

Task-specific modulation of adult humans' tool preferences: number of choices and size of the problem

Kathleen M. Silva · Thomas J. Gross · Francisco J. Silva

Published online: 9 December 2014
© Psychonomic Society, Inc. 2014

Abstract In two experiments, we examined the effect of modifications to the features of a stick-and-tube problem on the stick lengths that adult humans used to solve the problem. In Experiment 1, we examined whether people's tool preferences for retrieving an out-of-reach object in a tube might more closely resemble those reported with laboratory crows if people could modify a single stick to an ideal length to solve the problem. Contrary to when adult humans have selected a tool from a set of ten sticks, asking people to modify a single stick to retrieve an object did not generally result in a stick whose length was related to the object's distance. Consistent with the prior research, though, the working length of the stick was related to the object's distance. In Experiment 2, we examined the effect of increasing the scale of the stick-and-tube problem on people's tool preferences. Increasing the scale of the task influenced people to select relatively shorter tools than had selected in previous studies. Although the causal structures of the tasks used in the two experiments were identical, their results were not. This underscores the necessity of studying physical cognition in relation to a particular causal structure by using a variety of tasks and methods.

Keywords Physical cognition · Folk physics · Tool use · Stick-and-tube problem · Humans

When presented in a laboratory setting with a set of ten sticks of different lengths with which to drag an object out of a horizontal tube, crows frequently selected a stick that matched

the distance over which an object had to be retrieved or the length of the longest stick in the set (Chappell & Kacelnik, 2002). When tested in a naturalistic setting, this same species selected tools whose lengths were positively correlated with the depth of a hole that contained food, though the length of the tool did not reliably match the distance to the food (Bluff, Troscianko, Weir, Kacelnik, & Rutz, 2010). In another study conducted in a naturalistic setting, New Caledonian crows frequently selected on their first attempt twigs or leaf stems that were too short to reach the food, and only selected a twig or stem that was long enough to retrieve the food on subsequent attempts (Hunt, Rutledge, & Gray, 2006). This ability to choose a tool of sufficient length to retrieve an out-of-reach object has also been observed in woodpecker finches (Tebbich & Bshary, 2004) and nonhuman primates (Mulcahy, Call, & Dunbar, 2005).

Instead of a stick whose length matched the object's distance or the longest stick in a set, adult humans prefer sticks that are slightly longer than necessary to retrieve an object (Silva & Silva, 2010). This preference was evident on the first attempt to retrieve an object, and was unaffected by increased experience—that is, by having multiple opportunities to select a stick and retrieve an object (Silva & Silva, 2012). However, this preference was affected by providing fewer sticks from which to select a tool; fewer sticks in a set increased the likelihood of someone choosing the longest stick if the object was relatively far (Silva & Silva, 2010). This preference was also affected by increasing the ambiguity of the problem, in that having less time to study a problem caused people to choose longer sticks (Silva & Silva, 2012). In sum, although it depended somewhat on the distance over which an object had to be retrieved, adult humans shift their preference from a slightly longer-than-necessary stick to an even longer stick when there are fewer tools to choose from and when a tool-use problem is more ambiguous. Table 1 summarizes the key results of studies that have investigated subjects' physical cognition of tool length.

K. M. Silva (✉) · T. J. Gross · F. J. Silva
Department of Psychology, University of Redlands,
P.O. Box 3080, 1200 East Colton Avenue,
Redlands, CA 92373, USA
e-mail: kathleen_silva@redlands.edu

Table 1 Summary of studies examining physical cognition related to tool length

Study	Subjects	Object's Distance (cm)	Number of Tools and Lengths (cm) in Set	Results
Chappell & Kacelnik (2002), Exp. 1	New Caledonian crows	8–26	10 sticks, 8–26	Preferred stick often matched the object's distance
Chappell & Kacelnik (2002), Exp. 2	New Caledonian crow	8–26	10 sticks, 8–26	Preferred stick often the longest stick in the set ¹
Tebbich & Bshary (2004), Exp. 2	woodpecker finches	2–8	5 sticks, 1.5–7.5	Preferred stick slightly longer than object's distance ²
Mulcahy et al. (2005), Exp. 1	gorillas & orangutans	20	2 sticks, 10 vs. 30	Preferred 30-cm stick
Mulcahy et al. (2005), Exp. 1	gorillas & orangutans	25	2 sticks, 15 vs. 30	Preferred 30-cm stick
Mulcahy et al. (2005), Exp. 1	gorillas & orangutans	10	2 sticks, 15 vs. 30	No preference
Mulcahy et al. (2005), Exp. 2	gorillas & orangutans	25 or 35	2 sticks, 25 vs. 35	Preferred 35-cm stick ³
Mulcahy et al. (2005), Exp. 2	gorillas & orangutans	15 or 35	2 sticks, 15 vs. 35	Preferred 35-cm stick ³
Mulcahy et al. (2005), Exp. 2	gorillas & orangutans	15 or 25	2 sticks, 15 vs. 25	Preferred 25-cm stick ³
Mulcahy et al. (2005), Exp. 6	gorillas & orangutans	15 or 25	1 stick, 15 or 45	Refused tool when it was too short to reach object
Hunt et al. (2006), Exps. 1 and 2	wild New Caledonian crows	9 or 15	200+, approx. 6–16 ⁴	First stick's length independent of object's distance ⁵
Bluff et al. (2010)	wild New Caledonian crows	approx. 1–20	200+, approx. 2–30 ⁴	Object's distance a predictor of tool length
Silva & Silva (2010), Exp. 1	adult humans	8 or 16	10 sticks, 8–26	Preferred stick slightly longer than object's distance
Silva & Silva (2010), Exp. 2	adult humans	8	2 sticks, 8 vs. 26	Preferred 26-cm stick
Silva & Silva (2010), Exp. 2	adult humans	8	2 sticks, 8 vs. 12	Preferred 12-cm stick
Silva & Silva (2010), Exp. 2	adult humans	8	2 sticks, 12 vs. 26	No preference
Silva & Silva (2010), Exp. 2	adult humans	16	2 sticks, 16 vs. 26	Preferred 26-cm stick
Silva & Silva (2010), Exp. 2	adult humans	16	2 sticks, 16 vs. 20	Preferred 20-cm stick
Silva & Silva (2010), Exp. 2	adult humans	16	2 sticks, 20 vs. 26	Preferred 26-cm stick
Silva & Silva (2010), Exp. 3	adult humans	8	3 sticks, 8, 12, & 26	Preferred 12-cm stick
Silva & Silva (2010), Exp. 3	adult humans	16	3 sticks, 16, 20, & 26	Preferred 26-cm stick
Silva & Silva (2010), Exp. 3	adult humans	8	3 sticks, 12, 20, & 26	Preferred 20-cm stick
Silva & Silva (2012), Exps. 1a and 1b	adult humans	8 or 16	10 sticks, 8–26	Preferred stick slightly longer than object's distance ⁶
Silva & Silva (2012), Exp. 2	adult humans	8 or 16	10 sticks, 8–26	Preferred stick longer than object's distance ⁷

