

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Children's Cost-Benefit Analysis About Agents who Act for the Greater Good

Permalink

<https://escholarship.org/uc/item/46174478>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors

Finiasz, Zoe
Shore, Montana
Xu, Fei
et al.

Publication Date

2023

Peer reviewed

Children’s Cost-Benefit Analysis About Agents who Act for the Greater Good

Zoe Finiasz (zoe.finiasz@duke.edu)

Department of Psychology & Neuroscience, 417 Chapel Drive
Durham, NC 27708 USA

Montana Shore (montana.shore@duke.edu)

Department of Psychology & Neuroscience, 417 Chapel Drive
Durham, NC 27708 USA

Fei Xu (fei_xu@berkeley.edu)

Department of Psychology, 2121 Berkeley Way
Berkeley, CA 94720 USA

Tamar Kushnir (tamar.kushnir@duke.edu)

Department of Psychology & Neuroscience, 417 Chapel Drive
Durham, NC 27708 USA

Abstract

Acting for the greater good often involves paying a personal cost to benefit the collective. In two studies, we investigate how children ($N = 154$, $M_{age} = 7.94$ years, $SD = 1.13$, Range = 6.03 – 9.98 years) reason about cost and consequence. Children predicted how many agents would pay a personal cost to prevent a consequence for their entire community and judged agent(s) who refused to pay this cost. In Study 1, children expected more agents to pay a minor cost to prevent a major consequence and judged defection as less permissible than in the opposite case. Study 2 investigated the intermediate cases (Major/Major and Minor/Minor Cost/Consequence). Children expected agents to pay a minor cost regardless of consequence, and only expected agents to pay a major cost when consequence was major. In their judgments, children only considered consequence – defection was more permissible when consequence was minor, regardless of cost.

Keywords: child development, moral judgment, greater good, cost-benefit analysis, personal cost, collective consequence

Introduction

Sometimes our actions are only impactful when there are enough people behaving similarly. For example, a single person’s decision to recycle is inconsequential; however, if enough people behave similarly, we may notice a significant amount of recyclable material diverted from landfills. An important feature of these kinds of cases is that there is a threshold for consequence – if few perform a certain action, there is no impact for anyone, but if a certain number of actors behave similarly, the entire community will experience the consequences (Levine et al., 2020). Further, when acting for the greater good, one often needs to pay a personal cost (i.e., going out of your way to find a recycling bin) to prevent a societal consequence (wasted resources, climate change, etc.). Much like adults, young children learn that there are times that they must pay a personal cost (i.e., raising your hand) to prevent a negative consequence for the collective (everyone talking at once). In two studies, we investigate how

6- to 9-year-old children consider the trade-off between personal cost and collective consequence when predicting and judging agents’ actions for the greater good.

Research has found that in cases where there is a threshold for consequence, judgments about the permissibility of a single agent’s action are based on reasoning about what would happen if many people hypothetically engaged in the same action. A recent study by Levine et al (2020) found that if the number of people interested in performing a certain action (e.g., taking a stone from a path) exceeds the threshold for a negative consequence (e.g., no more stones left to make a path), adults and 4- to 11-year-old children believe the action is impermissible. Given the evidence that children can reason about hypothetical consequences in these cases, an open question remains about their considerations of the degree of personal cost and collective consequence. We ask whether children consider cost and consequence and whether they prioritize one over the other in their reasoning about agents who act for the greater good.

Previous research has shown that infants and children incorporate cost-benefit analyses into their social judgments (e.g., Jara-Ettinger et al., 2015, 2016; Zhao & Kushnir, 2022). In one prior study for example, when two agents refused to help, toddlers indicated that the agent who would have paid a higher cost to help was nicer (Jara-Ettinger et al., 2015). Similarly, 4 – 7-year-old children expect agents to maximize utilities and minimize costs in their actions and they extend this reasoning into their prosocial judgments (i.e., deciding who is nicer) (e.g., Bridgers et al., 2016, Jara-Ettinger et al., 2017). Further, children think differently about different *kinds* of costs – they start reasoning about overcoming physical costs (i.e., lifting something heavy) by 6 – 7 years, while their reasoning about overcoming psychological costs (i.e., resisting a desire) develops by 8 – 9 years (Zhao & Kushnir, 2022). In addition, there is evidence showing that children between ages 4 – 7 show an increasing understanding of conflicting desires between self and society (e.g., Lagutta, 2005; Starmans & Bloom, 2016). This prior

work indicates that children expect others to act when the cost of that action is minor, and that they will excuse inaction when the cost of that action is greater. Further, their reasoning about cost differs when considering psychological versus physical costs.

