Spring 2018

Model-Based Learning in Children and the Role of Proactive Control

Jennifer Ladouceur
Jennifer.Ladouceur@Colorado.EDU

MODEL-BASED LEARNING IN CHILDREN AND THE ROLE OF PROACTIVE CONTROL

Jennifer Ladouceur

Department of Psychology and Neuroscience

University of Colorado at Boulder

March 22, 2018

Honors Committee:

Dr. Yuko Munakata, Professor of Psychology (Thesis Advisor)

Dr. Eliana Colunga, Professor of Psychology (Honors Council Representative)

Dr. Iskra Fileva, Professor of Philosophy (Outside Reader)

Abstract

Everyone has to make reward-based decisions. For example, a child may decide to behave well in a grocery store because they know they will get rewarded with ice cream. Strategies for learning from previous outcomes include model-based (uses a model of the environment, such as generalizing the good behavior to the library to get further reward) and model-free (reacts habitually to previous reward). Previous research used a two-step decision-making task to distinguish between these two strategies, which consisted of two decision stages. Adolescents and adults used both model-free and model-based strategies, whereas children only used model-free. However, there was a limitation to this task: the more effortful, model-based strategy did not receive greater reward. We hypothesized that when children were incentivized they would use model-based strategies and that age would be positively correlated with these strategies. A modified two-step task that incentivized model-based strategies was used. We found that children utilized a mixture of both model-based and model-free strategies with no significant correlation between model-based strategies and age. Additionally, in the original task, children demonstrated anticipatory preparation for a second-stage decision that could indicate proactive control tendencies, which is the process of maintaining goal-relevant information in anticipation of needing it. We hypothesized that proactive control would be positively correlated with model-based strategies. To assess this, we used the cued task-switching paradigm, which had participants sort toys based on shape or color in a proactive possible and proactive impossible condition. We found a marginally significant correlation between proactive control and model-based learning. We also included a measure of need for cognition and found a nonsignificant correlation with model-based learning. Our study was the first to find that children are capable of utilizing model-based strategies.

Keywords: Reinforcement Learning, Proactive Control, Cognitive Development

Throughout life, everyone must make reward-based decisions. These types of decisions are usually made through reinforcement learning, which happens through receiving rewards or punishments after actions. Reinforcement learning consists of model-based and model-free strategies. An individual utilizing a model-based strategy proactively uses a complex cognitive model of the environment to flexibly dictate actions based on potential outcomes, requiring an effortful process. In contrast, a person utilizing a model-free strategy uses reflexive tendencies to reactively repeat the action most recently rewarded or punished, requiring low-effort (Potter, Bryce, & Hartley, 2016; Akam, Costa, & Dayan, 2015). For example, if someone is playing poker and wins the hand on the river, a model-free person would bet more money on the next hand, whereas a model-based person would not because they are utilizing a model of the game of poker to determine that they won based on a small probability of the river card going in their favor.

These two strategies have been distinguished using a two-step decision-making task, in which an individual is presented with a series of choices resulting in a reward or nothing. Tendencies to employ model-based or model-free decision-making strategies are assessed by analyzing whether individuals take into account the probability of receiving reward or if they are simply repeating previously rewarded actions. Adults typically utilize a mixture of both model-based strategies, that is taking into account the probability of reward for their decision, and model-free strategies, that is habitually repeating the previously rewarded action, as they are neither mutually exclusive cognitively nor from a neuroscience standpoint (Daw, Gershman, Seymour, Dayan & Dolan, 2011).

Decker and colleagues (2016) sought to determine whether children and adolescents exhibit model-based and model-free strategies in a similar manner as adults using a two-step decision-making task. In this task, participants make an initial decision between two different spaceships, each leading to either a red or purple pair of colored aliens. Set probabilities were in place regarding which initial decision (pair of spaceships) led to which subsequent decision (pair of aliens) either the majority of the time (common transition) or barely at all (rare transition). Specifically, one initial decision (e.g., a green spaceship) would lead to subsequent decision A (e.g., purple colored aliens) 70% of the time and subsequent decision B (e.g., red colored aliens) 30% of the time and vice-versa. Participants then made a subsequent decision between the two aliens that were the same color but had different physical characteristics. The selected alien then provided a reward or nothing. The reward probability of the aliens slowly and independently drifted between 0.2 and 0.8 across the task. An individual using a model-based strategy would leverage the transition probabilities to make a decision with the goal of obtaining the most reward possible, whereas an individual using a model-free strategy would make a decision only based on previous reward with the same goal in mind (Decker et al., 2016). For example, a participant may chose a green spaceship that makes a rare transition to red aliens, which then results in a reward. An individual using a model-based strategy would choose the other spaceship on the next trial because he/she knows that the other spaceship has a higher probability of returning to the same planet and reward as the previous trial. However, an individual using a model-free strategy would habitually choose the green spaceship again without leveraging the transition probabilities because that resulted in a reward on the previous trial.

