

# Problem solving in great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, and *Pongo abelii*): the effect of visual feedback

Christoph J. Völter · Josep Call

Received: 27 January 2012/Revised: 3 May 2012/Accepted: 15 May 2012/Published online: 30 May 2012  
© Springer-Verlag 2012

**Abstract** What kind of information animals use when solving problems is a controversial topic. Previous research suggests that, in some situations, great apes prefer to use causally relevant cues over arbitrary ones. To further examine to what extent great apes are able to use information about causal relations, we presented three different puzzle box problems to the four nonhuman great ape species. Of primary interest here was a comparison between one group of apes that received visual access to the functional mechanisms of the puzzle boxes and one group that did not. Apes' performance in the first two, less complex puzzle boxes revealed that they are able to solve such problems by means of trial-and-error learning, requiring no information about the causal structure of the problem. However, visual inspection of the functional mechanisms of the puzzle boxes reduced the amount of time needed to solve the problems. In the case of the most complex problem, which required the use of a crank, visual feedback about what happened when the handle of the crank was turned was necessary for the apes to solve the task. Once the solution was acquired, however, visual feedback was no longer required. We conclude that visual feedback about the consequences of their actions helps great apes to solve

complex problems. As the crank task matches the basic requirements of vertical string pulling in birds, the present results are discussed in light of recent findings with corvids.

**Keywords** Problem solving · Causal cognition · Puzzle box · Perceptual-motor feedback

## Introduction

The last decade has seen renewed efforts to investigate how much nonhuman animals know about causality (for a review see Seed and Call 2009). Problem-solving tasks, most of them based on tool using, have been a major component in this research effort (e.g., Cheke et al. 2011; Mulcahy et al. 2005; von Bayern et al. 2009).

In order to distinguish whether successful performance is based on causal knowledge or other sensorimotor processes, researchers have contrasted the outcomes of different experimental manipulations. In general, two main types of contrasts have been used: transfer tests and alternative tests. Transfer tests typically consist of a follow-up test conducted after subjects have mastered the original task. These follow-up tests usually involve changing the configuration of the elements of the task or even the type of stimuli presented. Passing such transfer tests is interpreted as an indication of cognitive abilities going beyond the scope of basic sensorimotor processes.

Alternative tests are the second type of contrast used by researchers in this area, and they are presented either as a follow-up to the original test (within-subject design) or in parallel to a different group of subjects (between-subject design). Alternative tests can in turn be subdivided into two basic types: the arbitrary cues type and the opaque type.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10071-012-0519-5) contains supplementary material, which is available to authorized users.

---

C. J. Völter (✉) · J. Call  
Department of Developmental and Comparative Psychology,  
Max Planck Institute for Evolutionary Anthropology,  
Deutscher Platz 6, 04103 Leipzig, Germany  
e-mail: christoph\_voelter@eva.mpg.de

C. J. Völter  
Department of Animal Behaviour, University of Bielefeld,  
Bielefeld, Germany

In the arbitrary cues type, subjects are confronted with cues that are perceptually identical to those of the original test but that hold an arbitrary rather than a causal relation with the presence of the reward. For instance, the inclination of an object can be caused by a piece of food being placed under it (e.g., an inclined board supported by a piece of banana) or by the shape of the object itself (e.g., a wedge). In both cases, the selection of the slanted stimulus (as opposed to a nonslanted stimulus) is invariably reinforced; however, the reward causes the cue only in one of them (cf. Call 2007).

In the opaque type of contrast, two groups of subjects are compared, one that has visual access to the inner workings of the apparatus and another one that does not. If subjects are solving a task by merely attending to the effect that their actions have on producing the reward (ignoring the mechanical causes for its delivery), those two approaches should not differ substantially. For instance, one may learn to operate a machine dispensing bubble gum not based on an understanding of its inner workings but because the handle leads one to grab and manipulate it. However, subjects seeing the inner workings may nevertheless have an advantage over subjects that are denied such access. Comparing subject's performance depending on visual access to the inner workings of an apparatus offers us a way to assess how much they gain from such privileged information.

Surprisingly, the latter approach has received less attention than the other two, even though it is perhaps a more obvious approach. In fact, there are very few studies that have systematically varied visual access to the inner workings of puzzle boxes. Rensch and Döhl (1967) found that a juvenile chimpanzee was able to solve a great variety of different puzzle boxes spontaneously and without trial and error. Interestingly, failure or overt trial-and-error learning was mainly observed in those puzzle boxes in which the causal relationships of the opening mechanisms were not visually accessible. Horner and Whiten (2005) systematically manipulated visual accessibility of a puzzle box in a social learning task. They found that chimpanzees only imitated irrelevant parts of a demonstration when visual information on causal relationships of the task was unavailable (opaque puzzle box). In contrast, when causal information was accessible (transparent box), chimpanzees only reproduced the relevant parts (from a causal point of view) of the demonstration while excluding irrelevant parts. These findings suggest that chimpanzees benefited from visual feedback of the effects of demonstrated actions on a puzzle box problem.

Despite the obvious differences in brain morphology and behavior between corvids and apes, both taxa have displayed remarkable problem-solving abilities (e.g., Weir et al. 2002; Bird and Emery 2009; Mendes et al. 2007;

Kenward et al. 2005; Boesch and Boesch 1990). An intriguing question is to what extent these problem-solving skills are supported by similar cognitive processes (Emery and Clayton 2004). With regard to the effects of visual feedback on corvids' problem solving, a vertical string pulling task was particularly informative. Several species of birds, mainly corvids and psittacids, can spontaneously use a continuous sequence of pulling up the string with the beak and stepping on loops of the string with their feet to obtain the food reward suspended by a string from a perch (Funk 2002; Heinrich 1995, 2000; Pepperberg 2004; Taylor et al. 2010; Werdenich and Huber 2006). However, as pointed out by Taylor et al. (2010), a positive perceptual-motor feedback cycle rather than some form of insightful problem solving may explain this result. Accordingly, pulling up the reward and stepping on the string move and hold the reward closer to the animal. As the movement of the reward is always within sight of the bird, this perceptual feedback may act as incremental reinforcement which shapes this complex behavior. And indeed, Taylor et al. found evidence for such an operant conditioning account: when visual feedback of the moving reward was restricted, naïve New Caledonian crows (*Corvus moneduloides*) either were not able to solve this problem or did so only by means of overt trial-and-error learning. Interestingly, even experienced crows that solved the problem spontaneously when visual feedback was available had much more problems with the visually restricted version of the task: the efficiency of their actions dropped, and the number of errors increased by a factor of 10. Therefore, Taylor and colleagues concluded that the most likely explanation for the behavior of the crows in the string-pulling task was not causal understanding of the string connectivity but rather operant conditioning.

Some authors have proposed that primates and corvids have undergone convergent evolution of cognitive abilities including causal reasoning, flexibility, imagination, and prospection (Emery and Clayton 2004; Seed et al. 2009). An important issue in this regard is to what extent corvids and primates benefit from visual feedback in a problem situation. As described above, the results by Taylor et al. (2010) suggest that crows heavily depend on perceptual feedback when solving the string-pulling task. In the present study, we examined—using three different apparatuses—whether great apes benefited from visual inspection of relevant task features and to what extent they also relied on a perceptual-motor feedback cycle.