Given that children are sensitive to the trade-off between costs and benefits to individuals, this raises the question of how (and whether) children weigh the cost to individuals against the benefits to the greater good. If cost matters, perhaps children will predict that agents will not be as ready to act for collective benefit when the personal cost is high and may even excuse inaction under these circumstances. If collective benefits are treated similarly to individual benefits, perhaps children will predict that agents will only pay a cost to act when the collective benefit is high and might excuse inaction when benefits to the collective are minimal. Finally, children might consider the trade-off between personal cost (low vs. high) and collective consequences (low vs. high) and weigh both as equally important in both their predictions and judgments of actions.

Reasoning about the greater good presents a unique case for considerations of cost and consequence because judgment of a single agent's action requires children to reason about hypothetical consequences. As such, a final consideration is how children reason about the hypothetical consequences (when only one agent has acted) versus the realized consequences (when enough agents have acted to incur a consequence). Prior work has shown that children do consider the hypothetical consequences in their permissibility judgments (Levine et al., 2020). We expand on these findings by asking whether children make these judgments not only in situations where they must reason about hypothetical consequences, but also in scenarios where enough other agents have acted similarly, such that the threshold for consequence has been met.

We investigate children's reasoning about agents who must pay a personal cost to prevent a harmful consequence to their community. In our studies, children learn that a single agent's action has no impact, but if enough agents act in the same way, a threshold for consequence is met. We chose to focus on ages 6 to 9 as prior work shows that children in this age range can reason about cost and consequence, and because existing research shows that children's understanding of psychological cost – namely, acting against one's own desires – develops during this age range (e.g., Zhao & Kushnir, 2019). We address a series of open questions about children's reasoning about acting for the greater good. First, we ask if children consider personal cost and collective consequence at all in these scenarios. Even further, we investigate whether they prioritize either cost or consequence in their reasoning. We also address the questions of whether children judge agents differently when they are reasoning about the hypothetical versus incurred consequences and whether they reason differently about physical versus psychological costs. We also address the questions of whether children judge agents differently when they are reasoning about the hypothetical versus incurred consequences and whether they

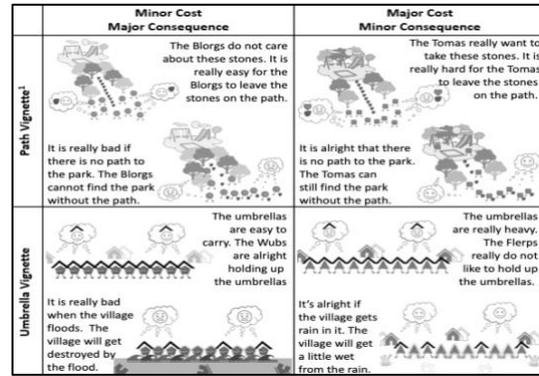


Figure 1. Script & still images of cost and consequence in the vignettes used for Study 1.

reason differently about physical versus psychological costs. We investigate these questions in terms of children's predictions and permissibility judgments of others' actions.

Study 1

The present study investigates whether children engage in cost-benefit analyses when making predictions and judgments about agents who must pay a personal cost to prevent a collective consequence. Here we compare children's responses in the extreme cases – Minor Cost/Major Consequence, and Major Cost/Minor Consequence. We chose to focus our questions on these cases because they provide the highest degree of contrast between cost and consequence. For each case, children were asked to *predict* the number of agents who would pay the cost and make *permissibility judgments* about agents' refusal to pay the cost. If children are sensitive to tradeoffs between personal cost and collective consequence, we expect that children will predict fewer agents to pay the cost and will rate defection as more permissible when the cost of doing an action is major but the consequence it prevents is minor. Study 1 also explores whether the kind of cost (psychological or physical) affects these predictions and judgments.