¹ Tools were placed out of sight of the tube. ² First stick selected was not always successful. ³ Object and tools were presented sequentially rather than simultaneously. ⁴ Tools consisted of leaf stems and twigs. ⁵ If first tool was too short, subsequent tool was likely longer. ⁶ Preference did not change with experience. ⁷ Time to study the problem was limited

On the basis of the results above, it seems that New Caledonian crows', other avian species', some nonhuman primates', and adult humans' understandings of physical causality include the ability to identify the distance to a desired object, to identify a tool whose length is sufficient to reach the object, and to anticipate how they will hold and use the tool to reach this object. But it also seems that adult humans' tool selections are influenced by the number and lengths of tools in a set, as well as perhaps by "cognitive effort," "margin of safety," and "usability." With regard to cognitive effort, a heuristic that favors selecting a tool that is slightly longer than necessary reduces the need for complicated evaluations or precise judgments of an object's distance and a tool's length, and for careful considerations of how a tool will be held and

used to retrieve the object in a particular environment. This same strategy produces a balance between margin of safety and usability. These constructs—cognitive effort, usability, and margin of safety—and related variables, processes, and rules may be why adult humans' tool selections differ from those of crows on similar tasks. In short, people's and other animals' behavior on tool-use tasks may be due to things (e.g., features of the task, associative learning, or procedural rules) other than simply the causal structure of a task and a subject's understanding of physical causality (Silva, Page, & Silva, 2005; Silva & Silva, 2006; Tecwyn, Thorpe, & Chappell, 2012).

In the present study, we examined the effects of two features of the task—number of choices and size of the problem—on

the stick lengths that people used to solve a stick-and-tube problem. Our general goal was to collect information about how these features affect the expression of people's physical cognition, which in turn may help interpret nonhumans' behavior on similar tasks or suggest future experiments with nonhuman subjects (e.g., Hachiga, Silberberg, Parker, & Sakagami, 2009; Silberberg et al., 2013; Silva et al., 2005; Silva & Silva, 2006).

Experiment 1

One paradigm used to investigate which features of tools animals attend to involves presenting the animals with a set of tools that vary in their functional (e.g., length) or nonfunctional (e.g., color) features, and then having experimenters observe the tools that the subjects select. A second paradigm used to investigate which features of tools influence subjects' preferences requires subjects to combine or modify potential tools in order to solve a problem, and having experimenters observe the characteristics of the manufactured tool (e.g., Visalberghi, Frigaszy, & Savage-Rumbaugh, 1995). In Experiment 1, we used a tool modification paradigm to examine whether adult humans' tool preferences might more closely resemble those of laboratory crows (e.g., Chappell & Kacelnik, 2002) if, instead of selecting a tool from a set of ten sticks, people could modify a single stick to what they deemed the ideal length to solve the problem. We chose to study this variable because prior research has shown that the number and lengths of available tools influence people's tool selections (e.g., Silva & Silva, 2010; see also Table 1). In this regard, people's tool preferences may reflect a compromise based on the available choices.

Three outcomes were possible: Participants might prefer sticks whose average length was about the same as, shorter than, or longer than the average length of sticks that people selected in previous studies to retrieve an object at the same distance. Given that the causal structure and dimensions of the task were identical to those used previously, the participants in Experiment 1 might construct stick tools that would be about the same length as those that participants had selected in previous studies (e.g., Silva & Silva, 2010, 2012). Alternatively, given that the number and lengths of available tools influenced people's tool selections (Silva & Silva, 2010), reducing the set of tools to a single stick that could be modified might affect people's tool length preferences. What was unknown was whether people would modify the stick to a length that more closely approximated the distance over which an object had to be retrieved (cf. Chappell & Kacelnik, 2002) or to a length that was longer than had been reported. Although we previously outlined a model of tool length preference (see Silva & Silva, 2010), this model requires at least two tools from which to choose. With only a

single tool that could be modified, several factors could influence the participants' preferences (e.g., the initial length of the stick; the distance of the object; the participant's estimate of the object's distance; or the participant's analysis, assessment, and management of risk). Given that adult humans shift their preferences from a slightly longer-than-necessary stick to an even longer stick when there are fewer tools to choose from, it seemed most likely that the participants in Experiment 1 would construct tools that were longer than those selected by people presented with a set of ten stick tools (e.g., Silva & Silva, 2010).

Method

Subjects A group of 32 undergraduate students (26 females, six males) attending a small liberal arts university in southern California volunteered for the study in partial fulfillment of a course's requirements.

Materials A 0.25-cm-diameter wooden dowel served as the tool that participants could modify and then use to retrieve a candy (M&M) from a transparent tube (4-cm diameter, 30 cm long) that was closed at one end. The dowel was 26 cm long, the same as the length of the longest stick used in previous studies (e.g., Chappell & Kacelnik, 2002; Silva & Silva, 2010, 2012). The tube was secured to a wooden frame, and this apparatus sat on a table. The experimenter placed the object randomly either 8 or 16 cm from the open end of the tube. The dowel was laid flat on the table and was centered and perpendicular to the tube. Figure 1 shows a schematic of the apparatus. For participants who opted to modify the dowel, a marker was used to draw a line on the dowel, and a Stanley 10-in. (25.4-cm) mini hacksaw was used to cut it.

