

Methodology and Research Practice

# Moving Developmental Research Online: Comparing In-Lab and Web-Based Studies of Model-Based Reinforcement Learning

Kate Nussenbaum<sup>1</sup>  Maximilian Scheuplein<sup>1</sup>, Camille V. Phaneuf<sup>1</sup>, Michael D. Evans<sup>1</sup>, Catherine A. Hartley<sup>1</sup>  <sup>a</sup>

<sup>1</sup> New York University, NY, US

Keywords: replication, decision-making, online testing, reinforcement learning, cognitive development

<https://doi.org/10.1525/collabra.17213>

---

## Collabra: Psychology

Vol. 6, Issue 1, 2020

---

For years, adult psychological research has benefitted from web-based data collection. There is growing interest in harnessing this approach to facilitate data collection from children and adolescents to address foundational questions about cognitive development. To date, however, few studies have directly tested whether findings from in-lab developmental psychology tasks can be replicated online, particularly in the domain of value-based learning and decision-making. To address this question, we set up a pipeline for online data collection with children, adolescents, and adults, and conducted a replication of Decker et al. (2016). The original in-lab study employed a sequential decision-making paradigm to examine shifts in value-learning strategies from childhood to adulthood. Here, we used the same paradigm in a sample of 151 children ( $N = 50$ ; ages 8 - 12 years), adolescents ( $N = 50$ ; ages 13 - 17 years), and adults ( $N = 51$ ; ages 18 - 25 years) and replicated the main finding that the use of a “model-based” learning strategy increases with age. In addition, we adapted a new index of abstract reasoning (MaRs-IB; Chierchia et al. 2019) for use online, and replicated a key result from Potter et al. (2017), which found that abstract reasoning ability mediated the relation between age and model-based learning. Our re-analyses of two previous in-lab datasets alongside our analysis of our online dataset revealed few qualitative differences across task administrations. These findings suggest that with appropriate precautions, researchers can effectively examine developmental differences in learning computations through unmoderated, online experiments.

### Introduction

For years, psychological research on value-based learning and decision-making has benefited from the large and convenient samples made possible through running experiments online (Buhrmester et al., 2016). The proliferation of coding tools, hosting platforms, and participant recruitment mechanisms, has made launching adult studies fairly straightforward (Anwyl-Irvine et al., 2020; de Leeuw, 2015; de Leeuw et al., 2014; Sauter et al., 2020). More critically, with the appropriate precautions, data collected online has been shown to be of comparable quality to data collected in the lab (Crump et al., 2013). However, to date, the vast majority of online psychological research has been conducted with adults. It is thus unclear whether it is possible to collect equally high-quality decision-making data from children and adolescents via remote, browser-based experiments.

Several research groups have recently demonstrated the feasibility of collecting data online from both infants and young children (Rhodes et al., 2020; Scott et al., 2017), without a “live” experimenter present via video chat (e.g., Chuey et al., 2020). However, most of these approaches have involved collecting responses to short animations, in tasks that typically take between 10 and 20 minutes. These experiments are well-suited to address a wide range of questions,

including, for example, those centered on children’s beliefs about others’ expertise or characteristics (Chuey et al., 2020; Leshin et al., 2020), or their ability to map novel verbs onto relevant actions (Scott et al., 2017). However, the computational characterization of children’s value-based learning and decision-making strategies often requires that participants make many repeated choices, in tasks that last between 30 minutes to 1 hour (Decker et al., 2015, 2016; van den Bos et al., 2012). While lengthier decision-making tasks have been used extensively in online studies of adults (e.g., Coenen et al., 2015; Dorfman et al., 2019; Garrett & Daw, 2020; Kool et al., 2017), to the best of our knowledge, they have not been used in online studies of children as young as 8 years old.

Here, our goal was to develop a pipeline to efficiently collect decision-making data remotely from a large number of child, adolescent, and adult participants (ages 8 - 25 years), and compare the quality of the data collected online to the quality of data previously collected in the lab. To accomplish these aims, we conducted an online replication of Decker et al. (2016). This prior study, conducted in our lab, examined age-related changes in decision-making strategies. Specifically, the study adapted a task (the “two-step task”) originally developed for use with adults (Daw et al., 2011). Briefly, the task requires participants to make many repeated decisions with “two steps” to gain as much reward as possible. On each trial, participants first select one of two

---

a Corresponding author: [cate@nyu.edu](mailto:cate@nyu.edu)

spaceships in which to travel. The spaceships both travel to the same two planets, but each has a different preferred planet which it visits 70% of the time, and a different non-preferred planet which it visits 30% of the time. Each planet is inhabited by two alien treasure-miners, who share treasure based on independent, slowly drifting reward probabilities. After reaching a planet, participants select one of the two aliens to ask for space treasure, and are either rewarded with a piece of treasure, or given nothing. The “two-step” design enables measurement of the extent to which participants use a mental model of the task’s transition structure to guide action selection. Participants who rely on a more habitual, “model-free” learning strategy tend to repeat first-stage spaceship selections that were recently rewarded and avoid those that did not lead to reward, regardless of whether they traveled to the preferred or non-preferred planet. Participants who rely on a “model-based” learning strategy use a mental model of the task structure to guide action selection (Daw et al., 2011), such that they tend to repeat first-stage spaceship selections that were recently rewarded and avoid those that did not lead to reward *only if* the spaceship traveled to the preferred planet and is therefore likely to go there again.

Because the two-step task enables the disentanglement of these two forms of value-based decision-making, it is well-suited to address questions about developmental changes in the use of complex, mental models to guide action selection. By employing this task in a sample of participants ages 8 – 25 years, Decker et al. (2016) found that with increasing age, individuals demonstrated greater use of a model-based learning strategy, which suggests that children did not recruit their knowledge of the structure of their environment to guide their decisions to the same extent as adults. A follow-up study (Potter et al., 2017) probed the cognitive processes that may underlie this developmental shift and found that age-related increases in model-based learning in the two-step task were mediated by developmental improvements in fluid reasoning. To examine this relation, we adapted a relatively novel measure of fluid reasoning — the matrix reasoning item bank (MaRs-IB; Chierchia et al., 2019) — for use online.

We chose to try to replicate the two-step task findings online for several reasons. First, we have already replicated the primary effect of interest — the greater use of model-based learning with increasing age — in a separate, in-lab, developmental sample (Potter et al., 2017), suggesting that the effect itself can be replicated under certain conditions. Second, other research groups have used variants of this task and sought to extend our prior developmental findings (Bolenz & Eppinger, 2020; Smid et al., 2020), suggesting that a second, robust replication attempt is useful for the field as a whole. In addition, there is an extensive adult literature involving variants of the two-step task that have been administered online, which has addressed questions about how individuals arbitrate between different learning systems when seeking reward (Kool et al., 2017), whether the use of model-based learning relates to habit formation (Gillan et al., 2015), and how biases in learning and decision processes may predict psychiatric symptomatology (Gillan et al., 2016). If these types of task variants can be effectively administered online in children and adolescents, future studies can efficiently leverage them to characterize the ontogeny of foundational learning and decision-making mechanisms. Finally, the two-step task shares many features with those administered in our typical in-lab studies, in that it seeks to characterize value-based learning and decision-making strategies through both computational mod-

el-fitting and simpler, regression analyses, by having participants spend about 45 minutes making repeated choices between options. Thus, we expect our assessment of data quality from this first online study to generalize to similar studies we aim to conduct online in the future.