Both model-based and model-free strategies were identified in adults and adolescents, but only a model-free strategy was identified in children aged eight to twelve years (Potter & Bryce, 2017; Decker et al., 2016). Additionally, response times at the second-stage choice (i.e., selecting an alien) were slower for all age groups after a rare transition relative to common transitions, indicating even young children had implicit knowledge of the task structure. Children also indicated that they had knowledge of the transition structure during post-task questioning, which means they formed a cognitive model, or mental map, of the rare and common spaceship transitions characteristic of model-based learning.

Further, adults did not acquire any more reward than adolescents and children, even though they utilized the more effortful, model-based strategy (Kool et al., 2016). Further, Kool and colleagues (2016) ran a computational model of this two-step decision-making task simulating the use of a model-based versus a model-free strategy and found no advantage for using either strategy. As such, the two-step spaceship task was unable to determine whether children had the ability to be model-based and were adaptively avoiding unnecessary effort due to the lack of reward in return. To address these limitations, we used a modified two-step task developed by Kool et al. (2016) that provided higher rewards for using the more effortful, model-based strategy. We sought to determine whether children between the ages of eight- and twelve-years-old, previously shown to not exert any model-based decision-making, engage in model-based strategies when better incentivized.

Although children did not demonstrate evidence of model-based strategy use in the original task, the slowed response times at the second-stage choice after unlikely transitions suggests that they were engaging in proactive preparation while performing the task. The preparation for the subsequent decision may also be reflective of emerging proactive control processes in children. Interestingly, difference in response times after common and rare transitions correlated with model-based strategies in adolescents and adults (Decker et al., 2016).

Proactive control is the process of maintaining relevant information in anticipation of needing it, whereas reactive control is the cognitive process that occurs in response to being cued for the information (Braver, 2012). Further, proactive control is believed to be more effortful than reactive control in children (Braver, 2012). For example, younger children need to be encouraged to use proactive control, but older children typically use it whenever they can (Lucenet & Blaye, 2014; Chevalier, Martis, Curran, & Munakata, 2015). Children aged six to eight years and older are capable of proactive control (Chatham, Frank, & Munakata, 2009); however, children between the ages of eight and twelve years old still vary in their proactive control tendencies (Chevalier, Martis, Curran, & Munakata, 2015).

We predict that (1) children will demonstrate the ability to utilize model-based strategies when incentivized; (2) the utilization of model-based strategies will be positively correlated with age; (3) proactive control and model-based strategies will correlate because participants must use proactive control in the modified two-step task to anticipate relevant information regarding which choice will lead to the desired outcome and earn the highest reward.

Methods

Participants

We recruited 47 children between the age of eight and twelve years old to voluntarily complete a 60-minute session that included the modified spaceship task, a cued task-switching paradigm, and a child-adapted version of the short-form Need for Cognition Scale (NfC) (Cacioppo, Petty, Feinstein, & Jarvis, 1996; Chevalier, 2017). Four participants were excluded due to equipment malfunction, and two other participants were excluded for not completing one or more tasks ($N = 41$; $M_{Age} = 10.27$, $SD_{Age} = 1.21$; 29 male). The age range of eight to twelve was selected to directly compare our findings to prior literature and because proactive control begins to emerge in children aged six years old and continues to develop throughout childhood (Chatham et al., 2009; Chevalier, Martin, Curran, & Munakata, 2015). Informed consent was obtained from parents/guardians, and verbal and written informed assent was obtained from the child prior to participating.

