Yuniarto, Laras S., Gerson, Sarah A. and Seed, Amanda M. 2020. Better all by myself: Gaining personal experience, not watching others, improves 3-year-olds’ performance in a causal trap-task. Journal of Experimental Child Psychology 194 , 104792. 10.1016/j.jecp.2019.104792

Publishers page: http://doi.org/10.1016/j.jecp.2019.104792

Better all by myself: Gaining personal experience, not watching others, improves 3-year-olds' performance in a causal trap-task

Laras S. Yuniarto [a], Sarah A. Gerson [a,1,2,*], Amanda M. Seed [a,2]

[a] School of Psychology & Neuroscience, University of St Andrews, St Andrews KY16 9JP, Scotland, UK. E-mail addresses: ly21@st-andrews.ac.uk (L.S. Yuniarto), ams18@st-andrews.ac.uk (A.M. Seed).

[1] Present address: School of Psychology, Cardiff University, Cardiff CF10 3AT, Wales, UK. E-mail address: GersonS@cardiff.ac.uk (S.A. Gerson).

[2] S.A.G. and A.M.S. contributed equally to the work.

* Corresponding author at: School of Psychology, Cardiff University, Cardiff CF10 3AT, Wales, UK. E-mail address: GersonS@cardiff.ac.uk (S.A. Gerson).

Declarations of interest: None.

Abstract

Children often learn from others' demonstrations, but in the causal domain, evidence acquired from observing others may be more ambiguous than evidence generated for oneself. Prior work involving tool-using tasks suggests that observational learning may not provide sufficient information about the causal relations involved, but it remains unclear whether these limitations can by mitigated by providing demonstrations using familiar manual actions rather than unfamiliar tools. We provided 2.5- to 3.5-year-old children (n = 67) with the opportunity to acquire experience with a causal trap-task by hand or by tool, actively or from observing others. Initially, children either generated their own experience or watched a yoked demonstration; all children then attempted the trap-task with the tool. Children who generated their own experience outperformed those who watched demonstration. Hand- or tool-use had no effect on performance with a tool. The implications of these findings for scaffolding self-guided learning and for demonstrations involving errors are discussed.

Keywords: Children, Tool Use, Observational Learning, Active Learning, Causal Learning

Highlights

- Children received different types of initial experience with a causal trap-task.
- From age 3, generating one's own experience led to above-chance performance.
- If the same experience was given through demonstration, performance was at chance.
- Using a tool was associated with more perseveration than using one's hands.

Introduction

Children learn a great deal through faithfully copying the form of an action. Such high-fidelity copying has been suggested to bootstrap the process of cultural transmission, helping children acquire and reproduce the actions needed for complex behaviours such as tool-use without necessarily understanding the underlying mechanism (Csibra & Gergely, 2009). Yet action copying may hinder causal learning, because one must look beyond the specific form of a demonstrator's action to extract the causal principles driving that choice of action. From age two, children often copy actions that are superfluous to the causal structure of a task (for a review, see Hoehl et al., 2019). However, such inefficient copying can be mitigated by giving children personal experience with the task prior to demonstration (Wood, Kendal, & Flynn, 2013).

In contrast to personal experiences, demonstrations are intrinsically ambiguous. For instance, demonstrators may fail to highlight all of the critical causal relations, and learners may not fully understand even the most well-formed demonstration (Sobel & Sommerville, 2010). Learner-directed exploration offers a comparably more powerful route to causal learning. For instance, it allows learners to confirm that a causal relation between two events is not influenced by some unobserved third variable (Kushnir, Wellman, & Gelman, 2009; Sobel & Sommerville, 2010), or to systematically isolate the task features most relevant to the causal structure (Sim, Mahal, & Xu, 2017). Finally, through attempting a task themselves, learners also gain experience with the actions needed to solve the task, and increased familiarity with those actions may make it easier to focus on the underlying causal structure. The advantage of such interventions over observing demonstrations can be seen throughout development. In infants, active personal experience with a task is more helpful than observational experience when learning new actions (Gerson, Mahajan, Sommerville, Matz, & Woodward, 2015; Gerson & Woodward, 2014). From 2.5 years, children learn causal rules equally efficiently from self-generated experience and direct instruction (Sim & Xu, 2017), but by 3–4 years, self-generated intervention trumps passive observation (Kushnir et al., 2009; Sim et al., 2017). Finally, school-aged children learn causal structures better if they intervene before watching an experimenter act, rather than the other way around (Kuhn & Ho, 1980; Sobel & Sommerville, 2010).