In the simplest problem (Experiment 1), the apes had to push a stick that was pre-inserted inside a horizontal tube to displace an out-of-reach food reward outside of the tube. We manipulated the visibility of the stick across subjects. In the tower problem (Experiment 2), four sticks that supported the food reward had to be removed from a

Plexiglas cuboid. Removing the sticks made the reward fall down to the ground where it became accessible to the subject. We manipulated the visibility of the relation between the sticks and the reward across subjects. Finally, in the crank apparatus (Experiment 3), the apes had to use a crank mechanism to extract a reward located inside a box. We varied the visual information available about both the crank mechanism and the progress of the reward. Since this task closely matched the basic requirements of the string-pulling task (i.e., sequential, continuous actions requiring the coordination of different body parts), it afforded a comparison between our results and those of Taylor et al. (2010).

## Experiment 1

### Subjects

Four bonobos (*Pan paniscus*), twelve chimpanzees (*Pan troglodytes*), three gorillas (*Gorilla gorilla*), and six orangutans (*Pongo abelii*) participated in this experiment. The subjects were housed at the Wolfgang Köhler Research Center, Leipzig Zoo (Leipzig, Germany). There were 17 females and eight males aged between 7 and 30 years. Eight subjects were nursery-reared and 17 mother-reared. Subjects had participated in various cognitive tasks prior to the study (see Table 1 for detailed information on each subject). Subjects lived in social groups of different sizes and had access to indoor (175–430 m<sup>2</sup>) and outdoor areas (1,400–4,000 m<sup>2</sup>). They were tested individually in special testing rooms (5.1–7.3 m<sup>2</sup>). Subjects were not deprived of food, and water was available ad libitum during testing.

### Materials

The apparatus consisted of a tube opened on both sides (length 40 cm, inner diameter 3 cm) (see Fig. 1). The tube was fixed horizontally on an acrylic glass panel (12 cm × 67 cm). In the clear version of the apparatus, the whole tube was transparent, whereas in the opaque version, 22 cm of the tube on one end was opaque. Rectangular barriers fixed at a distance of 5 cm from each end of the tube prevented the apes from looking inside the tube through the openings. The apparatus was mounted on the mesh inside the testing enclosure.

In the middle of the tube, a food reward (i.e., a piece of banana) was placed via a baiting hole in the back of the panel. The reward was visible to the subject in both conditions. Crucially, the subject could not reach the reward through the opening which was closest to the reward. On the other side of the tube, there was a stick (length 17 cm,

diameter 2.5 cm) that could be pushed further inside the tube in order to move the reward within range of the subject on the opposite side of the tube. In the clear condition, the stick was fully visible, whereas in the opaque condition, the stick was completely hidden.

### Procedure and design

Due to the high variability in subjects' finger length and diameter, we measured how far each individual could reach inside the tube before the first experimental trial. Therefore, all subjects were confronted with the clear tube without a stick inside. In the middle of the tube, there was a reward (a piece of monkey chow) out of reach of the subject. As the subjects tried to obtain the reward with their fingers, the maximal reach of each subject could be assessed. The position of the piece of banana that served as reward in the experiment was adjusted accordingly.

Half of the subjects received the clear apparatus, while the other half received the opaque apparatus. Assignment to the groups was random with the restriction that both groups were counterbalanced as much as possible for species, age, and sex (clear group: mean age 16.7, females 69 %; opaque group: mean age 15.8, females 68 %). In total, subjects got two trials on two test days (mean intertrial interval 9 days, range 7–10 days). Each trial lasted a maximum of 10 min and started when the subject approached the baited apparatus.

### Scoring and analysis

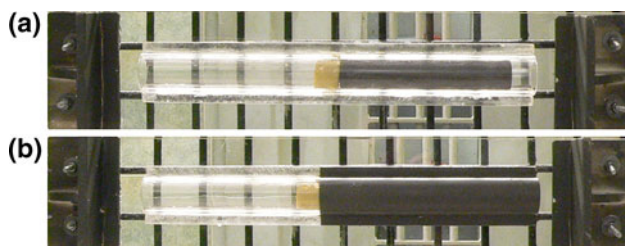
We coded the time the subject spent at the apparatus before she first pushed the stick inside the tube ( $T_S$ : time to solve the problem). We scored the trial as successful only if the subject extracted some of the reward after having pushed the stick inside. If the subjects failed to push the stick within 10 min, the trial was coded as not successful. All sessions were videotaped. A second coder (JC) scored 20 % of the trials to assess interobserver reliability, which was excellent (Spearman's correlation:  $r = 0.94$ ;  $N = 12$ ). Since the normality assumption was violated, we used non-parametric statistics (two-tailed) to analyze the effect of condition (clear vs. opaque), species, sex, and age on the  $T_S$ .

## Results

All subjects pushed the stick inside the tube and extracted the reward, except for one chimpanzee (Annett, clear condition) who approached the apparatus three times in the beginning of the first trial (T1) and after that completely lost interest (i.e., she never approached the apparatus any more in the course of trials 1 and 2). Additionally, one gorilla (Gorgo, clear condition) solved the task in T1 but

**Table 1** Species, age, sex, rearing history, and the experiments in which the subjects participated

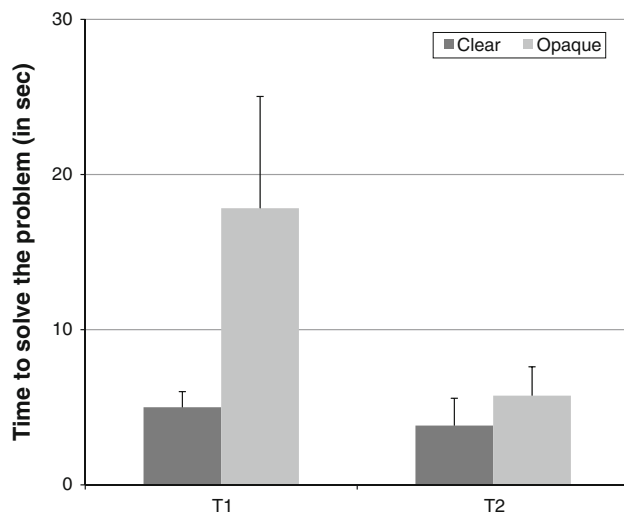
Name	Species	Sex	Age	Rearing history	Experiment participation
Joey	Bonobo	M	28	Nursery	1, 2, 3
Kuno	Bonobo	M	14	Nursery	1, 2, 3
Ulindi	Bonobo	F	17	Mother	1, 2, 3
Yasa	Bonobo	F	13	Mother	1, 2, 3
Alex	Chimpanzee	M	10	Nursery	1, 2, 3
Alexandra	Chimpanzee	F	11	Nursery	1, 2, 3
Annett	Chimpanzee	F	11	Nursery	1, 2
Corry	Chimpanzee	F	34	Nursery	2, 3
Dorien	Chimpanzee	M	30	Nursery	2
Fifi	Chimpanzee	F	17	Mother	1, 2, 3
Fraukje	Chimpanzee	F	34	Nursery	2
Frodo	Chimpanzee	M	17	Mother	2, 3
Gertrudia	Chimpanzee	F	17	Mother	1, 2, 3
Jahaga	Chimpanzee	F	18	Mother	1, 2, 3
Lome	Chimpanzee	M	9	Mother	1, 2, 3
Natascha	Chimpanzee	F	30	Nursery	1, 2, 3
Patrick	Chimpanzee	M	13	Mother	1, 3
Pia	Chimpanzee	F	11	Mother	1, 2, 3
Riet	Chimpanzee	M	33	Nursery	2, 3
Sandra	Chimpanzee	F	17	Mother	1, 2, 3
Ulla	Chimpanzee	F	33	Nursery	2, 3
Robert	Chimpanzee	M	35	Nursery	3
Tai	Chimpanzee	F	8	Mother	3
Gorgo	Gorilla	M	29	Nursery	1, 2, 3
Kibara	Gorilla	F	7	Mother	1, 2, 3
Viringika	Gorilla	F	15	Mother	1, 2, 3
Bimbo	Orangutan	M	30	Nursery	1, 2, 3
Dokana	Orangutan	F	22	Mother	1, 2, 3
Kila	Orangutan	F	10	Mother	1, 2
Padana	Orangutan	F	13	Mother	1, 2, 3
Pini	Orangutan	F	22	Mother	1, 2, 3
Raja	Orangutan	F	7	Mother	1, 2

**Fig. 1** Tube and stick problem: the stick could be further pushed inside the tube in order to move the reward (a piece of banana) closer to the opening on the *left side*. **a** Clear version (stick visible), **b** opaque version (stick not visible)

did not do so in the second trial (T2). These subjects were dropped from the subsequent analyses (remaining subjects:  $N = 23$ ).