Procedure The participants were run individually. After the experimenter had randomly assigned a participant to retrieve the candy, placed either 8 or 16 cm from the opening of the tube ($n = 16$ for each distance), a trial began with the

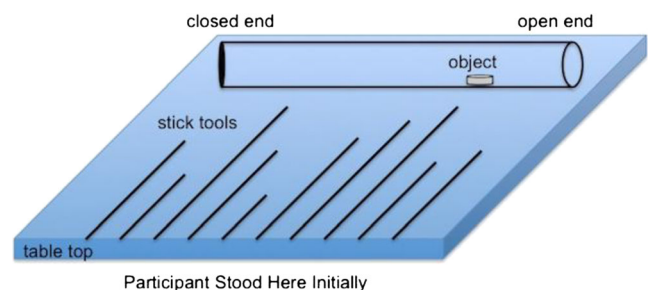


Fig. 1 A schematic (not drawn to scale) of the stick-and-tube problem used in Experiments 1 and 2. Note that in Experiment 1, only a single 26-cm stick that could be modified was available. In Experiment 2, participants could select a tool from a set of ten sticks (as shown). The distance of the object from the open end of the tube was either 8 or 16 cm (Exp. 1) or 68 or 136 cm (Exp. 2)

experimenter asking a participant to enter the room and stand in front of the materials while the experimenter read the instructions. The experimenter told participants that they were participating in a study about how people use tools to solve problems, and that they had one attempt to use the stick to retrieve the candy from the tube. Participants were told that, if they wished, they could shorten the stick by cutting it. In addition, everyone was told explicitly that they did not have to cut the stick if they did not want to. If someone wanted to cut the stick, the experimenter instructed him or her to draw a line on the stick at the location where it would be cut. After drawing the line, a participant was asked to cut the stick with the saw. Once someone cut the stick, he or she was instructed to hold the stick as if ready to retrieve the candy. Participants who did not cut the stick were told the same. When a participant was ready, the experimenter measured the distance between the working end of the stick's tip (i.e., the tip that would contact the candy) and the tip of the participant's finger closest to the working tip. This distance was the working length of the stick. The experimenter then instructed the participant to retrieve the object. The participant could move around to any side of the table to retrieve the candy. After retrieving the object, participants explained their reason for using the stick length that they had used. All statistical tests were two-tailed with $\alpha = .05$.

Results and discussion

Thirteen of 16 participants cut their sticks before attempting to retrieve the candy at the 8-cm distance; a further 14 of 16 participants did the same at the 16-cm distance. Figure 2 shows the mean lengths of the sticks and their working lengths when the candy was 8 or 16 cm from the open end of the tube. For the 8-cm distance, the mean length of the stick was 18.82 cm ($SD = 5.22$); for the 16-cm distance, the mean length was 21.15 cm ($SD = 3.12$). The difference between these mean

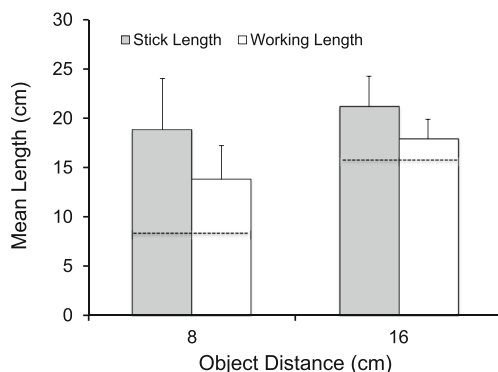


Fig. 2 Mean lengths of the sticks (gray bars) and their working lengths (white bars) in Experiment 1 when the candy was 8 or 16 cm from the open end of the tube. Error bars show the standard deviations. The dashed lines across the bars show the distances of the candy from the open end of the tube

lengths was not statistically significant [independent-samples t test: $t(30) = 1.53$, $p = .136$]. As holding a stick tool requires, the working lengths of the sticks were shorter than their actual lengths. However, the mean working lengths of the sticks were significantly shorter at the 8-cm distance than at the 16-cm distance, 13.78 cm ($SD = 3.43$) versus 17.88 cm ($SD = 2.01$), respectively [independent-samples t test: $t(30) = 4.11$, $p < .001$, Cohen's $d = 1.45$]. Regardless of the object's distance, the portions of the stick held in a participant's hand were not significantly different: 5.04 cm ($SD = 3.68$) at the 8-cm distance and 3.27 cm ($SD = 2.26$) at the 16-cm distance [independent-samples t test: $t(30) = 1.63$, $p = .114$].

The results also showed that both the actual and working lengths of the sticks were significantly longer than the object's distance [single-sample t tests: 8-cm distance, $t(15) > 6.72$, $ps < .0001$, Cohen's $ds > 1.68$; 16-cm distance, $t(15) > 3.72$, $ps < .003$, Cohen's $ds > 0.93$]. Thus, neither the mean length of the sticks nor their working lengths matched the distance over which an object was retrieved.

When asked to explain their reasons for using the stick length that they had used, participants who retrieved the candy over the 8-cm distance often said something about the stick's length and usability (e.g., "I wanted it [the stick] to be as long as possible, but not as long as it was." "If it was too long, the stick would hit the side of the tube, but I needed it to get around the object to pull it in." "I would rather cut and leave the stick longer than shorter to reach it, and I wanted to angle it."). People who retrieved the object over the 16-cm distance also referred to the stick's length and usability (e.g., "The longer the better, and it seemed safer to stay longer." "If I cut the stick too short I may not be able to reach it, but if I left it too long it won't make a difference." "I thought it would be long enough to overreach and pull toward me." "It would just reach the M&M without being too long, and it would be awkward to use if it was too long."). The participants who did not cut the stick reasoned that they could simply adjust how they held it (e.g., "If I wanted a shorter stick, I could hold it up higher to make a shorter stick.>").

