Report

Current Biology

Children's and Apes' Preparatory Responses to Two Mutually Exclusive Possibilities

Highlights

- Representing alternative future events is an important facet of effective foresight
- Children and apes had the chance to catch a target falling from one of two locations
- 2-year-olds and apes prepared for the target's emergence from only one location
- Many 3- to 4-year-olds simultaneously and consistently prepared for both possibilities

Authors

Jonathan Redshaw,
Thomas Suddendorf

Correspondence

j.redshaw@uq.edu.au

In Brief

Redshaw and Suddendorf gave children and apes the chance to catch a target dropped into an inverted Y-shaped tube. Many 3- to 4-year-olds, but no 2-year-olds or apes, spontaneously and consistently covered both exits from which the target could fall. This is a nonverbal demonstration of the potent capacity to prepare for alternative future events.

Redshaw & Suddendorf, 2016, Current Biology *26*, 1758–1762
July 11, 2016
http://dx.doi.org/10.1016/j.cub.2016.04.062

CellPress

Current Biology
Report

Children's and Apes' Preparatory Responses to Two Mutually Exclusive Possibilities

Jonathan Redshaw[1,*] and Thomas Suddendorf[1]
[1]School of Psychology, University of Queensland, Queensland 4072, Australia
*Correspondence: j.redshaw@uq.edu.au
http://dx.doi.org/10.1016/j.cub.2016.04.062

SUMMARY

Animal brains have evolved to predict outcomes of events in the immediate environment [1–5]. Adult humans are particularly adept at dealing with environmental uncertainty, being able to mentally represent multiple, even mutually exclusive versions of the future and prepare accordingly. This capacity is fundamental to many complex future-oriented behaviors [6, 7], yet little is known about when it develops in children [8] and whether it is shared with non-human animals [9]. Here we show that children become able to insightfully prepare for two mutually exclusive versions of an undetermined future event during the middle preschool years, whereas we find no evidence for such a capacity in a sample of chimpanzees and orangutans. We gave 90 preschool children and 8 great apes the opportunity to catch an item dropped into a forked tube with two bottom openings. Children's performance improved linearly across age groups (2, 2.5, 3, 3.5, and 4 years), with none of the youngest group but most of the oldest group spontaneously covering both openings the first time they prepared to catch the item. The apes performed like 2-year-olds on the first trial, with none of them covering both openings. Some apes and 2-year-olds eventually passed the task, but only in a manner consistent with trial-and-error learning. Our results reveal the developmental trajectory of a critical cognitive ability that allows humans to prepare for future uncertainty, and they also raise the possibility that this ability is not shared with other hominids.

RESULTS AND DISCUSSION

To investigate the capacity to represent and prepare for future uncertainty, we constructed a minimalist forked tube apparatus that had one opening at the top but two openings at the bottom (see Figures 1, S1, and S2). The experimenter could drop a ball or grape into the top of the tube and surreptitiously control which bottom opening it would fall from (in a pseudorandom order). After six observation trials—in which the subject could see the item fall but not catch it—children (n = 90; 18 each from age groups 2, 2.5, 3, 3.5, and 4 years; see Table S1 for details) and great apes (n = 8; 3 chimpanzees, *Pan troglodytes*, and 5 orangutans, *Pongo abelii*) were given the opportunity to catch the item for 12 test trials, and each time they failed to do so, it fell on a ramp and rolled out of reach. We were interested in whether the children and apes would cover one or both bottom openings with their hands when preparing to catch the item. Covering two holes would suggest an understanding that immediate future events can have more than one possible outcome. Of particular interest was first-trial performance, because over the experience of multiple trials, success may be shaped through simple trial-and-error learning.

Figure 2 shows the cumulative percentage of children (by age group) and great apes that covered both bottom openings of the tube for the first time over 12 trials. None of the 2-year-olds, few 2.5-year-olds, many 3- and 3.5-year-olds, and most 4-year-olds passed the first trial. A significant Cochran-Armitage χ^2 test revealed that children were more likely to cover both openings on the first trial as they increased in age, χ^2 (1) = 25.74, p < 0.001. Post hoc Pearson's χ^2 tests revealed that 4-year-olds were significantly more likely to cover both openings on the first trial than 3- and 3.5-year-olds combined, χ^2 (1) = 4.58, p = 0.032, who in turn were significantly more likely to cover both openings on the first trial than 2- and 2.5-year-olds combined, χ^2 (1) = 13.57, p < 0.001. This pattern of first-trial results suggests that the ability to prepare insightfully for mutually exclusive outcomes of a single future event emerges during the third and fourth years.

Despite their poor spontaneous performance on the first trial, however, many 2- and 2.5-year-olds covered both bottom openings at least once over the later trials, and nearly all 3-, 3.5- and 4-year-olds had done so by the end of the experiment (see Figure 2). Yet interestingly, not all of the children who showed this behavior maintained it across all subsequent trials. Figure 3 (cf. yellow segments) shows that many 2- to 3.5-year-olds (but no 4-year-olds) regressed to covering only one opening on at least one later trial. Of the children who covered both openings at least once (n = 67), regressing to covering one opening was significantly less likely among the 4-year-olds (0 out of 17) than among the younger children combined (19 out of 50), post hoc Pearson's χ^2 (1) = 9.02, p = 0.003, even though the 4-year-olds tended to cover both openings much earlier. This pattern of results substantiates the claim that most 4-year-olds and some 3-year-olds were solving the problem by accounting for future uncertainty, whereas many of the younger children were perhaps weakly conditioned into using the optimal response through trial-and-error learning and thus remained susceptible to covering only one opening on some trials.