There were no significant differences between species in the  $T_s$  for T1 and T2 (Kruskal–Wallis test: T1:  $\chi^2 = 4.84$ ,  $df = 3$ ,  $P = 0.18$ ; T2:  $\chi^2 = 5.38$ ,  $df = 3$ ,  $P = 0.14$ ). Splitting the data by condition produced similar results (clear condition: T1:  $\chi^2 = 4.63$ ,  $df = 3$ ,  $P = 0.20$ ; T2:  $\chi^2 = 4.23$ ,  $df = 3$ ,  $P = 0.24$ ; opaque condition: T1:  $\chi^2 = 5.49$ ,  $df = 3$ ,  $P = 0.10$ ; T2:  $\chi^2 = 3.50$ ,  $df = 3$ ,  $P = 0.33$ ). Thus, for further analyses, the data were collapsed across species.

Figure 2 shows the mean  $T_s$  for each trial as a function of condition. In T1, subjects in the clear group were on average more than three times as fast as those in the opaque group; however, this difference did not turn out to be statistically significant (Mann–Whitney test:  $U = 47.5$ ,  $P = 0.26$ ). In contrast, in T2, we found that the clear group tended to be faster than the opaque group (Mann–Whitney



**Fig. 2** Mean time to solve the tube and stick problem (+SEM) across species as a function of condition (clear vs. opaque). T1: first trial, T2: second trial

test:  $U = 36.0$ ,  $P = 0.06$ ). Overall, subjects were regardless of condition significantly faster in T2 compared to T1 (Wilcoxon's test:  $Z = 2.17$ ,  $P < 0.05$ ). Analyzing each condition separately, we found that the apes became faster in T2 than in T1 in the opaque condition ( $Z = 1.94$ ,  $P = 0.05$ ), but not in the clear condition ( $Z = 1.22$ ,  $P = 0.25$ ). The latter finding might be attributed to the fact that subjects in the clear condition required on average only 5 s in T1; thus, there was not much room for improvement in the clear condition.

## Discussion

Great apes of all species found the solution for the tube problem easily irrespective of the visual access to relevant parts of the task. Moreover, there were no strong differences in the speed of solving the task between the opaque and the clear conditions. Only in the second trial was there some indication that the subjects who could see the stick were faster at solving the task than those who could not see it. Unlike subjects in the clear condition, subjects in the opaque condition became faster at solving the task in the second trial compared to the first one.

One problem for interpreting these findings is that the latency  $T_s$  in the first trial of the clear condition was already close to optimum, which is indicative of a ceiling effect. Nevertheless, the improvement of latencies between T1 and T2 in the opaque condition suggests that apes remembered the presence and relevance of the stick inside the tube. However, these results are not very informative about the aspects that apes may encode regarding the functioning of the apparatus (e.g., perceptual vs. structural

knowledge about the function of the stick). Besides, it is conceivable that previous experience with the use of sticks to dislodge food located inside tubes (e.g., trap-tube task: Mulcahy and Call 2006; Martin-Ordas et al. 2008, 2009) may have contributed to the ceiling effect. In the next experiment, we once again investigated the effect of visible versus opaque presentation with an apparatus that required more complex operations to access the rewards, operations that could not be solely ascribed to experiencing a similar task in a previous study.

## Experiment 2

### Subjects

All subjects included in Experiment 1 plus four additional chimpanzees participated in this experiment (see Table 1 for additional details). Although all subjects had participated in a number of cognitive tasks prior to the current study, none of these tasks involved removing multiple sticks from a puzzle box.

### Materials

The apparatus consisted of a cuboid (79 cm  $\times$  10 cm  $\times$  10 cm) fixed on an acrylic glass panel (32.5 cm  $\times$  79 cm) and attached to the mesh in a vertical orientation (see Fig. 3). Approximately 15 cm separated the bottom of the mounted apparatus from the ground. Either the top or the bottom side of the cuboid was closed by an opaque cover plate fixed with a screw. The other side remained open. A set of four circular holes (diameter 1.6 cm) forming a straight line were drilled 25 cm from the top and the bottom of the apparatus so that four dowels (length 12 cm, diameter 1.4 cm) could be inserted laterally into the holes resulting in two platforms of four sticks each. The apparatus could be mounted at the mesh either with the opening facing upwards (up condition) or turned upside down, with the opening facing to the ground (down condition). A wooden, rectangular block (height 13.5 cm, base area 6  $\times$  6 cm) with a circular depression on its upper side (depth 0.5 cm) placed on the lower dowel-formed platform acted as the reward holder and kept the reward equidistant (i.e., 39.5 cm) between each of the ends of the apparatus.

There were two different versions of the apparatus. The clear version was made of transparent acrylic glass so that the subjects could see the inner workings of the apparatus including the inserted sticks making up the two platforms. In contrast, the opaque version was made of opaque acrylic glass except for a 3-cm clear acrylic glass area where the reward was located. In both apparatuses, the acrylic glass



**Fig. 3** Tower problem: **a** clear version, **b** opaque version. The apparatuses are shown in the down condition (opening on *bottom side*). Sticks that form the platforms (indicated by the *arrows*) could be removed laterally. Location of the reward (a grape) indicated by the *asterisk*

around this area was cut into six thin layers that only allowed looking through the glass when taking a perpendicular visual angle. The layers inside the apparatus prevented an observer from seeing through the acrylic glass when looking from other visual angles. Thereby, we could prevent the subjects from seeing the platforms in the opaque condition even though they could easily see the reward. In addition, an opaque screen was mounted above the apparatus at a distance of 15–20 cm that was to prevent the apes from looking inside the apparatus through the opening on top (up condition). Grapes were used as rewards. In order to prevent the grape from falling down from the wooden block (e.g., when the apes hit the apparatus), honey was put inside the depression to ‘glue’ the grape.

In the up condition (with the opening facing upwards), the apes could obtain the reward by removing the sticks of the upper platform and grabbing the grape from above. However, if they removed the lower platform first, the

grape fell down and was trapped at the bottom of the apparatus out of reach of the subject. In the down condition (with the opening facing the ground), the subjects could obtain the reward by removing the sticks of the lower platform so that the wooden block with the grape on it fell to the ground through the opening at the bottom of the apparatus.

### Procedure and design

Subjects were randomly assigned to the opaque or clear condition. We counterbalanced these groups for species, age, and sex (clear group: mean age 19.7, females 67 %; opaque group: mean age 19.3, females 71 %). In total, each subject received four trials on four testing days: two trials with the down condition and two trials with the up condition (intertrial interval 5 days, range 3–11 days). The order of administration of the up/down condition was counterbalanced within each group. Each trial lasted a maximum of 10 min and started as soon as the subject approached and was within arm’s reach of the apparatus.

