

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

"Because I want to": Valuing goals for their own sake

Permalink

<https://escholarship.org/uc/item/6dg5w6cf>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Chu, Junyi

Schulz, Laura

Publication Date

2022

Peer reviewed

”Because I want to”: Valuing goals for their own sake

Junyi Chu

Dept. of Brain and Cognitive Sciences, MIT
jchu@mit.edu

Laura E. Schulz

Dept. of Brain and Cognitive Sciences, MIT
lschulz@mit.edu

Abstract

People are often reluctant to reconsider their choices, sticking with their goals even when it is clear that they would be better off abandoning them. Explanations for this abound, including loss aversion, sunk costs, social and reputational pressures, and resource rational consideration of the costs of re-planning. Here we propose another hypothesis: In adopting a goal, you immediately reap the rewards of gaining information about what to do and how to act. Insofar as goals are rewarding in themselves, we predict that unless a goal is specifically devalued or the costs associated with it are very high, the default is not to engage in any reconsideration at all. We test this hypothesis by creating a stripped down scenario involving choices between two goals with transparently obvious cost differentials. The task is designed to minimize other factors that might contribute to goal persistence and indeed, we test both adults and very young children on virtually the same task to ensure that the cognitive load for adults is negligible. Both adults (Experiments 1-2) and 4-6-year-old children (Experiments 3-4) choose the less costly of two goals when shown the costs and goals together. However, when participants are shown the goals first, and only then shown that their chosen goal is more costly than the alternative, participants stick with higher cost goals, unless the goals are explicitly devalued.

Keywords: Decision making; Goals; Persistence; Rational Choice

Both adults and young children are sensitive to the costs and rewards of actions. Cost-benefit analyses guides adults’ choices not just in laboratory settings or economic decisions but also about health, crime (Becker, 1968), and socio-political choices about voting and alliances (Whiteley, 1995). Children also are sensitive to expected utilities: they prefer small immediate rewards to later larger ones but rationally modulate this preference according to expectations of environmental reliability (Kidd, Palmeri, & Aslin, 2013); they balance costs and rewards when exploring for information (Ruggeri & Lombrozo, 2015; Kidd, Piantadosi, & Aslin, 2012, 2014) and they expect others to maximize utilities as well (Liu & Spelke, 2017; Jara-Ettinger et al., 2015)

This sensitivity to utilities would suggest that, given a choice between two goals of equivalent value but different costs, both adults and children should choose the goal that is easier to obtain. However, both abundant research and everyday experience suggest that people do not always make the prima facie rational decision (Kahneman & Tversky, 1982). A striking instance of people’s failure to maximize utilities is that after people have chosen a goal, they are often reluctant to reconsider it, even when it is clear that it would be

advantageous to do so (Arkes & Blumer, 1985). Investors continue pouring money into projects even once it is clear that they are unprofitable (Garland, 1990); experienced pilots continue on their flight path even when the signs of danger are evident (O’Hare & Smitheram, 1995), and doctors persevere on treatment regimens even when better alternatives are available (Phillips et al., 2001; Okonofua et al., 2006). Indeed, faced with bad outcomes from an initial choice, people often paradoxically escalate their commitments (Staw, 1976).

Much of the work in economics and psychology has focused on the reasons why people deviate from the predictions of rational models (including misplaced optimism about the probability of success; Arkes & Hutzel, 2000); a willingness to take risks to avoid losses (Pope & Schweitzer, 2011); a sense of personal responsibility (McCarthy, Schoorman, & Cooper, 1993); social and reputational pressures and a failure to recognize alternative possibilities (Harvey & Victoravich, 2009). However, other work has focused instead on respects in which it may be rational to commit to a goal, even when seemingly better choices are available. Philosophers have suggested that tying ourselves to the mast of a thoughtful, committed decision allows us to fulfill our intentions despite temptations that might otherwise undermine our will. Among the benefits of ”rational resolve” and ”rational non-reconsideration” (Holton, 2004; Bratman, 1987) is avoiding the cognitive costs associated with weighing alternatives and changing courses of action. Relatedly, work in psychology has suggested that apparent deviations from optimal choice can be explained by resource-rational analyses that take into account the costs of acquiring and processing information and limits on time, attention, and memory (Lieder & Griffiths, 2020).