The advantage of generating personal experience during causal learning may be especially pronounced when the task's solutions cannot be learnt by copying the form of the actions involved, because the actions themselves must be related to some external goal. One such task is the trap-task, in which subjects must push a reward out of a horizontal tube, typically using a tool, while avoiding obstructions that would trap the reward (Horner & Whiten, 2007; Seed & Call, 2014; Visalberghi & Limongelli, 1994; Want & Harris, 2001). Because the tool can be inserted from either end of the apparatus, and the reward's location relative to the trap varies across trials, any strategy other than consistently avoiding the trap results in chance performance. As a result, the trap-task is ideal for comparing active and observational experience because it cannot be solved by copying a demonstrator's actions alone, but only by understanding those actions' spatial relation to the trap. Prior studies found that 2- to 4-year-olds failed to solve the trap-task independently, and performed only at chance even following correct, incorrect, and mixed-correctness demonstrations (Horner & Whiten, 2007; Want & Harris, 2001). However, neither study controlled for how the amount of task experience differed between the self-generated and demonstration groups. A more meaningful comparison would control the amount of task experience, while varying whether that experience was generated by one's own actions or given through demonstration.

Moreover, children may struggle to learn the trap-task from demonstration because its causal principles are complicated by the element of tool-use. Even the simple lateral motions called for in the trap-task require children to relate the tool, the reward, and the trap in terms of their spatial, physical, and causal properties (Fragaszy & Cummins-Sebree, 2005; Völter & Call, 2014). Simultaneously and continuously maintaining these relations may drain attentional and executive resources, thereby interfering with the selection of appropriate actions for a particular goal (Smitsman & Cox, 2008). Indeed, such a cognitive load would explain why children under 3 years often perseverate to one side when solving trap-tasks (Seed & Call, 2014; Want & Harris, 2001). Reducing the number of causal elements involved – namely, by replacing tool-use with manual actions – improves trap-task performance not only in 2.5-year-old children (Seed & Call, 2014), but also in chimpanzees (Seed, Call, Emery, & Clayton, 2009).

By extension, the difficulty of learning the trap-task through observation may be compounded by having to parse the goal of a demonstrator who is using a novel, unfamiliar tool. Observational learning of tool-use requires observers to chain multiple actions with proximal targets (e.g., grasping the tool, inserting the tool, pushing the tool against the reward) in service of an overarching distal goal (retrieving the reward by avoiding the trap). Such distal goals may therefore be more difficult to grasp when tool-use replaces manual actions. There is evidence for this difficulty in infants and pre-schoolers: Infants can identify an actor's goal from a direct manual reach towards a toy from the age of 6 months, but do not infer the goal of a tool-using action until later in development (Sommerville & Woodward, 2005; Woodward, 1998). Similarly, 2- to 5-year-olds are more likely to copy unecessary tool-using actions on an apparatus than unnecessary manual actions, perhaps because of the increased difficulty in identifying the relationship between the action and the goal (Taniguchi & Sanefuji, 2017).

The distal goals of tool-using actions could be made more apparent through comparison with a familiar action, by highlighting the analogical links between structurally similar situations (Gentner, 2010). In infancy, Gerson and Woodward (2012) found that 10-month-olds that used familiar manual actions to grasp for a toy alongside an experimenter's tool-use demonstration were more capable of goal understanding and imitation than infants that just observed the tool-using demonstration. A similar logic might apply to parsing more complex relational goals in older children: Seeing a familiar manual action might make it easier for children to understand the causal relationship between the action, the reward and the obstacle, and therefore the experimenter's goal (to extract the reward and avoid the trap). In this study, alongside exploring whether children learned better from self-generated experience compared to demonstration, we also aimed to explore if the type of action experienced would make a difference. We predicted that children would find it easier to extract a demonstrator's goal from manual demonstrations than from watching tool use, and that this might mitigate any difference between self-generated experience and demonstration.