### Scoring and analysis

Just like we did in Experiment 1, we coded the success and the latency  $T_s$  (time the subject spent at the apparatus before removing the reward). Additionally, we scored manipulation efficiency defined as whether the subject first targeted the dowels belonging to the platform that had to be removed to gain the reward (efficient) or the other set of dowels (inefficient).

All sessions were videotaped. A second coder (JC) scored 20 % of the trials to assess interobserver reliability, which was perfect (Spearman’s correlation:  $r = 1.0$ ;  $N = 10$ ). We used two-tailed nonparametric statistics to analyze the effect of condition (clear vs. opaque), species, sex, and age on success and  $T_s$ . For  $2 \times 4$  contingency tables, the Freeman–Halton extension of Fisher’s exact test was used (Freeman and Halton 1951). Since most of the successful individuals in the opaque condition managed to look inside the apparatus through the opening on top—despite the screen that was mounted above the apparatus—we did not analyze the data obtained in the up condition. Note that this effectively turned the opaque condition into a clear condition. Therefore, we restricted our analysis to the down condition.

## Results

### Success

One subject (Jahaga) was excluded from analysis as she could not solve the apparatus due to apparatus malfunction (i.e., a tilted stick that could not easily be removed from the

apparatus). Table 2 shows the number of successful individuals as a function of species and condition. There was neither an effect of species (Freeman–Halton extension of Fisher’s exact test:  $P = 0.92$ ) nor an effect of visual information available (clear vs. opaque) (Fisher’s exact test:  $P = 0.68$ ) on the number of successful apes across species. In subsequent analyses, we only analyzed those subjects who solved the apparatus in both trials ( $N = 20$ ).

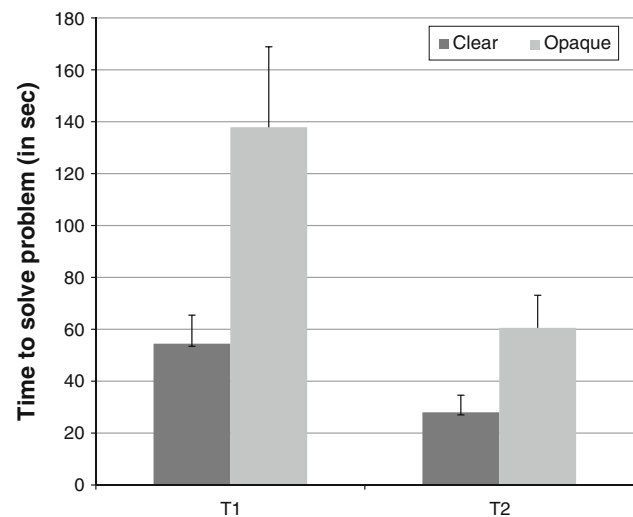
#### Latency ( $T_s$ )

We found no significant effect of order of the up/down condition (Mann–Whitney tests: overall: T1:  $U = 41.0$ ,  $P = 0.53$ ; T2:  $U = 32.5$ ,  $P = 0.20$ ; clear condition: T1:  $U = 8.0$ ,  $P = 0.25$ ; T2:  $U = 9.5$ ,  $P = 0.35$ ; opaque condition: T1:  $U = 10.0$ ,  $P = 1.0$ ; T2:  $U = 10.0$ ,  $P = 1.0$ ). Similarly, we found no effect of species overall (Kruskal–Wallis tests: T1:  $\chi^2 = 0.38$ ,  $df = 3$ ,  $P = 0.95$ ; T2:  $\chi^2 = 1.16$ ,  $df = 3$ ,  $P = 0.79$ ) or within the clear condition (T1:  $\chi^2 = 4.56$ ,  $df = 3$ ,  $P = 0.21$ ; T2:  $\chi^2 = 5.57$ ,  $df = 3$ ,  $P = 0.08$ ) or the opaque condition (T1:  $\chi^2 = 2.48$ ,  $df = 3$ ,  $P = 0.62$ ; T2:  $\chi^2 = 3.28$ ,  $df = 3$ ,  $P = 0.42$ ). Therefore, we collapsed the data across species and order of condition in our subsequent analyses.

Figure 4 shows  $T_s$  as a function of condition. Subjects in the clear condition were significantly faster in solving the task than those in the opaque condition (Mann–Whitney tests: T1:  $U = 18.0$ ,  $P < 0.05$ ; T2:  $U = 21.0$ ,  $P < 0.05$ ). In addition, we found that the apes in the clear condition got significantly faster in T2 compared to T1 ( $Z = 2.13$ ,  $P < 0.05$ ); in the opaque condition, such improvement was only found on trend level ( $Z = 1.96$ ,  $P = 0.06$ ).

#### Efficiency of the first action

Eight out of ten successful subjects in the clear condition directed their first stick removal attempts toward the relevant sticks (i.e., lower platform). In contrast, only four out of eight individuals confronted with the opaque



**Fig. 4** Mean time to solve the tower apparatus (+SEM) in the down condition as a function of visual information available (clear vs. opaque). T1: first trial, T2: second trial

apparatus did so (two subjects were removed from this analysis due to interference from a juvenile ape on his mother’s performance and first actions on the apparatus that could not be categorized as efficient/inefficient, respectively). However, there was no statistically significant effect of visual information on the efficiency of the first manipulation in T1 (Fisher’s exact test:  $P = 0.32$ ) or T2 ( $P = 0.48$ ).

Interestingly, some individuals spontaneously used the removed sticks to poke out the remaining sticks still inserted in the apparatus. Three of five successful orangutans, one of three bonobos, one of two gorillas, and one of ten chimpanzees repeatedly used such a technique. Four of the six tool-using apes were in the clear condition and two in the opaque condition. Neither species (Freeman–Halton extension of Fisher’s exact test:  $P = 0.15$ ) nor conditions (Fisher’s exact test:  $P = 0.64$ ) differed significantly in the number of tool-using individuals.

## Discussion

Most subjects regardless of their species were able to solve this problem irrespective of the degree of visual information about the mechanisms of the apparatus. Similarly, we detected no differences in efficiency between the clear and opaque groups. In contrast, we found that visual information affected the latency to solve the problem: subjects in the clear condition solved the problem significantly faster than those in the opaque condition. Additionally, subjects in the clear group, unlike those in the opaque group, got significantly faster in the second trial compared to the first one.

**Table 2** Number of subjects that solved the tower problem in both trials in the down condition as a function of species and visual information available (clear vs. opaque)

	Clear		Opaque	
	Pass	Fail	Pass	Fail
Bonobo	2	0	1	1
Chimpanzee	6	1	4	4
Orangutan	2	1	3	0
Gorilla	1	1	1	0
All subjects	11	3	9	5

This improvement in performance in the clear condition might be due to the fact that when causal relations are visually accessible (here: sticks that support the reward have to be removed to obtain the reward), the apes can learn the requirements of the task faster than when the apes have to rely on more abstract information like in the opaque condition (i.e., removing sticks from black box yields reward). In the same vein, the efficiency of the first operations on the apparatus was higher—even though not statistically significant—in the clear compared to the opaque condition. Such an advantage of causal compared to arbitrary cues would also be supported by previous research (Hanus and Call 2008, 2011).

These findings might however be confounded by the perceptual salience of the sticks inserted laterally into the apparatus, which is higher in the clear than in the opaque condition. Also, there are differences in the visibility of the reward: in contrast to the opaque condition, the reward could also be seen when looking from above at a very steep angle in the clear condition. Such differences might have affected subjects' motivation.

Even though this task was harder than the tube task in Experiment 1, the fact that nearly all subjects solved the task also points to a ceiling effect. In the next experiment, we presented a more demanding task, partly because it required more complex manipulations and partly because the perceptual salience of the relevant parts of an apparatus and the visibility of the reward were equalized across conditions.