Given the myriad accounts already advanced to explain people’s tendency to stick with their initial choices in the face of seemingly preferable alternatives, it might seem unnecessary to propose yet another hypothesis. However, our interest in this topic stems not from a primary interest in decision-making, but from our interest in the value of goals as constraints on planning and hypothesis generation. We suggest that we may value our goals not only for their particular content or the potential reward associated with achieving them but because goals are structured representations that support thought and action. Having a goal gives us information about which actions are worth taking and which ideas are worth

thinking about. Whether those ideas and plans actually result in the achievement of the goal or not, they may be valuable: plans generated in the service of one goal can be decoupled from that aim and repurposed to other ends.

If the goals we choose are rewarding as soon as we adopt them (because as soon as we entertain them, they allow us to think the next thought and plan the next plan), people might tend to stick with a chosen goal at cost and might do so even when the myriad other factors that can contribute to inertia in decision-making are unlikely to apply. This is not to say that we believe people will never change their minds: if a goal becomes meaningless (e.g., because the problem it was aiming to solve no longer exists) or if the cost differential between a chosen goal and an alternative becomes extreme, we expect people to seek out and adopt alternative goals. The idea that goals have an intrinsic value regardless of whether they are fulfilled is intended to supplement ordinary considerations of utility, not supplant them.

Here we test this hypothesis by giving participants a stripped-down task designed to minimize the many other factors that contribute to goal persistence, including loss aversion, uncertainty, misplaced optimism, social pressure, and the neural and cognitive costs associated with re-planning. We give participants a choice of two goals, designed to be equally compelling, so that at baseline, each will be adopted by roughly half the participants. In the baseline condition (Goals + Costs), participants choose their goal with full knowledge of both the goals and the associated costs. In the critical test condition (Goals First), and a second control condition (Goals Devalued), participants first choose their goal in the absence of any other information, and then see the costs associated with each goal, manipulated such that their chosen goal always has a higher cost. In the Goals Devalued condition additionally, participants are told that their chosen goal is no longer especially valuable (because the problem has disappeared or because others have solved it). In both the Goals First and Goals Devalued conditions, participants are asked if they want to stick with their original goal or switch. We predict that participants will opt for the easier goal at both baseline (Goals + Cost) and the Goals Devalued condition but will stick with their original costly goals in the Goals First condition.

We run experiments in both adults and young children for two reasons. First, although there has been a lengthy literature on stickiness in adults, we are unaware of literature on children's tendency to persist on costly goals. This is interesting to test in its own right because children might both be more likely to persist because they cannot represent action costs, or more willing to abandon one plan in favor of another. By running the same task in 4-6-year-olds and adults, we can be confident that the relative cognitive costs for switching or processing are negligible for adults. We focused on 4-6-year-olds because abundant evidence exists that they are sensitive to costs and rewards during decision-making (Jara-Ettinger et al., 2015, 2016).

Critically, our experiment is designed to mitigate against many existing explanations for why participants might stick with costly goals. It cannot be the case that participants are committed to their goal because of sunk costs or loss aversion – at the moment of choice, participants have not engaged in any work towards the goal at all. Similarly, it cannot be the case that participants are uncertain about the relevant costs or unrealistically optimistic about the probability of success – the costs are transparent and although the costs are relatively higher in one case than the other, both are eminently surmountable. The participants are not subject to any group dynamics or reputational threats – the choices of goals are closely matched and arbitrary so deviating from them is unlikely to trigger threats to identity or self concept. On similar grounds, philosophical arguments about the virtues of rational resolve and resistance to temptation are unlikely to apply; both goals are virtuous and neither has any implications for the participants' well-being. Finally, although we cannot rule out the possibility that there are always cognitive and neural costs associated with changing plans, the task is designed to be almost trivially easy. Participants have a forced choice of two options and the difference in the costs of the two options can be seen literally at a glance (see Figure 1B-C).