There are many potential pitfalls when collecting online data from children and adolescents, particularly with long tasks that require sustained focus. Prior to launching our online study, we identified four areas of concern, the impact of which we hoped to assess and potentially mitigate. Here, we lay out our concerns and discuss our strategies for reducing their impact *a priori*. We address whether these concerns were well-founded in our discussion of our results. First, we were concerned that children and adolescents would get bored and either stop attending to the task (as evidenced through random key pressing, clicking in and out of the browser window, etc.), or quit the task without completing it. To incentivize attention within the task, we told participants they would be paid a bonus based on their performance, as we did in the in-lab version of the study. In addition, to monitor participant attention to the task, beyond recording participants’ responses and reaction times on every trial, we also logged all browser interactions to determine when participants clicked out of the main task window. To incentivize task completion, we instructed participants that they would get paid only if they finished the entire experiment. Second, though our task already had extensive, interactive instructions and a lengthy tutorial, we were concerned that without an experimenter available to go through the instructions and answer questions in person, participants, and particularly young children, might struggle to comprehend how to perform the task. To address this potential concern, we added an audio track that read the instructions aloud to all participants. We also added comprehension questions at the end of the instructions to assess participants’ understanding of the task. To ensure the verbal instructions were audible, we also added a “test” of participants’ audio, in which they had to click on a picture of an animal named aloud prior to the start of the study. Third, we worried about participants encountering technological difficulties while trying to complete the experiment, including attempting to complete the experiment using an incompatible device or browser, or accidentally closing or refreshing their browser window (which would have aborted the task). We provided participants with clear instructions about the technological requirements for the study, and we ensured that the use of an incompatible device or browser would prevent them from starting the experiment at all (as opposed to causing problems mid-way through). To reduce the likelihood of participants accidentally quitting the task, we added a pop-up message that would allow participants to “cancel” any screen refresh or window exit.

Finally, we were concerned that other people might interfere with participants’ completion of the task — parents, for example, might “help” their children make choices, preventing us from accurately assessing children’s decision-making strategies. Other studies have addressed this concern by recording video data from all participants via their webcams (Rhodes et al., 2020), which experimenters can then code for instances of interference. This tactic, however, in addition to requiring experimenter time, imposes greater demands on the hosting server and raises potential privacy concerns. Thus, we simply provided parents with clear instructions (both in our initial recruitment email and on our consent form) that asked them to let their children complete the tasks on their own. By comparing the data we

collected from children online to the data we collected in the lab, we could examine whether there were differences in behavior that could be explained by increased interference from parents at home, and determine whether stronger mitigation strategies are needed in the future.

To determine whether we could successfully replicate our previous, in-lab findings in an online sample of children, adolescents, and adults, we created browser-based versions of our two experimental tasks with jsPsych (de Leeuw, 2015) and hosted them on Pavlovia (<https://pavlovia.org>). We recruited 151 8- to 25-year-old children, adolescents, and adults to complete the experiment. Briefly, we found that we were able to replicate all the key two-step task findings from prior, in-lab studies, including the mediation of the effect of age on model-based learning by fluid reasoning. Further, we did not encounter obvious problems with participant attrition, task comprehensibility, technological difficulties, or, to the best of our knowledge, interference from parents. However, the MaRs-IB fluid reasoning data that we collected look qualitatively different than the data collected in a previous, in-person experiment (Chierchia et al., 2019), suggesting that participants may have approached the task differently when completing it on their own. Despite these qualitative differences, our findings suggest that under some conditions and with appropriate precautions, it is possible to acquire high-quality decision-making data from children, adolescents, and adults via unmoderated, online studies.

## Methods

### Participants

One-hundred and fifty-one participants ( $N = 50$  children (8 - 12 years, mean age = 10.51 years, 25 females);  $N = 50$  adolescents (13 - 17 years, mean age = 15.58 years, 25 females);  $N = 51$  adults (18 - 25 years, mean age = 21.83 years, 26 females) completed the study. An additional 15 participants (9 children, 3 adolescents, and 3 adults) filled out the consent form but quit the experiment prior to completing the two-step task. In our in-lab studies of decision-making with participants in this age range, we have typically included 50 - 90 participants, which has given us adequate power to reveal age-related change in decision-making strategies (Cohen et al., in press; Decker et al., 2016; Nussenbaum et al., 2020; Potter et al., 2017). Here, because we assumed there might be more “noise” in our online sample, and due to the relative ease of collecting data from more participants, we *a priori* decided to recruit 50 participants in each age group for a total target  $N$  of 150.

**Recruitment and consenting.** Participants were primarily recruited from Facebook ads ( $n = 53$ ) and our database for in-lab studies ( $n = 76$ ), for which we have previously solicited sign-ups at local New York City science fairs and events. Some participants were also recruited via word-of-mouth ( $n = 8$ ), our lab website ( $n = 4$ ), and from [ChildrenHelping-Science.com](https://www.childrenhelping-science.com) ( $n = 2$ ) a website that lists online psychology research studies across different labs in which children can participate. All potential participants first registered for our online database, which asks adult participants and parents of child participants to report demographic information, including their age, race, ethnicity, household income, any visual impairments, and any history of psychiatric and learning disorders. For this study, we implemented the same inclusion criteria as Decker et al. (2016) and recruited participants between the ages of 8 and 25 with normal or corrected-to-normal vision and no reported history of psychiatric or learning disorders. We include more information about

our recruitment process in the supplement.

Participants were informed they would be compensated with a \$15 Amazon gift card if they completed the full study. They were also told they could receive a bonus based on their performance in the task; in reality, all participants received a \$5 bonus such that they were compensated with a \$20 Amazon gift card. Further, participants were informed that this study needed to be completed on a laptop or desktop computer (as opposed to a tablet or smartphone) with Chrome, Safari, or Firefox. If participants tried to launch the task on a tablet or smartphone, or with an incompatible browser, they would be unable to proceed past the first instructions page.

For participants under the age of 18, the Qualtrics consent form included clearly labeled sections for parents and children. If participants (and their parents) gave consent to participate and reported that their computers met the technological requirements, the consent form directed them to a link to launch the first task (the two-step task). If participants did not consent to the terms in the Qualtrics form, the consent form directed them to an end page with our lab’s contact information. We include a full description of how many participants we emailed to achieve our target of 150 participants in the supplement.

**Participant demographics.** One advantage of online testing is that it offers the opportunity to include participants who may be unable to visit the lab in person. This may promote the inclusion of a more diverse and representative group of participants (Leshin et al., 2020; Rhodes et al., 2020). However, there is also concern that online testing may exacerbate existing barriers to research inclusion (Lourenco & Tasimi, 2020). We examined distributions of participant race and ethnicity within our online sample and compared them to a recent in-lab study (Potter et al., 2017), as well as national and New York City demographics (Figure 1A). In addition, we report annual household income ranges for child, adolescent, and adult (Figure 1B) participants in our sample.

### Tasks

Participants completed two experimental tasks, each of which was hosted as its own Pavlovia project. Tasks were coded in jsPsych (de Leeuw, 2015) and are publicly available, along with all de-identified data files, and analysis code, on the Open Science Framework: <https://osf.io/we89v/>.

**Two-step task.** We used the version of the two-step task (Daw et al., 2011) designed to dissociate “model-free” and “model-based” learning, which Decker et al. (2016) originally adapted for use in developmental populations. Participants’ goal was to collect as much “space treasure” as they could, by traveling to different planets inhabited by different alien treasure-miners. On each trial, participants first chose between two different spaceships (first-stage choice) by pressing 1 or 0 on the keyboard to select the left or right option. Each spaceship had a 70% probability of traveling to one planet (e.g., the red planet), and a 30% probability of traveling to the other planet (e.g., the purple planet). On each planet, participants could ask one of two aliens for treasure (second-stage choice). The alien would either give them treasure or nothing (Fig. 2), depending on slowly drifting reward probabilities. The side of the screen where the spaceships and aliens appeared stayed constant throughout the task and was randomized for each participant. The planet that each spaceship visited most frequently was also randomized for each participant. As in Decker