### Experiment 3

#### Subjects

All subjects who participated in Experiment 2 except for two orangutans and three chimpanzees (who were not available for testing at the time of the study) plus three naïve chimpanzees participated in this experiment (see Table 1 for additional details). None of the subjects had previous experience with a crank mechanism.

#### Materials

The crank apparatus consisted of a rectangular cuboid (height 45 cm, inner base area  $8.8 \times 9.0$  cm) made of transparent acrylic glass that was fixed on a panel ( $46 \times 31$  cm). The apparatus was mounted on the cage mesh. The crank was located at the upper part of the cuboid (crank zone, height 5.7 cm, see Fig. 5). It consisted of a horizontal tube (diameter 3.5 cm) inside the apparatus that was connected to two rectangular acrylic glass pieces

( $6.0 \times 4.0 \times 2.0$  cm) on each side of the apparatus. These outer acrylic glass pieces formed the handles of the crank: by turning (rolling) the handles of the crank, a white string (diameter 0.4 cm) fixed onto the cylinder inside the apparatus was coiled up. This string ran behind the back side of the apparatus, out of reach of the subjects. The back side had a gap (width 0.8 cm) in the middle across the length of the apparatus. At the bottom part of the apparatus, the string was attached to a metallic sliding shuttle. The reward, a piece of monkey chow, was located inside the cuboid. Crucially, the reward was connected to this sliding shuttle through the gap in the back side by means of a jute thread. When the subjects turned the handle of the crank, the reward moved upwards. After three full turns of the crank, the reward reached its maximal height. On this level, there was an oval hole (height 3.8 cm, width 5.8 cm) in the front side of the apparatus through which the reward could be obtained. Due to the weight of the sliding shuttle, the reward would always fall back into its original position when the handle was released. The area between the oval hole and the original position of the reward was the so-called progress zone (height 25.0 cm) as the subject could witness how the sliding shuttle moved upwards, bringing the reward closer to the opening. Above the oval hole, there was an opaque partition that prevented the apes from witnessing the crank mechanism from below. In addition, the lateral surface of the crank zone and the upper and bottom sides of the apparatus were opaque. Hence, the crank mechanism could only be witnessed through the transparent front side.

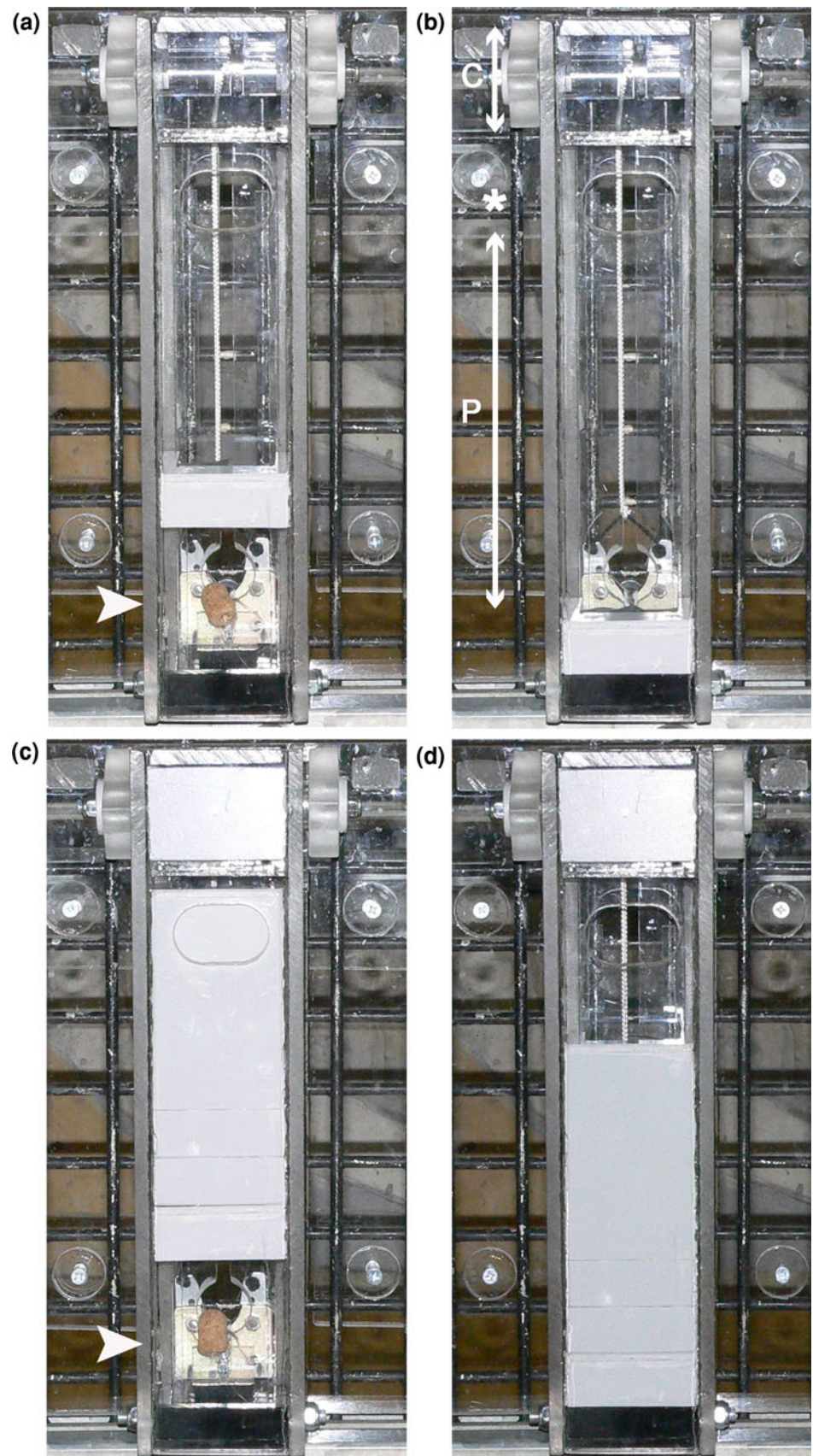
In all conditions, there was a U-shaped opaque screen (reward screen: height  $\times$  width  $\times$  depth:  $3.5 \times 8.6 \times 6.0$  cm) made of gray plastic material inside the cuboid that covered the reward at its starting position (bottom). Flexible rubber lamellae on the upper side of the screen allowed the reward to pass through when the crank was operated, but prevented the subjects from seeing the reward from above in its starting position. The progress zone could be occluded by means of second U-shaped opaque screen (progress screen: height  $\times$  width  $\times$  depth:  $20.0 \times 8.6 \times 6.0$  cm) that was positioned above the base screen. The crank zone could be occluded through an opaque plate (height  $\times$  width  $5.7 \times 8.8$  cm) that was inserted directly behind the transparent front side. The upper and lower sides of the apparatus were removable in order to insert and remove the above-mentioned screens.

#### Procedure and design

There were two phases. In phase 1, subjects were randomly assigned to one of four groups. Visual access to the apparatus was manipulated between these four groups. The reward screen was present in all groups. The clear group



**Fig. 5** Crank apparatus: **a** clear version, initial position (reward visible); **b** clear version, testing position (reward out of sight); **c** opaque version, initial position; **d** opaque version, testing position. *Arrow* location of the reward (piece of monkey chow), *asterisk* location of the hole in the front side, crank zone and progress zone indicated by the letters *C* and *P*, respectively. In the beginning of each trial, the base screen (and when present also the progress screen) was released so that it fell down and covered the reward (testing position)



received a completely transparent apparatus (progress and crank zones visible), whereas the opaque group received an apparatus with both zones occluded. The progress and the crank groups received an apparatus with the progress and crank zones occluded, respectively. We counterbalanced these groups for species, age, and sex (clear group: mean age 20.2, females 57 %; progress group: mean age 20.4, females 71 %; crank group: mean age 18.7, females 67 %; opaque group: mean age 19.3, females 71 %).