Materials

Modified Two-Step Task. A modified version of the Decker et al. (2016) two-step task that incentivizes model-based strategies was administered to children aged eight to twelve years to assess the utilization of model-based and model-free strategies (Kool et al., 2016). On each trial, participants were presented randomly with one of two pairs of spaceships and then given 2000 milliseconds to choose the left spaceship using the “F” key or the right spaceship using the “J” key. The spaceships changed position randomly on each trial to control for potential right or left side preference. Each spaceship always led to the same one of two planets, a red planet or a purple planet (Figure 1). After the selection, the chosen spaceship was highlighted with a box, and the corresponding planet appeared with an alien on it. The participant then had to press the spacebar to get the alien to ‘mine’ for “space treasure”. The alien presented between one and five pieces of “space treasure”, “antimatter”, or nothing. Each piece of space treasure resulted in one point and each piece of antimatter resulted in the deduction of one point from the total score that was displayed after each trial. The payoffs of each mine changed according to a Gaussian random walk ($\sigma = 2$) with reflecting bounds at -5 and +5. Each participant completed five practice trials with no time limit and 125 test trials with the 2000 millisecond time limit.

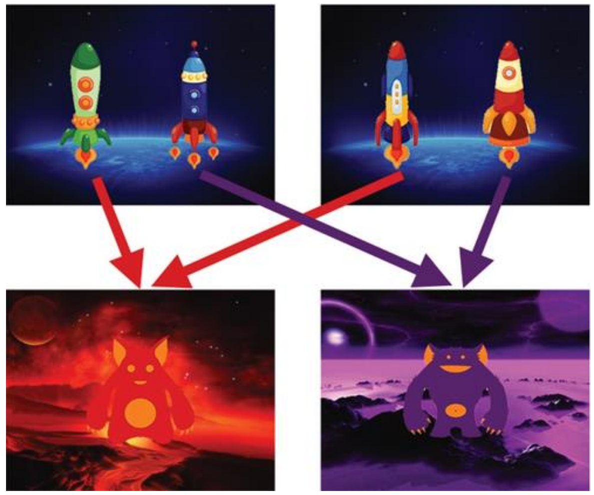


Fig. 1: Illustration of the deterministic spaceship transitions to each planet in the modified spaceship task (adapted from Kool et al., 2016).

Before the 125 test trials began, participants were trained extensively on the different aspects of the game. First, the participant learned about the values of space treasure and antimatter and were instructed to earn as much space treasure as possible. They learned that the mines change slowly over time by asking one alien to mine for them and observing how their mine started out good (e.g., with more space treasure) and then became worse over time. They then asked another alien to mine for them and observed the opposite. Participants were also trained on the deterministic transition of the spaceships and planets by first being told which spaceships went to each planet and then by identifying which spaceship went to each planet until they completed it with at least 60% accuracy. The colors of the planets and aliens in the training portion were different from the test trials.

The modified task differed from the original two-step decision-making task used by Decker et al. (2016) as follows: initial choices (which spaceship leads to which alien) were deterministic instead of probabilistic, two different first-stage choices could be presented, the subsequent choice between two aliens was removed by only presenting one alien, reward probabilities changed more quickly, and reward amounts varied. These changes have been shown to incentivize a model-based strategy for reward trade-off by instantiating a strong correlation between reward and utilizing a model-based strategy in computational models and in adults who completed this modified task (Kool et al., 2016).

Cued-Task Switching Paradigm. Proactive and reactive control tendencies were evaluated using a cued task-switching paradigm called the “Santa Claus Game”, which was composed of two rules and two conditions (Chevalier et al., 2015). The participant was presented with a brown box and then a colored toy they had to sort (Figure 2). Training for each rule occurred prior to completing condition one of the game and again before condition two. The first sorting rule was color. Participants were asked to press the “F” key if the toy was green or the “H” key if the toy was orange. The second sorting rule was shape. They were asked to press the “G” key if the toy was a doll or the “J” key if the toy was a plane. They then had to use both rules together by sorting the toy by either shape or color depending on the cue that was around the toy. If the cue was color swatches, they would sort by color, and if the cue was random shapes, they would sort by shape. After each explanation of the sorting rules, four practice trials occurred until they were completed with 100% accuracy for all three cue types. They then completed a second practice trial with no accuracy requirement before the test trials. Small icons of the shapes and colors were presented at the bottom of the screen during each trial in the same order as the keys to remind participants of the response options (e.g., from left to right the icons were a green patch, a doll, an orange patch, and a plane, which corresponded directly with the keys that generate that response: F, G, H, J, respectively).