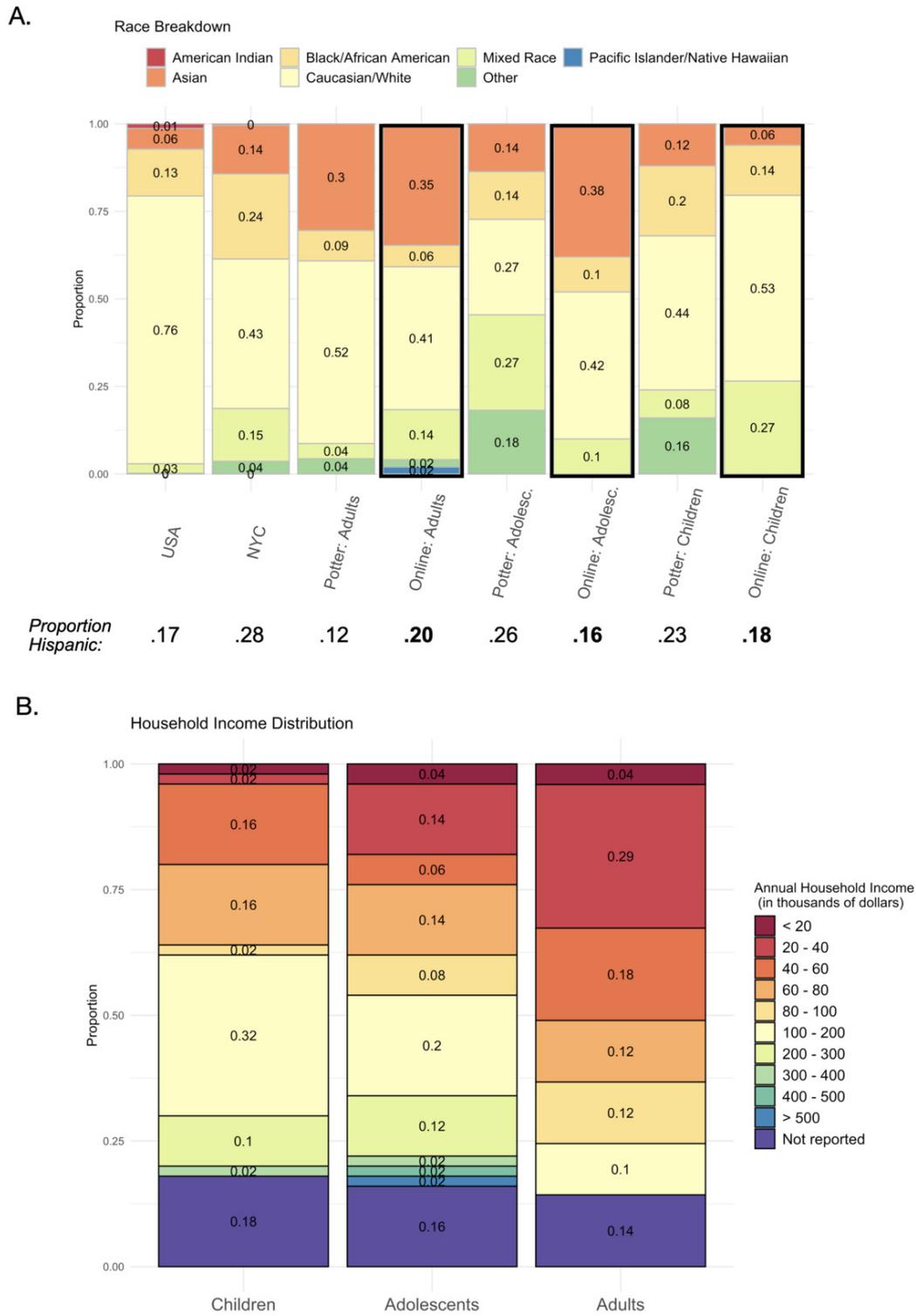


Figure 1. (A) Race and ethnicity of participants our online study (outlined in black) and those in a previous in-lab study (Potter et al., 2017) next to national ('USA') and New York City ('NYC') data for comparison. (B) Annual household income of child, adolescent, and (C) adult participants in our online study. For comparison, the median annual household in NYC from 2014 - 2018 was \$60,762 and nationally in 2018 was \$61,937 (U.S. Census Bureau).

et al. (2016), participants had three seconds to make each choice. Participants completed 200 trials, separated into four blocks of 50 trials. At each break screen, participants were told what proportion of the task they had completed (e.g., "You are halfway done!"). These messages were not included on the break screens in the in-lab version of the

task; we added them to increase motivation to complete the online task, in the absence of a live experimenter who could answer questions from children like, "Am I almost done?"

The task enables measurement of the extent to which participants relied on "model-free" vs. "model-based" learning and decision-making strategies. In this task, a

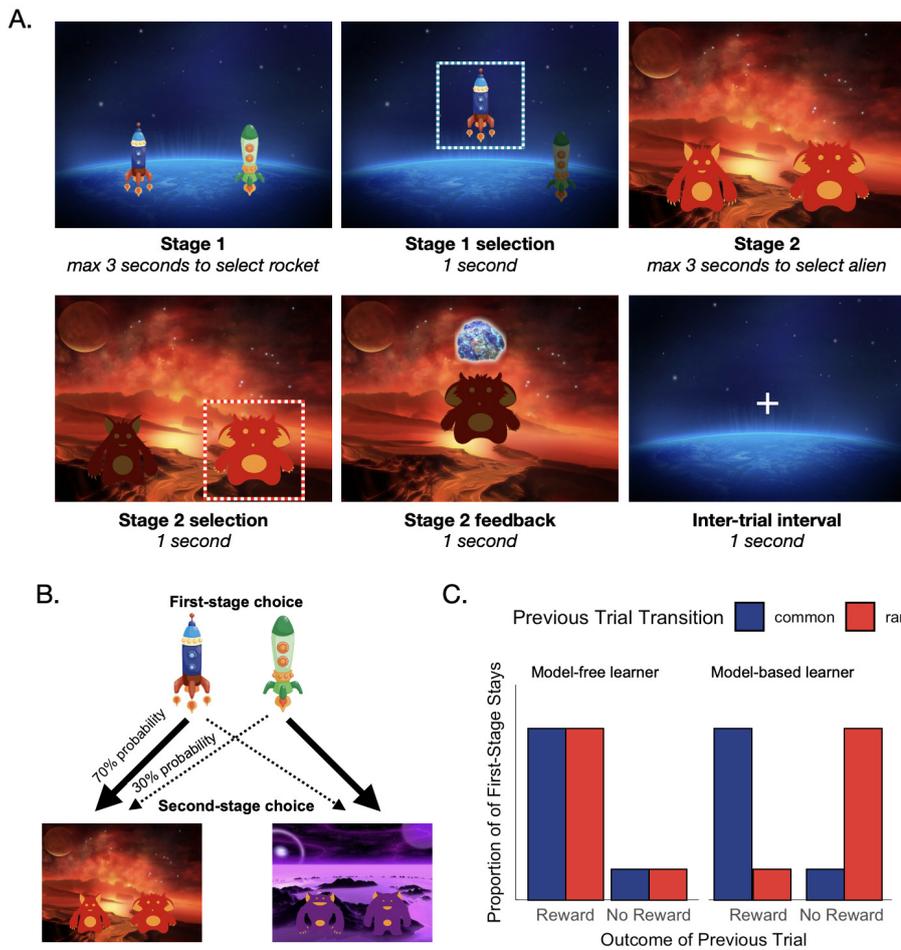


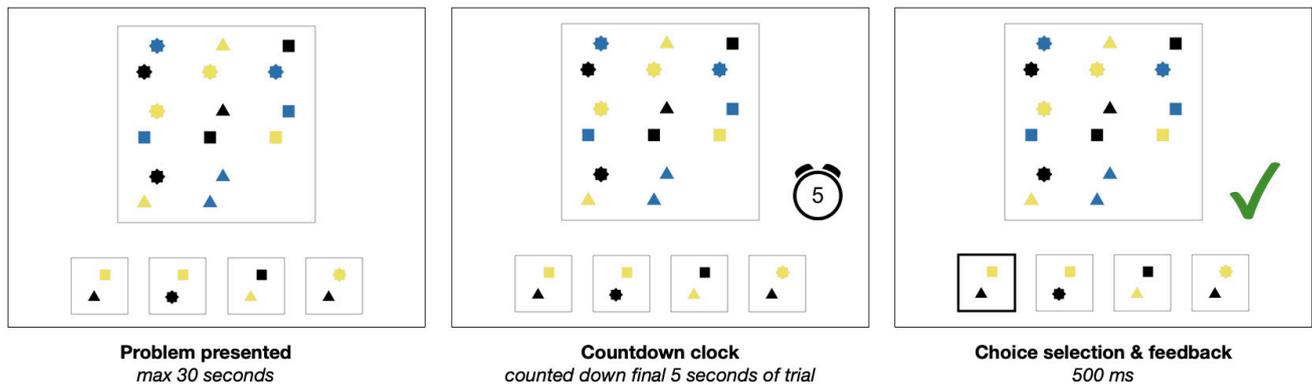
Figure 2. Participants completed an online version of the two-step task used in Decker et al. (2016). Participants completed 200 choice trials. On each trial (A), participants selected one of two spaceships, which took them to one of two planets according to a probabilistic transition structure (e.g., spaceship A went to the red planet on 70% of trials, and the purple planet on 30% of trials) (B). Participants encountered two aliens on each planet, each of which distributed treasure with a probability that slowly drifted throughout the experiment, and had to select which to ask for treasure. The task can dissociate model-free from model-based learning strategies because each strategy predicts a different influence of the previous trial’s reward and transition type on the likelihood that a participant repeats their first-stage choice (C).

