Older but Not Younger Preschoolers Use Efficiency to Evaluate How Well Goals Are Completed

Claudia G. Sehl, Nikki Tokat, Stephanie Denison, and Ori Friedman
Department of Psychology, University of Waterloo

Instrumental rationality dictates that goals should be completed efficiently rather than inefficiently. Here, we examined how young children come to use efficiency to evaluate how well goals are completed. Across a series of experiments, we showed children aged 3–7 (total $N = 618$; 48% female and 52% male; tested in a predominantly White region) scenarios where pairs of agents approached objects using paths that varied in length and efficiency, and then asked children which agent did a better job. In Experiments 1A and 1B, children aged 3 and 4 evaluated the agents using path length, but not efficiency. In Experiments 2A, 2B, and 3, children used path length to assess which agent did a better job from around age 4. But children only used efficiency from age 5 or later. These findings suggest that children's explicit judgments are not guided by the norm for efficiency, and Experiment 4 provides evidence that this late development is not an artifact of methodological problems. This developmental pattern might arise from limits in children's ability to make counterfactual judgments about what agents could have done.

Keywords: efficiency, normativity, counterfactuals

Suppose two people both need to drive across town. Even though they both manage to arrive, one does a better job. Why? Because one driver is efficient while the other is inefficient. The efficient driver takes a direct route and expends minimal time and effort getting to the destination. The inefficient driver, by contrast, takes a meandering and indirect route and ends up expending twice as much time and effort as needed.

In this article, we examine how children come to consider the efficiency of actions when evaluating how well goals are completed. For instance, suppose children see the two drivers. Will children conclude that the efficient driver did a better job? Investigating this question can advance knowledge about how young children recognize efficiency, the normative standards they use to evaluate how well goals are completed, and their thinking about counterfactuals. It may also speak to children's perception of rationality, since rationality is characterized by securing goals using the most efficient means available (Gergely & Csibra, 2003; Gergely & Jacob, 2012).

Considerations of efficiency underlie social judgments from infancy onward. The importance of efficiency for early social judgments was first discovered in work on how infants attribute intentionality and the capacity for goal-directed actions (Gergely et al., 1995). In one experiment, 12-month-olds saw an agent jump over a wall to reach a target. Infants' looking times to subsequent displays suggested they expected this agent would behave efficiently in other situations—with the wall removed, infants expected the agent to take the most direct route to the target, rather than jumping as before. By contrast, infants did not expect this if the agent's initial actions were inefficient—that is, if the initial display

Claudia G. Sehl https://orcid.org/0000-0003-4283-1139

The materials, data, and analytic code for all experiments are available online on the Open Science Framework and can be accessed at https://osf.io/ybsz9/. These experiments were funded by grants to Stephanie Denison and Ori Friedman from the Natural Sciences and Engineering Research Council of Canada.

Claudia G. Sehl played a lead role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing–original draft, and writing–review and editing. Nikki Tokat played a supporting role in investigation and project administration. Stephanie Denison played an equal role in conceptualization, funding acquisition, methodology, supervision, and writing–review and editing. Ori Friedman played an equal role in conceptualization, funding acquisition, methodology, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Claudia G. Sehl, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. Email: claudia.sehl@uwaterloo.ca

did not include a wall and the agent jumped for no reason. Since this discovery, much other work has provided evidence that infants are sensitive to efficiency when interpreting actions (e.g., Csibra et al., 1999; Liu & Spelke, 2017; Phillips & Wellman, 2005; Scott & Baillargeon, 2013; Skerry et al., 2013).

Young children also appear to consider efficiency in their social judgments and in their judgments about objects, and these judgments can take somewhat different forms. From age 4, children predict that an agent will take the shorter path to a goal over a longer one (Gönül & Paulus, 2021; also see Schuwerk & Paulus, 2016). Similarly, children aged 5–6 predict that an agent would rather obtain an object from a hill that is easy to climb rather than from one that is difficult to climb (Sehl et al., 2021). These judgments could reflect expectations that agents will efficiently minimize effort in obtaining goals. Young children likewise use such cost–benefit analyses to infer preferences for objects (Jara-Ettinger et al., 2015) and in other social judgments (e.g., Chu & Schulz, 2023; Jara-Ettinger et al., 2017). Consideration of efficiency may also inform children's thinking about the design of artifacts. In choosing which of two tools was designed to perform a function, 3- to 4-year-olds mostly choose an artifact without any extra parts over one with unnecessary parts, even when both artifacts are suited for the function (Kelemen et al., 2012; also see Pesowski et al., 2020).

Efficiency could also drive children's normative judgments about completing goals. The idea that goals should be completed efficiently is sometimes seen as a component of *instrumental rationality* (i.e., rationality in the means chosen to pursue goals; e.g., Hitchcock, 2011). While much research has examined children's understanding of norms, including moral and conventional norms, comparatively less work has looked at norms of instrumental rationality (for a review, see Schmidt & Rakoczy, 2023). What is more, work on these norms has not examined efficiency and instead looked at children's recognition that agents should use tools that can produce desired outcomes over tools that cannot—for instance, that driving a nail into wood should be accomplished with a hammer rather than a sponge (e.g., Dahl & Schmidt, 2018; Schmidt et al., 2012). Broadly consistent with the possibility of early emerging efficiency norms, infants expect others to preferentially approach efficient agents over inefficient ones (Colomer et al., 2020; also see Colomer & Woodward, 2023). Also, children may use efficiency to infer competence, which connects with doing things the right way. In one study, 4- to 5-year-olds judged that a person who quickly built a block tower was better at this task than a person who took longer to build an identical-looking tower (Leonard et al., 2019).

At the same time, a norm for completing goals efficiently might develop relatively late. Many findings suggesting that children recognize efficiency could also reflect their attention to other factors. To illustrate, consider again children's predictions that agents will take shorter routes over longer ones (Gönül & Paulus, 2021), their predictions that agents will fetch objects from shorter hills than higher ones (Sehl et al., 2021), and their judgments that a person who quickly builds a block tower is more competent than someone who takes more time (Leonard et al., 2019). While these findings could reflect beliefs about efficiency, they might also arise from expectation that agents prefer completing goals quickly rather than slowly, or via shorter routes rather than long ones. Strong evidence that children consider efficiency would be provided by finding that they can judge whether goals were successfully completed using the minimum time and effort possible.[1] These assessments might involve counterfactual reasoning (i.e., reasoning about how things could have gone differently): To judge whether goals could have been completed using less time and effort, children might consider how *else* the goals could have been completed. Examining the development of a norm for completing goals efficiently, then, may also be informative about children's sensitivity to efficiency in general and the development of counterfactual reasoning (for a review, see Nyhout et al., 2019a).