In such a context, we suggest that the reason participants stick with their chosen goal – despite its relatively higher cost – is that as soon as you've chosen the goal, you've reaped some of its rewards: you know what you are going to do and you know something about how you're going to do it (indeed in our simple case, you know almost everything about how to achieve it). That is, merely having the goal has set up a well-defined space for thinking, planning, and acting. We suggest that in this kind of context, the default is not to engage in any reconsideration at all. Unless, as in our control condition, the goal is specifically devalued or (as in a condition whose outcome seem sufficiently certain that we need not run it) the absolute cost of achieving the initial goal makes it actively aversive, we predict that people will be inclined to ignore the cost differential and stick with harder goals.

Experiment 1

We began by comparing adults' choices on the baseline condition (Goals + Cost) and critical test condition (Goals First). While the same goals and costs were presented in both conditions, if merely choosing a goal makes it more likely that participants will stick to it, then participants in the Goals First condition would complete costlier actions more often than participants in the Goals + Cost condition.

Methods

Participants Fifty-six adults were recruited on Amazon Mechanical Turk and received \$1.25 for participating. Participants were randomly assigned to condition (29 Goals+Cost and 27 Goals First). Fifteen additional adults participated but were excluded from analysis for failing attention check questions (n=4) or self-reporting that they repeated the study or have previously seen the stimuli used (n=11).

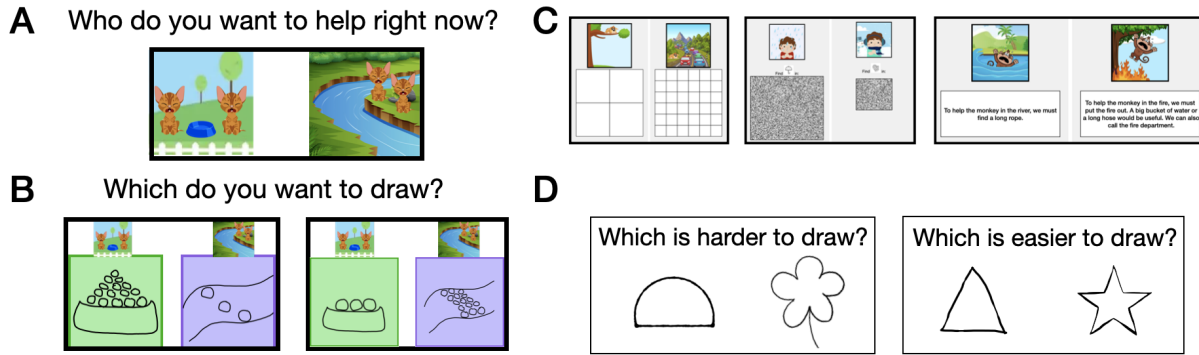


Figure 1: Experimental design. On each trial we presented two helping goals, matched for emotional intensity (e.g. sad kittens who were hungry vs. lost.) Each goal could be resolved by an accompanying action. In the Goals First and Devalued conditions, participants chose a goal (A) before seeing the actions (B). Their chosen goal was always accompanied by the costlier action. In the Goals + Cost condition, we counterbalanced across participants which goal had the harder action. Participants chose a goal only after seeing the target actions (B). In Experiments 1-2, we presented adults with different action types across trials (C). In Experiments 3-4, we presented children with easier or harder drawings to copy, and excluded children who responded inaccurately to any difficulty rating question (D).