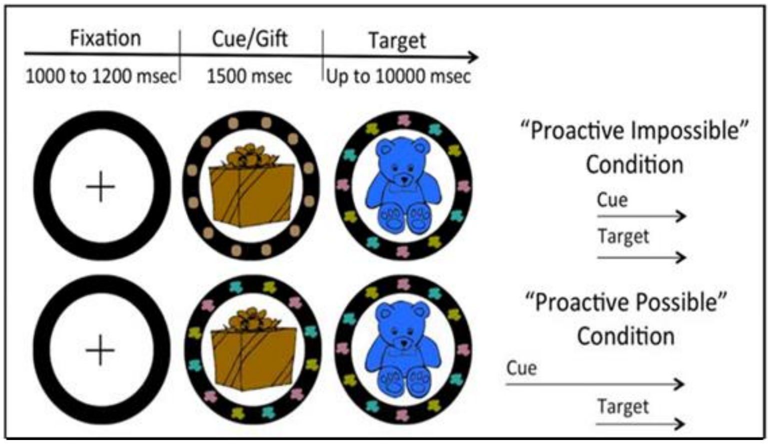


Fig. 2: Diagram of Cued-Task Switching Paradigm demonstrating proactive possible and impossible conditions with color swatches as the cue to sort by color (adapted from Doebel et al., 2017).

In the first condition, the child was shown the sorting cue prior to viewing the stimulus, making proactive control possible. In the second condition, the cue and stimulus were presented simultaneously, making it such that proactive control was impossible and participants were only able to utilize reactive control. The second condition had different shape and color stimuli. The cue (proactive possible) or just the gift (proactive impossible) were presented for 1500 ms before the toy was revealed. The participant then had up to 10,000 ms to respond but were encouraged to respond as quickly and accurately as possible. This task allowed us to assess to what degree children are able to utilize information to proactively prepare a correct response by comparing the mean response times of the two conditions. Those who are engaging in proactive control should have a faster proactive possible response time on average (Doebel, Barker, Chevalier, Michaelson, Fisher, & Munakata, 2017).

Need for Cognition Scale (NfC). A child-adapted version of the short-form NfC Scale was included to measure participants' preference to partake in and enjoy complex thought (Sadowski, 1993). The short-form version contains 18 items (Cacioppo et al., 1996) and is adapted from Chevalier (2017). Each item was read to the participant, which they then rated for personal accuracy using a thumb scale of one to five with five (thumb up) being “describes me perfectly”, four (thumb between up and middle) being “mostly describes me”, three (thumb middle) being “kind of describes me”, two (thumb between middle and down) being “describes me a little” and one (thumb down) being “doesn’t describe me at all.” The items were then summed after reverse-coding half of the responses. The modified short-form NfC was used to explore whether children with a greater need for cognition exhibit the more effortful learning strategy.

Procedure

Participants were recruited using a database they voluntarily entered into based on prior interest in participating in child development studies. Participants were taken into a room in the Cognitive Development Center that had a table, computer screen, and millisecond-accurate keyboard. Parents were given the choice to either sit in the room with their child or stay in the waiting area until their child was finished. Participants were asked to take a seat in front of the computer and told that they would play two computer games and then answer a few questions. A trained experimenter sat to the right of the child with their own keyboard and mouse. The spaceship task and Santa Claus Game were administered consecutively on the same computer with the experimenter reading instructions for both tasks aloud to the participant. Immediately after the two tasks were completed, the short NfC was administered to participants using paper and pen. The experimenter instructed the participant on how to use the thumb-scale and then read each statement to the participant, marking the correlating score (one-five) next to each statement. Finally, the participant received a small token regardless of performance, and the parent was compensated $5 for travel. The parent also received a one-page debrief of the study.