Like the 2-year-old children, none of the great apes spontaneously prepared for two potential outcomes on their first trial (see

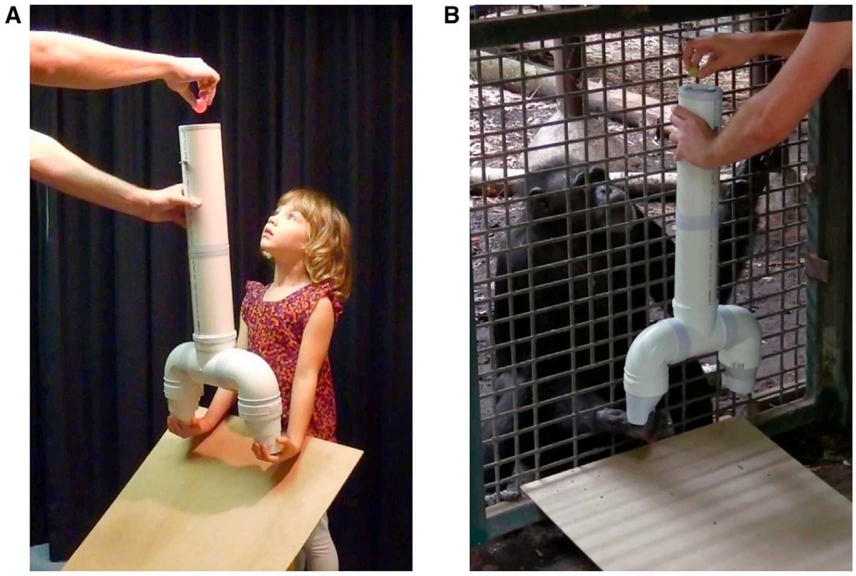

Figure 1. Depiction of the Forked-Tube Task Given to Children and Great Apes

Participants had the opportunity to cover one or both bottom openings of the forked tube when preparing to catch a ball or grape that would be dropped into the top opening. See also Figures S1 and S2.

(A) The child is spontaneously covering both bottom openings.

(B) The ape is covering only one opening.

Figure 2). Only one ape, the chimpanzee Holly, covered both bottom openings at all during the 12 trials. She did so on trials 9 and 11 but regressed to covering only one opening on trials 10 and 12, much like many of the younger children (cf. yellow segments in Figure 3).

Because of the apes' poor performance over these first 12 trials, we were interested in whether they could eventually learn to consistently cover both bottom openings over extra trial blocks (see Table 1). Of the five subjects (three chimpanzees and two orangutans) that were given extra trials, one chimpanzee (Samantha) and one orangutan (Dinar) learned to pass at least six trials in a row (see Figure S3, Table S2, and Supplemental Experimental Procedures for more details). Like the younger children, however, both subjects regressed to covering only one bottom opening at least once after first covering both openings. Such a response pattern—an initial appearance of the target behavior after many unsuccessful trials, followed by brief regression, then reappearance of the behavior and eventual maintenance—is consistent with simple operant conditioning principles rather than insightful behavior. Nonetheless, these results do show that great apes (like 2-year-old children) can learn to pass the task consistently, and so their lack of spontaneous and sustained success cannot be attributed to a basic limitation in the ability to synchronize the actions of two hands through mesh to solve the problem.

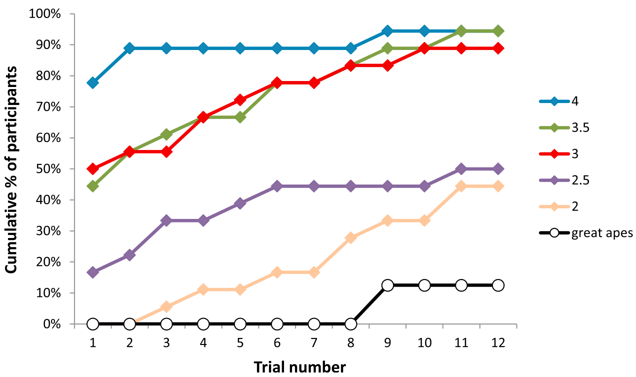

Figure 2. Cumulative Percentage of Children by Age Group and Great Apes that Covered Both Bottom Openings of the Forked Tube for the First Time across the Twelve Trials

Each increment represents a unique participant covering both bottom openings for the first time.

The results of the current study show that, by the middle preschool years, many children have the capacity to spontaneously and consistently prepare for two mutually exclusive versions of an immediate future event. We found limited evidence for this capacity in children younger than 3 years and no evidence in a sample of great apes. Interestingly, our results contrast with those of the only two previous studies of children's ability to prepare for alternative future event outcomes, which found strong positive evidence for the behavior only during the fifth year [10, 11]. Unlike our minimalist task, however, the tasks used in these previous studies relied heavily on language comprehension and included complex intermediate steps between the preparatory behavior and the future outcome. Our results more comfortably place the capacity to prepare for immediate alternative futures on a similar developmental trajectory to other future-oriented behaviors [8, 12], such as the abilities to delay gratification [13, 14], to select an appropriate object to solve a future problem [15, 16], and to save resources for the future [17].

The improvements on our task with increasing age may have been driven by developments in various cognitive components of foresight that typically occur around the third and fourth years [8]. One fundamental component is the capacity to form metarepresentations, which allow an agent to reflect on the relationship between their representation of a given event and the event itself [9, 18, 19]. An effective way to solve our task, for instance, would be to reflect on a representation of the item coming out of one bottom opening of the tube, to recognize that this representation could be incorrect (i.e., to metarepresent), and thus to simultaneously prepare for this possibility as well as for the alternative by covering both openings. This interpretation is consistent with the overall pattern of results from various domains suggesting that a fundamental shift in the representational mind occurs during the middle preschool years [18, 20]. Nevertheless, nonverbal behavioral patterns can often be explained in rich and lean manners [20], and future research may examine

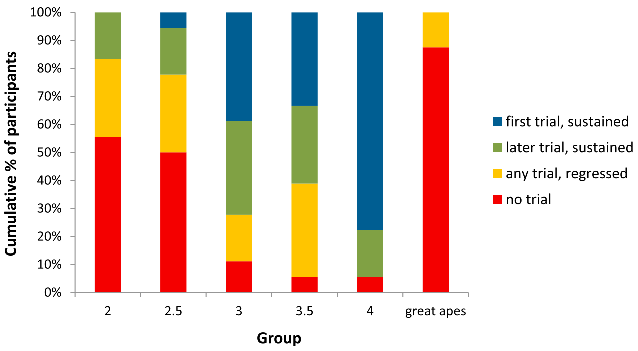

Figure 3. Grouping of Children by Age Group and Great Apes According to When They Covered Both Bottom Openings of the Forked Tube across Twelve Trials

Blue segments: covered both bottom openings on the first trial and all subsequent trials. Green segments: did not cover both openings on the first trial, but eventually covered both openings and sustained that response across all subsequent trials. Yellow segments: covered both openings on at least one trial and subsequently regressed to covering one opening on at least one trial. Red segments: did not cover both openings on any trial.

the feasibility of leaner accounts of the older children's performance.