Before each trial, the base screen was pushed upwards by 10 cm and was fixed by means of a screw so that the reward was visible to the subjects (initial position, see Fig. 5a, c). Each trial started when the subject approached the apparatus. After approximately 10 s (or earlier if the subjects already touched the handle of the crank), the base screen was released and fell down (testing position, see Fig. 5b, d). At that point, the reward was not visible any more in its starting position. Subjects received two 10-min trials to obtain the reward. The intertrial interval was on average 7 days (range 6–11 days).

In phase 2, we reshuffled group assignment to test the impact of having visual access to the progress zone on the subjects' performance. Subjects who witnessed the progress zone in phase 1 received an apparatus with occluded progress zone in phase 2 (clear group → progress-blocked condition, crank group → opaque condition). The subjects who had not witnessed the progress zone in phase 1 were now allowed to witness the progress zone (progress group → clear condition, opaque group → crank-blocked condition). Subjects that were successful in phase 1 got one refresher trial with their phase 1 condition one day before the first trial of phase 2 in order to make sure that they remembered the solution. Such a refresher trial was not given to unsuccessful individuals in phase 1 as we did not want to reduce their motivation to operate on the apparatus before entering phase 2.

### Scoring and analysis

Like in Experiments 1 and 2, we coded success and the latency  $T_S$ . A given trial was coded as successful when the reward could be reached by the ape through the frontal hole. In case the reward was dropped in the apparatus after having it removed from the sliding shuttle, the subjects were rewarded immediately by the experimenter with a piece of monkey chow.  $T_S$  was defined as time the apes spent at the apparatus before he or she could touch the reward for the first time. All sessions were videotaped. A second coder (JC) scored twenty percent of the trials to assess interobserver reliability, which was excellent (Spearman's correlation:  $r = 0.91$ ;  $N = 10$ ). We used two-tailed nonparametric statistics to analyze the effect of

condition (clear, progress blocked, crank blocked, opaque), species, sex, and age on success and  $T_S$ .

## Results

### Phase 1

One of four bonobos, two of sixteen chimpanzees, one of three gorillas, and two of four orangutans solved the problem in the first session (for an example, see online resource *esm\_1.mpg*). Two additional chimpanzees solved the problem in their second session. The number of successful individuals did not differ significantly between species (Freeman–Halton extension of Fisher's exact test:  $P = 0.27$ ). For the subsequent analyses, we collapsed the data obtained from the different species.

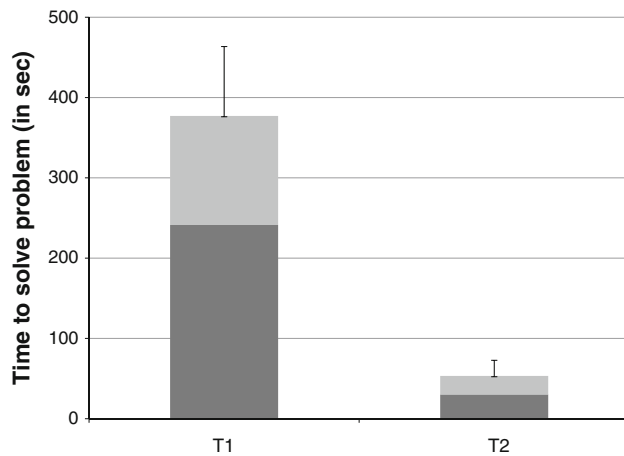
Table 3 shows the number of successful individuals across conditions. We found a significant effect of condition on the number of successful individuals (Freeman–Halton extension of Fisher's exact test:  $P < 0.01$ ). Only subjects in the clear and crank-blocked condition were able to solve the task. Thus, success seemed to depend on having visual access to the progress zone, which was occluded both in the progress-blocked and in the opaque conditions. Statistical analysis confirmed that significantly more subjects that could see the progress zone (clear and crank-blocked condition) solved the apparatus in the first two trials than those that could not see the progress (opaque and progress-blocked condition, Fisher's exact test:  $P < 0.01$ ).

Focusing on the performance of successful individuals ( $N = 8$ ; for the two individuals that solved the problem only in their second session, the latencies of T1 and T2 were combined) showed that there was neither a significant effect of species (Kruskal–Wallis tests: T1:  $\chi^2 = 4.58$ ,  $df = 3$ ,  $P = 0.17$ ; T2:  $\chi^2 = 2.83$ ,  $df = 3$ ,  $P = 0.52$ ) nor significant effect of condition (clear vs. crank blocked,

**Table 3** Number of successful individuals in the crank apparatus in phases I and II

Conditions	Phase I		Phase II	
	Pass	Fail	Pass	Fail
Clear	3	4	1	6
Crank blocked	5	1	3	4
Progress blocked	0	7	0	4
Opaque	0	7	0	1

In phase II, only those subjects are shown that failed to solve the problem in phase I. Group assignment was reshuffled between phase I and phase II according to the following schema: clear → progress blocked, crank blocked → opaque, progress blocked → clear, opaque → crank blocked



**Fig. 6** Mean time to solve the apparatus in phase I (+SEM) collapsed across crank-blocked and clear condition. T1: first trial, T2: second trial. The transition of *gray shades* marks the time when the reward became visible

Mann–Whitney test: T1:  $U = 4.0$ ,  $P = 0.39$ ; T2:  $U = 5.0$ ,  $P = 0.57$ ) on the latency to reach the reward.

Figure 6 shows the mean time to solve the problem across trials. Subjects got significantly faster in the second trial compared to the first one (Wilcoxon’s signed rank test:  $Z = 2.52$ ,  $P < 0.01$ ).

## Phase 2

Focusing on the subjects who failed to solve the problem in phase 1, one out of two orangutans, one of three bonobos, and two of twelve chimpanzees solved the task in phase two. Neither of the two gorillas solved the task. However, there was no significant effect of species (Freeman–Halton extension of Fisher’s exact test:  $P = 0.47$ ) on the number of successful individuals. Of these four subjects who solved the problem in phase 2, three were in the crank-blocked condition and one in the clear condition (see Table 3). Although there was no significant effect of condition ( $P = 0.44$ ), again only subjects for which the progress zone was visible solved the task.

All individuals who solved the problem in phase 1 (progress zone visible) also succeeded in phase 2 (progress zone occluded; for an example, see online resource *esm\_2.mpg*). There was no effect of condition (progress blocked vs. opaque) on the latency to solve the problem (Mann–Whitney test: T1:  $U = 7.0$ ,  $P = 1.0$ ; T2:  $U = 6.0$ ,  $P = 0.79$ ). In addition, there was no significant difference between latencies in T2 of phase 1 and the refresher trial in phase 2 (Wilcoxon’s signed rank test:  $Z = 0.31$ ,  $P = 0.84$ ) or between the refresher trial and T1 (phase 2) ( $Z = 0.22$ ,  $P = 0.74$ ) or between T1 and T2 ( $Z = 1.26$ ,  $P = 0.25$ ). The only significant difference was observed between the refresher trial and T2 of phase 2 ( $Z = 2.24$ ,  $P < 0.05$ ).