model-free learner is likely to repeat rewarded first-stage choices, regardless of the transition they experienced. A model-based learner is more likely to use a mental model of the task structure to guide their decisions, such that they are likely to repeat rewarded first-stage choices after common transitions but switch first-stage choices on trials where they received reward after a rare transition. In other words, if a model-based learner takes the blue spaceship to the red planet and receives reward, then on the next trial, they are likely to choose whichever spaceship goes to the red planet most often.

Prior to beginning the 200 “real” choice trials, participants completed the same, extensive tutorial that was used in Decker et al. (2016). In the tutorial, participants were introduced to the task cover story, and completed a set of practice trials that illustrated how to ask aliens for treasure, the probabilistic nature of the rewards, and the full trial structure (20 full practice trials). The tutorial used different spaceship, alien, and planet stimuli from the main task trials. After the 200 “real” choice trials, participants saw a screen that told them how much treasure they won. They then saw both spaceships again and were asked, “Which

spaceship went mostly to the red planet?” As in the choice trials, participants used the 1 and 0 keys to select the left or right spaceship.

We made several modifications to the task instructions to make them better suited to online administration. First, prior to the first instruction screen, we had participants click a single button that made the experiment window fill their entire screen. Second, we recorded audio to play over all instruction slides, so that all participants heard the instructions read aloud, as they would have in the in-lab study. Prior to the first instruction screen, participants completed two “audio tests” in which they heard an animal named aloud and had to click on the appropriate animal picture to proceed with the experiment. Participants received an error if they clicked incorrectly and had the option to replay the sound as many times as needed. To ensure that participants did not accidentally skip over instruction screens, we required them to both click a small red circle at the bottom of each page to turn it green *and* press a key on the keyboard to advance to the next page. Finally, we also added three true/false comprehension questions to the end of the instructions that re-iterated key information (the probabilistic na-



**Figure 3.** Participants completed an online version of the MaRs-IB task described in Chierchia et al. (2019). On each trial, participants viewed a 3x3 grid of images with the bottom right image missing. They had 30 seconds to select the correct “missing” image from a selection of four options.

ture of the transition structure, the slowly drifting reward probabilities, and the time limit to make each response). Participants received correct or incorrect feedback based on their response to each question, and they heard and saw a written statement reiterating the information. Because our task instructions and tutorial took between 10 and 15 minutes, we did *not* require participants to re-do them if they answered the comprehension questions incorrectly.

For both the instructions and task, we also implemented a pop-up warning that occurred whenever participants tried to exit or refresh a page. The warning asked if they were sure they wanted to leave the page, and they could click “cancel” to return to the task.

**Matrix reasoning item bank (MaRs-IB).** After the two-step task, participants were directed to the MaRs-IB, an open-access task that provides an index of fluid reasoning ability (Chierchia et al., 2019). The task involved a series of matrix reasoning puzzles. On each trial, participants were presented with a 3x3 grid of abstract shapes, with a blank square in the lower right-hand corner (Fig 3). Participants had 30 seconds to select the missing shape from one of four possible answers (the target and three distractors) by clicking on it. After 25 seconds, a 5-second count-down clock counted down the remaining time left in the trial. Upon making their selection, participants saw feedback — either a green check mark for correct responses or a red X for incorrect responses — for 500 ms.

Each puzzle comprised eight abstract shapes that could vary across four features: color, size, position, and shape. The dimensionality of each item ranged from 1 to 8, corresponding to the number of features or feature combinations that changed across items in the matrix. All participants saw the same sequence of puzzles that was administered in the previous study (Chierchia et al., 2019), though we used puzzles of “test form 1” from the color-vision deficient friendly version of the task. The sequence contained a scrambled mix of easy, medium, and hard puzzles. Because Chierchia et al. (2019) did not find differences in participant accuracy based on the algorithm used to generate the distractors, for all items, we used distractor items generated by the “minimal difference” algorithm, in which distractors are slight variations of the target stimulus to prevent pop-out effects. The position of the target and distractor items were randomized on each trial for each participant.

Participants completed either 8 minutes of puzzles or all 80 puzzles, whichever occurred first. In the original administration of the task (Chierchia et al., 2019), if participants

completed all 80 items within 8 minutes, they saw the sequence of puzzles again. However, their responses to the second presentation were not analyzed. Given that their responses to the second presentation of items were not analyzed, we terminated the task after participants viewed all 80 puzzles. Prior to beginning the real trials, participants went through a series of short instructions. As with the two-step task, we modified the original task by adding audio to the instructions. In addition, participants completed three practice trials of “easy” puzzles. Each practice trial was repeated until the participant answered it correctly.

### Previously collected datasets

As mentioned earlier, our lab has previously conducted two experiments in 8 - 25 year-olds using the spaceship version of the two-step task (see Decker et al., 2016; Potter et al., 2017). Each of these studies applied different criteria to determine inclusion in the final, analyzed sample. Here, we include *all* participants from both studies who completed the task. As such, the results from our analyses differ from those reported in the published manuscripts, but all the main findings and interpretations remain the same.

The dataset from Decker et al. (2016) (hereafter referred to as the “Decker dataset”) includes 80 participants (N = 30 children (8 - 12 years old; mean age = 9.87 years; 15 females), N = 28 adolescents (13 - 17 years old, mean age = 15.12 years, 15 females), N = 22 adults (18 - 25 years old; mean age = 21.18 years; 13 females)). Participants completed 200 choice trials (broken into blocks of 50 trials) on a computer in a testing room at Weill Cornell Medical College. An experimenter went through the instructions and tutorial with all participants and remained in the room while they completed the study.

The dataset from Potter et al. (2017) (the “Potter dataset”) includes 74 participants (N = 26 children (8 - 12 years old; mean age = 10.48 years; 14 females), N = 23 adolescents (13 - 17 years old; mean age = 15.27 years; 13 females), N = 25 adults (18 - 25 years old; mean age = 22.07 years; 14 females)). Participants completed 150 choice trials (broken into blocks of 50 trials) while undergoing functional magnetic resonance imaging (fMRI) at Weill Cornell Medical College. Participants could communicate with the experimenter via an intercom system in between task blocks.