The Current Experiments

In our experiments, children saw scenarios where pairs of agents went to target objects and were then asked which agent did a better job (for other developmental papers using similar "better job" test questions, see Foushee et al., 2017; Gweon & Asaba, 2018; Schleihauf et al., 2023). The experiments included two types of trials. One type contrasted an agent who took a direct path with one who took a long and *unnecessarily* circuitous path. In these trials, children might choose the agent taking the shorter path because of preferences for efficiency or because of preferences for shorter paths. In the other trial type, efficiency- and length-based reasoning led to differing predictions; the design of these trials, though, varied substantially across experiments. Experiments 1–3 included both trial types, while Experiment 4 only included the first kind (i.e., contrasting direct paths with unnecessarily circuitous ones). In our first experiments, we tested children aged 3–4 because we had initially expected that children at these ages would already prefer efficient agents over inefficient ones in our vignettes. This turned out to be wrong, and so the later experiments included older children.

General Methods

In Experiments 1 and 2, we aimed to test 20 children per age-in-years in each between-subjects condition, though exclusions and oversights led to slight departures from this goal. As explained below, we doubled this stopping rule in Experiment 3. Experiment 4 used a total sample of 30 children; it examines some methodological concerns with the other experiments but without examining age effects.

Most children were tested in person at childcare centers and schools in the region of Waterloo, Ontario (≥83% in each experiment); the remainder were tested online via Zoom. Demographic information was not collected from each child as per the allowances of our Institutional Review Board application. However, a recent census indicated that 64% of residents in the region are White, and South Asians are the largest visible minority; these figures were corroborated by a census of students in the local school board where most children tested.

Children tested in person saw images on a laptop computer. For children tested over Zoom, the experimenter shared their screen. Most experiments began with two warm-up trials. In these trials, children saw a yellow box and an orange one side by side. A fish

[1] Evidence for sensitivity to this was provided in early research on infants. For instance, when agents approach a goal, infants distinguish jumping for no reason from jumping to get over a wall (e.g., Gergely et al., 1995). While such findings made us think preschoolers might be sensitive to efficiency, infants' competence as assessed by looking-time studies often does not predict success in children when the same abilities are tested using explicit verbal measures.

appeared in one box, and children were asked what color box it was in. These trials were included so that children tested over Zoom would understand they could choose between agents in the main experiment by saying the color of the box with the agent they wanted to indicate. We did not include this warm-up in Experiments 2B and 4 because they were limited to children tested in person. Each experiment also included four test trials. These were administered in one of four different presentation orders, except in Experiment 4, which used two orders.

Responses were analyzed using generalized estimating equation (GEE) models for binary data with exchangeable correlation matrixes. We ran the models in R using *geepack* (Højsgaard et al., 2006) and used the joint_tests function in *emmeans* (Lenth et al., 2019) to yield Type 3 outputs; *emmeans* was also used to run post hoc tests. Finally, we used *ggeffects* (Lüdecke, 2018) to examine predicted confidence intervals.

Transparency and Openness

Materials, data, and analytic code are available at https://osf.io/ybsz9/overview. These studies were not preregistered. These studies were approved by the Office of Research Ethics at the University of Waterloo.

Experiments 1A and 1B

Children aged 3 and 4 saw two types of trials. In one type, one agent took a short direct path to the object while the other took an unnecessarily circuitous path. Here, preferences for both efficiency and shorter paths should lead children to say the short path agent did a better job. The other type of trial also contrasted short direct path with a longer circuitous one. However, here, both agents were efficient: With the long circuitous path, obstacles prevented taking a more direct route. In these trials, a preference for efficiency should *not* lead children to choose the short path agent. As a secondary aim, we also examined if children consider efficiency in another judgment. We showed some children the same scenarios but asked them about which agent had more fun. Whereas children may link efficiency with doing a better job, they might use inefficiency as a cue for fun (Chu & Schulz, 2023; also see Schachner & Carey, 2013, for work in adults).

Both experiments used the same design as one another, but as detailed below, they used different slides and scripts and conveyed the agents' movements in different ways. When we conducted Experiment 1A, we were surprised to find that 3- to 4-year-olds did not consider efficiency in their judgments. We conducted Experiment 1B using slightly different methods intended to make things easier. If anything, though, the changes made the task harder for children to follow.

Method

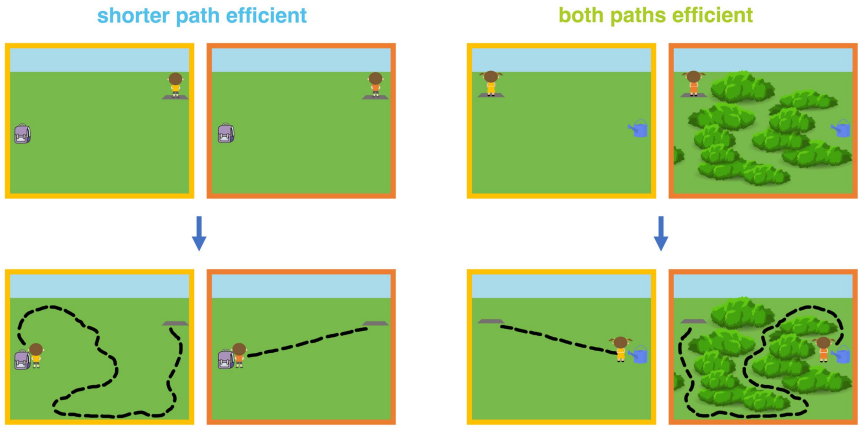
Participants

In Experiment 1A, we tested eighty-one 3- to 4-year-olds (M_{age} = 3;11, range = 3;0–4;10, 37 girls and 44 boys, 89% tested in person). In Experiment 1B, we tested eighty 3- to 4-year-olds (M_{age} = 3;10, range = 3;0–4;11, 33 girls and 46 boys, 94% tested in person). We also tested four other 3-year-olds (two in each experiment), but they did not respond to any of the test questions.

Procedure

Children completed four test trials; see Figure 1 for sample slides from Experiment 1A. In each trial, two agents were in separate fields and had to walk to an object in their field. Across the four trials, these objects were water bottles, backpacks, umbrellas, and watering cans. Children then saw the route each agent took. One agent walked straight to the desired object, while the other agent walked in a longer winding path.

Figure 1
Sample Trial Slides From Experiment 1A



Note. Children saw scenarios where one agent took a short direct path to an object and the other took a long circuitous path. In some trials, the agent taking the shorter path behaved more efficiently than the agent taking the longer one (left panel). In other trials, both agents behaved efficiently (right panel). The top row shows the first slide in each story, and the bottom row shows the second slide. See the online article for the color version of this figure.

The experiment used a 2 × 2 design with the factors trial type and judgment. Trial type was manipulated within-subjects: Each child saw two types of trials and completed two of each type. In one trial type, the agent who took the direct path was efficient whereas the agent taking the longer path was inefficient. In the other trial type, both agents were efficient—although one agent took a longer path than the other, it was the shortest path available given the constraints. Judgment was manipulated between-subjects. At the end of each trial, children were asked either "Which kid did a better job?" or "Which kid had more fun?"