Method

Participants. Sixty-four 6- to 9-year-old children in Durham, NC ($M_{age} = 7.98$ years, $SD = 1.16$ years, Range = 6.03 – 9.94 years) participated in the study. An additional 5 participants were tested but excluded from our analyses (1 failed memory checks, 4 did not complete study). Children were recruited from a developmental research participant database, a local museum, and through social media (i.e., Facebook, Instagram, and Twitter).

Vignettes. Using a 2x2 mixed design, children were randomly assigned to a Cost condition (Psychological, Physical; between-subjects). In the *Psychological Cost* condition, children heard about agents who had to pay a psychological cost by resisting a desire to take stones from a path that led to a park (adapted from Levine et al., 2020). We

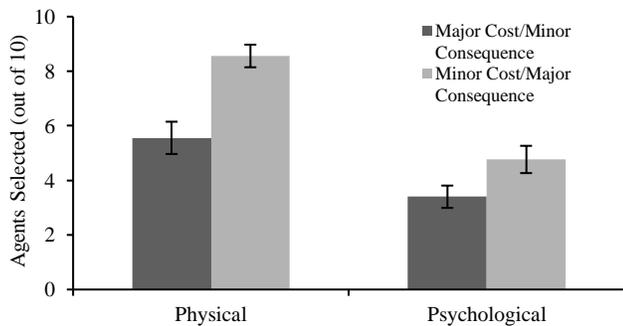


Figure 2. Children's responses to Prediction Questions by Vignette & Cost-Consequence Conditions, Study 1.

also designed a similar vignette for the *Physical Cost* condition: children heard about agents who had to pay a physical cost by holding up an umbrella to protect their village from rain. In both conditions, children heard two vignettes that manipulated the degree of personal cost and collective consequence (Major Cost/Minor Consequence, Minor Cost/Major Consequence; within-subjects). Thus, each child heard two stories, one where the personal cost was major and the collective consequence was minor, and a second vignette where the personal cost was minor, and the collective consequence was major (See Figure 1 for details). To illustrate the threshold for consequence in each vignette, children learned that, if one agent alone did not pay the cost, no consequence would occur, but if "too many" agents did not pay the cost, everyone would experience the consequence. The number of agents that needed to defect was never explicitly stated, all vignettes, showed 6 agents defecting when the threshold for consequence was met.

Test Questions. After each vignette, children answered four test questions: one Prediction question and three Permissibility Judgment questions. The prediction question asked children to estimate how many agents (out of 10) would pay the specified cost. The Permissibility Judgment questions asked children to judge an agent (or agents) who did not pay the cost (Likert scale: Really not okay – Really okay). In the *Lone Defector* question, children judged a single agent who did not pay act for the greater good. In the *Many Defectors* questions, children learned about a group of agents who did not pay the cost – here, enough agents did not pay the cost, and the consequence was realized. Children were asked to make two judgments in this case – one about a single agent from the group of defectors (*Single of Many*), and one about the entire group of defectors (*Group of Many*).

Coding. For the Prediction Questions, children's responses (0 – 10) were recorded. Permissibility Judgments were given on a 0 – 3 Likert scale with lower scores representing less permissibility (0: Really not okay, 1: A little not okay) and higher scores indicating more permissibility (2: A little okay, 3: Really okay). Each child answered two prediction questions (one for each Cost-Consequence Condition) and 6 Permissibility Judgments (three for each condition).

Procedure. Children were interviewed via Zoom or in-person. Prior to the interview, children received training to familiarize them with the scales used in both kinds of test questions with a set of unrelated examples. After training questions were completed, children heard one vignette, answered two memory questions, and then answered the four test questions. This was repeated a second time with the next vignette. The order of the vignettes was counterbalanced, and the order of the memory & test questions was fixed.