**Table 1. Logistic regression coefficients indicating the effects of previous reward, previous transition, and age on first-stage choice repetition.**

Predictors	Decker et al. (2016)			Potter et al. (2017)			Online		
	<i>Est.</i>	<i>CI (95%)</i>	<i>p</i>	<i>Est.</i>	<i>CI (95%)</i>	<i>p</i>	<i>Est.</i>	<i>CI (95%)</i>	<i>p</i>
Intercept	1.02	0.84–1.20	<0.001	0.82	0.63–1.01	<0.001	1.39	1.23–1.55	<0.001
Reward	0.27	0.17–0.36	<0.001	0.24	0.15–0.34	<0.001	0.35	0.29–0.41	<0.001
Transition	0.03	-0.02–0.09	0.216	0.00	-0.07–0.07	0.979	0.00	-0.05–0.05	0.942
Age	0.49	0.31–0.67	<0.001	0.45	0.27–0.64	<0.001	0.42	0.26–0.58	<0.001
Reward x Transition	0.22	0.15–0.30	<0.001	0.28	0.18–0.39	<0.001	0.41	0.33–0.48	<0.001
Reward x Age	0.13	0.03–0.22	0.011	0.04	-0.06–0.14	0.440	0.10	0.04–0.16	0.001
Transition x Age	0.02	-0.03–0.08	0.344	0.05	-0.02–0.13	0.124	0.03	-0.01–0.07	0.187
Reward x Transition x Age	0.16	0.08–0.24	<0.001	0.17	0.07–0.28	0.001	0.18	0.11–0.26	<0.001

## Analysis approach

Though we binned age into groups for data visualization purposes, we treated age as a continuous variable in all analyses. In our analysis of the two-step task data, we excluded the first 9 trials for each participant (Decker et al., 2016; Potter et al., 2017), as well as all trials in which participants failed to make either a first- or second-stage choice within the 3-second time limit. In our analysis of the MaRS-IB data, we excluded trials in which participants made responses in less than 250 ms as well as trials in which participants failed to respond before the 30-second time limit (Chierchia et al., 2019).

All regression analyses were conducted using the “afex” package (Singmann et al., 2020) in R version 3.5.1 (R Core Team, 2018). For logistic mixed-effects regressions, we assessed the significance of fixed effects via likelihood ratio tests. For linear mixed-effects regressions, we used F tests with Satterthwaite approximations for degrees of freedom. Our mediation analysis was conducted with the “mediation” R package (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014), and significance of the mediation effects was assessed via 1,000 bootstrapped samples. Computational model-fitting was conducted in Matlab 2020a (Mathworks Inc, 2020). Full details of the model-fitting procedure are included in the supplement. All analysis code is publicly available on our OSF page: <https://osf.io/we89v/>.

## Results

### Two-step task

**Online data quality.** We extracted four measures of data quality from our online dataset for each participant: their number of “browser interactions,” which include entering or exiting full screen or clicking in and out of the browser window with the task (participants who did the task perfectly should have had one browser interaction), the number of comprehension questions they answered correctly (out of three total), the number of first- and second-stage choices (out of 400) in which they failed to respond before the 3-second time limit (missed trials), and the number of choice trials in which they made a response faster than 150 ms (fast RTs). In general, the majority of participants across

age groups had few browser interactions (Medians: children = 1.5, adolescents = 1, adults = 1), responded to the comprehension questions correctly (Median for all age groups = 3), missed few responses (Medians: children = 5, adolescents = 1, adults = 1), and generally took more than 150 ms to make each choice (Median number of fast RTs: children = 19; adolescents = 9; adults = 13) (Figure 4). We have included detailed information about the number of participants in each age group who met particular data quality thresholds in our supplement.

**Learning: Regression analyses.** We examined participants’ use of model-free and model-based learning strategies by running a mixed-effects logistic regression examining the influence of the previous trial’s transition type and reward outcome, as well as participant age, on repeated first-stage choices. If participants used a model-free learning strategy, then they should be more likely to repeat first-stage choices after rewarded trials relative to unrewarded trials. If they used a model-based learning strategy, they should repeat first-stage choices after rewarded trials with common transitions *or* unrewarded trials with rare transitions more than after rewarded trials with rare transitions *or* unrewarded trials with common transitions. This would be reflected in a reward x transition interaction effect. Across all three datasets (Table 1), we observed a significant main effect of reward and a significant reward x transition interaction effect ( $ps < .001$ ; Figure 5), indicating that participants used both learning strategies throughout the task. We also observed a main effect of continuous age ( $ps < .001$ ), such that in all three datasets, older participants were increasingly likely to repeat first-stage choices, regardless of their outcomes. Most relevant to our primary question of interest, we also observed a significant reward x transition x age interaction effect, indicating an age-related increase in model-based learning ( $ps < .002$ ). In the Decker dataset and the online dataset, we further observed an age x reward interaction effect (Decker:  $p = .011$ ; online:  $p = .001$ ), which suggests that model-free learning *also* increased with age. This effect was not significant in the Potter dataset ( $p = .44$ ).

**Knowledge of task structure.** Next, we examined whether participants had explicit knowledge of the transition structure of the task, by examining their responses to the final, explicit question (“Which spaceship went mostly to the red

**Table 2. Mean accuracy (and standard deviation) of transition structure knowledge across age groups and datasets**

Age group	Decker et al. (2016)	Potter et al. (2017)	Online
Children	.75 (.45)	1.00 (0)	.74 (.44)
Adolescents	.82 (.39)	.83 (.39)	.86 (.35)
Adults	.86 (.36)	.84 (.37)	.80 (.40)

**Table 3. Linear regression coefficients indicating the effect of age and transition type on second-stage choice reaction times**

Predictors	Decker et al. (2016)			Potter et al. (2017)			Online		
	Est.	CI (95%)	p	Est.	CI (95%)	p	Est.	CI (95%)	p
Intercept	739	685–792	<0.001	973	927–1020	<0.001	659	629–690	<0.001
Age	-85	-138--32	0.003	-88	-134--41	<0.001	-33	-63--2	0.036
Transition	-40	-52--29	<0.001	-65	-83--47	<0.001	-65	-75--54	<0.001
Age x Transition	-10	-21--1	0.081	-27	-44.--9	0.005	-12	-23--2	0.024

**Table 4. Linear regression coefficients indicating the effect of age and RT difference on model-based learning**

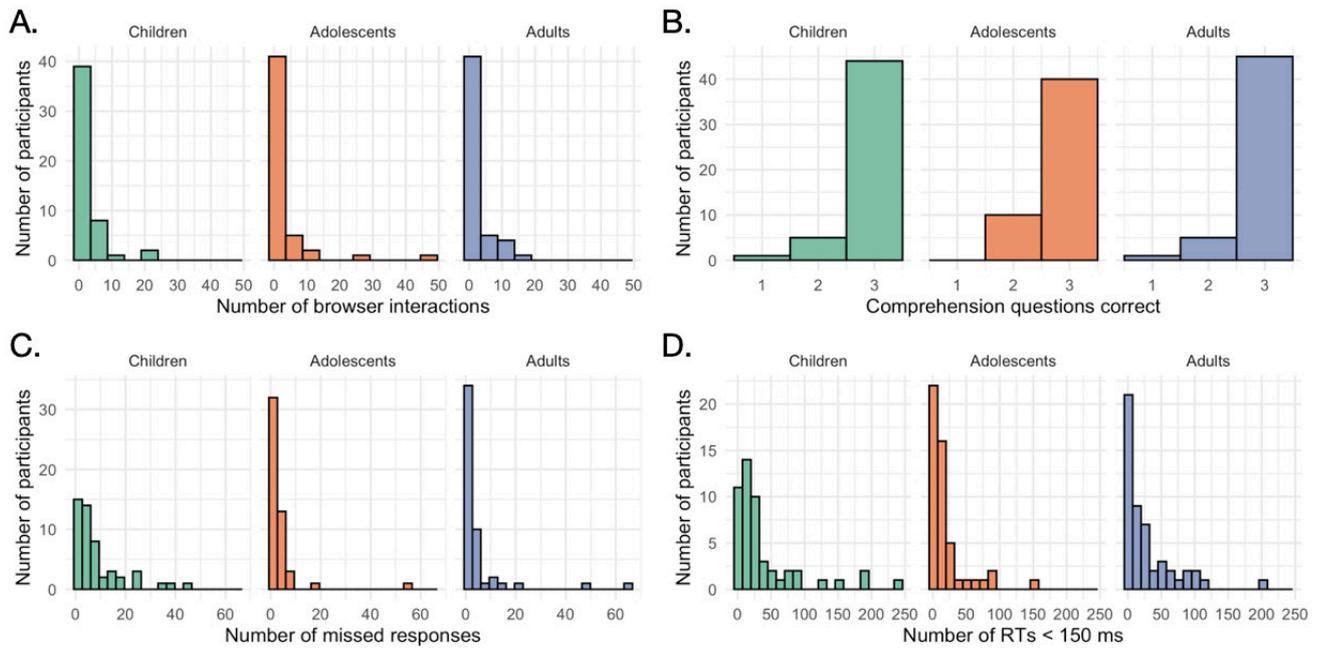
Predictors	Decker et al. (2016)			Potter et al. (2017)			Online		
	Est.	CI (95%)	p	Est.	CI (95%)	p	Est.	CI (95%)	p
Intercept	0.21	.16-.26	<0.001	0.27	.20-.35	<0.001	0.38	.33-.42	<0.001
RT Difference	0.11	.06-.16	<0.001	0.14	.06-.23	0.001	0.16	.11-.20	<0.001
Age	0.11	.06-.16	<0.001	0.10	.02-.18	0.016	0.13	.08-.18	<0.001
RT Difference x Age	0.04	-.00-0.09	0.074	0.00	-.09-.09	0.954	0.05	.00-.10	0.032