The present study investigated the effect of prior experience (self- or other-generated) on children's ability to use a tool to solve a trap-task, while varying whether this experience was generated with a tool or by hand. We tested children aged 2.5–3.5 years because previous research has shown this to be the age range over which children's ability to solve the manual task emerges, although they perform more poorly when using a tool at this age (Seed & Call, 2014; Völter & Call, 2014). Thus we expected this to be the most sensitive age range over which any effect of prior experience on performance on the tool-using task could be detected. Using a

2×2 between-subjects design, children gained 10 trials of experience with the trap-task in one of four between-subjects learning conditions: Self–Hand, Self–Tool, Demonstration–Hand, or Demonstration–Tool. All four groups then received 10 test trials with the tool version of the trap-task. Based on the research outlined above, we predicted that children in the Self conditions would outperform those in the Demonstration conditions at test, having actively generated more personal experience with the task to support causal learning. Furthermore, we predicted that the task would be easier to learn after practising manually, rather than with a tool, resulting in higher scores in the Hand than Tool conditions when using a tool at test. Finally, we predicted that watching a hand demonstration would result in better test performance than watching a tool demonstration, because familiar actions would make the demonstrator's distal goals easier to parse.

Materials and Methods

Participants

Sixty-seven children aged 30 to 42 months (*M* age = 36.2 months, 35 males) were recruited using opportunity sampling and tested on-site at a science museum and café in a small city in the UK. A power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) was used to determine that a sample size of 67 would be sufficient to detect a large effect ($f = 0.45$), assuming $\alpha = 0.05$ and power = 0.95. Five additional children were recruited but excluded from the final sample because they did not complete testing. All parents completed written informed consent, and all children gave verbal assent; participation was rewarded with stickers. Ethical approval was granted by a university ethics committee (PS11903).

Design

The 2×2 (experience type × action type) design yielded four between-subjects conditions: Self–Hand (n = 17), Self–Tool (n = 17), Demonstration–Hand (n = 17), and Demonstration–Tool (n = 16). Each child first completed an *initial phase* (10 trials) according to their assigned condition, followed by a *test phase* (10 further trials) that was identical across conditions, in which each child attempted to solve the task themselves with the tool. The participant recruitment process focused first on the Self–Hand and Self–Tool conditions, in which the 10 trials of initial experience were generated by the children themselves using the appropriate action type. Those children's patterns of successes and failures was then used to generate scripts for the 10 trials of demonstration given by the experimenter (E) in the corresponding Demonstration–Hand and Demonstration–Tool conditions. Prospective participants who matched a child in the Self-conditions (matching criteria: same sex, age ±2 months) were assigned to be their yoked partner in the corresponding Demonstration-condition. This ensured that those in the demonstration-conditions received experience typical for their sex and age group, approximating the performance they themselves might have achieved.

Materials

The trap-task was a plastic box (43 cm long, 10 cm deep, 24 cm tall) mounted above a base, with a central shelf and two exits at the bottom. The box back, sides, and shelf were opaque white, and the top and front were transparent. Sliding doors in the back allowed E to place a transparent plastic ball (5 cm diameter, containing a sticker reward) on the shelf, and to retrieve it from trapped exits. The ball could travel along the shelf (18 cm long, 6 cm deep, 15 cm above the base) and fall off either end to the bottom exits below. See Figure 1.

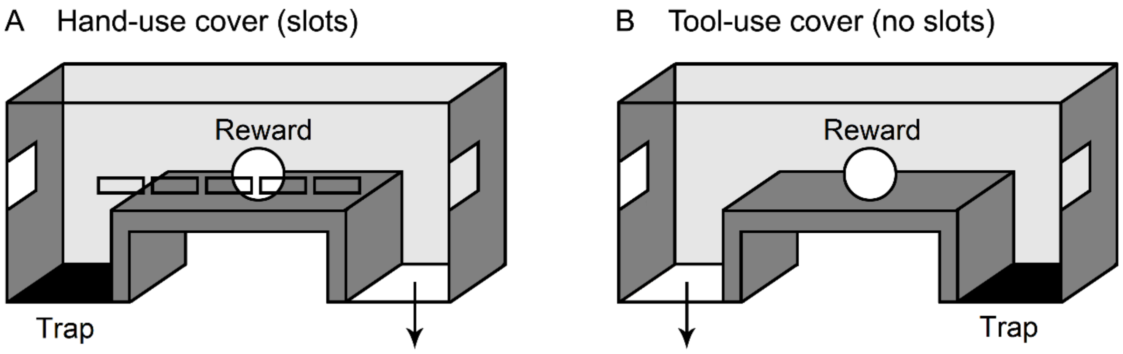

Figure 1. The trap-task as configured for manual (A) and tool (B) use, with sample trap locations. Arrows show exit locations of reward, given the location of the trap in that particular trial. Trap locations varied between trials.