Results

Prediction Questions. To investigate children's predictions about how many agents would pay a personal cost to prevent a societal consequence, we ran a linear mixed effects model with Vignette (Psychological, Physical; between-subjects), Cost-Consequence (Minor Cost/Major Consequence, Major Cost/Minor Consequence; within-subjects), and the two-way interaction between Cost-Consequence and Vignette as predictors and Age (in months) as a covariate.

We found a main effect of Cost-Consequence ($F(1, 123) = 21.22, p < .001$): children expected fewer agents to pay a major personal cost to prevent a minor collective consequence ($M_{MajMin} = 4.51/10$ agents, $SD = 3.17$) than in the opposite case ($M_{MajMin} = 6.97/10$ agents, $SD = 3.04$; $t(62) = 2.45, p < .001$; See Figure 2). We also found a main effect of Vignette condition ($F(1, 123) = 36.83, p < .001$): regardless of the degree of trade-off between cost and consequence, children predicted fewer agents would pay the psychological cost in the Psychological Condition ($M_{Psych} = 4.08/10$ agents, $SD = 2.57$) than the physical cost in the Physical condition ($M_{Physical} = 7.06/10, SD = 3.32$; $t(62) = 2.98, p < .001$). There was no significant interaction between these two factors, and no effect of age (all $ps > .11$).

Permissibility Judgments. To investigate children's permissibility judgments, we again ran a linear mixed effects model with Vignette (Psychological vs. Physical) & Cost-Consequence (High Cost/Low Consequence, Low Cost/High Consequence), Question Type (Lone, Single, Many; within-subjects) and all two-way interactions with Cost-Consequence as predictors, and Age (in months) as a covariate. We found main effects of Cost-Consequence ($F(1, 373) = 70.01, p < .001$) and Question type ($F(2, 373) = 63.70, p < .001$) as well as a significant interaction between these two factors ($F(2, 373) = 2.97, p = .049$). We also found an interaction between Vignette and Cost-Consequence ($F(2, 261) = 4.91, p = .03$). There was no main effect of Vignette, and Vignette did not interact with Question Type. There was also no effect of age on children's permissibility judgments (all $ps > .001$).

Post-hoc comparisons revealed that children judged defection as less permissible when agents were paying a minor cost to prevent a major consequence ($M_{MinMaj} = 0.98, SD = 1.00$) than in the opposite case ($M_{MajMin} = 1.78, SD = 1.06$; $t(180) = 7.74, p < .001$). Further, when there was only one defector and a consequence had not been realized,

children judged the defection as more permissible ($M_{Lone} = 2.09$, $SD = 0.77$) than when they were a part of a large enough group of defectors and thus the consequence had been realized ($M_{SingleOfMany} = 1.16$, $SD = 1.07$; $t(318) = 7.86$, $p < .001$). They also judged a single defection as more permissible than a group of defections ($M_{GroupOfMany} = 0.89$, $SD = 1.07$; $t(318) = 10.11$, $p < .001$; See Figure 3). Children did not differ in their judgments of a single defector from the group versus the group as a whole ($p = .07$).

Next, we compared children's responses to the median possible ranking (1.5) to determine if children rated defection as permissible in each case. When there was only one defector and a consequence had not yet been realized, children judged defection as permissible, regardless of cost or consequence [$(M_{MinMaj-Lone} = 2.33$, $SD = 0.78$; one-sample- $t(63) = 9.58$, $p < .001$), ($M_{MajMin-Lone} = 1.84$, $SD = 0.69$; $t(63) = 3.52$, $p < .001$)]. When enough agents defected to exceed the threshold for consequence, children rated defection as impermissible when the personal cost was minor and the consequence major [$(M_{MinMaj-Single} = 1.64$, $SD = 0.84$; $t(63) = 1.06$, *ns*), ($M_{MinMaj-Group} = 1.36$, $SD = 0.73$; $t(63) = -0.98$, *ns*)]. When the personal cost was major and a minor consequence was realized, children's average responses were no different from the median score [$(M_{MajMin-Single} = 0.67$, $SD = 1.06$; $t(63) = -7.92$, $p < .001$), ($M_{MajMin-Group} = 0.42$, $SD = 1.15$; one-sample- $t(63) = -11.81$, $p < .001$)].

