognize the causal relations within each task (Gentner, 1998, 2002; Gentner & Kurtz 2006; Gentner & Markman, 1997). But the reverse is not true. Thompson and Oden (2000) argued that generalized performance to novel problems or situations is a necessary, but not sufficient criterion for concluding that an animal's performance is mediated by analogical reasoning. One of the strongest evidence that an individual has developed knowledge at a relational level involves first trial transfer to novel and same class samples, but which differ physically from the original problem (Gentner, 1983; Oden, Thompson, & Premack, 1988, 1990; Roitblat & Fersen 1992; Thompson & Oden 2000; Thompson, Oden, & Boysen, 1997). Our results showed that subjects only solved the trap-table and the gap-task in the first trial. In the case of the trap-platform, subjects were above chance in the first session and no evidence of improvement across sessions was found (Martin-Ordas et al., 2008).

In conclusion, we found some evidence suggesting that apes possess some knowledge about the effects that traps have on slow moving unsupported objects. Such knowledge was apparent both in tasks in which subjects had to displace the objects as well as those in which they simply observed the displacement to make an inference about the reward location based on object interactions. However, this knowledge was not robust enough to prevent the influence of certain practice and task effects. In fact, the subjects' responses can be best explained as a combination of epistemic (knowledge), practical (actions required) and situational factors (task setup), not the unique influence of one of them. Moreover, subjects' knowledge may not have been abstract enough to allow them to establish broad analogies between tasks. Future studies are needed to assess the relative contribution of each of these factors in the solution to these problems and the nature of the causal knowledge that may support analogical capabilities in nonhuman animals.

References

Brown, A.L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science, 14*, 107–133.

Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus*). *Journal of Comparative Psychology*, *118*, 232-241.

Call, J. (2006). Inferences by exclusion in the great apes: The effect of age and species. *Animal Cognition, 9,* 393-403.

Call, J. & Carpenter, M. (2001). Do chimpanzees and children know what they have seen? *Animal Cognition, 4,* 207-220.

Gentner, D. (1983). Structure Mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155-170.

Gentner, D. (1998). Analogy. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 107-113). Oxford: Blackwell.

Gentner, D. (2002). Analogical reasoning, psychology of. *Encyclopedia of Cognitive Science*. London: Nature Publishing Group.

Gentner, D. & Markman, A.B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45-56

Gentner, D. & Kurtz, K. (2006). Relations, objects, and the composition of analogies. *Cognitive Science*, *30*, 609-642.

Girndt, A., Meier, T., & Call, J. (2008). Task constraints mask great apes' ability to solve the trap table task. *Journal of Experimental Psychology: Animal Behavior Processes, 34,* 54-62.

Henrich, B. (2000). Testing insight in ravens. In C. Heyes & L. Huber (Eds.), *The Evolution of Cognition* (pp. 289–305), Cambridge, Massachusetts: MIT Press.

Köhler, W. (1925). *The Mentality of Apes*. New York: Vintage Books.

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

Luchins, A. S. & Luchins, E. H. (1959*). Rigidity of Behaviour: A Variational Approach to Einstellung*. Eugene, Oregon: University of Oregon Press.

Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development*, *1*, 3–36.

Martin-Ordas, G., Call, J., & Colmenares, F. (2008). Tubes, tables and traps: Great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Animal Cognition, 11,* 423-430.

Mulcahy, N.J. & Call, J. (2006). How great apes perform on a modified trap-tube task. *Animal Cognition*, *9*, 193-199.

Oden, D. L., Thompson, R. K. R., & Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes)*. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 140 –145.

Oden, D. L., Thompson, R. K. R., & Premack, D. (1990). Infant chimpanzees (*Pan troglodytes)* spontaneously perceive both concrete and abstract same/different relations. *Child Development, 61,* 621– 631.

Povinelli, D.J. (2000). *Folk Physics for Apes: A Chimpanzee's Theory of How the Mind Works.* Oxford: Oxford University Press.

Premack, D. & Premack, A. (1994). Levels of causal understanding in chimpanzees and children. *Cognition*, *50*, 347–362.

Roitblat, H. L. & von Fersen, L. (1992). Comparative cognition: Representations and processes in learning and memory. *Annual Review of Psychology*, 43, 671-710.

Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2007, March). *Was there a spanner in the works? Chimpanzees' performance on the trap problem revisited*. Poster presented at The Mind of Chimpanzee Conference, Chicago, Illinois.

Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition in rooks (*Corvus frugilegus*). *Current Biology*, 16, 697–701.

Silva, F.J., Page, D.M., & Silva, K.M. (2005). Methodological-conceptual problems on the study of chimpanzees' folk physics: How studies with adult humans can help. *Learning & Behavior*, *33*, 47–58.

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

Tebbich, S., Seed, A. M., Emery, N. J., & Clayton, N. S. (2007). Non-tool-using rooks (*Corvus frugilegus*) solve the trap-tube task. *Animal Cognition*, *10,* 225-231.

Thompson, R .K. R. & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, *24*, 363–396.

Thompson, R .K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 31– 43.

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

Visalberghi, E. & Limongelli, L. (1996). Acting and understanding: Tool use revisited through the minds of capuchin monkeys. In A. E. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching Into Thought: The Minds of Great Apes* (pp. 57-79), Cambridge: Cambridge University Press.