Experiment 1 showed that the working length, but not the stick's length, was related to the candy's distance. These results differ from those obtained in previous studies of tool selectivity, in which both the mean length of the stick and its working length were related to an object's distance (Silva & Silva, 2010, 2012). However, these results might reflect a bias similar to those shown by gorillas and orangutans; these animals prefer the longer of two tools, regardless of whether both tools or only the longer one can reach an object (Mulcahy et al., 2005). Relative to our previous studies with adult humans, it seems that the length of the stick used to retrieve an object placed 8 cm from the opening of the tube was most different (see Table 2). In Experiment 1, the mean length of this stick was longer than in the previous studies. The length of the stick used to drag the object 16 cm was more similar to

Table 2 Mean stick lengths from Experiment 1 and previous studies with adult humans

Object's Distance	Present Exp. 1 Mean (SD) <i>n</i>	Silva & Silva (2010), Exp. 1 Mean (SD) <i>n</i>	Silva & Silva (2012), Exp. 2 Mean (SD) <i>n</i>
8 cm	18.82 (5.22) <i>n</i> = 16	15.50 (5.40) <i>n</i> = 13	16.40 (3.23) <i>n</i> = 10
16 cm	21.15 (3.12) <i>n</i> = 16	22.60 (6.60) <i>n</i> = 13	20.80 (2.35) <i>n</i> = 10

the lengths of sticks that had been selected to do the same in previous studies.

A general implication of these results is that they highlight the importance of the task, in terms of both its causal structure and its nonfunctional features, in the study of physical cognition (see Silva et al., 2005; Silva & Silva, 2006; Tecwyn et al., 2012). Despite the fact that we can be reasonably certain that adult humans' physical cognition consists of an intuitive understanding of constructs such as "gravity" and "transfer of force," and that the causal structure and dimensions of the problem used in Experiment 1 were identical to those used previously to study people's folk physics (e.g., Silva & Silva, 2010), modifying a single stick to an ideal length produced different results from selecting a stick from a set of ten. That this difference, which is related to the "evaluation and choice" phase of the problem rather than the "execution and solving" phase, influenced the length of the sticks that people preferred underscores the necessity of studying physical cognition in relation to a particular causal structure by using a variety of tasks and methods (Girndt, Meier, & Call, 2008; Martin-Ordas, Jaeck, & Call, 2012; Seed, Call, Emery, & Clayton, 2009; Tecwyn et al., 2012; Teschke & Tebbich, 2011).

A more specific implication of these results is that they suggest that people's tool selections may be modulated by factors similar to those that influence people's decisions in other situations. For example, the initial length of the tool used in Experiment 1 could anchor people's judgments of how long of a stick they would need to retrieve the object, just as people estimating the percentage of African countries in the United Nations are influenced by giving them an initial value before they estimate the real value. A lower initial value of 10 % results in a lower estimate (25 %) than a higher initial value of 65 %, which produces an estimate of 45 % (see Tversky & Kahneman, 1974). It seems likely that anchoring and other context effects, such as shifts in preferences due to the available choices (e.g., Trueblood, Brown, Heathcote, & Bussemeyer, 2013), influence what people consider the ideal stick length to retrieve an out-of-reach object.

Another set of influences during the evaluation phase may be related to risk—analysis (i.e., identifying unwanted outcomes), assessment (i.e., assigning probabilities and values to unwanted outcomes), and management (i.e., selecting a course of action to reduce the likelihood or intensity of unwanted outcomes). In Experiment 1 and previous studies (e.g., Silva

& Silva, 2010, 2012), several participants had explained that they modified a stick to a particular length or selected a particular stick because that length provided a margin of safety related to estimating the object's distance or potential problems that could occur when retrieving the object. These analyses, assessments, and management of risk are themselves influenced by heuristics and biases, such as the ease with which examples of unwanted events come to mind (i.e., the availability heuristic) or the fact that people are risk-averse and thus attach more value to losses than to comparable gains (Tversky & Kahneman, 1974). Viewed in this manner, it may be difficult to distinguish between "Does the subjects' behavior reflect a conservative strategy related to risk?" versus "Does the subjects' behavior reflect a failure to accurately encode the object's distance?" (e.g., Mulcahy et al., 2005). Perhaps a conservative strategy of using a longer-than-necessary tool reduces the need to accurately encode an object's distance. That every participant in Experiment 1 preferred a stick that was long enough to reach the object shows that any heuristics and biases were not detrimental and may have actually been beneficial (Gigerenzer & Gaissmaier, 2011).

Experiment 2

An anonymous reviewer of a previous study (i.e., Silva & Silva, 2012) had commented that differences between adult humans' and crows' tool selections might be due to the scale of the stick-and-tube problem. What is a body-sized problem for a crow is only a forearm-sized task for an adult human. Sabbatini et al. (2012) made a similar comment after analyzing the differences between chimpanzees' and capuchin monkeys' behavior on a tool-use task. Using the same-sized tools and task to study species of disparate body sizes imposes different demands on the use of the tools, independent of the animals' physical cognitions, that might produce differences in the species' behavior.

A 30-cm tube and a set of 8- to 26-cm stick tools were designed to examine crows' physical cognition, and the scale of this stick-and-tube task was comparable to the bodily dimensions of these birds. Perhaps a crow's tool selection is mediated by mechanisms and processes related to minimizing energy expenditure and procurement time while maximizing

the likelihood of obtaining the out-of-reach food. No such tool selection pressure exists when adult humans are using crow-sized sticks and tubes. A more comparable stick-and-tube task for adult humans would be one whose dimensions were proportionally similar to the size of adult humans just as stick-and-tube tasks used with crows are proportional to the size of their bodies. In Experiment 2, we used a tool selection paradigm to examine the effect that increasing the scale of the stick-and-tube problem from crow-sized to human-sized would have on people's tool preferences.

A larger, human-sized task should be more challenging to use than a smaller, crow-sized task. For this reason, the selection pressure may favor choosing a stick tool that is relatively shorter, lighter, and easier to maneuver than those chosen in previous studies with a crow-sized tube and sticks. However, if the causal structure of the task trumps the influence of its scale, then the stick selected should be proportionally equivalent in length (i.e., of the same ordinal value within the tool set) to those selections reported previously (e.g., Silva & Silva, 2010, 2012).

Method

Subjects A group of 64 undergraduate students (43 females, 21 males) attending a small liberal arts university in southern California volunteered for the study in partial fulfillment of a course's requirements.

Materials Ten 1.6-cm-diameter dowels, ranging in length from 68 to 221 cm in increments of 17 cm, served as the tools. Participants selected one of the sticks to retrieve a puck-shaped foam object (10.0-cm diameter, 1.50-cm height) from a transparent tube (34-cm diameter, 244 cm long) closed at one end. The tube was secured to a wooden frame, and the whole apparatus sat on several tables pulled together.