During testing, the front of the box was covered by one of two transparent, removable panels. One panel had five slots cut into it (each slot 1.75 cm high, 5 cm long, spaced 1 cm apart), allowing the ball to be pushed along the shelf using a finger. The other lacked slots, requiring children to push the ball by inserting a tool through gaps (4 cm tall, 3.5 cm wide) on either side of the box. The tool (35 cm long, 2.5 cm diameter) was made of purple plastic with a handle at one end. See Figure 1.

From the child's perspective, the ball could only exit the box through the bottom exits. All other possibilities were either too narrow (the side gaps) or on the side of the box controlled by E (the sliding doors). In each trial, E obstructed one of the bottom exits with an orange plastic trap. The trap's location varied pseudo-randomly throughout the experiment, appearing equally often on both sides but never more than twice consecutively on a particular side. Trap locations were matched for yoked pairs.

Procedure

Pre-testing. Children sat in front of the trap-task, either by themselves or on a parent's lap. Parents were asked to refrain from giving cues. E sat directly behind the box. After affixing the appropriate front panel, E held the ball at each of the vertical exits and encouraged the child to take it, before placing a sticker in the ball and placing the ball on the central shelf. E then came to the child's side of the table to demonstrate how to move the ball, without fully moving it into an exit; in the Hand-conditions, she used a finger to push it slightly in both directions, and in the Tool-conditions, she briefly inserted the tool in each side of the box without touching the ball. Finally, E encouraged the child to look ("Watch!") as she slid the trap piece into place.

Phase 1: Initial experience. The child received 10 trials of active (self-generated) or observational (demonstration) experience with the trap-task, using either a hand or the tool to extract the ball. In the Self-conditions, E told the child, "Now it's your turn. Try to get the ball out of the box." E always gave the tool to the child orientated vertically, to avoid favouring either side of the trap-task. If the child requested help or did nothing, E responded neutrally, and reminded them how to move the ball by repeating the pre-testing demonstration. In the Demonstration-conditions, E said, "First, I'm going to show you a few times how to get the ball out of the box, then it'll be your turn." She then repeated the failure or success of the matched child on that trial, using the appropriate action type; the child was not permitted to retrieve the ball during Phase 1.

A trial ended when the child/E pushed the ball off the shelf into one of the vertical channels. Whenever the ball was successfully retrieved, E celebrated the success and gave the child the sticker. Whenever the ball became trapped, E noted that the child would not get the sticker but reaffirmed that whoever was trying (child or E) could try again. E then placed another sticker in the ball, placed it on the shelf, and drew the child’s attention to the trap. She said, “Try to get the ball out of the box!” or “Let’s see if I can get this ball out of the box…” as appropriate, and the next trial began. In the Demonstration-conditions, the 10 trials of initial experience therefore consisted of the same sequence of successful and unsuccessful trials as had been produced by the participant’s yoked partner.

Phase 2: Test. After the initial phase, E removed the front panel and replaced it with the slotless panel (in all conditions, to control for disruption). She held the ball at each vertical exit and encouraged the child to take it, before briefly inserting the tool in both sides of the box to show how the ball could be retrieved. The child then attempted to solve the trap-task using the tool for 10 further trials. The procedure for this portion was identical to that described above for the self-generated conditions. Testing continued until the child had completed 10 test trials, for a total of 20 trials; every child received a final sticker regardless of performance.

Scoring and Analysis

At the end of each trial, E documented which channel the ball had been pushed into and whether that choice was correct. A tripod-mounted videocamera captured children’s actions on the front of the box. Twenty percent of all videos were randomly selected for re-coding; observers agreed on the outcomes of over 98% of trials (Cohen’s kappa = .96).