Materials and Procedures Participants completed an on-line survey taking approximately 10 minutes. Trials began with a brief cover story describing two characters who were equally worthy of receiving help (Fig. 1A): each character faced a different problem resulting in identically intense emotional states (e.g. sad kittens who were hungry vs. lost, shivering children who were cold vs wet from rain, scared monkeys trapped in fire vs. river, puppies stuck in a tree vs. on the road). Each character could be helped by performing some repetitive action; the action type was identical within a story but one task was always more effortful (e.g. clicking 5 vs 20 times, typing a short vs. long paragraph, searching a small vs. larger scene; see Fig 1C). This allowed us compare the generality of action cost across different materials. We randomized the order of stories and characters within stories, and counterbalanced which character required a harder task.

In the Goals First condition, we first presented both characters without their target actions and asked participants to choose a helping goal, without any knowledge of action costs: “*This family of kittens are hungry/lost... Who you want to help?*” (Fig. 1A). Next, participants saw the required actions (Fig. 1B). Critically, participants’ initially chosen goal was always paired with the harder action and the alternative goal turned out to be easier. We measured participants’ choice to stick with their original goal or switch to the easier task: “*You wanted to help the lost kittens, so you’ll need to click 20 times to make cat food. Are you ready to make cat food, or do you want to switch to the other one?*”.

In the Goals + Costs condition, participants made a single choice of who to help after receiving full information about both characters and the required actions: “*This family of kittens... To help them, you need to... (x2). Who do you want to help right now?*”

Results

Our primary effect of interest is whether the likelihood of choosing the harder action differed by condition. To test this we conducted a mixed-effects regression predicting action choice from condition, with random intercepts for subject and story. We obtained a significant effect of condition (likelihood ratio test $\chi^2(1)=22.94, p < .001$; OR=27.1, 95%CI=[5.51-134]), with participants choosing the harder task more often in the Goals First condition (M=2.89 trials, SD=1.48) than in the Goals+Cost condition (M=1.14, SD=1.03).

To assess responses against chance responding, we calculated estimated marginal means per condition (i.e. model-predicted probability of choosing the harder task on any given trial). Participants in the Goals First condition chose the harder task more often than chance (M=.86, 95%CI=[0.60-0.96], $z = 2.53, p = .011$), with few adults always choosing the easier drawings (n=three or 11%, not different than chance of 6.25%). In the Goals+Cost condition however, the harder drawing was chosen less often than chance (M= 0.18, 95%CI=[.06-0.43], $z = -2.38, p = .018$), with 10 adults (34%) always choosing the easier drawing (significantly more often than chance, binomial $p < .001$).

Discussion

We found that adults preferentially persist and take on costs to achieve their initially chosen goals. Given the objective tasks demands (to help one of the characters), this additional effort was unnecessary. Indeed participants in the Goals + Costs condition preferred the easier goal. However, given participants’ personally adopted goals (to help a particular character in the Goals First condition), the cost differential seemed to matter much less.

One limitation of Experiment 1 is that we cannot tell if

participants in the Goals First condition might have persisted with their original goals despite the higher action cost due to some pragmatic demand from being prompted to choose whether to switch or maintain their goals. In Experiment 2 we control for this possibility by allowing participants to switch on all trials. Instead, we manipulate the value of participant's goals by either resolving the chosen goal (thus devaluing it) or leaving it unresolved.

Experiment 2

Here we test Goals First and Goals Devalued trials within the same participants. The two trial types are identical except that in the Goals Devalued condition, participants learn that their chosen goal is no longer especially valuable (because the problem has disappeared or because others have solved it) before they decide whether to switch or stay with the initial goal. If participants in Experiment 1 persisted on Goals First trials for reasons other than valuing their chosen goal above and beyond its' extrinsic reward, then the same factors should apply in the Goals Devalued condition and we should find no condition difference. If however participants were motivated simply by a difference in *wanting* after having chosen a goal, then devaluing that goal should reduce observed persistence. An a priori power analysis based on pilot data indicated that a sample of $n=41$ would provide 80% power to detect a medium effect size.