The experimenter placed the object randomly either 68 or 136 cm from the open end of the tube. These distances were measured from the edge of the object farthest from the opening of the tube and were proportionally similar to the 8- and 16-cm distances used with the smaller apparatuses in Experiment 1 and previous studies. The sticks were laid flat on several tables and were centered and perpendicular to the tube. Each stick was separated by 15 cm from an adjacent stick, the bottoms of the sticks were aligned, and the positions of the sticks varied randomly between participants. Figure 1 shows a schematic of the apparatus and the setup.

Procedure The participants were run individually. After the experimenter had randomly assigned a participant to retrieve the object placed either 68 or 136 cm from the opening of the tube ($n = 32$ for each distance), a trial began with the experimenter asking a participant to enter the room and stand in front of the materials while the experimenter read the

instructions. The experimenter told participants that they were participating in a study about how people use tools to solve problems, and that they had one attempt to select and use a stick to retrieve the object from the tube. When a participant was ready, the experimenter asked him or her to select a stick and to hold it as if he or she was ready to retrieve the object. The experimenter recorded whether a participant held the stick with one or two hands and measured the working length of the stick (as in Exp. 1). The participant could move around to any side of the tables to retrieve the object. Participants were not restricted from inserting their hands into the tube. If a participant inserted his or her hand(s) into the tube, the experimenter measured the distance between the open end of the tube and the tip of the participant's finger that was inserted farthest into the tube. After retrieving the object, participants were asked to explain their reason for selecting the stick that they had used. All statistical tests were two-tailed with $\alpha = .05$.

Results and discussion

Seventeen of 32 participants (53 %) inserted at least one of their hands into the tube when the object was 68 cm from the opening; a further 21 of 32 participants (66 %) did the same at the 136-cm distance. These proportions were not significantly different ($z = 1.02$, $p = .31$). Nine of the participants (28 %) used two hands to hold the stick and retrieve the object when it was 68 cm from the opening of the tube; 14 participants (44 %) did the same at the 136-cm distance. These proportions also were not significantly different ($z = 1.30$, $p = .19$). Finally, the mean distances that the participants inserted their hand(s) into the tube were not significantly different: 4.38 cm ($SD = 6.47$) and 7.59 cm ($SD = 8.67$), for the 68- and 136-cm distances, respectively [independent-samples t test: $t(62) = 1.66$, $p = .102$]. In sum, the manners in which participants used the stick tools were similar regardless of the object's distance.

Figure 3 shows the mean lengths of the sticks selected by the participants and the sticks' working lengths when the object was 68 or 136 cm from the open end of the tube. For

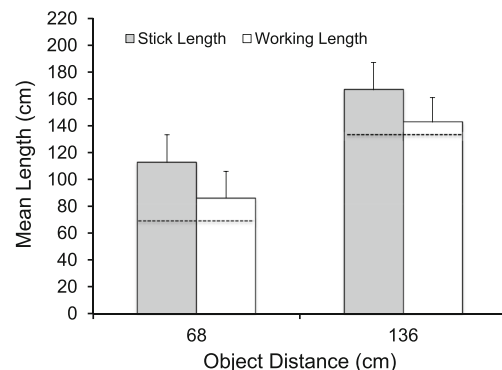


Fig. 3 Mean lengths of the sticks (gray bars) and their working lengths (white bars) in Experiment 2 when the object was 68 or 136 cm from the open end of the tube. The details are the same as in Fig. 2

the 68-cm distance, the mean length of the sticks selected was significantly shorter ($M = 112.63$ cm, $SD = 20.60$) than that of the sticks selected at the 136-cm distance ($M = 166.82$ cm, $SD = 20.90$) [independent-samples t test: $t(62) = 10.45$, $p < .001$, Cohen's $d = 2.61$]. As holding a stick tool requires, the working lengths of the sticks were shorter than their actual lengths. However, the mean working lengths of the sticks were significantly shorter at the 68-cm distance than at the 136-cm distance: 86.03 cm ($SD = 19.97$) versus 143.03 cm ($SD = 18.79$), respectively [independent-samples t test: $t(62) = 11.76$, $p < .0001$, Cohen's $d = 2.94$]. Regardless of the object's distance, the portions of the sticks held in a participant's hand(s) were not significantly different, 26.59 cm ($SD = 10.26$) at the 68-cm distance and 23.78 cm at the 136-cm distance ($SD = 14.89$) [independent-samples t test: $t(62) = 0.88$, $p = .382$].

The results also showed that both the actual and working lengths of the sticks were significantly longer than the object's distance [single-sample t tests: 68-cm distance, $t_s(31) > 5.10$, $ps < .001$, Cohen's $ds > 0.90$; 136-cm distance, $t_s(31) > 2.12$, $ps < .05$, Cohen's $ds > 0.37$]. Thus, neither the mean length of the sticks nor their working lengths matched the distance over which the object had to be retrieved.

Although the mean lengths of the sticks and their working lengths did not match the retrieval distance, we could check whether the matching-length stick was selected more or less often than would be expected by chance. It was not. No one selected a matching stick at the 68-cm distance [binomial parameters: $n = 32$, $k = 0$, $p = .1$, P of 0 or more out of 32 = 1.0], and only three participants selected the matching stick at the 136-cm distance [binomial parameters: $n = 32$, $k = 3$, $p = .1$, P of 3 or more out of 32 = .63]. None of the 64 participants selected the longest stick in the set [binomial parameters: $n = 64$, $k = 0$, $p = .1$, P of 0 or more out of 64 = 1.0].

When asked to explain their selections, most people said something about the stick's length in relation to its usability. For example, some participants who retrieved the object over the 68-cm distance said, "The stick was long enough to reach, but not so long that I couldn't maneuver it" and "I didn't choose the longer because it would be hard to control. I would have more control with a shorter stick." Participants who retrieved the object over the 136-cm distance provided similar reasons: "It [the stick] was long enough but I still have control" and "I chose on the longer side in case it was too short, but not too long because it would be too difficult." These reasons were similar to those of people who solved the smaller, crow-sized task (e.g., Silva & Silva, 2010, 2012, and Exp. 1 above).