Whatever the cognitive underpinnings, our results clearly show that 4-year-olds and some 3-year-olds have the capacity to spontaneously and consistently prepare for mutually exclusive future event outcomes. The 2-year-olds and apes, on the other hand, did not provide such evidence. As in other developmental and comparative studies, these negative results have to be interpreted cautiously, as performance may be affected by factors other than the capacity of interest. One may, for instance, be concerned about the role of children's and apes' gravity bias, whereby they expect a dropped item to fall perpendicularly to the ground [21]. Our subjects' performance appeared largely unaffected by this bias, however, as nearly all of them covered at least one bottom opening of the tube on all trials and none of them placed their hands directly underneath the top opening (see Supplemental Experimental Procedures). Another concern may be that the younger children and apes expected the experimenter to cooperate at the beginning of the task (and force the ball/grape to come out of the single covered opening), or that these subjects simply lacked the behavioral flexibility required to switch to using two hands after the experience of using one hand during the practice trials (see Experimental Procedures). Neither of these possibilities, however, can explain why all three of the apes and most of the 2-year-olds that eventually covered two openings on any trial subsequently switched back to covering only one opening (see Figure 3 and Table 1). We also note that most apes and 2-year-olds switched between covering the left and right sides of the forked tube (see Table S2 and Supplemental Experimental Procedures), indicating that they were genuinely considering the contingencies of the task rather than simply perseverating with a single, sub-optimal response that was reinforced 50% of the time.

There may still be reasons other than lack of an ability to spontaneously and consistently prepare for alternative futures that account for why the 2-year-olds and great apes failed our task. One can only provide subjects with opportunities to demonstrate capacities, and although we devised our minimalist paradigm with the aim of giving young children and apes the best possible chance to pass, we acknowledge that future studies with other subjects and/or paradigms may discover some competence. For now, however, we only have compelling evidence for prudent preparatory responses to two mutually exclusive possibilities in 3- to 4-year-old children. Our pattern of results thus points to a gradual development of this capacity over the middle preschool years. The results also raise the possibility that humans' closest extant relatives lack the capacity altogether, despite being capable of responding flexibly to new risks [22] and possibly solving problems with some foresight [23–25]. This would suggest that a new cognitive ability for dealing with environmental uncertainty evolved after the split of the human and chimpanzee lineages approximately 6–8 million years ago [26]. Our study therefore contributes a new perspective to the ongoing debate about what, if anything, is unique to human foresight [7, 9, 27–30] and cognition more generally [20, 31].

The capacity to consider and prepare for multiple, even mutually exclusive versions of future events is central to human adaptability and prudent decision making. Here we have introduced a new minimalist paradigm to chart the ontogeny of this ability and

Table 1. Summary of Great Apes' Performance across All Trials of the Forked-Tube Task

Subject	Species	Sex	Age	Total Number of Trials	First Trial Success?	Any Trial Success?	First Trial Passed	Passed All Subsequent Trials?	Learned to Pass Consistently?
Holly	*Pan troglodytes*	F	25	96	no	yes	9	no	no
Samantha	*Pan troglodytes*	F	30	84	no	yes	15	no	yes
Cassie	*Pan troglodytes*	M	42	48	no	no	–	–	–
Sekara	*Pongo abelii*	F	25	12	no	no	–	–	–
Dinar	*Pongo abelii*	M	28	24	no	yes	16	no	yes
Hsing Hsing	*Pongo abelii*	M	40	24	no	no	–	–	–
Teliti	*Pongo abelii*	F	5	12	no	no	–	–	–
Pulang	*Pongo abelii*	F	21	12	no	no	–	–	–

Subjects received different total numbers of trials because of time constraints and varying levels of cooperativeness. Passing consistently was defined as passing at least six trials in a row. See also Figure S3 and Table S2.

commence investigating its phylogeny. Our findings augur well for future research. Variations of the paradigm may prove valuable not only for establishing the capacities and limits of children and animals, but also for nonverbally examining basic prospection difficulties in certain clinical populations [32].

EXPERIMENTAL PROCEDURES

Subjects

Participants included 90 children (18 each from age groups 2, 2.5, 3, 3.5, and 4 years), three common chimpanzees (*Pan troglodytes*), and five Sumatran orangutans (*Pongo abelii*; mean age of all apes = 27 years, range = 5–42 years). All parents and/or guardians gave informed consent for their children to participate, as approved by the ethics board of the School of Psychology at the University of Queensland. The ape study was approved by the relevant animal ethics boards of the University of Queensland, Rockhampton Zoo, and Perth Zoo.

Materials and Procedure

Participants were first taught how to catch a ball or grape from a basic tube that contained only one opening at the top and one opening at the bottom, and were considered to pass this practice phase after catching three consecutive items. Subjects were then introduced to the forked tube and given six observation trials (in which they could not catch the item), with the item coming out of the bottom openings in the following pseudorandom order: right, left, left, right, left, right (from the experimenter's perspective). Twelve test trials immediately followed, with the subjects given the opportunity to catch the item as it exited the bottom openings in the following pseudorandom order: right, left, left, right, left, right, right, left, right, left, left, right (from the experimenter's perspective). Participants were scored for whether they covered one or two bottom openings when preparing to catch the item on each trial. Further methodological details are described in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.04.062.

AUTHOR CONTRIBUTIONS

J.R. conceptualized, designed, and carried out the experiment, and also conducted the statistical analyses and wrote the manuscript. T.S. assisted in designing the experiment, testing the chimpanzees, and writing the manuscript.

ACKNOWLEDGMENTS

We sincerely thank the parents and children who participated at the Early Cognitive Development Centre, Graeme Strachan and the Rockhampton Zoo for allowing us to work with the chimpanzees, and Holly Thompson and the Perth Zoo for allowing us to work with the orangutans. We also thank Melissa Brinums, Karri Neldner, and Emma Prater for assisting with the testing of the apes and scoring their responses.