Results

Children’s performance was analysed in terms of the proportion of correct trials over the 10 trials of the test phase; the full data is available online (DOI: 10.6084/m9.figshare.9964169). Because our proportion data was non-normally distributed, we modelled children’s performance during the test phase using a generalized linear model (GLZM) with multinomial probability distribution and cumulative logit link function. Ordinal data takes a multinomial distribution, and the use of a cumulative logit link function is what assumes ordered data. Age was included as a covariate, and experience type and action type were considered between-subjects predictors; the interaction term was experience type × action type. There was a significant main effect of experience type when controlling for age, Wald $\chi_1^2 = 6.22, p = .013$, such that children in the Demonstration-conditions performed worse than children in the Self-conditions, β = -1.18, 95% Wald CI [-2.47, 0.12]. Contrary to our predictions, there was no effect of action type, Wald $\chi_1^2 = 0.01, p = .915$, nor was there an interaction between action type and experience type, Wald $\chi_1^2 < 0.01, p = .976$. Finally, age in months was significantly associated with test performance, Wald $\chi_1^2 = 14.73, p < .001$, such that children performed better with increasing age, β = 0.26, 95% Wald CI [0.13, 0.40]. See Figure 2.

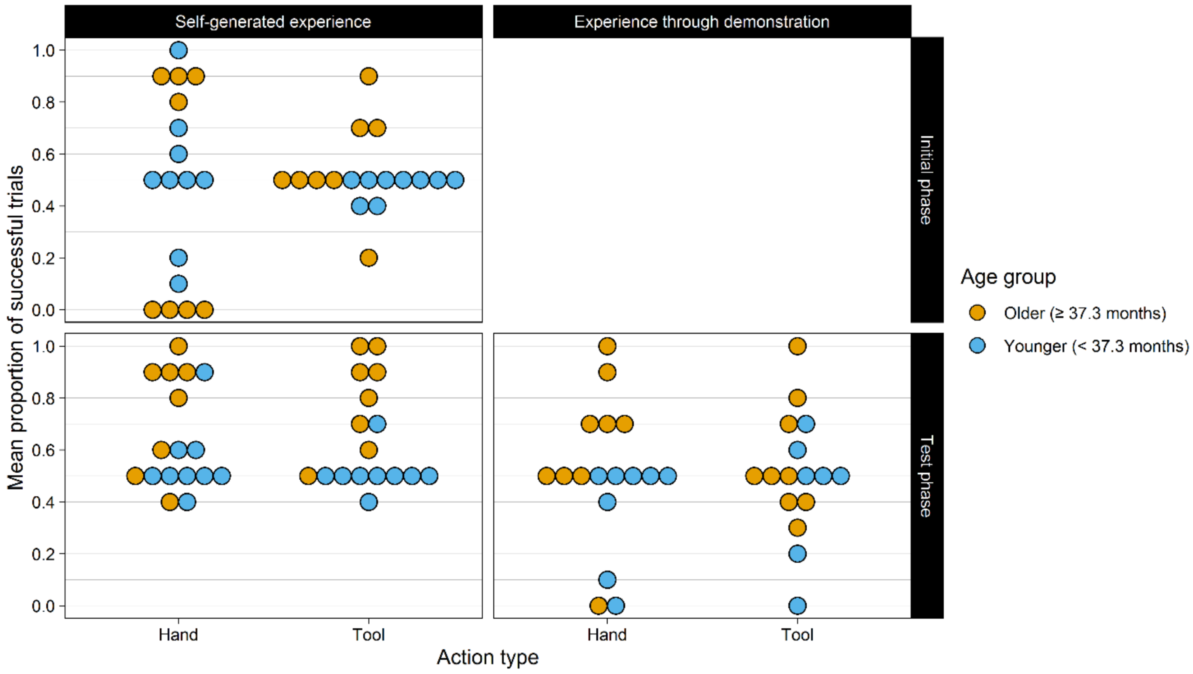

Figure 2. Dotplot showing children’s performance, split by action type (hand or tool), experience type (self-generated or demonstration), and phase (initial or test). Each dot represents an individual child. For ease of interpretation, age in months is

represented here using a median split (*Mdn* = 37.3 mos), but note that age was treated as a continuous variable in the analyses. Initial phase data from the demonstration-groups is omitted because these children did not generate their own initial experience.