planet?”). Across all three datasets, the majority of participants across age groups demonstrated awareness of the transition structure (Table 2). In the Decker and online datasets, there was not a significant effect of participant age on response accuracy ( $p_s > .46$ ). In the Potter dataset, children’s responses were more accurate, such that there was a negative relation between age and accuracy ( $\beta = -.81$ ,  $SE = .41$ ,  $p = .046$ ).

We also examined whether participants’ RTs in selecting a second-stage choice option were influenced by the transition they experienced. If participants had no awareness of the task structure, then we would expect to see no differences in RTs after common vs. rare transitions. However, if they had knowledge of the transition structure, then we would expect them to react more slowly after unexpected — or rare — transitions. And indeed, across all three datasets, participants made second-stage choices more slowly following rare vs. common transitions ( $p_s < .001$ ; Table 3; Figure 6A). Across datasets, we also observed a main effect of age on RTs, such that older participants made faster choices ( $p_s < .037$ ). Finally, in both the Potter and online dataset, we observed an age x transition interaction effect, such that the influence of transition type on RTs increased with increasing age ( $p_s < .025$ ; Figure 6A). This effect was only marginal in the Decker dataset ( $p = .081$ ).

**Relation between knowledge of task structure and learning.** Decker et al. (2016) found that participants’ knowledge of the transition structure of the task — as revealed through their slower RTs after rare transitions — predicted their use of a model-based learning strategy. To test for this effect across datasets, we computed an RT difference score for each participant by subtracting their mean second-stage choice RT following common transitions from their mean RT following rare transitions. We then extracted their reward x transition random slope from a mixed-effects model examining repeated first-stage choice repetition (run without age) and tested how these slopes varied as a function of RT difference scores and age in a linear regression. Across all three datasets, we found that RT difference scores predicted model-based learning ( $p_s < .002$ ; Table 4; Figure 6B). In the online dataset, we further observed an RT difference x age interaction effect, such that the relation between RT difference scores and model-based learning increased with increasing age ( $p = .032$ ).

**Reinforcement learning computational modeling.** While the regression analyses provide insight into participant choice behavior, they only consider the influence of the previous trial on participants’ decisions. To characterize how participants used a longer learning history to drive their choices, we fit a variant of a computational reinforcement



**Figure 4.** The majority of participants across age groups (A) had only 1 browser interaction during the task, (B) responded to all three comprehension questions correctly, (C) failed to respond on only a small number of trials, and (D) generally took more than 150 ms to make their choices.

learning model that has previously been used to quantify the recruitment of model-free and model-based learning strategies (Daw et al., 2011; Decker et al., 2016; Otto et al., 2013). The model consists of both a “model-free” and “model-based” learning algorithm, which separately compute first-stage action values on each trial. Critically, value estimates from each algorithm are scaled by separate free parameters fit to participants’ choices: a model-free inverse temperature ( $\beta_{MF}$ ) and a model-based inverse temperature ( $\beta_{MB}$ ). Higher values of ( $\beta_{MF}$ ) and ( $\beta_{MB}$ ) indicate greater recruitment of a model-free and model-based learning strategy, respectively. We include full details of the model and fitting procedure in the supplement. To address our primary question about changes in learning and decision strategies across development, we used linear regressions to examine how the two inverse temperatures that controlled the extent to which participants “weighted” the model-free and model-based value estimates ( $\beta_{MF}$  and  $\beta_{MB}$ ) varied as a function of age. Across all three datasets, we observed an increase in  $\beta_{MB}$  with increasing age (Decker:  $\beta = .32$ ,  $SE = .15$ ,  $p = .032$ ; Potter:  $\beta = .39$ ,  $SE = .14$ ,  $p = .008$ ; Online:  $\beta = .42$ ,  $SE = .12$ ,  $p < .001$ ). This relation was specific to  $\beta_{MB}$ ; we did not observe a significant relation between age and  $\beta_{MF}$  in any of the three datasets (Decker:  $\beta = .13$ ,  $SE = .12$ ,  $p = .28$ ; Potter:  $\beta = -.20$ ,  $SE = .15$ ,  $p = .19$ ; Online:  $\beta = .11$ ,  $SE = .09$ ,  $p = .23$ ). Thus, our computational modeling results align with our logistic regression results in that they suggest that from childhood to adulthood, participants’ use of a model-based learning strategy increased. While the logistic regression results suggest that within the Decker and online datasets, the use of a model-free learning strategy may increase with increasing age, the computational modeling results do not provide evidence to support this interpretation, suggesting that model-free learning may not vary robustly as a function of age.

**Power comparison.** In conducting our online study, we aimed to collect a larger sample than in either of our two

previous in-lab studies in anticipation of increased “noise” in participant choices. Was this larger sample size necessary to see the emergence of the effects we hypothesized? To determine whether each administration of the task yielded equally robust results, we conducted a post-hoc power analysis and examined how many participants from each sample were necessary to include to reliably produce our primary result of interest: the reward  $\times$  transition  $\times$  age interaction effect on repeated first-stage choices. We simulated different versions of each data set by randomly sampling (with replacement) the choice data from 10 - 30 children, adolescents, and adults. For each sample size, we performed 100 different simulations in which we randomly sampled a subset of participants and ran the first-stage choice logistic regression. We were interested in determining the minimum number of participants per age group that would produce a significant reward  $\times$  transition  $\times$  age interaction effect on at least 80% of simulations.

We found that when sampling from the Decker dataset, we could reliably detect a significant interaction effect (85% power) when we included 15 participants from each age group (total  $N = 45$ ). This effect remained reliably detectable at all sample sizes between 16 and 30 participants per age group (81 - 100% power) and was not as reliably detected at smaller sample sizes of 10 - 14 participants per age group (57 - 78% power). When sampling from the Potter dataset, we could reliably detect a significant interaction effect (83% of simulations) when we included 25 participants from each age group (total  $N = 75$ ). This effect remained reliable at all sample sizes between 26 and 30 participants per age group (83 - 91% power) and was not as reliably detected at smaller sample sizes of 10 - 24 participants per age group (44 - 79% power). Finally, when sampling from the online dataset, we found that we needed to include 21 participants per age group (total  $N = 63$ ) to detect our effect of interest on at least 80% of simulations (80% power). We could consistently detect this effect at all sample sizes above 21 par-

Table 5. MaRs-IB descriptive statistics

Measure	Chierchia et al. (2019)				
	Younger adolescents	Mid adolescents	Older adolescents	Adults	
N	185	184	184	106	
Items completed (mean)	33.7	32.3	29.0	33.4	
Accuracy (mean)	0.61	0.68	0.73	0.81	
RT (s; median)	6.94	7.55	7.95	9.45	
Inverse efficiency (median)	11.84	10.94	11.06	11.41	
	Online				
	Children	Younger adolescents	Mid adolescents	Older adolescents	Adults
N	29	21	29	21	51
Items completed (mean)	58.7	65.1	51.1	40.6	50.0
Accuracy (mean)	0.39	0.39	0.52	0.63	0.59
RT (s; median)	5.61	4.86	7.49	8.88	8.18
Inverse efficiency (median)	15.68	13.81	16.06	15.41	13.64

participants per age group (83 - 97% power). At sample sizes below 21 participants per age group, this effect was only sometimes reliable (83% power with 18 participants per age group but 55 - 78% power at all other sample sizes between 10 and 20 participants per age group).