Experiment 2 showed that the lengths of the sticks selected by the participants were influenced by the object's distance. Relative to previous studies in which adult humans had selected a stick to retrieve an object that was 8 or 16 cm from the opening of the tube, using a human-sized tube and sticks

resulted in the participants choosing a stick whose length was more similar to the object's distance. Table 3 shows the mean standardized lengths of the sticks selected in Experiment 2 and in previous studies in which the object's distance was 8 or 16 cm. To obtain the standardized means and standard deviations, we divided the actual means and standard deviations by the length of the longest stick in an experiment (221 cm in Exp. 2 and 26 cm in the previous studies); thus, the greater the standardized mean, the more similar the length of the average stick was to the length of the longest stick in a set.

As is shown in Table 3, people selected relatively shorter sticks when the stick-and-tube task was more than 8 times larger than in previous studies (e.g., Silva & Silva, 2010, 2012). It seems that increasing the scale of the task did make adult humans' tool selections more similar to those of laboratory crows, in that the preferred stick length was closer to the object's distance. Although this outcome could be a direct result of the scale of the task, it could also be an indirect result of this scale. Specifically, because participants could insert their hands and arms into the large tube, they could select a relatively shorter stick and still successfully retrieve the object. This is similar to a crow that inserts its bill into a tube to retrieve an object with a stick. In these circumstances, the functional length of a stick is made longer by inserting a hand (in the case of humans) or a bill (in the case of crows) into a tube to get the object. Despite this similarity in people's and laboratory crows' tool selections and use, people continued to select tools that were longer than the objects' distances, but not the stick that matched this distance or the longest stick in the set, two strategies that have been observed with some crows (Chappell & Kacelnik, 2002; Wimpenny, Weir, Clayton, Rutz, & Kacelnik, 2009; but see Hunt et al., 2006). Overall, it appears that the reasons for selecting a stick that is longer than necessary, but not the same length as the object's distance or the longest stick in a set, were the same as in previous studies: Sticks that are longer than the distance to the object are long enough to reach the object while being maneuverable enough to have control while retrieving the object.

Table 3 Mean standardized stick lengths from Experiment 2 and previous studies with adult humans

Object's Distance	Present Exp. 2 Mean (SD) <i>n</i>	Silva & Silva (2010) Exp. 1 Mean (SD) <i>n</i>	Silva & Silva (2012) Exp. 2 Mean (SD) <i>n</i>
Near	.52 (.09) <i>n</i> = 32	.60 (.21) <i>n</i> = 13	.63 (.12) <i>n</i> = 10
Far	.75 (.09) <i>n</i> = 32	.87 (.25) <i>n</i> = 13	.80 (.09) <i>n</i> = 10

In Experiment 2, the near and far distances were 68 and 136 cm, respectively, from the opening of the tube. In the previous experiments, these distances were 8 and 16 cm, respectively

General discussion

The purpose of this study was to examine the effect of modifications to the features, but not to the causal structure, of a stick-and-tube problem on the stick lengths that adult humans would use to solve the problem. In Experiment 1, contrary to results obtained when adult humans selected a tool from a set of ten sticks (e.g., Silva & Silva, 2010, 2012), asking participants to modify a single stick to what they considered its ideal length for retrieving an out-of-reach object in a tube did not yield a stick length that was generally related to the object's distance. Consistent with prior research, though, the average working length of the stick was related to the object's distance: An object positioned 8 cm from the opening of the tube resulted in people using sticks with shorter working lengths than they did when the object was 16 cm from the opening of the tube.

It is unclear why providing participants with only a single stick that they could modify instead of ten sticks from which to select a tool produced different results. We mentioned above that people's tool selections may be modulated by factors similar to those that influence people's judgments and decisions in other situations—anchoring, heuristics, biases, the number and types of available choices, and risk analysis, assessment, and management (see Gigerenzer & Gaissmaier, 2011; Kahneman & Tversky, 1979; Trueblood et al., 2013; Tversky & Kahneman, 1974). In relation to risk, something more permanent or uncorrectable about cutting an object—whether that object is a dowel, a piece of lumber, or a sheet of paper—might cause people to err on the side of caution. Carpenters and builders are aware of the adage, “Measure twice, cut once.” The core of this adage is that if you cut wood improperly short, it may be unusable. To protect against cutting something too short, one should leave it a little longer. The influence of the object's distance also might be modulated by past experiences and rules similar to “cut no more than one-third of your lawn's height.” Given the initial length of the stick and the objects' distances, it seems that participants did not want to remove more than $x\%$ of the stick's length. Or perhaps the difference between the results of Experiment 1 and those of related studies with adult humans (e.g., Silva & Silva, 2010, 2012) is that the absence of a set of sticks from which to select a tool might have resulted in a shift from using relational rules (“select the longest stick” or “select the stick that is just longer than the object's distance”) to more absolute rules (e.g., “cut no more than $x\%$ of the stick's length”).

Regardless of the reason, it seems that the lengths of the sticks used to retrieve the object over the 8-cm distance in Experiment 1 differed from the lengths of the sticks that had been selected to do the same in previous studies (see Table 2). Had the objects' distances in Experiment 1 been more dissimilar (e.g., 4 vs. 24 cm), the lengths of the actual sticks modified

by the participants might have been related to the objects' distances. Indeed, Mulcahy et al. (2005) showed that gorillas' and orangutans' preferences for longer tools were more pronounced when the differences between the tools' lengths were more pronounced. Although evaluating the correctness of this possibility and others—especially those related to objective quantities (e.g., the object's distance, the length of the tools, or the number of tools in a set)—requires parametric investigations, a potentially fruitful line of research is the development of specific quantitative models to guide such investigations and provide a context in which to interpret results (see Allen, 2014; Chappell & Hawes, 2012).

Note that there was not a major cost to leaving a stick unnecessarily long by a few centimeters (see also Mulcahy et al., 2005). A satisficing strategy (Simon, 1956) that involves selecting or manufacturing a stick that is long enough to reach the object but not too heavy or unwieldy is “good enough” if a participant adjusts how she or he holds the stick in order to compensate for any subjective imperfection in its actual length. The different average working lengths corresponding to the objects' distances shows that this is indeed what the participants did. Such a strategy also reduces the cognitive effort related to estimating lengths, distances, and use. It is worth noting that other primates sometimes prefer longer tools even when a shorter tool is good enough (Mulcahy et al., 2005). Whether these nonhumans' behavior simply accords with a satisficing strategy or whether these animals problem-solve with deliberate attention to satisficing may be impossible to say.

