

## Response to Reviews

### Editor's comments:

Also, please consider these two minor points:

-I don't know if it would be possible to convey the expected Phi value if participants always switched, but responded randomly besides this (e.g., after choosing A, randomly choose between B, C, or D). But if a Phi value is associated with this kind of strategy, expressing it might help readers better interpret the findings. (The more general point is that the extreme Phi and Beta values are clearly expressed, but it might help to have some indication of what at least one other value might imply).

Unfortunately, it is hard to associate a specific value of phi with this type of strategy since it would appear more like random responding, resulting in the model being best fit with a very low beta value, at which point phi becomes unidentifiable. To test whether this was the case, we ran simulations using this strategy and fit the model to the simulations. The model fit this strategy poorly (much worse than to participants' data), and the resulting phi values were highly variable, with the whole range of possible values (from close to 0 to close to 1) occurring across different simulations. We agree, however, that it would be useful to include what another value of phi might represent to facilitate readers' understanding of the model. We have noted in the text that a value of 0.5 represents approximately equal weighting of reward values and choice lag in producing responses.

As a side note, in a previous project with a task equivalent to the Baseline condition, we tried a model that encapsulates the random switching strategy that is described above and found that for almost all children it was a worse fit than the model that we use in the current work.

-For papers to be accepted at Cognition, the data need to be made publicly available (e.g., in a public repository like OSF). So it may save time later if you were to provide the link in the next version of the paper.

We have added the data to a repository on OSF (<https://osf.io/ph9kz/>) and have included a link in the manuscript.

### Reviewer #1:

My concern continues to be that I'm not sure this data actually speak to a deep attentional shift/disruption; it's still not entirely clear how the children are interpreting the task, and so I don't believe there needs to be a story about attention at all here (as much as I'm sympathetic to this account and also believe it to be a deeply interesting story). Specifically, consider that children just have a heuristic, that when a task is

about exploration/sampling, the job is to try each one in succession. This doesn't have to be because their attention is broader (indeed, you could have perfectly focused attention and still decide to follow a rule to sample in an order, such as clockwise around the quadrants.) The question is -- what happens when the task becomes less clear -- because there is suddenly one object (salient) that is not like the others. The children must then figure out why that object is different and how it is supposed to relate to the task. In the case when it is convergent with the reward, the inference is fairly apparent and young learners should quickly converge on maximizing the salient cue. When the cue is not the highest reward, the cue becomes confusing in that the purpose of the task may be less clear -- leading to more random (but not systematic) exploration. My point is simply that this explanation for the behavior does not depend on a claim of children having more distributed attention; and these results cannot thus be taken as evidence that a salient cue thus disrupts this broad attention.

The reviewer brings up an interesting alternative explanation. The reason that we doubt the likelihood of this specific alternative is that it suggests that young children are using an explicit rule. However, because such exploration decreases with age (in fact, we have yet unpublished data indicating that systematic exploration increases in infancy to about 3-years of age, and then decreases between ages of 3 and 7). Given these data, we should conclude that the rule use decreases with age, which appears highly unlikely. Additionally, this alternative requires three separate explanations to account for the results of the three conditions, whereas an attentional explanation can account for the levels of systematic exploration in all three conditions with a single mechanism. We have updated the Discussion (pg. 17) to better highlight this point. We agree, however, that since we did not measure attention directly, we cannot rule out the possibility that disrupting distributed attention is not the primary factor driving differences and between the Baseline condition and other conditions in children. In light of these concerns, we note this possibility in Discussion and the importance of directly measuring attention (e.g. with eye-tracking) in future studies (pg. 18).

Indeed, the adult data speak more strongly to the claim that the saliency "breaks" the broader attentional pattern, as adults over-select the wrong item in the conflict condition, suggesting it is the adults (and not the children) that have a hard time inhibiting the saliency. This is almost counter to the broader story the authors are making, but the fact that I could tell a just-so story either way leads me to be concerned that we cannot speak clearly to the broader, developmental-cognitive theory posited here to help motivate and interpret this study.