In Experiment 1A, the paths taken by each agent were conveyed via dotted lines. Children first saw the agents at their starting places and were told they had to go to the target objects (e.g., "Here are two kids. The kids have to get to the watering cans. Let's see how they get to the watering cans."). Next, they saw both agents by their goal objects with dotted lines depicting their paths.

In Experiment 1B, the scenarios were animated and displayed serially. Children were first introduced to the agents and told they had to get their possessions (e.g., "Here are two kids. The kids have to get to their watering cans."). The experimenter indicated the agent on the left and explained that children would see this agent get their item ("First, we'll see how the kid in the yellow box gets their watering can."). An animation then showed this agent going to the object. Then, the experimenter introduced the agent on the right ("Next, we'll see how the kid in the orange box gets to their watering can."), and an animation showed this agent going to their object. The animations did not include a dotted line showing the path.

In all experiments except 1B, paths were matched across trial types. For instance, in Experiment 1A, trials where the shorter path was efficient used horizontally or vertically flipped versions of the paths in trials where both were efficient.

Results

We entered children's choices of the short direct path (scored 1) and the long circuitous path (scored 0) into a GEE model with trial type, judgment, and age in months as predictors. See Table 1 for the outputs of each model and Figure 2 for plots of the results.

In Experiment 1A, the only significant effect was the main effect of judgment, $p < .001$. Children selected the short path agent more often when judging who did better than when judging who had more fun. Single sample tests conducted using intercept-only models found that children's responses differed from chance with both judgments. Children asked about who did a better job mostly chose the agent who took the shorter path, $p = .001$. By contrast, children asked about who had more fun mostly chose the agent who took the longer path, $p = .044$.

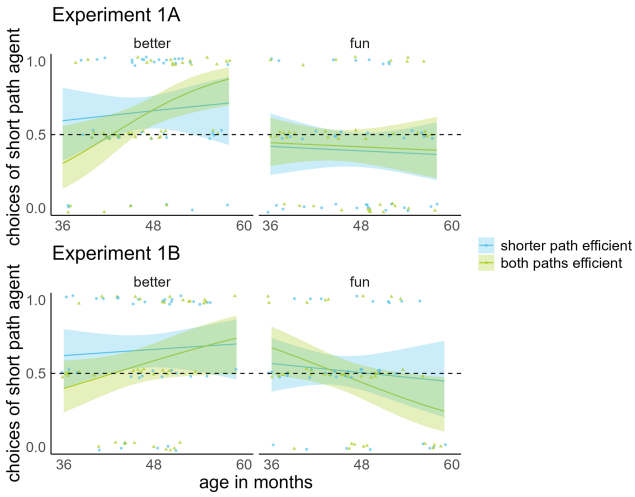
In Experiment 1B, the only significant effect was the two-way interaction between age and judgment, $p = .020$. To unpack this interaction, we examined whether responses to each question type were differently affected by age. However, the effect of age was nonsignificant for both, $ps \geq .089$.

Table 1
Output of the Generalized Estimating Equation Models in Experiments 1A and 1B

	Experiment 1A		Experiment 1B	
Term	$F(1, 312)$	p	$F(1, 305)$	p
Trial	0.02	.883	1.81	.179
Judgment	14.86	<.001	3.67	.056
Age	1.93	.166	0.09	.763
Trial × Judgment	0.15	.701	0.28	.596
Trial × Age	2.43	.120	0.04	.852
Judgment × Age	3.29	.071	5.49	.020
Trial × Judgment × Age	2.35	.127	2.66	.104

Note. Both models included trial, judgment, and age in months as predictors.

Figure 2
Choices of the Short Path Agent in Experiments 1A and 1B



Note. Children judged which agent did a better job (left) or had more fun (right) in scenarios where one agent took a short direct path to an object and the other took a long circuitous path. In some trials, the agent taking the shorter path behaved more efficiently than the agent taking the longer one. In other trials, both agents behaved efficiently. In this and other plots, regression lines and 95% confidence intervals were predicted from generalized estimating equation models. Jittered points indicate children's mean scores averaged across trials. See the online article for the color version of this figure.

Discussion

Children aged 3–4 were insensitive to efficiency (i.e., there was no effect of trial type), both when judging which agent did a better job and which had more fun. In Experiment 1A, children's responses differed across these judgments—they mostly said the short path agent did a better job, and mostly said the long path agent had more fun. When considering these responses in only the trial where one agent takes a long path for no apparent reason and the other takes a direct path, this could appear to be a judgment based on efficiency. But children also responded this way when the long path was taken by necessity, suggesting these responses did not reflect reasoning about efficiency. Responses were noisier in Experiment 1B, which provided only weak evidence of children responding differently to the two questions. Despite these seeming differences between Experiments 1A and 1B, exploratory analyses comparing their results did not yield significant differences (see the Supplemental Materials).

Experiments 2A and 2B

We continued to investigate children's normative evaluations (i.e., which agent did a better job) but did not further pursue their attributions of fun. Given children's weak performance in the first experiments, we extended the age range to include children aged 3–7. We also used a different design. In our first experiments (especially 1A), children said the short path agent did a better job regardless of whether the long path agent behaved efficiently or inefficiently. While this suggests insensitivity to efficiency, it is possible that choices of the short path agent were based on an additional preference for short paths. A more sensitive test, then, might avoid pitting efficiency against path length.

In the next experiments, we again included trials where one agent took a short direct path and the other needlessly took a long path. In these trials, both efficiency and quickness could bring children to prefer the short path agent. However, we also included a new kind of trial where *both* agents took long circuitous paths: One agent was efficient because obstacles prevented taking a shorter path. The other agent was inefficient because their path was unconstrained, and so a shorter path was possible. In these trials, a preference for efficiency supports choosing the agent forced to take a long path.

Both experiments use the same design and testing slides, but they used different testing scripts. In Experiment 2A, even the oldest children did not appear to consider efficiency, and so we slightly modified the script in Experiment 2B to help make some aspects of the task clearer.

Method

Participants

In Experiment 2A, we tested one hundred 3- to 7-year-olds (M_{age} = 5;6, range = 3;0–7;11, 49 girls and 51 boys, 86% tested in person). In Experiment 2B, we tested one hundred twenty-seven 3- to 7-year-olds (M_{age} = 5;5, range = 3;0–7;11, 59 girls and 68 boys, 100% in person).[2] Two other children were tested but not included. One did not answer any questions, and the other did not speak English well enough to complete the procedure.