In Experiment 2, substantial changes to the scale of the tools and the problem produced results that differed from those reported in previous studies with adult humans (Silva & Silva, 2010, 2012). However, when asked to justify their selections, people's explanations were similar to those reported previously. It seems that people chose a stick that was long enough to reach the object, but not so long as to impair its usability. This length-to-usability ratio is important enough that people consider it when evaluating both a crow-sized stick-and-tube problem and a similar, human-sized problem. Despite this similarity of reasoning, increasing the scale of the problem influenced people to select a relatively shorter tool, perhaps because increasing the size of the sticks to human size altered the usability portion of the length-to-usability ratio: Larger and heavier sticks may have resulted in greater perceptible stick-to-stick usability than when crow-sized sticks were used. Because people could insert their hands and arms into the tube, the selection pressure may have favored choosing a slightly shorter stick tool that was easier to use.

In the case of adult humans, the emerging picture is that their tool selections are influenced by their understanding of the causal structure of a task and its specific features, such as the number and lengths of stick tools, how much time people have to study a problem, and the scale of the task. The effect of

these features may be mediated by variables and processes related to judgment and decision-making, such as cognitive effort, heuristics, biases, usability, and risk. It remains for future research to determine whether these and other factors similarly influence the expression of nonhuman animals' physical cognition.

To some extent, the precise reasons for the absence of a clear relation between an object's distance and the lengths of the sticks that people modified (Exp. 1) or why a larger stick-and-tube problem influenced people to select relatively shorter sticks (Exp. 2) may be less important than the fact that changes to the features of the task (e.g., whether subjects selected or modified a stick, or the size of the sticks and tube) affected the lengths of the sticks that people preferred when solving the problem. In Experiment 1, the causal structure of the task and its scale were identical to those used in previous studies (e.g., Silva & Silva, 2010, 2012). In Experiment 2, the causal structure and procedure were identical to those used in previous studies (e.g., Silva & Silva, 2010, 2012). The results of these experiments illustrate that the specific features of the task used to study physical cognition cannot be overlooked or devalued; people are sensitive to more than the casual structure of the task. Changing these nonfunctional features changes people's behavior, even when their folk physics or understanding of the problem may be unchanged (see also Silva et al., 2005; Silva & Silva, 2006).

The same rationale applies to the study of nonhumans' physical cognition. In studies designed to pit a high-level, generalizable, cognitive view of an animal's folk physics against a low-level, task-specific, perception-based view (e.g., Povinelli, 2000), the outcome of this competition is relatively unimportant if a change in the nonfunctional features of the task, but not in its causal structure, produces different results (e.g., Martin-Ordas et al., 2012). This is one reason why there are no widely agreed-upon litmus tests for a particular cognitive ability (Seed, Seddon, Greene, & Call, 2012), why researchers are calling for an analysis of the component parts of "insightful" tool-use behavior (Cheke, Bird, & Clayton, 2011; Seed & Boogert, 2013; Shettleworth, 2009), and why it is necessary to elucidate causes in terms of a multitude of mechanisms and processes (e.g., inhibitory control, attention shifting, or working memory capabilities) and variables (e.g., the size of the problem and number of tools) before concluding that an animal does or does not possess a particular cognitive ability (Seed et al., 2012; Tecwyn et al., 2012). Experiments with adult humans may help researchers by providing a context in which to interpret nonhumans' behavior on similar tasks and by identifying key variables that should be controlled and understood before forming conclusions about a species's physical cognition (see also Boogert, Arbilly, Muth, & Seed, 2013).

The real challenge, though, is sketching the conditions for a comparative study of physical causality. As a reviewer asked:

What are the critical conditions for making valid cross-species comparisons? How many variations are enough? What are the most important or revealing manipulations? Specific answers are elusive, but studies of concept learning have provided general guidelines (e.g., Katz, Wright, & Bodily, 2007; Vonk & MacDonald, 2002; Zentall, Galizio, & Critchfield, 2002). For example, what conditions are necessary before a researcher can claim that a subject understands "shapes" or "colors"? Presumably, the answer is that the subject responds appropriately to certain specifiable characteristics across a wide range of situations. A subject that understands shapes might differentially respond to keys with a square, a triangle, and a circle on them. Or perhaps a subject that understands shapes might correctly point to different shapes when prompted to do so. Or perhaps a subject that understands shapes can say "square," "triangle," or "circle" when asked the question, "What shape is that?" The same is true when one is trying to determine whether a subject understands the numerosity concepts "more" versus "less," the temporal concepts "long duration" versus "short duration," or paintings by Monet versus Picasso. How many variations are enough before we conclude that a subject understands shape, color, numerosity, time, art, or the relationship between tool length and an object's distance? No specific criteria exist, and most if not all understanding breaks down at boundary conditions. Even something as simple for most adult humans as distinguishing between "red" and "not red" becomes impossible in the dark or if the stimuli are presented so rapidly as to be almost imperceptible. The best that we can say is that demonstrations of understanding require that subjects make a common response to a class of stimuli that have certain specifiable characteristics. The more varied the responses and the features of the task, the more evidence we have of a subject's understanding of these specifiable characteristics. Along the way, we will collect information about what influences the expression of this understanding.

Author note We thank Thomas F. Gross for helping to construct the apparatus used in the experiments.