Received: February 25, 2016
Revised: April 1, 2016
Accepted: April 29, 2016
Published: June 23, 2016

REFERENCES

1. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. *36*, 181–204.
2. Friston, K. (2010). The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. *11*, 127–138.
3. Behrens, T.E., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221.
4. Bach, D.R., and Dolan, R.J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. Nat. Rev. Neurosci. *13*, 572–586.
5. Bar, M. (2011). Predictions in the Brain: Using Our Past to Generate a Future (Oxford University Press).
6. Suddendorf, T. (2006). Behavior. Foresight and evolution of the human mind. Science *312*, 1006–1007.
7. Suddendorf, T., and Corballis, M.C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? Behav. Brain Sci. *30*, 299–313, discussion 313–351.
8. Suddendorf, T., and Redshaw, J. (2013). The development of mental scenario building and episodic foresight. Ann. N Y Acad. Sci. *1296*, 135–153.
9. Redshaw, J. (2014). Does metarepresentation make human mental time travel unique? Wiley Interdiscip. Rev. Cogn. Sci. *5*, 519–531.
10. Beck, S.R., Robinson, E.J., Carroll, D.J., and Apperly, I.A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. Child Dev. *77*, 413–426.
11. Robinson, E.J., Rowley, M.G., Beck, S.R., Carroll, D.J., and Apperly, I.A. (2006). Children's sensitivity to their own relative ignorance: handling of possibilities under epistemic and physical uncertainty. Child Dev. *77*, 1642–1655.
12. McCormack, T., and Atance, C.M. (2011). Planning in young children: A review and synthesis. Dev. Rev. *31*, 1–31.
13. Mischel, W., Shoda, Y., and Rodriguez, M.I. (1989). Delay of gratification in children. Science *244*, 933–938.
14. Garon, N.M., Longard, J., Bryson, S.E., and Moore, C. (2012). Making decisions about now and later: Development of future-oriented self-control. Cogn. Dev. *27*, 314–322.
15. Suddendorf, T., Nielsen, M., and von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. Dev. Sci. *14*, 26–33.
16. Scarf, D., Gross, J., Colombo, M., and Hayne, H. (2013). To have and to hold: episodic memory in 3- and 4-year-old children. Dev. Psychobiol. *55*, 125–132.
17. Metcalf, J.L., and Atance, C.M. (2011). Do preschoolers save to benefit their future selves? Cogn. Dev. *26*, 371–382.
18. Perner, J. (1991). Understanding the Representational Mind (MIT Press).
19. Pylyshyn, Z. (1978). When is attribution of beliefs justified? Behav. Brain Sci. *1*, 592–593.
20. Suddendorf, T. (2013). The Gap: The Science of What Separates Us from Other Animals (Basic Books).
21. Tomonaga, M., Imura, T., Mizuno, Y., and Tanaka, M. (2007). Gravity bias in young and adult chimpanzees (Pan troglodytes): tests with a modified opaque-tubes task. Dev. Sci. *10*, 411–421.
22. Hockings, K.J., McLennan, M.R., Carvalho, S., Ancrenaz, M., Bobe, R., Byrne, R.W., Dunbar, R.I., Matsuzawa, T., McGrew, W.C., Williamson, E.A., et al. (2015). Apes in the Anthropocene: flexibility and survival. Trends Ecol. Evol. *30*, 215–222.
23. Mulcahy, N.J., and Call, J. (2006). Apes save tools for future use. Science *312*, 1038–1040.
24. Janmaat, K.R., Polansky, L., Ban, S.D., and Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. Proc. Natl. Acad. Sci. USA *111*, 16343–16348.
25. Bräuer, J., and Call, J. (2015). Apes produce tools for future use. Am. J. Primatol. *77*, 254–263.
26. Langergraber, K.E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., Fawcett, K., Inoue, E., Inoue-Muruyama, M., Mitani, J.C., Muller, M.N., et al. (2012). Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. Proc. Natl. Acad. Sci. USA *109*, 15716–15721.

27. Corballis, M.C. (2013). Mental time travel: a case for evolutionary continuity. Trends Cogn. Sci. *17*, 5–6.

28. Suddendorf, T. (2013). Mental time travel: continuities and discontinuities. Trends Cogn. Sci. *17*, 151–152.

29. Osvath, M., and Martin-Ordas, G. (2014). The future of future-oriented cognition in non-humans: theory and the empirical case of the great apes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130486.

30. Clayton, N.S., Bussey, T.J., and Dickinson, A. (2003). Can animals recall the past and plan for the future? Nat. Rev. Neurosci. *4*, 685–691.

31. Penn, D.C., Holyoak, K.J., and Povinelli, D.J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. Behav. Brain Sci. *31*, 109–130, discussion 130–178.

32. Henry, J.D., Addis, D.R., Suddendorf, T., and Rendell, P.G. (2016). Introduction to the Special Issue: Prospection difficulties in clinical populations. Br. J. Clin. Psychol. *55*, 1–3.

Current Biology, Volume *26*

Supplemental Information

Children's and Apes' Preparatory Responses to Two Mutually Exclusive Possibilities