Methods

Participants Forty-one adults were recruited via Amazon Mechanical Turk and paid \$1.25 for participating. Thirteen additional adults participated but were excluded from analysis for failing attention check questions ($n=2$) or self-reporting that they repeated the study or have previously seen the stimuli used ($n=11$).

Materials and Procedures We used the same materials and procedure as in Experiment 1, Goals First condition. However, two trials were modified ("Goals Devalued" trials) to include an additional piece of information immediately after participants chose their goal. Specifically, participants saw an image and sentence describing their chosen goal already being resolved (e.g., "*You wanted to help the lost kittens, so you need to... Oh! The lost kittens already got help*"), before being prompted to choose an action to complete ("*Are you ready to help... or do you want to switch to the... ?*"). Thus, all participants completed two Goals First trials and two Goals Devalued trials. We randomized story order and which story was assigned to which trial type.

Results

We conducted a mixed-effects logistic regression predicting action choice from trial type (Goals First or Devalued), with random intercepts for subject and story. We obtained a significant effect of trial type ($\chi^2(1) = 20.3, p < .001; OR=0.18, 95\%CI=[.08-.41]$), with adults choosing to stick with the

harder task more often on Goals First trials ($M=1.41$ of 2 trials, $SD=0.77$) than Devalued trials ($M=0.85, SD=0.91$).

Consistent with Experiment 1, inspection of estimated marginal means indicated that participants chose the harder task more often than chance on Goals First trials ($M=.78, 95\%CI=[.59-.89], z=2.82, p = .005$). In the Devalued condition however, participants' choices did not differ from chance ($M=.39, 95\%CI=[.22-.59], z=-1.06, p = .29$).

Discussion

Notably, the additional information provided on Goals Devalued trials did not change either the action cost or the action outcomes (i.e. copying a large bowl of kibble still helped to make cat food for the hungry kittens, even if they were no longer hungry) but it does change the value of the problem that participants have initially chosen and thus the value of the plan. This result supports the notion that the preference to *not* reconsider alternative goals stems more from concern with the value of ones' goals than from concern about the affiliated action costs.

Experiment 3

There is abundant evidence that even infants and children are sensitive to the action costs and rewards. Here we test if the observed propensity to persist with initial goals despite higher costs might generalize to younger children. We pre-registered¹ a target sample of $n=60$ based on an a priori power analysis for a medium effect of condition ($h=0.5; OR=2.5$). Data collection is ongoing; here we report only pre-registered analyses.

Methods

Participants Forty-two 4-6-year-olds ($M = 5.56$, range = 4.50-6.42 years) were tested on Zoom with an experimenter and given a \$5USD Amazon gift card for participating. An additional four children were tested but excluded for inaccurately identifying drawing difficulty during practice.

Participants were randomly assigned to condition: Goals+Cost ($n=21$ children, $M_{age}=5.58$ years) or Goals First ($n=21$ children, $M_{age}=5.54$ years).

Materials and procedure Testing sessions were conducted via the Zoom video calling platform and lasted approximately 25 minutes. The experimenter displayed slides through screen share and presented children with a series of binary choices. To minimize experimenter variability, we used pre-recorded audio clips for any prompts that children had to respond to. Otherwise, experimenters followed a standard script to transition between different trials and experimental phases.