Finally, to explore the interaction between Vignette and Cost-Consequence, we compared children's permissibility judgments within each Vignette condition. Children rated defection as less permissible when agents had to pay a minor psychological cost to prevent a major consequence ($M_{MinMaj-Psych} = 0.87$, $SD = 0.95$) than in the opposite case ($M_{MajMin-Psych} = 1.88$, $SD = 1.08$; $t(318) = 6.76$, $p < .001$). The same pattern of results emerged for children who heard about agents who had to pay a physical cost (Physical Vignette) where children rated defection as less permissible when agents had to pay a minor cost to prevent a major consequence ($M_{MinMaj-Physical} = 1.69$, $SD = 1.03$) than in the opposite case ($M_{MajMin-Physical} = 1.08$, $SD = 1.04$; $t(318) = 4.32$, $p < .001$).

Discussion

Study 1 shows that, at least by age 6, children consider the trade-off between personal cost and collective consequence in their predictions and judgments of agents who act for the greater good. When agents must pay a minor cost to prevent a major consequence, children expect more agents to pay this cost, and judge defection as less permissible than in the opposite case. Further, children believe it is more permissible to defect in cases where the threshold for consequence has not been met – if only one agent defects, it is more okay than if multiple have defected. Taken together, these findings show that children are engaging in cost-benefit analyses when reasoning about agents who act for the greater good in both their predictions and judgments.

The current study also investigates children's consideration of the kind of cost – physical or psychological. While

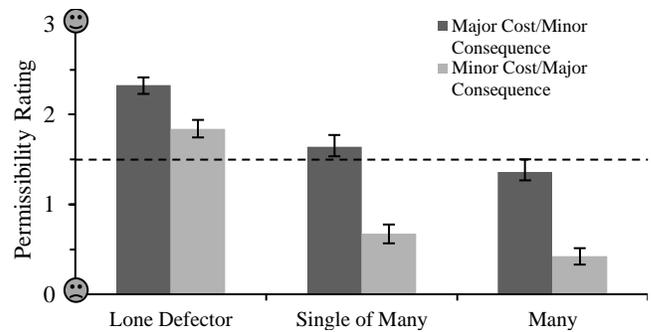


Figure 3. Children's responses to Permissibility Judgments by Cost-Consequence Condition, Study 1.

children predicted more agents to pay a physical cost over a psychological one, the pattern of responses was the same across vignette conditions; they expected more agents to pay a minor cost to prevent a major consequence. Further, there was no main effect of vignette on children's responses to the Permissibility Judgment questions. A possible explanation for this is that physical costs may be easier for children to understand. This possibility is supported by existing work comparing children's considerations of physical and psychological cost (e.g., Starmans & Bloom, 2016; Zhao & Kushnir, 2019). Because of this, we chose to focus on the physical cost condition for Study 2.

Study 2

Study 1 investigated children's cost-benefit analyses in the extreme cases – when cost is minor and consequence is major versus its opposite. From these findings, we can conclude that children consider *either* cost and consequence in their judgments and predictions but leaves an open question about which factor they prioritize in their reasoning. In Study 2, we manipulate degree of personal cost and collective consequence separately. We use the same test questions in a mixed design, where personal cost is varied between participants and consequence within.

One possibility is that children prioritize collective consequences over the personal cost to the individual for acting to mitigate those consequences. If this is the case, we would expect no difference in their responses across levels of personal cost, especially if the consequences are high and/or realized. On the other hand, children might be more focused on personal cost, especially if the cost is high. If this is the case, we would expect differences across cost conditions, regardless of the consequence. A third possibility is that children still consider both cost and consequence in their reasoning. If this is the case, we should see children's responses differing across cost and consequence.