Jonathan Redshaw and Thomas Suddendorf

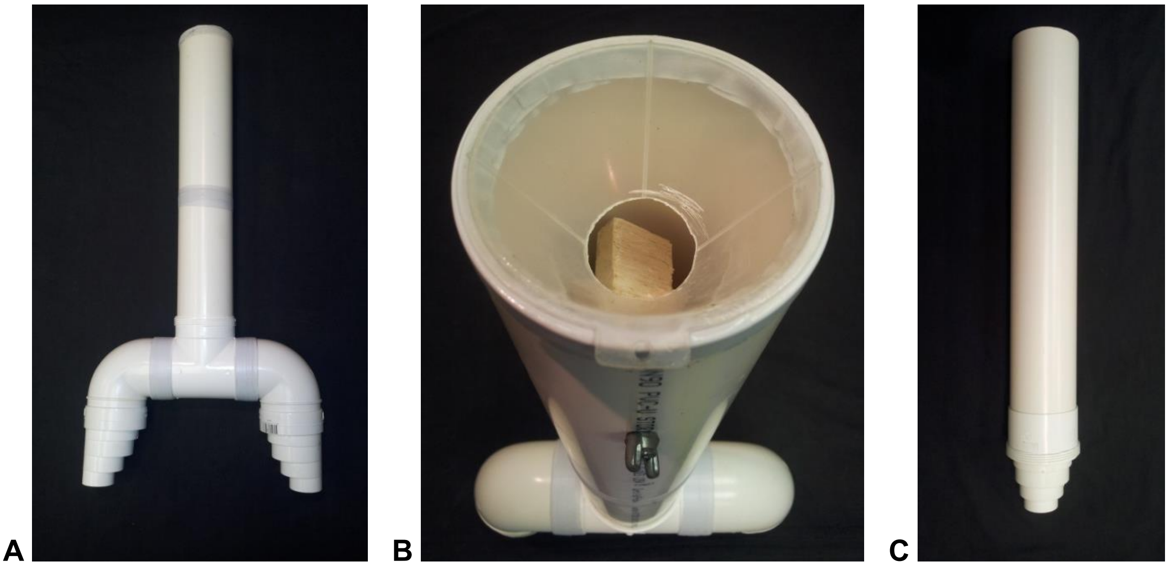

Figure S1 (Related to Figure 1). Depictions of the forked and straight tubes used in the experiment. Panel (**A**) shows the forked tube as it appeared from the children’s and apes’ perspective. Panel (**B**) shows the same forked tube as it appeared from the experimenter’s perspective, revealing the wingnut screw, wooden platform and funnel that allowed control over which bottom opening the balls would fall from. When testing the apes, the tapering bottom openings of the forked tube were changed from ‘stepped’ to ‘smooth’ to prevent the grapes from getting caught inside. The openings were further narrowed for the orangutans to allow them to completely cover both holes through the smaller mesh (see Figure S2B). Panel (**C**) shows the straight tube, which was used only in the practice phase of the experiment.

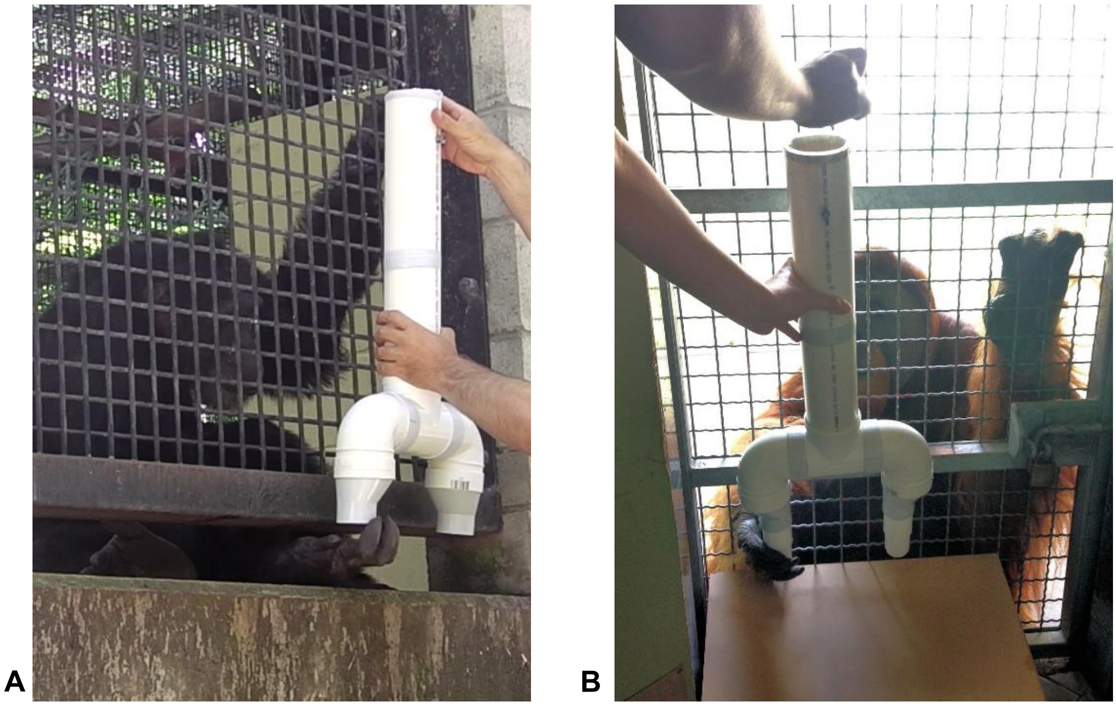

Figure S2 (Related to Figure 1). Experimental setting for (**A**) the chimpanzee Cassie and (**B**) the five orangutans. The experimental setting for the other two chimpanzees can be seen in Figure 1 of the main text.

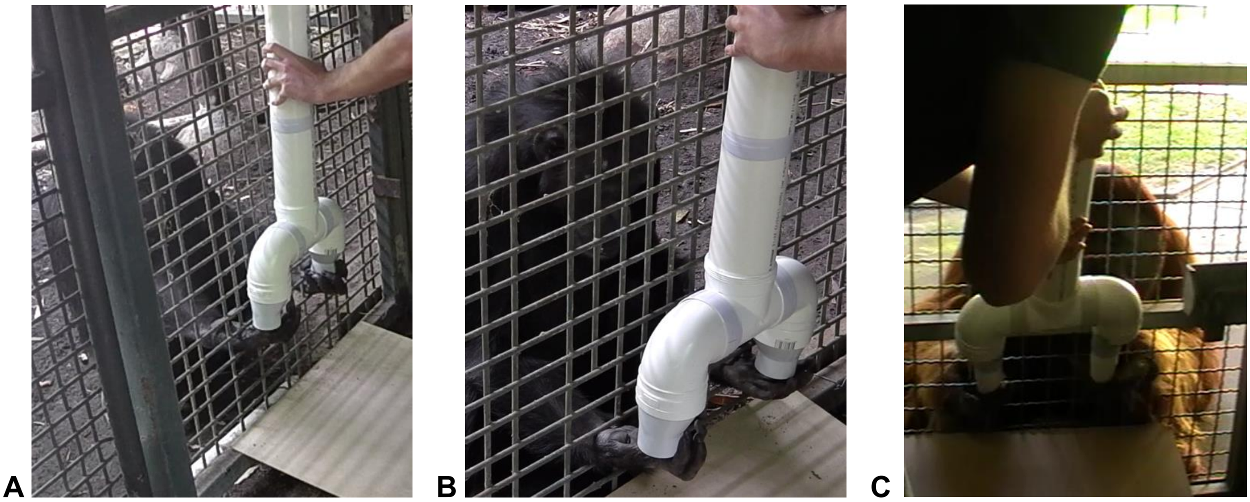

Figure S3 (Related to Table 1). Video stills of (**A**) Samantha, (**B**) Holly, and (**C**) Dinar eventually covering both bottom openings after many trials of covering only one opening.