Analysis

This project was preregistered with the Open Science Framework (https://osf.io/wheqd/), and analyses were conducted as proposed unless otherwise noted. Additional analyses will be described as exploratory. Response times during the Cued Task-Switching Task faster than 200 ms and slower than 10,000 ms were excluded (Chevalier et al., 2015), as well as responses times on incorrect trials. A generalized logistic mixed-effects regression analysis was performed using the lme4 package to examine relative use of model-based and model-free strategies in children (Decker et al., 2016). All analyses were performed with the open-source R software (https://www.rstudio.com/).

Results

Participants' learning strategies were assessed (model-based, model-free, or both) through a generalized logistic mixed-effects regression analysis detailed in Kool et al. (2016), mirroring prior analyses in Decker et al. (2016). Prior reward, prior first-stage spaceship pair, and prior reward differences between second-stage states, as well as all interactions, were used to predict staying behaviors, subsequent returns to the same second-stage state (i.e., same planet). In this analysis, prior reward was the parameter used to detect model-based strategies because if someone was model-based, they should have been able to return to the previously rewarding planet regardless of the other parameters. The prior reward-by-prior first-stage spaceship pair interaction was used to detect model-free strategies because if someone was model-free they would choose the same spaceship that last rewarded them so long as they are presented with the same pair of spaceships. Including the reward difference between second-stage states on the previous trial controls for the varying levels and drift of potential reward between second-stage states.

Contrary to previous literature, children engaged in both model-based and model-free learning strategies in a pattern similar to what has been previously observed in adults (Kool et al., 2016). Figure 3a depicts the learning strategies we observed in children, and figure 3b depicts the previously observed pattern in adults (Kool et al., 2016). Prior reward positively correlated with second-stage staying behaviors (model-based; reward estimate = 0.056, p = .044). As reward increased, participants were more likely to exhibit stay behaviors regardless of whether they were presented with the same pair of spaceships (model-based). The prior reward-by-prior first-stage spaceship pair interaction positively correlated with second-stage staying behaviors (model-free; reward estimate = 0.287, p < .001). As reward increased, participants were more likely to exhibit staying behaviors, especially if they were presented with the same pair of spaceships (model-free), suggesting that children remained primarily model-free (Decker et al., 2016). However, these results suggest that children have the ability to be model-based. The full logistic mixed-effects regression results are shown in Table 1.

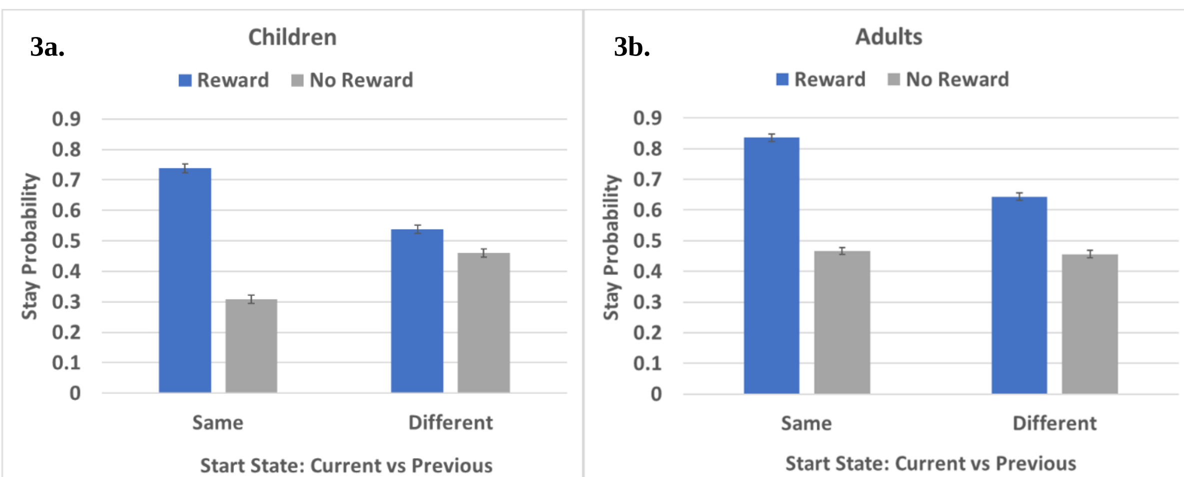


Fig. 3: Children (3a, current study) and adults (3b, Kool et al., 2016) both demonstrated model-based and model-free strategies in the modified spaceship task. The y-axis is the probability of the participants returning to the same planet. The x-axis is if the participants were presented with the same or different pair of spaceships as the previous state. Blue bars indicate that the participants were rewarded on the previous trial and grey bars indicate that they were either punished or not rewarded on the previous trial.