As in Experiment 1, our main dependent variable was children's adoption of action cost. Test trials used the same four cover stories as in Experiments 1 and 2. Instead of asking children to click or type, we asked children to copy either more or less complex drawings onto paper (see Fig. 1B),

¹<https://osf.io/et6gs>

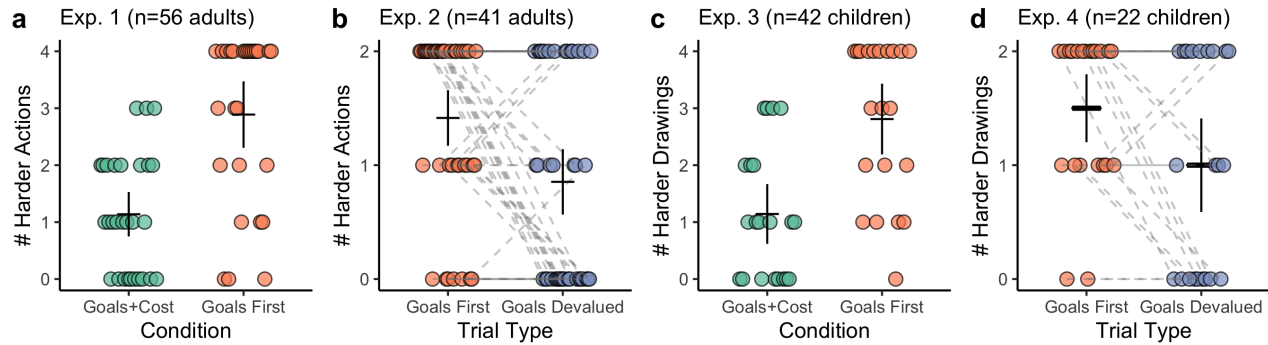


Figure 2: Frequency of choosing the harder action by experiment. Error bars show 95% CIs around condition means. When jointly choosing goals and action costs, participants preferred the lower-cost goal, more often than chance (Experiments 1 & 3, Goals+Cost condition). In the critical Goals First condition (all experiments), participants preferred to maintain their initial goal and completed higher-cost actions more often than chance. However, when participants' initially chosen goals were resolved (Experiments 2 & 4, Goals Devalued trials), we observed a decrease in goal-commitment behavior.

and told children that only accurate copies would “magically” transform into our story book to help the characters.

Prior to the test trials, we used an introduction phase to ensure fidelity of this action cost manipulation. First, as a check of general motor skill, we presented children with two simple shapes (semicircle / triangle) and asked them to copy one on paper. All children successfully drew their chosen shape, and the experimenter took a photo to transform it onto the slide deck (e.g., a semicircle into a rainbow). A second familiarization trial involved more complex shapes (star / flower). In order to reinforce children's expectation that target drawings must be copied exactly, without embellishments or simplification, we did not transform children's first attempt and instead told them: “*Hmm, the magic didn't work this time. The magic only works if you copy exactly what you see*”. Inaccurate drawings were given specific suggestions for improvement (e.g. “*Make sure to have exactly 5 petals*”) and accurate drawings were given general suggestions (e.g. “*Make it a bigger*”). Children's second attempt always transformed successfully.

Next, children answered two difficulty rating questions which served as inclusion criteria (Fig. 1D). These involved the same shapes from familiarization: children first judged whether the semicircle or flower was harder to draw and then judged whether a triangle or star was easier to draw.

Finally, children completed four test trials similar to Experiment 1. As with the adult participants, we measured whether children chose the harder or easier task on each trial, and whether these choices differed by condition. To help children verbally indicate their choices, we always introduced the first character in a green box on the left and the second character in a purple box on the right. Across participants, we counterbalanced which character was introduced first, which character required a harder drawing, and presented stories in one of two possible orders.

Results

We conducted a mixed-effects regression predicting action choice from condition, with random intercepts for subject². We obtained a significant effect of condition ($\chi^2(1)=16.3$, $p < .001$; OR=12.8, 95%CI=[3.25-50.3]). Children chose the harder drawing more often in the Goals First condition (M=2.81 trials, SD=1.36) than in the Goals+Cost condition (M=1.14, SD=1.15). Inspection of estimated marginal means indicated that in the Goals First condition, the harder drawing was chosen more often than chance (M=0.78, 95%CI = [.58-.90], $z = 2.66$, $p = .007$), with only 1 child always switching to the easier drawing. In the Goals+Cost condition however, the harder drawing was chosen less often than chance (M=0.21, 95%CI = [.10-.40], $z = -2.79$, $p = .005$, with eight children (38%) always choosing the easier drawing (significantly more than expected by chance of 6.25%, $p < .001$).