While it is true that both children and adults show effects of saliency on choice proportions, raw choice proportions alone cannot tell the full story because a variety of different strategies may lead to similar choice proportions. The computational modeling analyses of participants' choice pattern are critical to uncovering these differences in strategy. Children's overall strategy changes between conditions, whereas adults' does not. When saliency is manipulated, children show a reduction in response switching and an overall shift from

systematic toward more random exploration. This disruption of systematic exploration due to the saliency manipulation is informative and occurs only in children. So, while it may be possible to come up with a variety of different explanations based only on the choice proportions, the modeling analyses reveal important differences between adults and children that provide a clearer picture with important implications for developmental-cognitive theory.

In the revision we tried to better highlight the importance of the modeling analyses, the differences between conditions that they uncover, and the implications for cognitive development. We have made changes throughout the manuscript, notably on pages 6, 10, and 17.

More minor note:

On page 5, the authors note "In the same way that attention shifts over time and is less likely to return to recently focused items, less recently chosen or attended options may become increasingly appealing over time." The statement requires citation and even conflicts to some extent with recent work on attention and decision making (that more "attended to" items are more likely to be chosen in a singular decision; data from Krajbich & Rangel, 2011; modeling work from Callaway & Griffiths, 2019). It may strengthen the authors' argument to state why the less recently chosen/attended options become more appealing. Is this true for both children and adults? Intuitively it would feel correct for children, whose "meta-goal" within the game may be to learn about the world, such that appeal arises from opportunity for information gain.

We thank the reviewer for this observation and suggestion. We have expanded on this specific point in the Introduction (pg. 5) in order to clarify our hypothesis. The referenced work is consistent with our hypothesis, and in fact, work from Krajbich and colleagues was highly influential in motivating these ideas. The attentional mechanisms that the reviewer alludes to and those that we are referring to operate at different time-points in the decision process. It is true that within a single decision attended to items are more likely to be selected in adults, and this may also be true in young children (indeed, our hypothesis assumes that it is). But, after a choice is made (and the outcome for that choice observed) there may be little reason to continue to attend to the object and attention may shift. The way that it shifts is likely different between adults and children, wherein adults control attention to continue focusing it selectively on rewarding options, while children are more likely to shift to a new option. Attention would then influence choices in the way that is suggested by the referenced work. More specifically, we suggest that children's shift may be driven by something like graded novelty preference: the recently chosen option is no longer novel, but the unselected options have been increasing in novelty since last being chosen. We further describe and clarify this idea in the revision (pg. 5).

### **Reviewer #3:**

The authors have been highly responsive to the previous reviews. It is especially helpful to have the adult comparison group. The authors have clarified what is potentially novel about these results.

Their argument shares some similarities with the Hierarchical Competing Systems computational model of the effect of the number of "A" trials on A-not-B task performance in infants and toddlers (see Marcovitch & Zelazo, 2009), which might be worth mentioning in the introduction.

We thank the reviewer for pointing us toward this interesting work. We have added a brief discussion of this idea to the Introduction (pg. 5).

What is most surprising about the results to me is children's performance in the Baseline condition. Why don't they home in on the high value option and stick with it, like adults do?

We agree that children's behavior in the Baseline condition is extremely interesting, and that understanding could be important. We are very interested in this question. The current paper is a follow-up to our original study that reports this effect (Blanco & Sloutsky, 2019 *PsyArXiv*) and is one part of a larger line of ongoing research that attempts to better understand this behavior.