Table S1 (Related to Figures 2 and 3). Gender split and ages of each group of children.

Age group	Gender split (M/F)	Mean age (months)	SD age (months)
2-year-olds	10/8	24.30	.85
2.5-year-olds	10/8	30.38	.77
3-year-olds	10/8	36.63	.77
3.5-year-olds	8/10	41.97	.33
4-year-olds	9/9	48.78	.65

Table S2 (Related to Table 1). Great apes' performances across all trials of the forked tube task.

Subject	Species	Sex	Age	Total number of trials	Trials with left[a] hole covered only	Trials with right[a] hole covered only	Trials with both holes covered	Longest streak covering both holes
Holly	*Pan troglodytes*	F	25	96[b]	34	56	4	1
Samantha	*Pan troglodytes*	F	30	84	14	36	34	13
Cassie	*Pan troglodytes*	M	42	48	0	48[c]	0	-
Sekara	*Pongo abelii*	F	25	12	12	0	0	-
Dinar	*Pongo abelii*	M	28	24	0	16	8	7
Hsing Hsing	*Pongo abelii*	M	40	24	16	8	0	-
Teliti	*Pongo abelii*	F	5	12	2	10	0	-
Pulang	*Pongo abelii*	F	21	12	11	1	0	-

[a]From the apes' perspective.
[b]Includes the final two trials on which Holly refused to attempt to catch the grape.
[c]On the first trial Cassie initially covered the left hole but then moved to the right hole as the grape was dropped.

Supplemental Experimental Procedures

Methodology for children

Participants. Ninety preschool children participated in the experiment at the University of Queensland in Brisbane, Australia, between August 2013 and February 2014. The children were recruited from a database of parents and caregivers who had previously expressed an interest in participating in early childhood psychology experiments. All children participated individually, with a caregiver present in the experimental room at all times. There were five age groups (2-year-olds, 2.5-year-olds, 3-year-olds, 3.5-year-olds, and 4-year-olds), each consisting of 18 children aged within 2.5 months of the respective group label. Table S1 shows the descriptive statistics for each age group.

Apparatus. The main “forked tube” apparatus (see Figure S1A & S1B) consisted of a number of PVC pipe fittings connected in such a way that the apparatus had one opening at the top but two openings at the bottom. The top arm of the tube was cylindrical, approximately 50cm long, and approximately 9cm in diameter. Inside the top opening was a funnel fastened with tape, such that a ball (or grape) dropped into the tube would fall in approximately the same place inside the tube every time. Approximately 3cm underneath the funnel, on the inside of the tube, was a wooden platform approximately 8cm long x 3cm wide, fastened to the inside of the tube with a wingnut screw accessible on the outside. Underneath the wooden platform was a fixed piece of cardboard the same width as the inner diameter of the tube and running the rest of the length of top arm. The wingnut screw could be turned to rotate the wooden platform in such a way that a dropped item would be forced to run down one side of the cardboard, on either the left or right side of the tube. The bottom of the top arm of the tube connected to an inverse T-section tube, with each side of the T attached to a 90° curved pipe fitting. At the bottom of each curved fitting was a tapering fitting that allowed children to entirely cover each bottom opening of the apparatus with their hands. Once constructed, the apparatus allowed the experimenter to drop a ball into the top opening with full control over which bottom opening the ball would fall from.

Another, more basic pipe apparatus used only in the practice phase of the experiment was the “straight tube”. This apparatus consisted of a single cylindrical pipe, approximately 50cm long with a diameter of approximately 9cm, with the bottom opening connected to a tapering fitting that allowed children to entirely cover the opening with their hands (see Figure S1C). The balls dropped into the straight and forked tubes were spherical polybutadiene “bouncy balls” approximately 2.5cm in diameter. The “ramp” that the balls rolled down was a piece of plywood approximately 90cm long and 60cm across. This ramp was leaned against the top of a small wooden chair approximately 45cm high, such that an uncaught ball would land on the ramp and roll away from a child standing behind the chair.

Procedure.

Practice phase. The children were asked to stand behind the wooden chair, while their caregivers were asked to sit on the other side of the room (or in some cases, behind their child to calm them). The experimenter introduced the children to the straight tube and the bouncy balls, before asking them to place their hands behind their back. The experimenter then stood directly above the ramp and dropped three balls into the straight tube one at a time, such that each ball fell onto the ramp and rolled away from the children. The children were then invited to catch the next ball, with the experimenter demonstrating how to completely cover the narrow opening at the bottom of the straight tube with a single hand. The experimenter continued dropping balls into the straight tube until the children had caught three consecutively, which they were encouraged to place into a small bucket at their feet.

The experimenter then introduced the children to the forked tube, and again asked them to place their hands behind their back and observe. The experimenter again stood directly above the ramp and dropped six balls into the tube such that they fell onto the ramp and rolled away from the children. During this procedure, the experimenter surreptitiously turned the wingnut screw (not visible from the children’s perspective) such that the ball came out of the bottom openings of the tube in the following pseudorandom order: *right, left, left, right, left, right* (from the experimenter’s perspective). Some children brought their hands from behind their back and attempted to reach for the ball during this observation phase, but the experimenter held the tube far enough from them that they could not catch it. The experimenter touched the screw between each trial, whether he turned it or not, such that the children had no obvious external cue as to which side of the tube the ball would fall from on any given trial.

Test phase. After the six observation trials with the forked tube, the experimenter told the children that they could try to catch the balls again (without mentioning the opportunity to cover both holes). The experimenter then stood above the ramp and dropped balls into the forked tube for 12 trials, forcing them to come out of the bottom openings of the tube in the following pseudorandom order: *right, left, left, right, left, right, right, left, right, left, left, right* (from the experimenter’s perspective). In order to minimize ambiguous responses, the experimenter did not drop the ball until the children had settled their hand/s in a fixed position (typically underneath one or both bottom openings). Again, the experimenter touched the wingnut screw between each trial, whether or not he turned it to change the bottom opening that the ball would fall from.

Individual trials immediately followed each other, as long as the children were still interested in participating. Children were encouraged to place caught balls into the bucket at their feet, and all children were rewarded with stickers at the end of the experiment.