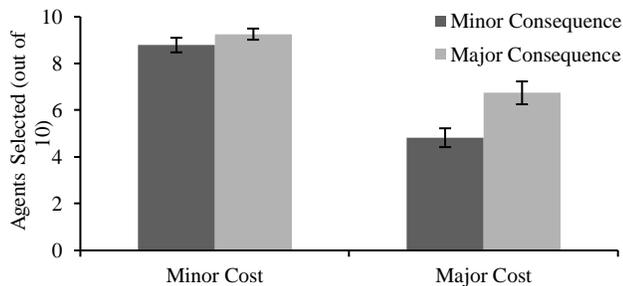


Figure 4. Children's responses to Prediction Questions by Cost & Consequence Conditions, Study 2.

Method

Participants. Ninety 6- to 9-year-old children in Durham, NC ($M_{age} = 7.93$ years, $SD = 1.10$ years, Range = 6.11 – 9.98 years; data collection ongoing) participated in the present study. A preregistered *a priori* power analysis indicated that we need 120 children in our final sample to achieve sufficient power.

An additional 8 participants were tested but excluded from our analyses (4 failed memory checks, 1 parent interference, 3 did not finish completing the study). Children were recruited from a developmental research participant database, a local museum, and through social media (i.e., Facebook, Instagram, and Twitter).

Procedure. Using a 2x2 mixed design, children were randomly assigned to a Cost condition (Minor Cost, Major Cost; between-subjects). Within each Cost condition, children heard two vignettes that manipulated consequence (Minor Consequence, Major Consequence; within-subjects). All participants heard the same Physical Cost vignette from Study 1 (holding umbrellas to protect their village from rain). Children were interviewed via Zoom and in-person. The procedure, familiarization questions, test questions, and coding scheme were the same as in Study 1.

Results

Prediction Questions. To investigate children's predictions about how many agents would pay a personal cost to prevent a collective consequence, we ran a linear mixed effects model with Cost (Minor, Major; between-subjects), Consequence (Minor, Major; within-subjects), and their two-way interaction as predictors, with Age (in months) as a covariate.

We found a main effect of Cost condition ($F(1, 87) = 48.61$, $p < .001$) where children predicted fewer agents would pay a major cost ($M_{MajCost} = 5.90$, $SD = 3.04$) than a minor one ($M_{MinCost} = 8.96$, $SD = 1.95$; $t(88) = 7.01$, $p < .001$). We also found a main effect of Consequence ($F(1, 87) = 16.46$, $p = .03$) where children expected fewer agents to act to prevent a minor consequence ($M_{MinCons} = 6.99$, $SD = 3.02$) than a major one ($M_{MajCons} = 8.00$, $SD = 2.81$; $t(88) = 3.48$, $p < .001$; See Figure 4). There was no significant interaction between these two factors, and no effect of age (all $ps > .08$).

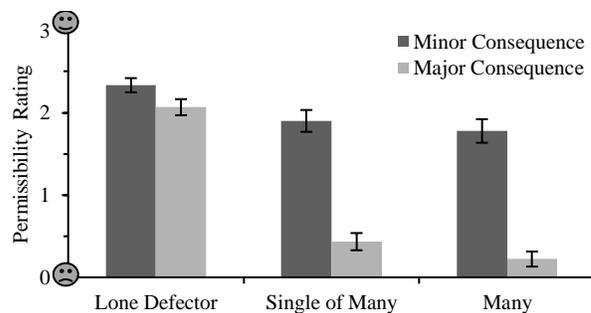


Figure 5. Children's responses to Permissibility Judgments by Consequence Condition, Study 2.

Permissibility Judgments. To analyze children's permissibility judgments, we again ran a linear mixed effects model with Cost (Major vs. Minor), Consequence (Major vs. Minor), and their interaction as predictors. In addition, we included Question Type (Lone, Single, Many; within-subjects) and its two-way interactions with Cost and Consequence as predictors. We again included Age (in months) as a covariate in our model.