## Discussion

All four species of great apes were able to solve the crank apparatus, a task requiring a sequence of coordinated actions: first, the handle of the crank had to be turned repeatedly with one or two hands on one or both sides of the apparatus. This turning action had to be repeated until the reward reached the appropriate height to remove it. Along the way, the crank had to be kept from falling back to its original position. Once the reward reached the appropriate height, the subjects had to hold the crank while grabbing the reward with the other hand.

Only subjects who had visual access to the progress of the reward while turning the handle solved the problem. In contrast, visual access to the crank zone that yielded information on the causal mechanism involved did not have any effect on the performance. In addition, successful subjects increased their performance dramatically from trial 1 to trial 2. Hence, visual feedback on the effects of their operations seems to be necessary and sufficient for the apes to solve this problem. Such perceptual feedback might act as a form of incremental reinforcement (Taylor et al. 2010): accordingly, the turning behavior would continuously be reinforced the closer the reward moved to the hole in the front side. Another—not mutually exclusive—account is that individuals required visual feedback to take advantage of their causal knowledge about the task.

After having found the solution in phase 1, the apes had no problem (not even an increased latency) to solve the task in phase 2 albeit the progress zone was now occluded. This shows that after limited experience (three trials in a time period of more than 3 weeks), the apes did not require perceptual-motor feedback anymore. This means that visual feedback may have been necessary to learn how to solve the task, but it was not necessary to maintain the solution.

## General discussion

Taken together, the present results highlight the remarkable problem-solving abilities of the great apes. All problems were solved by all four species of nonhuman great apes when visual inspection of the apparatuses was possible. Interestingly, the less complex first two problems were solved also by the apes that did not get any visual information on the mechanism involved, showing the power of trial-and-error learning. Yet, there were differences in performance between the clear and opaque conditions: in the tube and stick (Experiment 1) and the tower problem (Experiment 2), these differences were observed in terms of latencies: subjects confronted with the clear version of the problem solved the task more quickly than those in the

opaque condition. In the crank problem (Experiment 3), the differences were shown in terms of success (and latency). In particular, only subjects who had visual access to the effects of their actions on the reward were able to solve the task.

These findings indicate that apes benefit from visual inspection when solving such puzzle boxes. In the tube and stick task, previous experience of the apes with similar experimental setups (e.g., the trap-tube task) most likely influenced their performance. In the tower problem, perceptual salience of relevant parts of the apparatus (sticks inserted into the tower apparatus) possibly might have played an important role. However, these two explanations do not apply to the results obtained with the crank apparatus (Experiment 3). In particular, the apes had no previous experience with a crank mechanism (or similar object manipulation tasks involving sequential turning of an object), and the perceptual salience of the parts of the apparatus that were to be manipulated (the handle of the crank) was identical across conditions. Thus, these results suggest that the performance of the apes benefits from visual feedback on the effects of their actions.

Results from the crank experiment allow us to catch a glimpse on the learning mechanism at work: having visual access not to the crank mechanism but to the progress of the reward turned out to be necessary for them to solve the task. This finding raises the possibility that operant conditioning plays an important role here: accordingly, the progress of the reward in response to the turning of the crank acts as an intermediate reinforcer, which would make the recurrence of this action more likely. However, three aspects suggest that even though operant conditioning may be implicated in the solution process, it does not fully explain it.

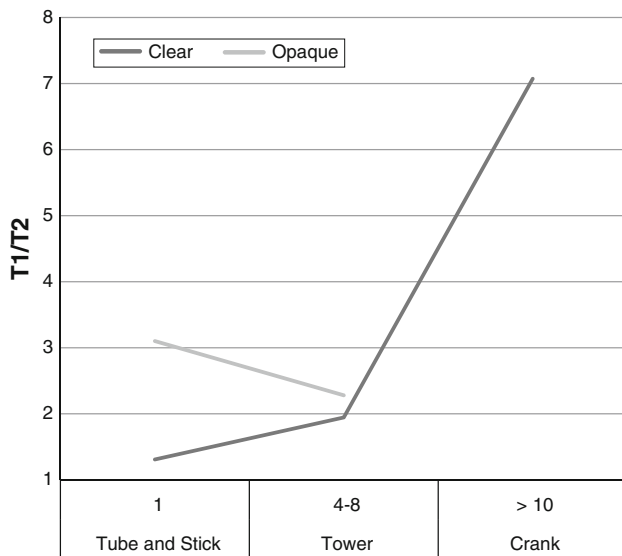
First, operant conditioning is controlled by information about spatiotemporal means–ends relations. It is unclear that operant conditioning would work if subjects did not see the action that brings the reward closer (e.g., the cranking action) as a means toward the end of bringing the reward within reach. In fact, one could argue that based on operant conditioning alone, actions that did not lead to a successful outcome should be extinguished. Since turning the crank only produced the reward if this action was repeated multiple times, it should have disappeared, which is precisely what happened in those conditions in which the means–ends relations were not visible.

Second, further support for the idea that the apes had some knowledge about the causal structure of the crank task is the fact that subjects with limited experience with the appropriate solving strategy had no problem whatsoever to solve the task when visual feedback of the progress of the reward was restricted. Once acquired, the response was not controlled by the visual feedback. This is a striking difference to results obtained by Taylor et al. (2010),

showing that experienced New Caledonian crows made much more errors in a vertical string-pulling task when visual feedback was restricted.

Third, there is the issue of the impressive leap in performance from the first to the second trials. Figure 7 shows the relative learning rate between the first (T1) and second trials (T2) across experiments 1–3. In the most simple task, the tube and stick problem, the apes performed close to maximum already in the very first trial when they could see the relevant parts of the apparatus (clear condition). Consequently, no significant improvement in performance was observed from T1 to T2. In contrast, in the opaque condition, learning occurred between the first and second trials: presumably, the apes remembered that the stick was existent in the opaque part of the tube in T1 and hence improved their performance in T2. In a more difficult task, the tower problem, the latency significantly decreased from T1 to T2 only in the clear condition. The latter finding suggests that visual access to causally and/or perceptually relevant information allows great apes to learn faster. Finally, in the most difficult experiment, the crank problem, only subjects who could witness the progress zone (clear and crank-blocked condition, both were coded here as ‘clear’) were able to solve the problem. The learning rate in this last task exceeded the previous experiments by far. In general, these results suggest that only in very simple situations (Experiment 1), great apes learn quickly when there is no visual feedback available. When the degree of difficulty increases (in terms of the number and complexity of actions necessary to solve the problem, Experiments 2 and 3) only apes that got visual access to the interior of the apparatuses significantly improved their performance from T1 to T2. Together, these findings are compatible with the view that the apes take at least some of the basic causal structures of the tasks into account. Accordingly, the leap in performance between T1 and T2 increases with task complexity in the clear condition as, once the apes solved the problem for the first time, they might have understood which actions were relevant and, therefore, can actively reproduce the actions necessary in the second trial.

How do the present findings fit to the data on corvids’ string pulling abilities? Comparing the results of Taylor et al. (2010) with the current data reveals that both great apes and New Caledonian crows initially rely on visual feedback of their actions on the apparatus. As noted above, a striking difference between the behaviors of corvids and primates is the performance of experienced subjects: while crows that had solved the standard problem 10 times before made a lot of errors and showed reduced efficiency when facing the visually restricted version of the task, great apes did not show—after a learning history of three successful trials—a drop in performance at all. This difference is even more remarkable when considering that the quality of



**Fig. 7** Relative decrease in latencies from trial one (T1) to trial two (T2) across experiments 1–3 and separated by condition (clear vs. opaque). The numbers on the abscissa indicate the minimal actions necessary to solve the problem as a measure of task complexity. Note, however, that the crank task requires sequential and motorically more complex actions than the tower or tube and stick task

actions necessary to solve the task varies between the string-pulling and crank tasks. While the crank mechanism requires rather indirect and unusual actions, string pulling is much more straightforward. Despite such differences in task complexity, great apes faster emancipate from perceptual-motor feedback, which might be explained by a deeper understanding of the causal structure of the task.