All sessions were videotaped, and for each trial the children were scored for whether they put at least part of their hand directly underneath at least part of one or two bottom openings before the ball exited the tube. On some trials the ball bounced off their hand and rolled away as it exited the tube, but as long as they met the above criterion they were still considered to be covering the opening. On early trials, a few younger children initially placed their hands on the tube itself rather than underneath the openings, but they nearly always settled their hands underneath one or both openings before the ball exited the tube. On the few occasions that a child did not cover any openings before the ball exited (e.g., they put their hand/s slightly to the side of the opening/s), they were scored as if they had covered one opening. Only 5 out of 90 children showed this response on any trial. Three 2-year-olds and one 2.5-year-old failed to cover a hole on 1 out of 12 trials; and one 2-year-old refused to attempt to catch the ball on the last 9 trials after she had covered the incorrect opening and missed catching it on the first 3 trials.

Methodology for great apes

Participants. Three captive born and raised adult chimpanzees (one male, two female) participated in the experiment at Rockhampton Zoo in Rockhampton, Australia, between 4 and 6 February, 2014. "Cassie" is a male aged 42 years at the time of testing, who had previously participated in experiments on object permanence understanding [S1-2], inferential reasoning [S3], imitation recognition [S4], metacognition [S5], and mirror and video self-recognition (unpublished). "Holly" is a female aged 25 years at the time of testing, who had previously participated in experiments on metacognition [S5], object permanence understanding and video self-recognition (unpublished). "Samantha" is a female aged 30 years at the time of testing, who had not previously participated in behavioral experiments. The chimpanzees are unrelated and live together in a large zoo enclosure with separate living and sleeping quarters. They are provided with food, medical care and enrichment activities at various times throughout the day by zookeepers. Water was available at all times throughout testing.

Five captive born and raised orangutans (two male, three female) participated in the experiment at Perth Zoo in Perth, Australia, between 1 and 3 October, 2015. At the time of testing, "Dinar" and "Hsing Hsing" (the males) were aged 28 and 40 years respectively; and "Sekara", "Teliti", and "Pulang" (the females) were aged 25, 5, and 21 years respectively. Hsing Hsing is the father of Teliti, but all others are unrelated. All of the orangutans at Perth Zoo are housed in pairs, but only one animal from each pair was tested. During testing, Sekara and Pulang were housed with their infants, Teliti was housed with her mother, and Dinar and Hsing Hsing were housed with unrelated adult females. The orangutans are provided with food, medical care and enrichment activities at various times throughout the day from zookeepers. Water was available at all times throughout testing.

Apparatus. The materials used with the chimpanzees were identical to those used with the children, with a few exceptions. Instead of bouncy balls, the items dropped into the tubes were grapes. The grapes tended to be smaller in diameter than the bouncy balls, and so the tapering fittings of the forked tube openings were changed from "stepped" to "smooth" to prevent the grapes from getting caught in the inner steps of the tube (see Figure 1 in the main text). The openings of the forked tube were further narrowed for the orangutans, as this allowed them to completely cover both openings with their hands through the smaller cage mesh (see Figure S2B). The plywood ramp was used when testing Holly, Samantha, and all of the orangutans, but not Cassie, because of the novel setting in which he was tested (see below for more information).

Experimental setting for the chimpanzees.

Holly and Samantha. The experimental setting for Holly and Samantha is depicted in Figure 1 of the main text. The plywood ramp was placed on the experimenter's side of a meshed wall, with one end sitting on top of a metal bar such that it was raised approximately 20cm from the ground. Approximately 30cm above the ramp were two large holes in the mesh that each allowed the chimpanzees to comfortably put one hand through to the side containing the ramp. These holes were horizontally separated by approximately the same distance that separated the two bottom openings of the forked tube, such that during the experiment the chimpanzees were easily able to cover both bottom openings (with two hands) if they wished. The ramp was raised high enough such that any uncaught grape would roll away from the chimpanzees before they had a chance to grab it. While each chimpanzee was being tested individually, the other two chimpanzees were distracted with attention from research assistants in other areas of the enclosure.

Cassie. The experimental setting had to be altered for Cassie because he refused to sit in the area where Holly and Samantha were tested. To reduce his apprehension, we locked Cassie away from the other chimpanzees in the sleeping area of the enclosure. This area contained a meshed wall with a gap at the bottom of approximately 10cm, such that Cassie was able to comfortably place both of his hands through the gap. On the outside of this wall was a drop to the ground of approximately 120cm, such that a researcher could stand and hold the tube in front of Cassie during the experiment (see Figure S2A). The plywood ramp was unnecessary in this setting, as any uncaught grapes simply fell to the ground on the researcher's side of the wall before Cassie could grab them.

Experimental setting for the orangutans.

The experimental setting for the orangutans is depicted in Figure S2B. The plywood ramp was placed on the experimenter's side of a meshed door, sitting on a crate such that the end of the ramp closest to the mesh was raised approximately 20cm from the ground. The mesh contained holes that were approximately 5.5cm (wide) x 4cm (high), allowing all orangutans to reach through and completely cover both openings with at least two fingers if they wished. All orangutans were tested away from the other individual in their enclosure, except on trials 2-4 for Pulang when her infant was clinging to her back (but not obstructing the use of her hands).

Procedure. The procedure was broadly similar to that used with the children. Like the children, the apes first each received three observation trials and three successful practice trials with the straight tube. They then each received six observation trials and twelve (initial) test trials with the forked tube, with the grape falling out of the bottom openings in the same pseudorandom orders that were used with the children (except on mistrials, see next section). During the observation trials, the tube was held far enough from the meshed wall that the apes could not reach through and catch the grape (observation trials were repeated if the apes were not looking when the grape fell). On test trials, however, the tube was held against the meshed wall such that the apes could easily cover both bottom openings if they wished. As for the child participants, the experimenter touched/turned the wingnut screw (not visible to the apes) after every trial, such that there were no obvious external cues as to the bottom opening that the grape would fall from. On a few test trials, the grape bounced off an ape's hand and fell out of reach even after they had covered the correct bottom opening of the tube. On these trials the ape was given the grape so as to not punish them for a response that would have been otherwise rewarded. Individual trials immediately followed each other, as long as the apes were willing to continue participating.