Table 1

Summary of Generalized Logistic Mixed-Effects Regression on Stay Behaviors

Parameter	Estimate	Std. Error	Z value	Pr (>|z|)
(Intercept)	0.009214	0.052863	0.174	0.8618
Prior Reward	0.056053	0.027864	2.012	0.0443 *
Prior First-Stage Spaceship Pair	-0.015583	0.079004	-0.197	0.8436
Reward Difference Between Second-Stage States	0.031940	0.016448	1.942	0.0522
Prior Reward:Prior First-Stage Spaceship Pair	0.286746	0.040630	7.058	1.69e-12 ***
Prior Reward:Reward Difference Between Second-Stage States	-0.006909	0.004358	-1.585	0.1129
Prior First-Stage Spaceship Pair:Reward Difference Between Second-Stage States	0.009710	0.025528	0.380	0.7037
Prior Reward: Prior First-Stage Spaceship Pair:Reward Difference Between Second-Stage States	-0.011974	0.006800	-1.761	0.0783

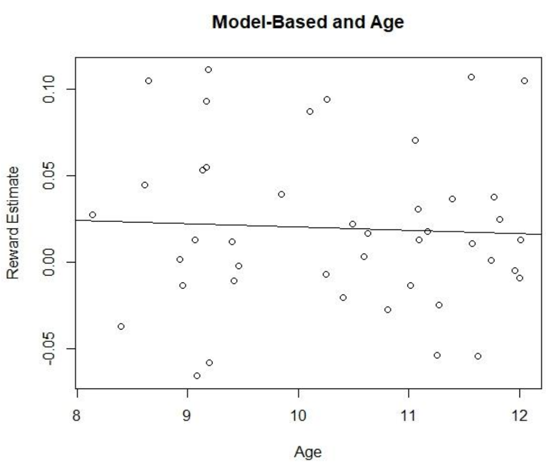
Note: * p < .05, ** p <.01, *** p < .001

Table 1: Summary of logistic mixed-effects regression on stay behaviors in the modified spaceship task. Prior reward denotes model-based behavior and prior reward:prior first-stage spaceship pair denotes model-free behavior.

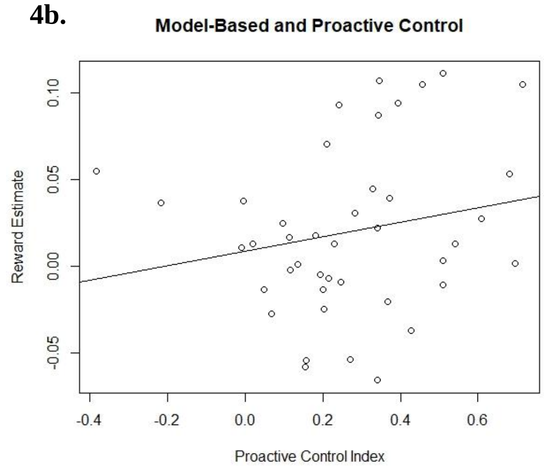
After determining that model-based strategies were used in the modified task, we assessed whether model-based learning increased with age. For the following individual differences analyses, we extracted each participant's reward parameter estimate, representing model-based strategy use, from the logistic regression analyses. Age and model-based learning were not significantly correlated ($r = -0.050$, $p = 0.750$; Figure 4a). Within the age range used, model-based learning did not significantly increase with age.

We then assessed whether proactive control tendencies were related to model-based strategy use. A proactive control index was calculated for each participant by subtracting their median reaction time on the proactive possible condition from their median reaction time on the proactive impossible condition. A multiple linear regression was used to predict reward estimates based on the proactive control index added with median response times on the proactive impossible condition. Median response times for the proactive impossible condition were included to control for slower median response times allowing for faster response time differences on the proactive possible condition, biasing the proactive index. Proactive control was marginally significantly correlated with model-based strategies (r = 0.105, p = 0.060; Figure 4b). These results suggest that proactive control may be an underlying factor in model-based learning.