Three non-exclusive alternative interpretations of the results for the authors to reconsider:

Children had difficulty learning to maximize rewards in this task, as shown in the Baseline. Salience scaffolded learning in the Congruent condition, but it disrupted learning in the Competition condition, where they knew the salient cue was low in value, yet could not resist it occasionally. It is a straw man argument to suggest that 5-year-olds would be lured 100% by salience alone, even when it conflicts with reward value (again, PFC is not absent at this age). On this account, it is not that their choices were "geared toward learning rather than maximizing reward" (pp.17-18), but instead, they are learning HOW to maximize rewards in the face of distracting stimuli.

This is an intriguing idea, and it may be that case that one thing that children are learning is how to maximize rewards in the face of distracting stimuli. This possibility does not contradict our interpretation or results, though. But, there is little evidence to support this interpretation in the current data, since it would suggest that children should maximize reward when there are no distracting stimuli (as in the Baseline condition). Few children do so. Children in the Baseline condition systematically explored instead of maximizing reward. Furthermore, the current study cannot distinguish between not learning to maximize reward, and simply preferring not to, so it would be misleading to suggest that children had difficulty learning to maximize reward in this task.

Indeed, in our previous study (Blanco & Sloutsky, 2019 *PsyArXiv*) we found that most children knew what the best option was despite not choosing it more often than the other options. In addition, we cannot conclude that salience disrupted *learning* in the Competition condition, since choice proportions were equivalent to the Baseline condition. The main difference between the conditions is revealed by the modeling analyses which instead show a substantial disruption of *systematic exploration* in the Competition condition. While children's choices are likely generally geared toward learning, it is hard to say what, if anything, children were learning in the Competition condition given their largely random responding in that condition.

It could also be that children were less motivated than adults, insofar as the rewards were quite distant; they were symbolized by drawings, which symbolized the value of candy in text, which was translated into a meter tracker, which symbolized not candy, but stickers (how many?) for every 180 "candies" earned. Although adults would understand this, and might be quite motivated to do well even without the cover story of candies or stickers but merely "points," it is possible this was too abstract for children. The salient cue in the Congruent condition helped them just enough to perform better, despite them not really understanding the symbolic goal. How would mature nonhuman primates perform in this experiment, if there were real food rewards? Based on the literature, would they be expected to perform more like adults and less like children? Relatedly, I wonder how rapidly children would learn if the rewards were more direct and appetitive.

The questions of how nonhuman animals would perform on this task is quite interesting, and could provide valuable insight into human development, but that question is outside of the scope of the current paper. For now, we can only speculate on how an adult primate would perform at this task, but we do have some unpublished data with pigeons (who are known to be neophobic, rather than novelty seeking). Pigeons show a pattern of local optimization. They typically try two, maybe three options, pick the best of those options, and then exploit it almost exclusively for hundreds of trials across many sessions spread over several days. Exploration after the initial few trials is very rare. But, it is important to note that the task for pigeons was more directly rewarding and was appetitive in that they received a food reward after every choice. It would be very interesting to see how an adult primate would perform on a version similar to what was run on humans. The potential limitation that the reviewer notes, however, that children's behavior may be different in this relatively abstract scenario than in a directly rewarding (e.g. immediate appetitive rewards) situation, is important to consider. We add discussion of this limitation and the importance of addressing this issue with future research to the Discussion (pg. 18-19).

In addition, we clarify that children earned one sticker for each 180 candies collected (pg. 8).

Lastly, it could be that the response pattern referred to as exploration in the Baseline condition is not merely due to an underdeveloped attention system, but also due to exploration itself being intrinsically rewarding. Especially if they did not fully understand the symbolism and goal of the task, this might be a default behavior. Would adults show this pattern if they were given a 4-choice selection task with no rewards associated with them? This idea is most intriguing.

We agree with the reviewer that exploration may be intrinsically rewarding (particularly to children), and we find this idea very intriguing as well. We've added text to the Discussion noting this possibility (pg. 19). This idea is completely consistent with our viewpoint, although we also note that the mechanisms driving the particular patterns in children's exploratory choices still require explanation that goes beyond this possibility. The current work represents an attempt to elucidate those mechanisms.