### MaRs-IB

**Online data quality.** As with the two-step task data, we extracted four measures of data quality from our online dataset for each participant (Figure 7): their number of “browser interactions,” the number of trials it took them to answer all three practice questions correctly, the number of choice trials in which they failed to respond before the 30-second time limit (missed trials), and the number of choice trials in which they made a response faster than 250 ms (fast RTs). As with the two-step task data, the majority of participants across age groups had few browser interactions (median for all age groups = 1), answered the practice questions correctly (median number of trials to answer all three practice questions correctly for all age groups = 3), missed few responses (medians for all age groups = 1), and generally took more than 250 ms to make each choice (medians for all age groups = 0).

**Task performance.** Chierchia et al. (2019) reported results from 659 participants aged 11 - 33 years (185 younger adolescents aged 11.27 - 13.39 years; 184 mid-adolescents aged 13.4 - 15.91; 184 older adolescents aged 15.93 - 17.99; 106 adults aged 18.00 - 33.15 years). They examined four indices of participant performance: the total number of items completed, task accuracy (number of items correct / number of items completed), median response times for correct items, and inverse efficiency (median response times for correct items / accuracy), which can account for differences in individual’s speed-accuracy tradeoffs. We first examined whether participants completing our online version of the MaRs-IB performed similarly to the participants who completed the in-person version in Chierchia et al. (2019). To report summary statistics, we used the same age groups as Chierchia et al. (2019) with an additional “children” age

group for participants aged 8.00 - 11.27 years (Table 5). As Table 5 indicates, across age groups, participants in our online experiment completed more items but were less accurate than participants in Chierchia et al. (2019). Though participants in the online experiment responded more quickly on correct trials, they still demonstrated poorer (higher) inverse efficiency relative to participants in the in-lab study.

Despite these apparent differences in participants’ task performance, we next sought to determine whether we observed similar effects of item dimensionality in our sample. Chierchia et al. (2019) found that increasing item dimensionality slowed response times and reduced performance accuracy, suggesting that this manipulation worked as intended and modulated item difficulty. We similarly observed a robust effect of item dimensionality on response accuracy,  $\beta = -.71$ ,  $\chi^2(1) = 761.6$ ,  $p < .001$ , and response times (log transformed, here and in subsequent analyses) to correct items,  $\beta = .23$ ,  $F(1, 3609.2) = 329.05$ ,  $p < .001$ , suggesting that the intended difficulty manipulation worked effectively online as well.

We next examined how performance changed across development. Chierchia et al. (2019) found that only accuracy varied across development. Specifically, they found that accuracy varied as a function of both linear and quadratic age, such that accuracy increased throughout childhood and early adolescence before leveling off into adulthood. We first examined whether a model with a quadratic age term provided a better fit to the data than a model with just a linear age term. It did not,  $\chi^2(1) = 2.58$ ,  $p = .11$ , so we removed this additional term from our model and examined accuracy as a function of linear age. Like Chierchia et al. (2019), we similarly found an effect of linear age on participant accuracy,  $\beta = .49$ ,  $\chi^2(1) = 55.69$ ,  $p < .001$  (Figure 8A).

We also observed an effect of age on RTs to correct items,  $\beta = .12$ ,  $F(1, 150.24) = 7.3$ ,  $p = .008$ , such that RTs increased with increasing age (Figure 8C). This is in line with the linear trend that Chierchia et al. (2019) observed. However, unlike the previous study, we also observed an effect of linear age on the number of items participants completed,  $\beta =$

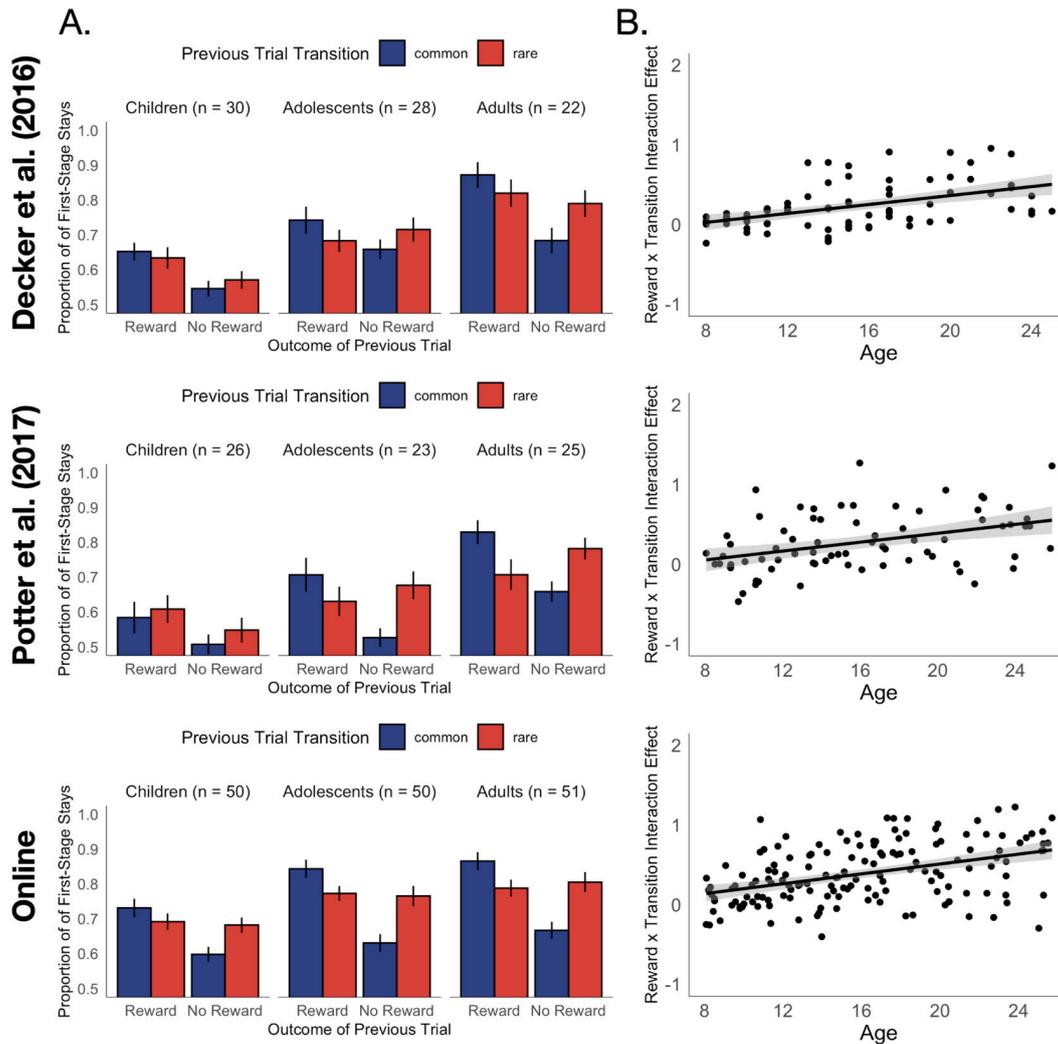


Figure 5. Across all three datasets, model-based learning increased from childhood to adulthood. (A) shows the proportion of first-stage choice repetitions ('stays') as a function of the outcome and transition experienced on the previous trial for each age group. (B) shows the random plus fixed reward x transition interaction effect from a model in which age was excluded, for each subject, plotted as a function of age. The line represents the best-fitting linear regression line +/- 1 standard error.

-5.3, SE = 1.4,  $p < .001$ , with younger participants completing more items relative to older participants (Figure 8B). Finally, we found that age was not related to inverse efficiency,  $\beta = -196.4$ , SE = 510.3,  $p = .70$ , which was also in line with Chierchia et al.'s findings (Figure 8D).