We found a main effect of Question type ($F(2, 527) = 80.19$, $p < .001$) as well as an interaction between Consequence and Question Type ($F(2, 527) = 42.63$, $p < .001$). There were no other significant main effects or interactions (all $ps > .06$; See Figure 5, results are collapsed across cost conditions). When there was only one defector, and a consequence had not been realized, children judged a single agent's defection as more permissible ($M_{Lone} = 2.22$, $SD = 0.66$) than when he was a part of a large enough group of defectors to cause a consequence ($M_{SingleOfMany} = 1.17$, $SD = 1.08$; $t(448) = 10.73$, $p < .001$). They also judged a single defection as more permissible than a group of defections ($M_{GroupOfMany} = 0.98$, $SD = 1.07$; $t(448) = 12.66$, $p < .001$). Children did not differ in their judgments of a single defector from the group versus the group as a whole ($p = .13$).

To investigate whether children rated an agent's action as permissible, we compared their responses to the median possible ranking (1.5). When there was only one defector and a consequence had not been realized (Lone Defector), children judged defection as permissible, regardless of cost or consequence [$(M_{Minor-Lone} = 2.34$, $SD = 0.64$; $t(89) = 12.54$, $p < .001$), $(M_{Major-Lone} = 2.09$, $SD = 0.66$; one-sample- $t(89) = 8.41$, $p < .001$)]. When enough agents defected to exceed the threshold for consequence (Single of Many, Group of Many) children also rated defection as permissible when the consequence was minor [$(M_{Minor-Single} = 1.88$, $SD = 0.83$; $t(89) = 4.31$, $p < .001$), $(M_{Minor-Group} = 1.76$, $SD = 0.92$; one-sample- $t(89) = 2.65$, $p = .009$)]. Children only rated defection as impermissible when there was a major consequence [$(M_{Major-Single} = 0.46$, $SD = 0.78$; $t(89) = -12.68$, $p < .001$), $(M_{Major-Group} = 0.20$, $SD = 0.50$; one-sample- $t(89) = -24.58$, $p < .001$)].

Discussion

The findings of Study 2 provide further evidence that children are engaging in cost-benefit analysis when reasoning about

agents who act for the greater good. We replicated the results of Study 1 in the extreme cases – children expect more agents to pay a minor cost to prevent a major consequence, and judge defection as less permissible. Importantly, we found that children consider the trade-off between the degree of personal cost and collective benefit in their reasoning. Children in our current sample judged defection as more permissible when no consequence was realized, but they believed it was more okay to be a lone defector when the hypothetical consequence to the greater good was minor than when it was major. When enough agents defected so that consequences were realized, children in our sample judge defection as less okay overall, but importantly they rated defection when a consequence was major as less permissible than when a consequence was minor.

We also found that children prioritize avoiding collective harm in their permissibility judgments. When the consequence is major, children judge defection more harshly than when the consequence is minor, regardless of the cost an agent would incur. In their action predictions, however, children considered both cost *and* consequence, generally expecting agents to pay a minor cost regardless of the degree of consequence, but only expecting agents to pay a major cost if the consequence was also major. As in Study 1, they also expected fewer agents to pay a major cost and fewer agents to act to prevent a minor consequence. Taken together, these findings indicate that children consider both cost and consequence to different degrees when they are reasoning about agents who act for the greater good.

General Discussion

The present studies investigate whether children engage in cost-benefit analysis when thinking about agents who act for the greater good. A unique feature of many acts for the greater good is the existence of a threshold for consequence – a certain number of people need to engage in the same behavior for there to be a noticeable effect on society. Further, acting for the greater good often requires an individual to pay a personal cost to benefit society. While previous work has shown that children do engage in cost-benefit analyses when thinking about other kinds of moral scenarios, questions about the greater good provided an unexplored avenue for this kind of reasoning. In two studies, our findings demonstrate that children do engage in cost benefit analysis when reasoning about the greater good.

To begin, it should be noted that, across all questions, children showed evidence of cost benefit analysis. They consistently expected fewer agents to pay a high cost to prevent a minor consequence, and judged defection in this case as more permissible. Further, this trend remained regardless of the kind of cost (physical vs. psychological) and the scale of defection (single vs. many defections).