Finally, we assessed whether need for cognition correlated with model-based learning. NfC Scale scores were not significantly correlated with the reward parameter estimate (r = -0.095, p = 0.549; Figure 4c). A need for cognition did not generalize to model-based strategies in the modified spaceship task.



4b.



4c.

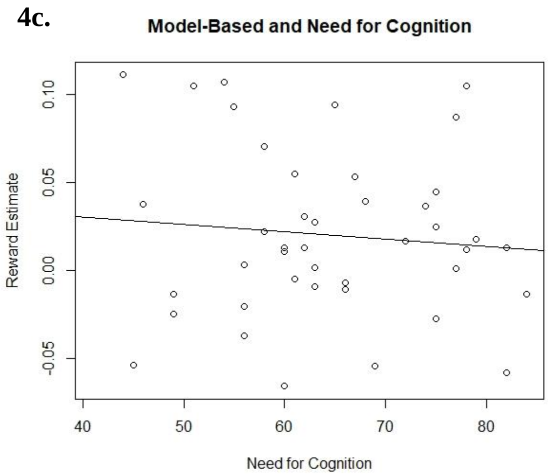


Figure 4: Reward estimate denotes use of model-based strategies. (a) Nonsignificant correlation between model-based strategies and age (r = -0.050, p = 0.750). (b) Marginally significant correlation between model-based strategies and proactive control (r = 0.105, p = 0.060). (c) Nonsignificant correlation between model-based strategies and need for cognition (r = -0.095, p = 0.549).

Discussion

These results support our first prediction that children would utilize model-based strategies when the more effortful strategy incentivized with reward. The larger effect and greater significance of the model-free parameter suggests children were naturally more model-free but did have the ability to utilize model-based strategies. Children may be operating under an effort-reward trade-off. Due to the model-based strategy requiring more effort than the model-free strategy, children may be adapting the amount of effort they put forth based on the amount of reward they will get in return. In the original spaceship task, the more effortful strategy did not provide greater reward. So, children did not utilize a model-based strategy at all (Decker et al., 2016; Potter et al., 2016). However, the modified spaceship task incentivized the model-based strategy a small amount, resulting in children demonstrating a small amount of the model-based learning strategy.

In everyday life, people of all ages consider the effort-reward trade off in their decision-making, but the burden of effort could be more salient for children than it is for adolescents and adults. Therefore, when children are asked to perform more effortful tasks, they may need to be offered more incentive than an adolescent or adult to complete it. Despite extensive changes to the original spaceship task, the increase in incentive for using a model-based strategy was still small. The predicted reward rate for the model-based parameter was .52 in the original task and was .58 in the modified task, which is only an increase in reward of .06 (Kool et al., 2016). Yet, the small increase in the incentive in the modified spaceship task was enough for children to engage in the more effortful, model-based strategy. We suggest future studies implement a greater increase in incentive, which will potentially result in a greater increase in the use of the more effortful, model-based strategy.

Our results did not support our prediction that age and model-based learning are significantly correlated. The nonsignificant correlation could have been due to the selected age range of 8-12 years old. This relatively limited age range could have contributed to the small variation in reward estimate from -0.1 to 0.1, which may not have been large enough to detect a correlation. The original spaceship task was administered to participants aged 8 to 25 years old, which resulted in a significant positive correlation between the model-based parameter and age. The variation in the model-based parameter for this larger age group was between -0.24 and 0.8 (Decker et al., 2016). For the modified spaceship task, an increase in age range to include younger children in future studies may lead to a significant correlation between these two parameters within the child population.

There was a marginally significant correlation between model-based learning and proactive control. A factor that could have affected the significance is the limited age range and subsequent lack of variance. Due to the 8-12-year-old age range that we selected based on prior literature, the amount of variance between subjects in the cued-task switching paradigm and modified spaceship task was not as large as it could have been. Most, if not all, 8-year-olds utilize proactive control, which limited the variation possible in the cued-task switching paradigm (Chatham, Frank, & Munakata, 2009). In Figure 4b, all but two participants fell above zero on the proactive control index meaning that most participants utilized some level of proactive control. Future research should either increase the age range to allow for more variance between proactive control and learning strategy or decrease the age range to 6-7-year-olds and simplify the tasks. Using a younger, more precise age range would be beneficial because there is more variance in proactive control tendencies within 6-7-year-olds than within 8-12-year-olds (Chevalier, Martis, Curran, & Munakata, 2015). Additionally, follow-up studies with more participants may be needed to achieve sufficient power to understand this relationship further.