The causal knowledge that apes might have used seems to be of the practical type. That is, they might have understood *what* caused the beneficial outcome but not necessarily *how* it was caused. In line with such an account is that the observation ‘turning the crank makes the reward go up’ was necessary and sufficient for the apes to solve the task; in contrast, visual information on the crank mechanism, that is, on how the apparatus works, had no impact on their performance. This is not to say that nonhuman great apes are generally unable to make inferences about how one event caused the other. Previous research revealed that great apes can understand functional features of tools (Bania et al. 2009) as well as physical properties such as weight (Hanus and Call 2008, 2011) or rigidity (Furlong et al. 2008; Manrique et al. 2010). Other properties such as connectivity, gravity, and solidity seem to be harder (Cacchione and Call 2010; Herrmann et al. 2008; Povinelli 2000) for them to assess. In the crank task, the apes might not have grasped the quite complex relations involved, that is, the connectivity between the reward, the sliding shuttle, the string, and the crank just from the inspection of the crank mechanism.

The current findings also have implications for other candidates of insightful problem solving: for example, in the floating peanut task (Mendes et al. 2007; Hanus et al. 2011), such a positive perceptual-motor feedback cycle exists, too. Therefore, an important issue for future research will be to examine whether great apes can also solve this task initially and after some experience, respectively, when no visual feedback is available.

In summary, the present results show that visual inspection helps great apes to solve different apparatus problems. In particular, visual feedback on the consequences of their actions seems to be important for their performance. This is especially true when the task is complex in nature or involves novel mechanisms. In the crank apparatus, the very fast rate of learning raises the possibility that besides operant conditioning, apes’ performance involved some understanding of the causal structure of the task. This notion gets additional support by the fact that individuals with limited experience could solve the crank problem without visual feedback on their actions. In contrast, no evidence was found that in the context of the crank mechanisms the apes understood the causal properties involved (i.e., the string connectivity). However, the extent and exact nature of great apes’ causal knowledge cannot be conclusively determined here and will require further research.

**Acknowledgments** Christoph Völter was supported by a German National Academic Foundation studentship. We thank Raik Pieszek for constructing the experimental apparatuses and the animal caretakers of the Leipzig Zoo. We also thank Matthias Allritz, Virginia Gonzalez, and Carolin Kade for their help with data collection.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Bania AE, Harris S, Kinsley HR, Boysen ST (2009) Constructive and deconstructive tool modification by chimpanzees (*Pan troglodytes*). *Anim Cogn* 12(1):85–95
- Bird CD, Emery NJ (2009) Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proc Natl Acad Sci* 106(25):10370
- Boesch C, Boesch H (1990) Tool use and tool making in wild chimpanzees. *Folia Primatol* 54:86–99
- Cacchione T, Call J (2010) Intuitions about gravity and solidity in great apes: the tubes task. *Dev Sci* 13(2):320–330
- Call J (2007) Apes know that hidden objects can affect the orientation of other objects. *Cognition* 105(1):1–25
- Cheke LG, Bird CD, Clayton NS (2011) Tool-use and instrumental learning in the Eurasian jay (*Garrulus glandarius*). *Anim Cogn* 14(3):441–455
- Emery NJ, Clayton NS (2004) The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306(5703):1903–1907

- Freeman G, Halton J (1951) Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika* 38(1–2):141
- Funk M (2002) Problem solving skills in young yellow-crowned parakeets (*Cyanoramphus auriceps*). *Anim Cogn* 5(3):167–176
- Furlong EE, Boose KJ, Boysen ST (2008) Raking it in: the impact of enculturation on chimpanzee tool use. *Anim Cogn* 11(1):83–97
- Hanus D, Call J (2008) Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Curr Biol* 18(9):R370–R372
- Hanus D, Call J (2011) Chimpanzee problem-solving: contrasting the use of causal and arbitrary cues. *Anim Cogn* 14(6):871–878. doi: [10.1007/s10071-011-0421-6](https://doi.org/10.1007/s10071-011-0421-6)
- Hanus D, Mendes N, Tennie C, Call J (2011) Comparing the performances of apes (*Gorilla gorilla*, *Pan troglodytes*, *Pongo pygmaeus*) and human children (*Homo sapiens*) in the floating peanut task. *PLoS ONE* 6(6):e19555
- Heinrich B (1995) An experimental investigation of insight in common ravens (*Corvus corax*). *Auk* 112(4):994–1003
- Heinrich B (2000) Testing insight in ravens. In: Heyes C, Huber L (eds) *The evolution of cognition Vienna series in theoretical biology*. The MIT Press, Cambridge, pp 289–305
- Herrmann E, Wobber V, Call J (2008) Great apes' (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, *Pongo pygmaeus*) understanding of tool functional properties after limited experience. *J Comp Psychol* 122(2):220–230
- Horner V, Whiten A (2005) Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Anim Cogn* 8(3):164–181
- Kenward B, Weir A, Rutz C, Kacelnik A (2005) Tool manufacture by naive juvenile crows. *Nature* 433:138
- Manrique HM, Gross ANM, Call J (2010) Great apes select tools on the basis of their rigidity. *J Exp Psychol Anim Behav Process* 36(4):409–422
- Martin-Ordas G, Call J (2009) Assessing generalization within and between trap tasks in the great apes. *Int J Comp Psychol* 22(1):43–60
- 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. *Anim Cogn* 11(3):423–430
- Mendes N, Hanus D, Call J (2007) Raising the level: orangutans use water as a tool. *Biol Lett* 3:453–455
- Mulcahy NJ, Call J (2006) How great apes perform on a modified trap-tube task. *Anim Cogn* 9(3):193–199
- Mulcahy NJ, Call J, Dunbar RIM (2005) Gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) encode relevant problem features in a tool-using task. *J Comp Psychol* 119(1):23–32
- Pepperberg IM (2004) “Insightful” string-pulling in Grey parrots (*Psittacus erithacus*) is affected by vocal competence. *Anim Cogn* 7(4):263–266
- Povinelli DJ (2000) *Folk physics for apes: the chimpanzee's theory of how the world works*. Oxford University Press, New York
- Rensch B, Döhl J (1967) Spontanes Öffnen verschiedener Kistenverschlüsse durch einen Schimpansen. *Zeitschrift für Tierpsychologie* 24(4):476–489
- Seed A, Call J (2009) Causal knowledge for events and objects in animals. In: Watanabe S, Blaisdell AP, Huber L, Young A (eds) *Rational animals, irrational humans*. Keio University Press, Tokyo, pp 173–187
- Seed A, Emery N, Clayton N (2009) Intelligence in corvids and apes: a case of convergent evolution? *Ethology* 115(5):401–420
- Taylor AH, Medina FS, Holzhaider JC, Hearne LJ, Hunt GR, Gray RD (2010) An investigation into the cognition behind spontaneous string pulling in New Caledonian crows. *PLoS ONE* 5(2):e9345
- von Bayern AMP, Heathcote RJP, Rutz C, Kacelnik A (2009) The role of experience in problem solving and innovative tool use in crows. *Curr Biol* 19(22):1965–1968
- Weir A, Chappell J, Kacelnik A (2002) Shaping of hooks in New Caledonian crows. *Science* 297:981
- Werdenich D, Huber L (2006) A case of quick problem solving in birds: string pulling in keas, *Nestor notabilis*. *Anim Behav* 71(4):855–863