We also explored possible age effects by including an additional fixed effect of age (in months). This model did not explain significant additional variance (likelihood ratio test $\chi^2(1) = 1.63$, $p = .20$), and including an age by condition interaction did not improve model fit compared to the condition-only model ($p = .37$) or condition and age models ($p = .29$).

Discussion

Experiment 3 suggests that children, like adults, tend to persist with goals despite less costly alternatives. Children did so despite recognizing and preferring the lower cost option in the Goals+Cost condition, reflecting an ability to evaluate costs and a motivation to reduce costs. However, it is possible that children might persist for different reasons than adults. For instance, children might be generally “stickier” or less willing to switch their minds when prompted. In Experiment 4, we ask: would children also be sensitive to the value of their goals, and persist less on goals that are already resolved?

²We excluded a random effect of story due to limited sample size and model convergence issues.

Experiment 4

We pre-registered³ a target sample of $n=41$ based on an a priori power analysis for a medium within-subjects effect of trial type ($h=0.5$; $OR=2.5$). Data collection is ongoing; here we report only pre-registered analyses.

Methods

Participants Twenty-one 4-6-year-olds ($M = 5.49$, $SD = .50$, range = 4.67-6.33 years) were tested on Zoom with an experimenter and given a \$5USD Amazon gift card for participating. An additional nine children were tested but excluded for inaccurately assessing drawing difficulty pictures during practice ($n=8$) or not being able to copy the practice pictures ($n=1$).

Results

As in Experiment 2, we conducted a mixed-effects regression predicting action choice from trial type (Goals First or Devalued), with random intercepts for subject. We obtained a significant effect of trial type ($\chi^2(1)=8.54$, $p = .003$; $OR=0.19$, $95\%CI=[.05-.65]$), with children choosing to stick with the harder task more often on Goals First trials ($M=1.5$ of 2 trials, $SD=0.67$) than Devalued trials ($M=1.0$, $SD=0.93$).

Consistent with Experiment 2 and 3, children chose the harder task more often than chance on Goals First trials ($M=.84$, $95\%CI=[.61-.95]$, $z=2.65$, $p=.008$). In the Devalued condition however, participants' choices did not differ from chance ($M=.50$, $95\%CI=[.26-.74]$, $z=.02$, $p=.99$). Exploratory analysis found no significant effects of age.

General discussion

Across four experiments, we show that both adults and children persist with costly goals, despite having no sunk costs in their initial choice, transparently less costly alternatives, no social pressures to maintain their choice, and negligible costs associated with comparing goals or re-planning. We suggest that this behavior is consistent with the idea that goals have a value independent of their content or the probability of achieving them: the value of supporting thoughts, plans, and actions. The default favors sticking with a chosen goal and mitigates against considering other plans, even at cost, because the goal itself is valuable.

The current study also shows that young children, like adults, both rationally consider expected utilities in deciding their goals (i.e., in choosing the less costly of two goals at baseline) and resist switching to less costly goals once they have made a choice. Arguably of course, children's reluctance to switch goals might well be due to the costs associated with evaluating other options. Many studies suggest that children struggle with cognitive control and switching tasks (e.g., Traut et al., 2021; Zelazo, 2006). However, we suspect that task demands are unlikely to account for children's performance here. Experiment 4 suggests that many children,

like adults, readily switch goals when their initial goals are devalued.