Procedure

Children completed four test trials; see Figure 3 for sample slides. In each trial, children saw scenarios where two agents in separate fields went to target objects (water bottles, backpacks, umbrellas, and watering cans). There were two types of trials, and children saw two of each type. In one trial type, one agent's path was direct while the other's path was long and unconstrained. In the other type of trial, both agents took long paths, with one constrained by obstacles and the other unconstrained. Each trial was conveyed using two slides. The first showed the agents at their starting places, while the second showed them by the target objects with dotted lines depicting their paths.

The two experiments used slightly different testing scripts. Here are sample wordings from the backpack trial:

2A. Here are two kids. The kids have to get to the backpacks. Let's see how they get to the backpacks. One of the kids did a better job. Which kid did better?

2B. Here are two kids. The kids have to get their backpacks. Here's how they got to their backpacks. So which kid did a better job?

In Experiment 2A, children were told the agents had to go to the target objects before the agents went to the objects, but not afterward. In Experiment 2B, children were told this both before and after the agents went to the objects. We thought that reminding children of the agents' goals immediately before the test question might improve performance. Also, the script for Experiment 2A mentioned that one agent did a better job before the test question was asked, while the script for Experiment 2B did not do this.[3]

Results

We entered children's choices of the efficient (scored 1) and inefficient (scored 0) agents into a GEE model with trial type and age in months as predictors; see Figure 4 for plots of the results.

In Experiment 2A, there was a main effect of trial type, $F(1, 393) = 8.02$, $p = .005$; a main effect of age, $F(1, 393) = 6.40$, $p = .012$; and no interaction, $F(1, 393) = 0.22$, $p = .642$.

Children chose the efficient agent more often in trials contrasting short and long paths than in trials contrasting constrained and unconstrained long paths. Also, choices of efficient agents increased with age. In trials where paths differed in length, children mostly chose the efficient agent from age 4;2, 95% CI [.51, .71]. In trials with constrained and unconstrained long paths, children never chose the efficient agent at rates exceeding chance.

In Experiment 2B, there was also a main effect of trial type, $F(1, 500) = 6.30$, $p = .012$; a main effect of age, $F(1, 500) = 38.20$, $p < .001$; and no interaction, $F(1, 500) = 0.10$, $p = .765$. As before, efficient agents were chosen more often in trials manipulating path length than in trials manipulating whether long paths were constrained. In trials where paths differed in length, children mostly chose the efficient agent from age 4;4, 95% CI [.51, .67]. In trials with constrained and unconstrained long paths, children mostly chose the efficient agent from age 5;11, 95% CI [.51, .66].

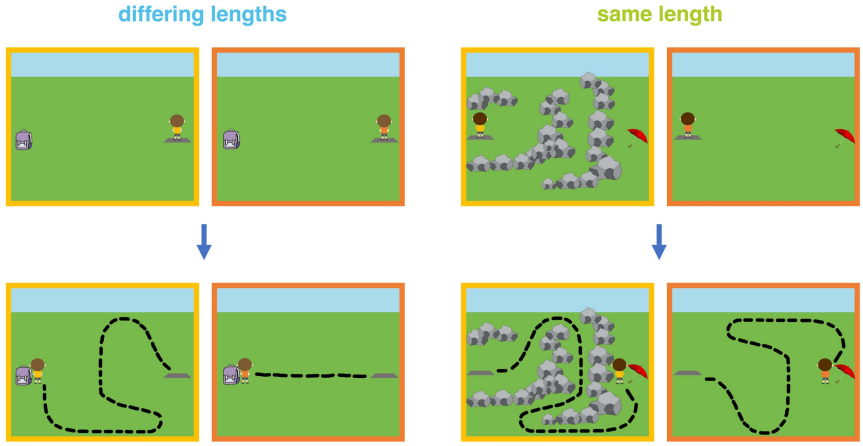
Discussion

Children more often chose the efficient agent when the agents' paths differed in length than in trials where both had the same path, but one agent could have used a shorter one. These findings suggest that a preference for efficiency may come relatively late, since responses in the differing length trials might reflect a more basic preference for shorter paths. In Experiment 2A, children never clearly preferred efficiency in the same length trials, whereas in Experiment 2B, this preference emerged at age 5. Despite this, exploratory analyses comparing the results of these experiments did not reveal significant differences between them (see the Supplemental Materials).

[2] This experiment has a larger sample size than the other experiments. Here is why: A new research assistant collected the data. After she had tested 34 children, some concerns were raised that she was speaking to children too quickly. She therefore restarted testing, speaking at a slower pace. However, we later looked at the results with and without the original children and found no difference. Hence, we retain the full sample here.

[3] We ran Experiment 2B only after completing Experiment 3. This is why the script for Experiment 3 more closely resembles the script from Experiment 2A.

Figure 3
Sample Trial Slides From Experiments 2A and 2B

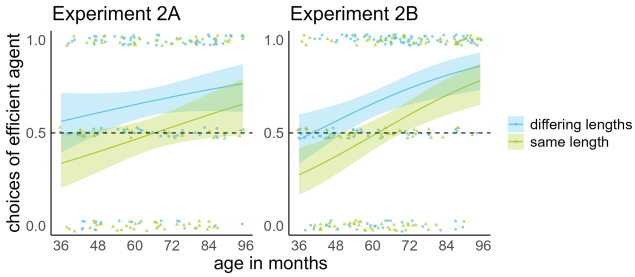


Note. Children saw scenarios where one agent was more efficient than the other. In some trials, the efficient agent took a shorter path than the inefficient agent (left panel). In other trials, both agents took long paths, but whereas the efficient agent's path was constrained by obstacles, the inefficient agent was unconstrained (right panel). The top row shows the first slide in each story, and the bottom row shows the second slide. See the online article for the color version of this figure.

Experiment 3

The experiments so far used scenarios varying in whether agents were impeded by obstacles. In this experiment, we took a different approach. Children saw scenarios where agents collected flowers. In some trials, the agents took the same path but differences in the positions of the flowers across the fields meant that one agent could have taken a shorter path. The agent unnecessarily bypassed closer flowers to obtain more distant ones. The other trials again contrasted short paths with long ones. The flowers in both fields were in the same positions, but one agent collected them using a short direct path while the other agent took an unnecessarily circuitous route.

Figure 4
Choices of the Efficient Agent in Experiments 2A and 2B



Note. Children judged which agent did a better job in scenarios where one agent took an efficient path to an object and another agent took an inefficient path. In some trials, the inefficient agent took a longer path than the efficient one but could have taken the same path. In other trials, the agents' paths were the same length, but the inefficient agent could have taken a shorter path. See the online article for the color version of this figure.