Extra trials. As described in the main text, some of the apes also received extra test trials (in blocks of twelve) with the forked tube. Each block of twelve trials followed the same pseudorandom order as for the initial twelve trials, with the exception of the special "opposite" trials (given to Holly on trial blocks 4, 5, and 8; Cassie on trial block 4; Dinar on trial block 2; and Hsing Hsing on trial block 2). These trials involved the experimenter waiting for the ape to place one hand underneath one of the bottom openings of the forked tube, before turning/touching the wingnut screw and forcing the grape to come out of the uncovered opening.

Mistrials. Because the grapes were relatively small compared to the bouncy balls used with children, the wingnut screw and wooden platform were not 100% effective at forcing the grapes to come out of the intended bottom opening of the forked tube. Thus, there were a few mistrials in which a grape came out of the unintended opening. With the chimpanzees, mistrials that occurred during the regular pseudorandom order trials were corrected for on later trials, such that the grapes still came out of each bottom opening approximately 50% of the time. Only three mistrials occurred during the opposite trials, all while testing Cassie (trials 37, 44 and 47). They were not considered strongly problematic for interpreting the results, however, given that Cassie did not cover two holes on any of his 48 total trials, and given that we had already demonstrated it was possible to condition chimpanzees to cover two holes over many trials (with Samantha). With the orangutans, mistrials were prevented by the experimenter removing the tube before the grape could come out of the unintended opening, and then proceeding with the trial as intended. Rarely, an ape covered the incorrect opening but still managed to get the grape after it bounced off the ramp and inadvertently landed within their reach; but this did not happen often enough to change the overall pattern of approximately 50% reinforcement on trials where the incorrect opening was covered.

Supplemental results from the great apes.

As described in the main text, two apes (one chimpanzee and one orangutan) covered both bottom openings on at least 6 consecutive trials. The chimpanzee (Samantha) covered both openings on trials 15, 36, 42-45 and 49-60 on the first day of testing; and on trials 61, 63-64, 67, and 73-84 on the second day (she received 84 total trials). Thus on each day she showed signs of regressing to covering only one opening before eventually learning to cover both openings consistently. The orangutan (Dinar) covered both openings on trials 16 and 18-24 on his only day of testing (he received 24 total trials), but regressed to covering only one opening on trial 17. He received no more formal trials, but when we returned briefly the following day to take a photograph of him performing the task he again regressed covering only one opening (see Figure S2B). The response pattern of both of these subjects suggests that they were solving the task in the manner of trial-and-error learning rather than insight. The only other ape to cover both bottom openings was the chimpanzee Holly, who did so on trials 9, 11, 39, and 82 across two days of testing (she received 96 total trials). Even though she had passed four total trials, however, she eventually refused to attempt the task on the last two trials of trial block 8 (which were opposite trials) after her response of covering only one opening had gone unrewarded for ten consecutive trials. Figure S3 depicts Samantha, Holly, and Dinar covering both bottom openings.

Table S2 shows which holes of the forked tube (left, right, or both) were covered by the apes across all trials (both pseudorandom and opposite). Only two apes (Cassie and Sekara) covered the same single opening across all trials (without switching sides or covering both openings), suggesting that most of the subjects were considering the contingencies of the task and not simply perseverating with one response. Across all regular pseudorandom order trials (i.e., not including opposite trials) on which the chimpanzees covered only one opening, Samantha covered the correct bottom opening 20 out of 50 times (40%), Holly 26/57 (45.6%), and Cassie 17/36 (47.2%; although he covered the right opening on all trials, the proportion is not 50% because of mistrials, see above). Among the orangutans, Sekara covered the correct opening 6 out of 12 times (50%), Dinar 6/12 (50%), Hsing Hsing 7/12 (58.3%), Teliti 6/12 (50%), and Pulang 7/12 (58.3%). None of these percentages significantly differed from chance (50%), all $p > .20$, suggesting that the apes were not using any hypothetical external cue that enabled them to cover the appropriate hole during the pseudorandom trials.

Supplemental results from the 2-year-old children.

Considering that the 2-year-old children also performed poorly, their responses were re-coded for which specific holes they covered on each trial. Video of one participant was missing, but among the 17 other 2-year-olds, only 3 of them covered the same single opening on all trials (without switching sides or covering both openings). Like the apes, this suggests that most of the 2-year-olds were genuinely considering the contingencies of the task and not perseverating with a single response. Across all trials on which these 17 2-year-olds covered one opening (160), they covered the correct opening 72 times (45%). This percentage does not significantly differ from chance (50%), $p = .236$, suggesting that the 2-year-olds were not using any hypothetical external cue that enabled them to cover the appropriate hole with their single handed responses.

Supplemental results from the children who passed the first trial.

Across age groups, 34 out of 90 children covered both bottom openings on the first trial. Of these participants, 29 also covered both openings on the second trial (and of this subsample, 28/29 covered both openings on all trials). The other five of these participants (two 2.5-year-olds, two 3-year-olds, one 3.5-year-old), however, subsequently failed the second trial, and all five of them also failed at least one more trial (although all of them did cover both openings again at some stage). The six children who passed the first trial but failed at least one subsequent trial were classified as ‘regressors’ in main text Figure 3 (cf. yellow bars).

Supplemental References

S1. Collier-Baker E., and Suddendorf T. (2006). Do chimpanzees (Pan troglodytes) and 2-year-old children (Homo sapiens) understand double invisible displacement? J. Comp. Psychol. *120*, 89-97.

S2. Collier-Baker E., Davis J.M., Nielsen M., and Suddendorf T. (2006). Do chimpanzees (Pan troglodytes) understand single invisible displacement? Anim. Cogn. *9*, 55-61.

S3. Hill A., Collier-Baker E., and Suddendorf T. (2011). Inferential reasoning by exclusion in great apes, lesser apes, and spider monkeys. J. Comp. Psychol. *125*, 91-103.

S4. Nielsen M., Collier-Baker E., Davis J.M., and Suddendorf T. (2005). Imitation recognition in a captive chimpanzee (Pan troglodytes). Anim. Cogn. *8*, 31-36.

S5. Neldner, K., Collier-Baker, E., and Nielsen, M. (2015). Chimpanzees (Pan troglodytes) and human children (Homo sapiens) know when they are ignorant about the location of food. Anim. Cogn. *18*, 683-699.