Performance characteristics

Because prior research associated tool-use with chance and side-biased performance, we also examined the characteristics of children's performance. Using binomial tests, each child's overall performance in the initial (Self-conditions only) and test phases was assigned one of four characteristics: significantly above chance (≥.9 successful trials), significantly below chance (≤.1 successful trials), side-biased (pushing the ball to a given side for ≥.9 trials), or other (see Table 1).

Table 1.
Performance of children in initial and test phases, by condition.

					Initial phase				Test phase			
		Age (months)				Fail				Fail		
Condition	*n*	*M*	range	SD	AC	BC	SB	Other	AC	BC	SB	Other
Self–Hand	17	36.9	30.8–42.4	3.2	4	5	4	4	5	0	7	5
Self–Tool	17	36.2	30.1–42.7	4.2	1	0	11	5	4	0	8	5
Demo–Hand	17	36.9	30.0–42.0	3.5					2	3	9	3
Demo–Tool	16	36.1	30.0–40.8	4.0					1	1	8	6

Note. For each block of 10 trials, each child's performance was categorised as above chance (AC, ≥9 trials correct), below chance (BC, ≤1 trial correct), side-biased (SB, chance performance with the ball pushed to one particular side for ≥9 trials), or other (chance performance without side bias). Initial phase data from the demonstration-groups is once again omitted because these children did not produce their own initial experience.

Chi-square analyses reveal that in the initial phase, non-chance performance was significantly associated with condition in the initial phase, $\chi_1^2 = 24.52, p < .001, V = 1.28$, as was side-bias, $\chi_1^2 = 23.00, p < .001, V = 1.24$. Half of the children in the Self-Hand condition (9/17) consistently directed the reward into the either the trap or the open exit, whereas the majority of children in the Self–Tool condition (11/17) performed exactly at chance due to side-bias. However, these differences did not carry over into the test phase: There was no association between condition and non-chance performance at test, $\chi^2_3 = 1.71, p = .634, V = 0.16$, nor between condition and side-bias, $\chi^2_3 = 0.51, p = .916, V = 0.09$.

Discussion

A key finding from the present study is that children who generated their own experience went on to significantly outperform those who watched yoked demonstration. These findings extend a growing literature on the benefits of self-generated experience in other causal tasks (Kushnir et al., 2009; Rakison & Krogh, 2012; Sim et al., 2017; Sim & Xu, 2014, 2017; Sobel & Sommerville, 2010) and in pedagogy (Mavilidi, Okely, Chandler, Domazet, & Paas, 2018), raising the question of the cognitive mechanisms that might explain this advantage.

Active experience may be a particularly powerful tool in a causal learning context because of the special status of action as intervention. Interventionist accounts of causation argue that an event, *X*, can only be said to cause another event, *Y*, if intervening on *X* subsequently changes *Y* (Gopnik & Schulz, 2007). In this view, self-generated action functions as an independent variable within the causal system, because only through intervention can truly causal relations be distinguished from those that merely correlate or covary. Our results show that children learnt the trap-task's underlying causal relations (as indexed by success at test) most effectively through acting on the variables involved, rather than through observing others' interventions; crucially, this was true despite the content of the interventions being held constant.

These results align with work from the adult literature proposing that active intervention improves performance on causal tasks through high-level mechanisms such as hypothesis testing (Sobel & Kushnir, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). However, it remains possible that these improvements were driven by lower-level mechanisms, such as differences in temporal cues between active and passive interventions (Lagnado & Sloman, 2004), or by increased automaticity of movement gained through repeated practice. In our task, all children's initial experience was seen in real-time and presented with cause preceding effect. Furthermore, practising with the necessary action type did not improve test performance, as the children who had had to transfer between action types were just as successful as those who had practised with a tool throughout.

Indeed, the present study found no effect of action type at test, but the initial phase of learning was markedly influenced by hand- or tool-use. Initial performance in the Self-Tool condition was characterized by high rates of side-bias, whereas Self-Hand tended to be either significantly above or below chance. In other words, learning the task by hand seemed to better equip children to causally relate the trap, their hands, and the reward, regardless of which location they chose to target. By comparison, tool-use encouraged random or perseverative responses, as might be expected given its cognitive load (Fragaszy & Cummins-Sebree, 2005; Smitsman & Cox, 2008; Völter & Call, 2014). Yet the lack of difference between action types at test could suggest that the Self–Hand children found it difficult to transfer their solution across the change in action type from hand to tool.