Method

Participants

We tested two hundred and one 3- to 7-year-olds (*M*age = 5;5, range = 3;0–7;11, 106 girls and 95 boys, 83% in person).[4]

Procedure

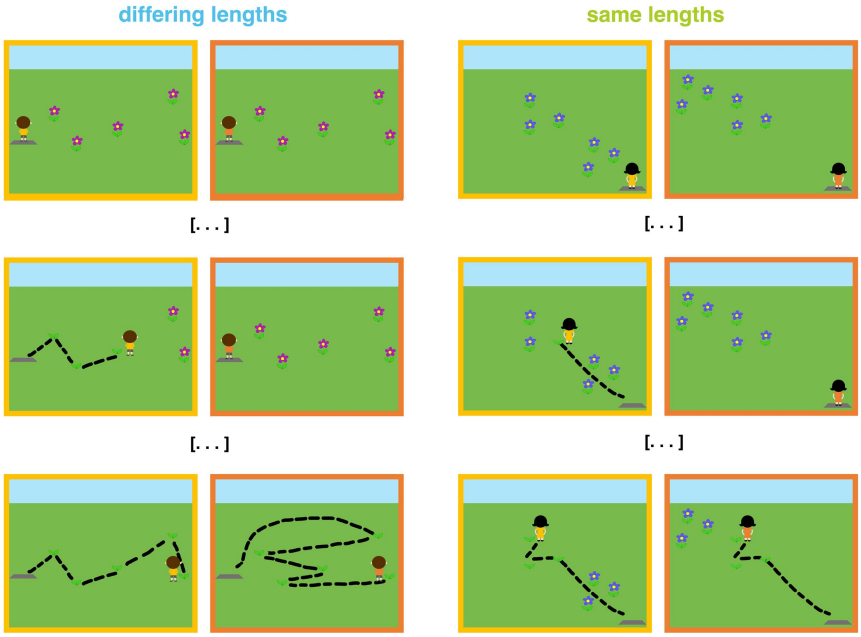
Children completed four trials where two agents collected flowers in separate fields; see Figure 5 for sample materials. There were two types of trials, and children completed two trials of each type. In one trial type, both agents took the same path in collecting three flowers, among others. The positions of other flowers varied across the fields, such that one agent collected the closest available flowers while the other needlessly bypassed flowers. In the other trials, the flowers in both fields were in the same positions, and one agent took a short direct path whereas the other agent's path was long and indirect.

In this experiment, children saw the agents move step by step as they collected each flower. With each move, dotted lines indicating their paths grew longer. Also, in each trial, children first saw the movements of the agent on the left and only afterward saw the movements of the agent on the right. Here is the script (text in square brackets varied across the two trial types):

> Here are two kids. The kids can get [all the/three] flowers. First, we'll see how the kid in the yellow box gets their flowers. Next, we'll see how the kid in the orange box gets their flowers. One of the kids did a better job. Which kid did better?

[4] In this experiment, we doubled our stopping rule of $N = 20$ per age per between-subjects condition. Initially, we only planned on testing children aged 3–4. After testing the initial planned sample of 80 children, we observed weak evidence for an effect of trial type (i.e., $p = .054$). To get a clearer picture, we decided to increase the sample size and extend the age range.

Figure 5
Sample Trial Slides From Experiment 3



Note. Children saw scenarios where one agent was more efficient than the other. In some trials, the efficient agent took a shorter path than the inefficient agent (left panel). In other trials, both agents took long paths, but whereas the efficient agent's path was constrained by obstacles, the inefficient agent was unconstrained (right panel). In this experiment, the scenarios used several slides. Here, the top row shows the first slide in each story, the middle row shows an intermediary slide, and the bottom row shows the final slide. See the online article for the color version of this figure.

Results

We entered choices of the more efficient (scored 1) and less efficient (scored 0) agent into a GEE with trial type and age as predictors; see Figure 6. There was a main effect of trial, $F(1, 797) = 14.00$, $p < .001$; a main effect of age, $F(1, 797) = 63.90$, $p < .001$; and no interaction, $F(1, 797) = 0.60$, $p = .452$. Children more often chose the efficient agent in trials where the agents' paths differed in length than in trials where they had the same path. Also, choices of the efficient agent increased with age. In trials with paths of differing lengths, children mostly chose the efficient agent (i.e., the one with the shorter path) from age 3;11, 95% CI [.51, .66]. In trials where both agents had the same path, children mostly chose the efficient agent (i.e., the one who did not needlessly pass flowers) from age 5;2, 95% CI [.51, .63].

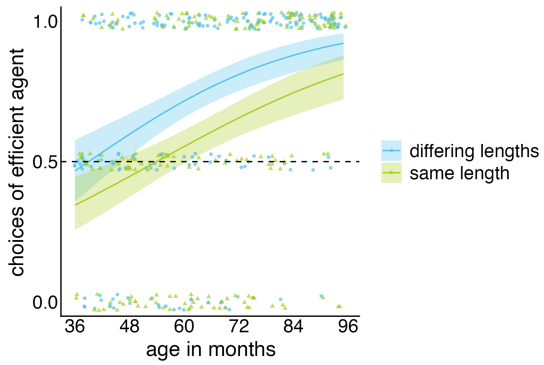
Discussion

Children again more often chose the efficient agent in trials contrasting short-versus-long paths than in trials where both agents had the same paths (but one agent could have used a shorter one). In this experiment, though, preferences for short paths and efficiency both emerged comparatively early; we discuss potential explanations below.

Experiment 4

We conducted a final experiment to address three potential problems with the stimuli from our experiments so far: (a) *Obstacles*. In experiments with obstacles, children might have thought the agents could have climbed over the obstacles or taken a shortcut by going around some of the obstacles. If so, children could have felt that *both* agents in each trial were inefficient (i.e., since both could have taken

Figure 6
Choices of the Efficient Agent in Experiment 3



Note. Children judged which agent did a better job: an efficient agent who completed a goal taking the shortest path possible or an inefficient agent whose path was unnecessarily long. In some trials, the inefficient agent took a longer path than the efficient one but could have taken the same path. In other trials, the agents' paths were the same length but the inefficient agent could have taken a shorter path. See the online article for the color version of this figure.

shorter routes). (b) *Perceptual access*. Because the agents were shown from behind in our slides, children might have assumed the agents could not see where they were going. That is, children might have felt there was a reasonable explanation for agents taking unnecessarily long paths, explaining away their seeming inefficiency. (c) *Motives*. In Experiment 3, where the inefficient agent bypassed nearby flowers to collect more distant ones, children might again have inferred there was a good reason for taking a long path. For instance, children could have concluded that the distant flowers looked nicer to that agent or that the agent wanted to continue walking past the flowers after collecting them.

To address these concerns, we updated slides and scripts from the earlier experiments, focusing only on trials where one agent took an unnecessarily long path; we did not revisit trials manipulating path length. In this final experiment, we tested children aged 3–5 to see if they would prefer efficient over inefficient agents with the new stimuli. Because the focus of this experiment was on addressing methodological concerns, we did not look for developmental effects.