To disambiguate cost and consequence, Study 2 examined this trend further by introducing intermediate cases where both cost and consequence were major, or both were minor. In children's predictions, they considered both cost and consequence, expecting agents to pay a minor cost regardless

of the degree of consequence, but only expecting agents to pay a major cost when the consequence was also major. However, we found that children prioritized consequence in their permissibility judgments – children rated defection as less permissible when the consequence was major, whether a defector acted alone or whether they acted as part of a group large enough for consequences to be realized.

An important feature of scenarios like the ones we presented in this study is that a single agent's action should not be enough to incur any kind of consequence. As such, judging a single actor in these cases requires one to consider what happens if enough people also acted this way. Children's responses to the Lone Defector Questions show that they are engaging in this kind of hypothetical reasoning. While they judged defection in this case as overall more permissible than in cases where the consequences were realized, children still rated defection as less permissible in cases where consequences were major. If they were not considering *hypothetical* consequences, we would not see any differences because the actual outcome in this question (i.e., village stays dry) is no different across conditions. Instead, their ability to think hypothetically about what would happen if too many agents behaved similarly provides further support for existing research showing that children think about the hypothetical consequences to the greater good in their moral reasoning (Levine et al., 2020).

While the current study demonstrates children's ability to engage in cost benefit analysis about the greater good, several open questions remain. First is the question of kinds of costs and consequences. Future work will explore this question further by manipulating the kind of cost and the kind of consequence to explore whether this factor influences children's judgments. A second open question is regarding developmental timelines. The current study investigated this question with 6- to 9-year-old children. We did not anticipate any age effects to emerge in this range and indeed, across both studies and all test questions, age was not a significant factor. Given the evidence that reasoning about hypothetical consequences is present in younger children (Levine et al., 2020) and that younger children also engage in cost-benefit analysis about other kinds of moral problems (e.g., Jara-Ettinger et al., 2015, 2016), we hope to explore the developmental timeline of this reasoning by adapting our existing methodology to be suitable for younger age groups.

Finally, future research should consider how children reason about real-world instances where they themselves must act for the greater good. The current study presented children with a novel scenario, but children encounter this kind of question regularly in their daily lives (i.e., deciding whether to take one piece of candy on Halloween). Further, the COVID-19 pandemic has presented children with a very salient (and high-stakes) situation in which they need to pay a personal cost (wearing an uncomfortable mask, getting a vaccine shot) to keep their community safe. Given its prevalence in young-children's lives, future work may investigate both how children think about these real-world cases and how children actually behave in these scenarios.

Acknowledgments

We thank Joanna Brooke, Alyssa Gabbidon, and Kyra Hoskin for help with data collection. We also thank the staff of the Durham Museum of Life and Science for their partnership and ongoing support of our research. Finally, we thank the children and families in the Durham community who participated in this project.

References

- Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2019). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*, *4*(2), 144–152. <https://doi.org/10.1038/s41562-019-0748-6>
- 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.2016.05.011>
- 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>
- Lagattuta, K. H. (2005). When you shouldn't do what you want to do: Young children's understanding of desires, rules, and emotions. *Child Development*, *76*(3), 713–733. <https://doi.org/10.1111/j.1467-8624.2005.00873.x>
- Levine, S., Kleiman-Weiner, M., Schulz, L., Tenenbaum, J., & Cushman, F. (2020). The logic of universalization guides moral judgment. *Proceedings of the National Academy of Sciences*, *117*(42), 26158–26169. <https://doi.org/10.1073/pnas.2014505117>
- Starmans, C., & Bloom, P. (2016). When the spirit is willing, but the flesh is weak. *Psychological Science*, *27*(11), 1498–1506. <https://doi.org/10.1177/0956797616665813>
- Zhao, X., & Kushnir, T. (2019). She helped even though she wanted to play: Children consider psychological cost in social evaluations. *CogSci*, 3199–3205.
- Zhao, X., & Kushnir, T. (2022). When it's not easy to do the right thing: Developmental changes in understanding cost drive evaluations of moral praiseworthiness. *Developmental Science*, *26*. <https://doi.org/10.1111/desc.13257>