In these tasks, we intentionally used goals with moral and emotional content (e.g., rescuing hungry kittens or monkeys stuck in trees). We did this to try to elicit something of the authentic attachment people have to real goals in the real world. Arguably however, participants were especially loyal to these goals because they involved altruistic acts for other agents. Insofar as participants felt beholden to the particular agents they had chosen to help, they might have been particularly unwilling to consider other options. Future research might replicate the current design with less affectively laden goals to see if this makes people more likely to opt for less costly alternatives. Consistent with this possibility however, some recent work in domains as neutral as navigation in 2-D grid worlds suggests that adults are slow to correct costly paths towards initially chosen goals (Cheng et al., 2021).

The current study is preliminary and our ideas about the intrinsic value of goals remain speculative. A task with more fine-grained, quantitative measures and graded manipulations of costs and rewards would allow us to assess the value of goals with more precision. For the moment however we will simply observe the paradox that in constraining our choices, goals motivate us to act; thus we will side with the philosophers in arguing for the rationality of non-reconsideration.

Acknowledgements

We are grateful to the participants and families for volunteering in this research. We thank Sofia Riskin, Ashley Lederman, and Bianca Santi for assistance with data collection and annotation, and members of the Early Childhood Cognition Lab for their feedback. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140.
- Arkes, H. R., & Hutzler, L. (2000). The role of probability of success estimates in the sunk cost effect. *Journal of Behavioral Decision Making*, 13(3), 295–306.
- Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13–68). Springer.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press, Cambridge.
- Cheng, S., Tang, N., An, W., Zhao, Y., Zhou, J., Shen, M., & Gao, T. (2021). Intention beyond desire: Humans spontaneously commit to future actions. In *Proceedings of the 43rd annual meeting of the cognitive science society*.
- Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology*, 75(6), 728.

³<https://osf.io/5skga>

- Harvey, P., & Victoravich, L. M. (2009). The influence of forward-looking antecedents, uncertainty, and anticipatory emotions on project escalation. *Decision Sciences*, 40(4), 759–782.
- Holton, R. (2004). Rational resolve. *The Philosophical Review*, 113(4), 507–535.
- 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.
- 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.
- Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246(1), 160–173.
- Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children’s decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, 126(1), 109–114.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, 7(5), e36399.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The goldilocks effect in infant auditory attention. *Child Development*, 85(5), 1795–1804.
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43.
- Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, 160, 35–42.
- McCarthy, A. M., Schoorman, F. D., & Cooper, A. C. (1993). Reinvestment decisions by entrepreneurs: rational decision-making or escalation of commitment? *Journal of Business Venturing*, 8(1), 9–24.
- O’Hare, D., & Smitheram, T. (1995). ”Pressing on” into deteriorating conditions: An application of behavioral decision theory to pilot decision making. *The International Journal of Aviation Psychology*, 5(4), 351–370.
- Okonofua, E. C., Simpson, K. N., Jesri, A., Rehman, S. U., Durkalski, V. L., & Egan, B. M. (2006). Therapeutic inertia is an impediment to achieving the healthy people 2010 blood pressure control goals. *Hypertension*, 47(3), 345–351.
- Phillips, L. S., Branch Jr, W. T., Cook, C. B., Doyle, J. P., El-Kebbi, I. M., Gallina, D. L., . . . Barnes, C. S. (2001). Clinical inertia. *Annals of Internal Medicine*, 135(9), 825–834.
- Pope, D. G., & Schweitzer, M. E. (2011). Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1), 129–57.
- Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, 143, 203–216.
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16(1), 27–44.
- Traut, H. J., Chevalier, N., Guild, R. M., & Munakata, Y. (2021). Understanding and supporting inhibitory control: Unique contributions from proactive monitoring and motoric stopping to children’s improvements with practice. *Child Development*, 92(6).
- Whiteley, P. F. (1995). Rational choice and political participation—evaluating the debate. *Political Research Quarterly*, 48(1), 211–233.
- Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, 1, 297–301. doi: 10.1038/nprot.2006.46