Method

Participants

We tested thirty 3- to 5-year-olds (M_{age} = 4;7, range = 3;0–5;10, 13 girls and 17 boys, 100% in person). Two other children were tested but not included. One did not answer any questions, and the other only answered two of the four test questions.

Procedure

Children completed four test trials; see Figure 7 for sample slides. There were two types of trials, and children saw two of each type. In one trial type, agents went to target objects (watering cans, umbrellas). Both agents took long paths, with one constrained by obstacles and the other unconstrained. In contrast with the corresponding scenarios from the earlier experiments, there were no gaps between the obstacles, and the agents' faces were visible. Also, the script explained that the agents could not climb over the obstacles and that they knew where the goal objects were. Here is the script from the scenario with watering cans:

> Here are two kids. The kids have to get their watering cans. They know where the watering cans are, and they can't climb over the bushes. Here's how they got to their watering cans. So which kid did a better job?

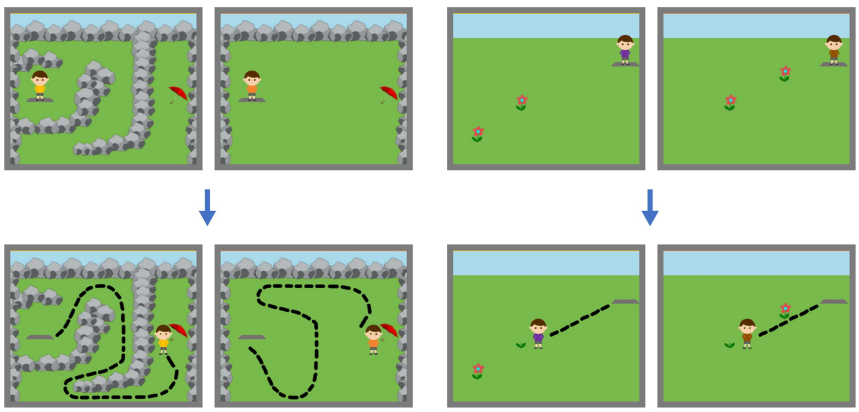
In the other type of trial, both agents took the same path in collecting a flower from a field. One agent collected the closest available flower while the other agent needlessly bypassed a closer flower. These trials were simplified versions of the scenarios from Experiment 3. The agents' faces were visible, and the script specified that the flowers were identical to one another, that agents knew where the flowers were, and that the agents had to return to their starting places:

> Here are two kids. The kids have to get a flower and bring it back to where they started. They each see both flowers and the flowers are all the same. Here are the flowers they got. So which kid did a better job?

Results

We entered choices of the more efficient (scored 1) and less efficient (scored 0) agent into a GEE with trial type as a predictor; see Figure 8. There was no main effect, $F(1, 118) = 0.13$, $p = .724$. Single sample tests conducted using intercept-only models found that children's responses did not differ from chance for either trial type: obstacles, $p = .590$; bypassing, $p > .999$. To examine whether performance changed with age, we also ran separate GEE models on children's responses for each trial type. For both trial types, children's choices did not vary with age: obstacles, $F(1, 58) = 0.13$, $p = .722$; bypassing, $F(1, 58) = 0.15$, $p = .701$.

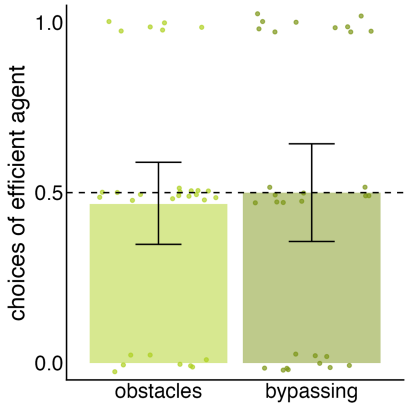
Figure 7
Sample Trial Slides From Experiment 4



Note. Children saw scenarios where both agents took the same paths, but where one agent's path was unnecessarily long. In one trial type, the efficient agent's path was constrained by obstacles while the inefficient agent was unconstrained (left panel). In the other trial type, the efficient agent went to the closest flower, while the inefficient agent unnecessarily bypassed a closer one (right panel). The top row shows the first slide in each story, and the bottom row shows the second slide. See the online article for the color version of this figure.

Figure 8

Choices of the Efficient Agent in Experiment 4



Note. Children judged whether an efficient or inefficient agent did a better job. Both agents took the same path. In some trials, their paths were circuitous, with the efficient agent constrained by obstacles and the inefficient agent unconstrained. In other trials, the efficient agent retrieved a close object while the inefficient agent unnecessarily bypassed a closer object. See the online article for the color version of this figure.

Discussion

Children aged 3–5 chose between efficient and inefficient agents at chance, even though this experiment used stimuli intended to remove potential ambiguities and to make inefficiency clearer. If anything, children may have performed worse in this experiment for the scenarios where the inefficient agent needlessly bypassed their goal object.

General Discussion

We investigated how children come to use efficiency to evaluate how well goals are completed. Children saw scenarios where pairs of agents approached objects using paths varying in length and efficiency, and then judged which agent did a better job. In Experiment 1A, we tested children aged 3–4 and found they evaluated goal completion using path length. They chose an agent who took a shorter path rather than a long one when judging who did a better job, and responded in the opposite way when instead judging which agent had more fun. Children's responses were unaffected, though, by whether the agents differed in efficiency. For instance, they mostly said the agent with the shorter path did a better job, even when the agent taking the longer path behaved efficiently as well. Results were less clear in Experiment 1B, though an analysis comparing the experiments suggested their results did not significantly differ.

In the next experiments, we tested children aged 3–7. In some trials, one agent took a shorter path than the other, and from around age 4, children mostly said the agent taking the shorter path did a better job (i.e., from 4;2 and 4;4 in Experiments 2A and 2B and from 3;11 in Experiment 3). In other trials, both agents took paths of the same length, but one agent efficiently took the shortest possible path whereas the other agent could have taken a shorter one. In these trials, children mostly chose the efficient agent from age 5 or later (never in Experiment 2A, from 5;11 in Experiment 2B, and from 5;2 in Experiment 3). Finally, in Experiment 4, we tested children aged 3–5 using stimuli intended to remove some potential ambiguities from those used in the other experiments. However, children in this experiment chose between the efficient and inefficient agents at chance rates.

Together, these findings suggest that a norm for completing goals efficiently first emerges at around age 5 or 6; before this, children also evaluate how well goals are completed but by attending to path length. A norm for completing goals efficiently does not imply that children think *all* actions should be efficient. One reason is that children might not see all activities as goal-directed. For instance, going for a walk or playing might not normally be viewed as goal-directed or at least not in the same way as instrumental actions like fetching a water bottle. In our first experiments, we started to investigate this distinction by asking children about which agent had more fun, and children responded differently than when judging who did a better job (for related work, see Chu & Schulz, 2023). Children seeing efficiency as normative also does not imply that they like efficient agents more than inefficient ones. People are not maximally efficient (Ongchoco & Jara-Ettinger, 2020) and might appear robotic or inhuman if they were. So even with instrumental actions, children might acknowledge that the more efficient agent did a better job while nonetheless feeling that they would rather be friends with the less efficient agent.