Lastly, our results did not support our exploratory prediction that a need for cognition would be significantly correlated with the use of model-based strategies. The nonsignificant correlation implies that a general desire to engage in more effortful and complex thought does not directly transcribe to engaging in model-based learning strategies, even though they are the more effortful strategy. However, the Need for Cognition Scale has not been validated for use in children because it has been used in a small number of studies. Further, some of the language used in the questionnaire was not easily understood by our age range. The lack of validation and potential lack of understanding could mean that the NfC does not accurately measure one's desire to engage in complex, effortful thought within the tested age group.

Overall our findings demonstrate that, contrary to previous literature, children do have the ability to utilize model-based decision-making strategies. This finding suggests that children are not biased towards pure exploration of the environment without the ability to implement it into decision-making as previously proposed (Decker et al., 2016). Additionally, children used a mixture of model-based and model-free strategies similar to adolescents and adults, which brings into question the claims that model-based learning has a strong developmental component (Decker et al., 2016). To further assess the developmental aspect, we suggest administering the modified spaceship task to an even wider and younger age range. It may also be beneficial to include other developmental measures, such as brain imaging of the regions associated with model-free and model-based learning, including the ventromedial prefrontal cortex and caudate and putaminal regions (Voon, Reiter, Sebold, & Groman, 2017).

Further, model-based and model-free decision-making processes have real-world implications. Those who are more model-based make decisions based off cognitive models of the world around them, whereas those who are more model-free are making decisions purely based off previous reward. Due to the reflexive, habitual nature of the model-free strategy it would be interesting to explore whether inhibition and delayed gratification play a role in which decision-making strategy an individual is more likely to use. The correlation between inhibition and learning strategy could be assessed by using a delayed gratification task, such as the marshmallow task, and the modified spaceship task. We predict that those with greater inhibition would use more model-based than model-free strategies.

Acknowledgements

I wish to express my sincere gratitude to Professor Yuko Munakata and Jesse Niebaum for their assistance and guidance in conducting this study. I would also like to thank Professor Iskra Fileva, and Professor Eliana Colunga for their time and support throughout this process. I am also very grateful for Tenzin Dorje’s assistance with scheduling, Grace Dostart’s assistance with revisions, and all other Cognitive Development Center members for their feedback. Finally, I would like to thank all of the families that participated in our study.

References

Akam T, Costa R, Dayan P (2015). Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task. *PLoS Computational Biology,* 11(12). doi:10.1371/journal.pcbi.1004648

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16* (2), 106–113. doi:10.1016/j.tics.2011.12.010

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological bulletin*, *119*(2), 197.

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, *106* (14), 5529-5533. doi:10.1073/pnas.0810002106

Chevalier, N. (2017). Willing to think hard? The subjective value of cognitive effort in children. *Child development*.

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: *The case of reactive and proactive control. Journal of Cognitive Neuroscience, 27*(6), 1125-1136. doi:10.1162/jocn_a_00782

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204-1215. doi:10.1016/j.neuron.2011.02.027

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological Science, 27*(6), 848-858. doi:10.1177/0956797616639301

Doebel, S., Barker, J. E., Chevalier, N., Michaelson, L. E., Fisher, A. V., & Munakata, Y. (2017). Getting ready to use control: Advances in the measurement of young children’s use of proactive control. *PloS One, 12*(4), doi:10.1371/journal.pone.0175072

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLoS Computational Biology*, *12*(8), e1005090.

Lucenet, J., & Blaye, A. (2014). Age-related changes in the temporal dynamics of executive control: A study in 5-and 6-year-old children. *Frontiers in psychology*, *5*, 831.

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science, 21*(2), 71-77. doi:10.1177/0963721412436807

Potter, T. C., Bryce, N. V., & Hartley, C. A. (2016). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, *25,* 272-280. doi:10.1016/j.dcn.2016.10.005

Sadowski, C. J. (1993). An Examination of the Short Need for Cognition Scale. *Journal of Psychology*, *127*(4), 451.