References

- Allen, C. (2014). Models, mechanisms, and animal minds. *The Southern Journal of Philosophy*, 52, 75–97. doi:10.1111/sjp.12072
- Bluff, L. A., Troscianko, J., Weir, A. A. S., Kacelnik, A., & Rutz, C. (2010). Tool use by wild New Caledonian crows *Corvus moneduloides* at natural foraging sites. *Proceedings of the Royal Society B*, 277, 1377–1385. doi:10.1098/rspb.2009.1953
- Boogert, N. J., Arbilly, M., Muth, F., & Seed, A. M. (2013). Do crows reason about causes or agents? The devil is in the controls. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E373. doi:10.1073/pnas.1219664110
- Chappell, J., & Hawes, N. (2012). Biological and artificial cognition: What can be learned about mechanisms by modeling physical

- cognition problems using artificial intelligence planning techniques? *Philosophical Transactions of the Royal Society B*, 367, 2723–2732. doi:10.1098/rstb.2012.0221
- Chappell, J., & Kacelnik, A. (2002). Tool selectivity in a non-primate, the New Caledonian crow (*Corvus moneduloides*). *Animal Cognition*, 5, 71–78. doi:10.1007/s10071-002-0130-2
- Cheke, L., Bird, C., & Clayton, N. (2011). Tool-use and instrumental learning in the Eurasian jay (*Garrulus glandarius*). *Animal Cognition*, 14, 441–455. doi:10.1007/s10071-011-0379-4
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482. doi:10.1146/annurev-psych-120709-145346
- Girndt, A., Meier, T., & Call, J. (2008). Task constraints mask great apes' ability to solve the trap-table task. *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 54–62. doi:10.1037/0097-7403.34.1.54
- Hachiga, Y., Silberberg, A., Parker, S., & Sakagami, T. (2009). Humans (*Homo sapiens*) fail to show an inequity effect in an “up-linkage” analog of the monkey inequity test. *Animal Cognition*, 12, 359–367. doi:10.1007/s10071-008-0195-7
- Hunt, G. R., Rutledge, R. B., & Gray, R. D. (2006). The right tool for the job: What strategies do wild New Caledonian crows use? *Animal Cognition*, 9, 307–316. doi:10.1007/s10071-006-0047-2
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291. doi:10.2307/1914185
- Katz, J. S., Wright, A. A., & Bodily, K. D. (2007). Issues in the comparative cognition of abstract-concept learning. *Comparative Cognition & Behavior Reviews*, 2, 79–92. doi:10.3819/ccbr.2008.20005
- Martin-Ordas, G., Jaeck, F., & Call, J. (2012). Barriers and traps: Great apes' performance in two functionally equivalent tasks. *Animal Cognition*, 15, 1007–1013. doi:10.1007/s10071-012-0504-z
- Mulcahy, N. J., Call, J., & Dunbar, R. I. M. (2005). Gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) encode relevant problem features in a tool-using task. *Journal of Comparative Psychology*, 119, 23–32. doi:10.1037/0735-7036.119.1.23
- Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. New York, NY: Oxford University Press.
- Sabbatini, G., Truppa, V., Hribar, A., Gambetta, B., Call, J., & Visalberghi, E. (2012). Understanding the functional properties of tools: Chimpanzees (*Pan troglodytes*) and capuchin monkeys (*Cebus apella*) attend to tool features differently. *Animal Cognition*, 15, 577–590. doi:10.1007/s10071-012-0486-x
- Seed, A. M., & Boogert, N. J. (2013). Animal cognition: An end to insight? *Current Biology*, 23, R67–R69. doi:10.1016/j.cub.2012.11.043
- Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 23–34. doi:10.1037/a0012925
- Seed, A. M., Seddon, E., Greene, B., & Call, J. (2012). Chimpanzee “folk physics”: Bringing failures into focus. *Philosophical Transactions of the Royal Society B*, 367, 2743–2752. doi:10.1098/rstb.2012.0222
- Shettleworth, S. J. (2009). Animal cognition: Deconstructing avian insight. *Current Biology*, 19, R1039–R1040. doi:10.1016/j.cub.2009.10.022
- Silberberg, A., Parker, S., Allouch, C., Fabos, M., Holberman, H., McDonald, L., . . . Wyatt, L. (2013). Human risky choice in a repeated-gambles procedure: An up-linkage replication of Lakshminarayanan, Chen and Santos (2011). *Animal Cognition*, 16, 907–914. doi:10.1007/s10071-013-0623-1
- Silva, F. J., Page, D. M., & Silva, K. M. (2005). Methodological-conceptual problems in the study of chimpanzees' folk physics: How studies with adult humans can help. *Learning & Behavior*, 33, 47–58. doi:10.3758/bf03196049
- Silva, F. J., & Silva, K. M. (2006). Humans' folk physics is not enough to explain variations in their tool-using behavior. *Psychonomic Bulletin & Review*, 13, 689–693. doi:10.3758/bf03193982
- Silva, F. J., & Silva, K. M. (2010). How do adult humans compare with New Caledonian crows in tool selectivity? *Learning & Behavior*, 39, 38–95. doi:10.3758/LB.38.1.87
- Silva, F. J., & Silva, K. M. (2012). More but not less uncertainty makes adult humans' tool selections more similar to those reported with crows. *Learning & Behavior*, 40, 494–506. doi:10.3758/s13420-012-0069-3
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138. doi:10.1037/h0042769
- Tebich, S., & Bshary, R. (2004). Cognitive abilities related to tool use in the woodpecker finch, *Cactospiza pallida*. *Animal Behaviour*, 67, 689–697. doi:10.1016/j.anbehav.2003.08.003
- Tecwyn, E. C., Thorpe, S. K. S., & Chappell, J. (2012). What cognitive strategies do orangutans (*Pongo pygmaeus*) use to solve a trial-unique puzzle-tube task incorporating multiple obstacles? *Animal Cognition*, 15, 121–133. doi:10.1007/s10071-011-0438-x
- Teschke, I., & Tebich, S. (2011). Physical cognition and tool-use: Performance of Darwin's finches in the two-trap tube task. *Animal Cognition*, 14, 555–563. doi:10.1007/s10071-011-0390-9
- Trueblood, J. S., Brown, S. D., Heathcote, A., & Bussemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, 24, 901–908. doi:10.1177/0956797612464241
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131. doi:10.1126/science.185.4157.1124
- Visalberghi, E., Frigaszy, D. M., & Savage-Rumbaugh, S. (1995). Performance in a tool-using task by common chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), an orangutan (*Pongo pygmaeus*), and capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 109, 52–60. doi:10.1037/0735-7036.109.1.52
- Vonk, J., & MacDonald, S. E. (2002). Natural concepts in a juvenile gorilla (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behavior*, 78, 315–332. doi:10.1901/jeab.2002.78-315
- Wimpenny, J. H., Weir, A. A. S., Clayton, L., Rutz, C., & Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in New Caledonian crows. *PLoS ONE*, 4, e6471. doi:10.1371/journal.pone.0006471
- Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior*, 78, 237–248. doi:10.1901/jeab.2002.78-237