Younger children evaluated goal completion based on path length rather than efficiency. Their responses could indicate the belief that it is better to complete goals quickly rather than slowly (also see Leonard et al., 2019). Their preference for shorter paths is also broadly consistent with findings that from age 4, children also predict that agents will take shorter paths rather than long ones (Gönül & Paulus, 2021). Even so, it was not obvious that younger children would equate taking a shorter path with doing better. Instead, they could have seen taking a longer path as involving more work and effort and concluded that this counted as doing better. Similarly, 3- to 4-year-olds often respond to questions by preferentially choosing options with more of a quantity (e.g., Sullivan et al., 2018; also see Stavy et al., 2002).

Young children's preference for shorter paths also helps address concerns about whether our participants understood the test question about which agent did a better job. While 3-year-olds might have had difficulty with the question, this is unlikely for 4-year-olds. If they had misunderstood the questions, they would not have preferentially chosen agents with shorter paths. Previous work on other topics also suggests that preschoolers can successfully answer questions about which of two agents did a better job (e.g., Foushee et al., 2017; Gweon & Asaba, 2018). Nonetheless, it could be informative to test children using other questions. For instance, perhaps children would respond differently if asked about which agent took a better path. This question might help children focus on the trajectory of motion and reduce chances of children considering other things, such as which agent worked harder.

What accounts for the developmental change in children's use of efficiency—why did older children use efficiency to guide their evaluations, and why did younger children neglect it? One explanation is that our younger participants were aware that the agents differed in efficiency, but did not see it as normative; they might not think it is any better to complete goals efficiently rather than inefficiently. On this account, younger children could base their responses on efficiency if asked some other question. For instance, they might succeed if asked to indicate which agent could have taken a shorter path. Younger children might be expected to succeed given previous work suggesting that both infants and preschoolers

recognize efficiency, though (as noted above) at least some of the findings with preschoolers could reflect attention to factors like path length and speed. Regardless, one question raised by this account is what brings children to come to see efficiency as normative.

Another possibility, though, is that 3- and 4-year-olds simply could not evaluate which actions were more efficient. This proposal is difficult to square with claims that infants and toddlers distinguish between efficient and inefficient goal completion and use this as a basis for recognizing goal-directed actions and imitating others (for an overview, see Gergely & Csibra, 2003). One explanation is that infants' and toddlers' recognition of efficient actions could hinge on a perceptual or core knowledge system. While such a system should presumably remain active in older children, their responses to explicit questions could incorporate other input including additional background beliefs (Scholl & Gao, 2013, see especially pp. 203–204). Another explanation is that our tasks incorporated many demands not present in the tasks for infants and toddlers, including the need to remember and compare scenarios and the need to respond to questions. These demands could have prevented the younger children from determining which action was more efficient.

Nonetheless, the proposal that younger children could not evaluate efficiency may fit with another body of work on the development of counterfactual reasoning—the ability to think about how things could have turned out differently (for reviews, see Beck & Riggs, 2014; Nyhout & Ganea, 2019a). Our experiments may have involved counterfactual reasoning because children may have asked themselves whether agents *could have* taken a shorter path. Counterfactual judgments about what could have happened matter in many socio-moral evaluations (Gautam & McAuliffe, 2024), as when children judge that someone who gave the best gift they could was being nicer than someone who could have given a nicer one. But social judgments based on what could have happened typically emerge at age 5–6 or later (Gautam et al., 2023; Wong et al., 2023; Zhao et al., 2021), though some designs allow for earlier success (e.g., Gweon & Asaba, 2018).[5]

A further explanation for younger children's neglect of efficiency in our experiments is that the scenarios did not do enough to convey the importance of acting efficiently. This might be clearer to children, for instance, if they were told that the agents had to retrieve the goal objects as quickly as possible. If changing the scripts in this way brought younger children to choose efficient over inefficient agents, it would show that these children can recognize efficiency (i.e., contrary to the proposal relating to counterfactual reasoning considered above). But such findings would not show that younger children normally see efficiency as normative (i.e., since the task script would basically tell them that the agents were supposed to be acting efficiently).

Our experiments raise other directions for future research. One is to better understand performance variations across the experiments, since this could be informative about conditions that help children assess efficiency and consider counterfactuals more generally. For instance, it looks like the efficiency norm emerged earlier in Experiment 3 than in the earlier ones. Several differences between the experiments could explain this difference (assuming it is robust). For instance, it could have resulted from agents in that experiment collecting multiple objects rather than just one or because that experiment looked at a different kind of efficiency (i.e., whether the agent bypassed potential targets). The findings of the final experiment lend some support to multiple objects explanation—that experiment used a version of the task where agents each collected just one object, and 5-year-olds no longer preferred the efficient agent.

Future research could also aim to explore children's understanding of efficiency with their thinking about costs (Jara-Ettinger et al., 2016). Although taking an unnecessarily circuitous path is inefficient, it is not necessarily especially costly. Norms of efficiency might emerge earlier, then, if children consider agents who incur costs in completing goals (i.e., either costs that are avoidable or unavoidable). Also, varying the costliness of inefficiency might affect children's inferences of agents' motives. With low costs, inefficiency might indicate that an agent is just having fun. But with high costs, it could indicate that the agent is either incompetent or missing information—they are unable to avoid the costs or are unaware the costs can be avoided.

[5] Children also use counterfactual reasoning in other ways. For instance, thinking about counterfactuals sometimes involves envisioning how things would be given some counterfactual premise, as when imagining what a muddy floor would look like if people had removed their shoes before walking on it (e.g., Harris et al., 1996; Kominsky et al., 2021; Nyhout & Ganea, 2019b). Assessing efficiency probably does not require this kind of detailed working out of what things would be like if a counterfactual premise were true.