Although we had anticipated that tool-demonstrations would not support learning, we did not find better performance following a hand-demonstration, perhaps because both demonstration groups were equally confused about the demonstrator's goals. Regardless of action type, the study's yoked design ensured that many children watched the demonstrator – an unfamiliar and authoritative adult – repeatedly and purposefully produce errors, yet express dismay whenever she trapped the reward. Children are sensitive to the intentions, knowledge status, and competence of adult models (Koenig & Jaswal, 2011; Schmidt, Rakoczy, & Tomasello, 2011), meaning that errorful demonstrations might conflict with children's assumption that adults will prioritise giving them relevant, useful knowledge (Csibra & Gergely, 2009). This could be explored further by demonstrating only correct solutions, or by providing more verbal support, such as explaining that the demonstrator intends to show both correct and incorrect solutions to the task. However it should be noted that previous work with the trap-task (Want & Harris, 2001) found that providing both correct and incorrect demonstrations was more effective in helping children to avoid side-bias than providing only correct demonstrations – though in this study, as in ours, only a small minority of the 2-3 year-olds that received demonstrations performed significantly above chance level (Want & Harris, 2001).

Another avenue for future work concerns our finding that performance in our sample improved with age. One contributing factor may be that, in order to successfully generate evidence about the trap-task, children in the Self-conditions had to select appropriate and informative learning strategies with minimal guidance. Doing so may have been particularly challenging for younger children because the metacognitive and executive functions necessary for self-regulated learning improve throughout childhood (Pauen & Bechtel-Kuehne, 2016; Roebers, 2017). Indeed, a comparable trend has been described earlier in development in motorically simpler tasks involving 'machines' activated by a particular class of block. Although the causal rules governing block choice can be learnt through free play by 30 months (Sim & Xu, 2017), 19-month-olds learnt best through play facilitated by an adult (Sim & Xu, 2015). For certain age groups or task types, self-generated and demonstrated evidence may be most helpful in combination. Future work with the trap-task and related tasks could investigate whether younger children's ability to self-generate evidence benefits from initial, explicit guidance.

Taken together, our findings highlight the advantage of generating experience for oneself, as doing so provides crucial confirmation of the causal, rather than correlational, nature of the relations involved. They also suggest potential disadvantages in watching others demonstrate the trap-task, possibly due to difficulty parsing demonstrators' goals, or to concerns about the demonstrator's perceived incompetence. However, the effectiveness of social or individual routes to learning also likely depends on the task at hand, the individual child's developmental stage and knowledge state, and the specific features of the social model.

Acknowledgements

We thank Geraldine Brown, Zeynep Civelek, and Solveig Jurkat for their help in data collection.

Funding sources

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 639072).

References

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fragaszy, D. M., & Cummins-Sebree, S. E. (2005). Relational spatial reasoning by a nonhuman: The example of capuchin monkeys. *Behavioral and Cognitive*

Neuroscience Reviews, *4*(4), 282–306. https://doi.org/10.1177/1534582306286573

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gerson, S. A., Mahajan, N., Sommerville, J. A., Matz, L., & Woodward, A. L. (2015). Shifting goals: Effects of active and observational experience on infants' understanding of higher order goals. *Frontiers in Psychology*, *6*, 1–13. https://doi.org/10.3389/fpsyg.2015.00310

Gerson, S. A., & Woodward, A. L. (2012). A claw is like my hand: Comparison supports goal analysis in infants. *Cognition*, *122*(2), 181–192. https://doi.org/10.1016/j.cognition.2011.10.014

Gerson, S. A., & Woodward, A. L. (2014). Learning from their own actions: The unique effect of producing actions on infants' action understanding. *Child Development*, *85*(1), 264–277. https://doi.org/10.1111/cdev.12115

Gopnik, A., & Schulz, L. (2007). *Causal learning: Psychology, philosophy, and computation*. (A. Gopnik & L. Schulz, Eds.). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195176803.001.0001

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). 'Over-imitation': A review and appraisal of a decade of research. *Developmental Review*, *51*(August 2018), 90–108. https://doi.org/10.1016/j.dr.2018.12.002

Horner, V., & Whiten, A. (2007). Learning from others' mistakes? Limits on understanding a trap-tube task by young chimpanzees (Pan troglodytes) and children (Homo sapiens). *Journal of Comparative Psychology*, *121*(1), 12–21. https://doi.org/10.1037/0735-7036.121.1.12

Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development*, *82*(5), 1634–1647. https://doi.org/10.1111/j.1467-8624.2011.01618.x

Kuhn, D., & Ho, V. (1980). Self-directed activity and cognitive development. *Journal of Applied Developmental Psychology*, *1*(2), 119–133. https://doi.org/10.1016/0193-3973(80)90003-9

Kushnir, T., Wellman, H. M., & Gelman, S. A. (2009). A self-agency bias in preschoolers' causal inferences. *Developmental Psychology*, *45*(2), 597–603. https://doi.org/10.1037/a0014727

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 856–876. https://doi.org/10.1037/0278-7393.30.4.856

Mavilidi, M.-F., Okely, A., Chandler, P., Domazet, S. L., & Paas, F. (2018). Immediate and delayed effects of integrating physical activity into preschool children's learning of numeracy skills. *Journal of Experimental Child Psychology*, *166*, 502–519. https://doi.org/10.1016/j.jecp.2017.09.009

Pauen, S., & Bechtel-Kuehne, S. (2016). How toddlers acquire and transfer tool knowledge: Developmental changes and the role of executive functions. *Child Development*, *87*(4), 1233–1249. https://doi.org/10.1111/cdev.12532

Rakison, D. H., & Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age? *Developmental Science*, *15*(1), 43–53. https://doi.org/10.1111/j.1467-7687.2011.01096.x

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science*, *14*(3), 530–539. https://doi.org/10.1111/j.1467-7687.2010.01000.x

Seed, A. M., & Call, J. (2014). Space or physics? Children use physical reasoning to solve the trap problem from 2.5 years of age. *Developmental Psychology*, *50*(7), 1951–1962. https://doi.org/10.1037/a0036695

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*(1), 23–34. https://doi.org/10.1037/a0012925

Sim, Z. L., Mahal, K. K., & Xu, F. (2017). Learning about causal systems through play. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 1078–1083).

Sim, Z. L., & Xu, F. (2014). Acquiring inductive constraints from self-generated evidence. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1431–1436).

Sim, Z. L., & Xu, F. (2015). Toddlers learn with facilitated play, not free play. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 2200–2205). Retrieved from http://docs.wixstatic.com/ugd/9f32e5_af61f0739723456fa7b0c1e7a3c18bff.pdf

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2- and 3-year-old children. *Developmental Psychology*, *53*(4), 642–651. https://doi.org/10.1037/dev0000278

Smitsman, A. W., & Cox, R. F. A. (2008). Perseveration in tool use: A window for understanding the dynamics of the action-selection process. *Infancy*, *13*(3), 249–269. https://doi.org/10.1080/1525000O8020379

Sobel, D. M., & Kushnir, T. (2006). The importance of decision-making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419. https://doi.org/10.3389/fpsyg.2010.00176

Sobel, D. M., & Sommerville, J. A. (2010). The importance of discovery in children's causal learning from interventions. *Frontiers in Psychology*, *1*(176), 1–7. https://doi.org/10.3389/fpsyg.2010.00176

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition*, *95*(1), 1–30. https://doi.org/10.1016/j.cognition.2003.12.004

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). *Inferring causal networks from observations and interventions*. *Cognitive Science* (Vol. 27). https://doi.org/10.1016/S0364-0213(03)00010-7

Taniguchi, Y., & Sanefuji, W. (2017). The boundaries of overimitation in preschool children: Effects of target and tool use on imitation of irrelevant actions. *Journal of Experimental Child Psychology*, *159*, 83–95. https://doi.org/10.1016/j.jecp.2017.01.014

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*, *108*(1), 15–22. https://doi.org/10.1037//0735-7036.108.1.15

Völter, C. J., & Call, J. (2014). The cognitive underpinnings of flexible tool use in great apes. *Journal of Experimental Psychology: Animal Behavior Processes*, *40*(3), 287–302. https://doi.org/10.1037/xan0000025

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development*, *72*(2), 431–443.

https://doi.org/10.1111/1467-8624.00288

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Copy me or copy you? The effect of prior experience on social learning. *Cognition*, *127*(2), 203–213. https://doi.org/10.1016/j.cognition.2013.01.002

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4