References

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives*, *8*(3), 175–179. https://doi.org/10.1111/cdep.12082

Chu, J., & Schulz, L. E. (2023). Not playing by the rules: Exploratory play, rational action, and efficient search. *Open Mind: Discoveries in Cognitive Science*, *7*, 294–317. https://doi.org/10.1162/opmi_a_00076

Colomer, M., Bas, J., & Sebastian-Galles, N. (2020). Efficiency as a principle for social preferences in infancy. *Journal of Experimental Child Psychology*, *194*, Article 104823. https://doi.org/10.1016/j.jecp.2020.104823

Colomer, M., & Woodward, A. (2023). Should I learn from you? Seeing expectancy violations about action efficiency hinders social learning in infancy. *Cognition*, *230*, Article 105293. https://doi.org/10.1016/j.cognition.2022.105293

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237–267. https://doi.org/10.1016/s0010-0277(99)00039-6

Dahl, A., & Schmidt, M. F. H. (2018). Preschoolers, but not adults, treat instrumental norms as categorical imperatives. *Journal of Experimental Child Psychology*, *165*, 85–100. https://doi.org/10.1016/j.jecp.2017.07.015

Foushee, R., Falkou, N., & Li, P. (2017). "Two-pound cookies" or "two pounds of cookies": Children's appreciation of quantity expressions. *Language Acquisition*, *24*(4), 400–416. https://doi.org/10.1080/10489223.2016.1225744

Gautam, S., & McAuliffe, K. (2024). Why imagining what could have happened matters for children's social cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *15*(1), Article e1663. https://doi.org/10.1002/wcs.1663

Gautam, S., Owen Hall, R., Suddendorf, T., & Redshaw, J. (2023). Counterfactual choices and moral judgments in children. *Child Development*, *94*(5), e296–e307. https://doi.org/10.1111/cdev.13943

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gergely, G., & Jacob, P. (2012). Reasoning about instrumental and communicative agency in human infancy. *Advances in Child Development and Behavior*, *43*, 59–94. https://doi.org/10.1016/B978-0-12-397919-3.00003-4

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-h

Gönül, G., & Paulus, M. (2021). Children's reasoning about the efficiency of others' actions: The development of rational action prediction. *Journal of Experimental Child Psychology*, *204*, Article 105035. https://doi.org/10.1016/j.jecp.2020.105035

Gweon, H., & Asaba, M. (2018). Order matters: Children's evaluation of underinformative teachers depends on context. *Child Development*, *89*(3), e278–e292. https://doi.org/10.1111/cdev.12825

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259. https://doi.org/10.1016/s0010-0277(96)00715-9

Hitchcock, D. (2011). Instrumental rationality. In P. McBurney, I. Rahwan, & S. Parsons (Eds.), *Argumentation in multi-agent systems* (pp. 1–11). Springer-Verlag. https://doi.org/10.1007/978-3-642-21940-5_1

Højsgaard, S., Halekoh, U., & Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software*, *15*(2), 1–11. https://doi.org/10.18637/jss.v015.i02

Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology. General*, *146*(11), 1574–1585. https://doi.org/10.1037/xge0000345

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604. https://doi.org/10.1016/j.tics.2016.08.007

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006

Kelemen, D., Seston, R., & Saint Georges, L. (2012). The designing mind: Children's reasoning about intended function and artifact structure. *Journal of Cognition and Development*, *13*(4), 439–453. https://doi.org/10.1080/15248372.2011.608200

Kominsky, J. F., Gerstenberg, T., Pelz, M., Sheskin, M., Singmann, H., Schulz, L., & Keil, F. C. (2021). The trajectory of counterfactual simulation in development. *Developmental Psychology*, *57*(2), 253–268. https://doi.org/10.1037/dev0001140

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *Package "emmeans": Estimated marginal means, aka least-squares means*. Comprehensive R Archive Network.

Leonard, J. A., Bennett-Pierre, G., & Gweon, H. (2019). Who is better? Preschoolers infer relative competence based on efficiency of process and quality of outcome. In A. Goel, C. Seifert, & C. Freska (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 639–645).

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, *160*, 35–42. https://doi.org/10.1016/j.cognition.2016.12.007

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, *3*(26), Article 772. https://doi.org/10.21105/joss.00772

Nyhout, A., & Ganea, P. A. (2019a). The development of the counterfactual imagination. *Child Development Perspectives*, *13*(4), 254–259. https://doi.org/10.1111/cdep.12348

Nyhout, A., & Ganea, P. A. (2019b). Mature counterfactual reasoning in 4- and 5-year-olds. *Cognition*, *183*, 57–66. https://doi.org/10.1016/j.cognition.2018.10.027

Ongchoco, J. D., & Jara-Ettinger, J. (2020). Beyond rationality: We infer other people's goals by learning agent-variable expectations of efficient action. In S. Denison, M. Mack, Y. Xu, & B. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*.

Pesowski, M. L., Quy, A. D., Lee, M., & Schachner, A. (2020). Children use inverse planning to detect social transmission in design of artifacts. In S. Denison, M. Mack, Y. Xu, & B. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 845–851).

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*(2), 137–155. https://doi.org/10.1016/j.cognition.2004.11.005

Schachner, A., & Carey, S. (2013). Reasoning about 'irrational' actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309–327. https://doi.org/10.1016/j.cognition.2013.07.006

Schleihauf, H., Zhang, Z., Gomez, A., & Engelmann, J. M. (2023). From outcome to process: A developmental shift in judgments of good reasoning. *Cognition*, *236*, Article 105425. https://doi.org/10.1016/j.cognition.2023.105425

Schmidt, M. F. H., & Rakoczy, H. (2023). Children's acquisition and application of norms. *Annual Review of Developmental Psychology*, *5*(1), 193–215. https://doi.org/10.1146/annurev-devpsych-120621-034731

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333. https://doi.org/10.1016/j.cognition.2012.06.004

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception* (pp. 139–170). The MIT Press. https://doi.org/10.7551/mitpress/9780262019279.003.0009

Schuwerk, T., & Paulus, M. (2016). Preschoolers, adolescents, and adults visually anticipate an agent's efficient action; but only after having observed it frequently. *Quarterly Journal of Experimental Psychology (2006)*, *69*(4), 800–816. https://doi.org/10.1080/17470218.2015.1061028

Scott, R. M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological Science*, *24*(4), 466–474. https://doi.org/10.1177/0956797612457395

Sehl, C. G., Friedman, O., & Denison, S. (2021). Blind to bias? Young children do not anticipate that sunk costs lead to irrational choices. *Cognitive Science*, *45*(11), Article e13063. https://doi.org/10.1111/cogs.13063

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(46), 18728–18733. https://doi.org/10.1073/pnas.1312322110

Stavy, R., Tsamir, P., & Tirosh, D. (2002). Intuitive rules: The case of "more A—More B." In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 217–231). Springer. https://doi.org/10.1007/0-306-47637-1_12

Sullivan, J., Bale, A., & Barner, D. (2018). Most preschoolers don't know most. *Language Learning and Development*, *14*(4), 320–338. https://doi.org/10.1080/15475441.2018.1489813

Wong, A., Cordes, S., Harris, P. L., & Chernyak, N. (2023). Being nice by choice: The effect of counterfactual reasoning on children's social evaluations. *Developmental Science*, *26*(6), Article e13394. https://doi.org/10.1111/desc.13394

Zhao, X., Zhao, X., Gweon, H., & Kushnir, T. (2021). Leaving a choice for others: Children's evaluations of considerate, socially-mindful actions. *Child Development*, *92*(4), 1238–1253. https://doi.org/10.1111/cdev.13480