**Correlation between MaRs-IB and WASI.** Forty-four participants (19 children, 19 adolescents, 6 adults) who participated in our online study had previously participated in an in-lab study in which we had administered the vocabulary and matrix reasoning subtests of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 2011). MaRs-IB accuracy has previously been demonstrated to correlate with other indices of fluid reasoning, including the matrix reasoning portion of the International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014). In an exploratory analysis, we examined whether MaRs-IB accuracy also related to raw scores on the matrix reasoning (MR) portion of the WASI. We ran a linear regression examining the interacting effects of age and raw WASI MR scores on MaRs-IB accuracy. We found a main effect of WASI MR scores on MaRs-IB accuracy,  $\beta = .06$ , SE = .03,  $p = .047$  (Figure 9). When we included WASI MR scores in the model, age was

no longer a significant predictor of MaRs-IB accuracy,  $\beta = .03$ , SE = .03,  $p = .248$ , suggesting that the age-related variance that we observed in MaRs-IB accuracy was accounted for by age-related differences in fluid reasoning, as indexed by the WASI. We did not observe a significant WASI score x age interaction effect on MaRs-IB accuracy,  $\beta = -.03$ , SE = .03,  $p = .306$ . To ensure the relation that we observed between WASI MR scores and MaRs-IB accuracy was not driven by the small group of participants with low WASI MR scores, we re-ran our analysis excluding those participants with raw WASI MR scores below 15 ( $n = 6$  children; mean age = 10.1 years). Even after excluding these participants, we continued to observe a relation between WASI MR scores and MaRs-IB accuracy,  $\beta = .05$ , SE = .02,  $p = .041$ .

### Relation between model-based learning and fluid reasoning

Potter et al. (2017) found that fluid reasoning ability, as indexed by raw WASI MR scores, fully mediated the relation between age and model-based learning. We examined

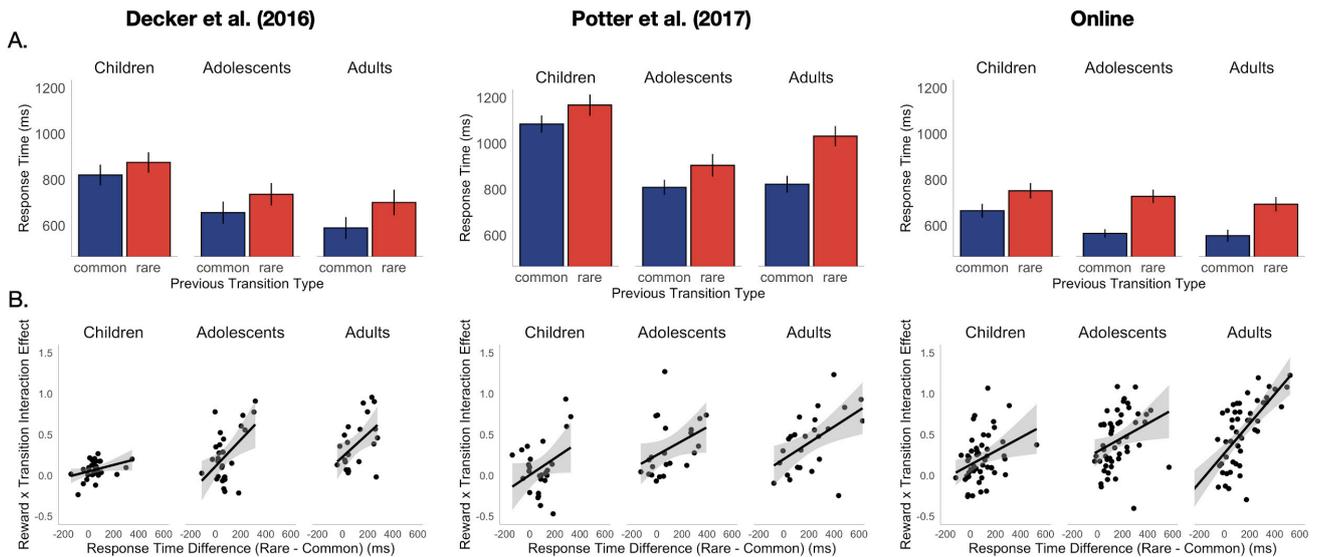


Figure 6. (A) Across all three datasets, participants across age groups were slower to make their second-stage choices following rare transitions than following common transitions. (B) The extent to which participants demonstrated RT-slowing predicted their use of a model-based learning strategy. The lines on the scatter plots represent the best-fitting linear regression line  $\pm$  1 standard error.

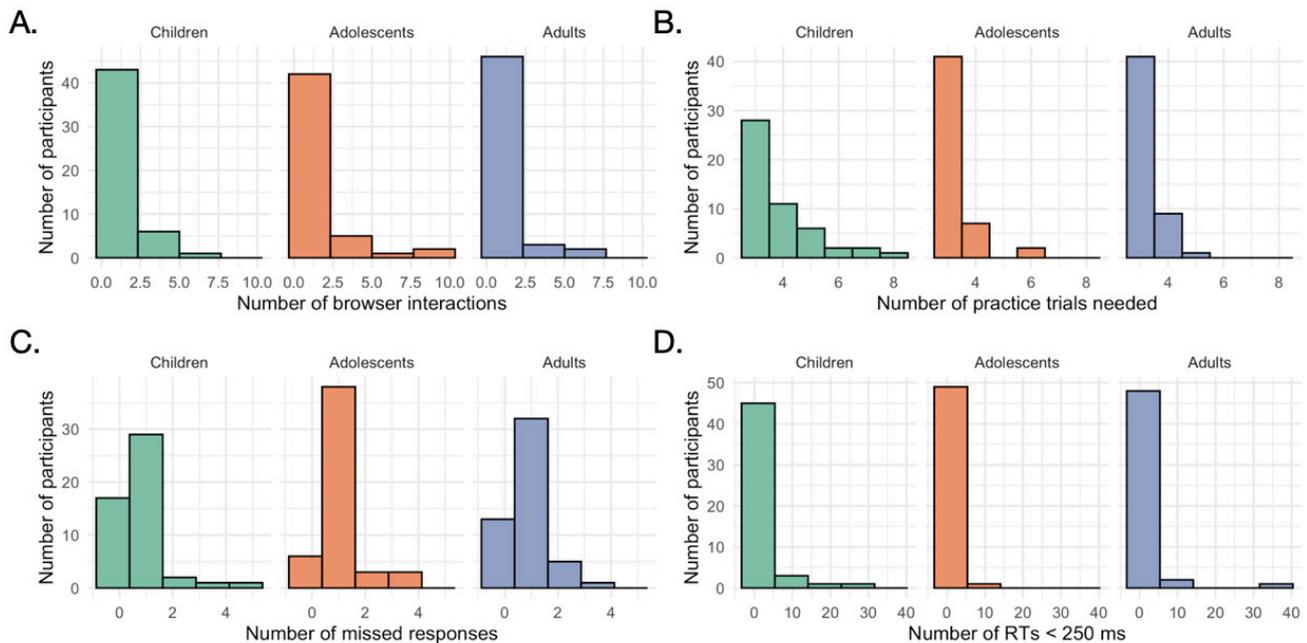


Figure 7. The majority of participants across age groups (A) had 1 browser interaction during the task, (B) answered all three practice questions correctly on the first try, (C) failed to respond on only a small number of trials, and (D) generally took more than 250 ms to make their choices.

whether we could replicate this mediation effect in our on-line sample using MaRs-IB accuracy. To examine model-based learning, we extracted each participant's reward x transition random effect from our repeated choices regression (without age). MaRs-IB accuracy positively related to this index of model-based learning,  $\beta = .48$ ,  $SE = .07$ ,  $t = 6.8$ ,  $p < .0001$ . A mediation analysis revealed that MaRs-IB accuracy partially mediated the relation between age and model-based learning (Figure 10). The standardized indirect effect was .20 (95% confidence interval: [.09, .31];  $p < .001$ )

and the standardized direct effect was .22 (95% CI = [.03, .42],  $p = .016$ ). These results align with those reported in Potter et al. (2017), and suggest that improvements in fluid reasoning across development support the increased recruitment of a model-based learning strategy. Potter et al. (2017) observed a full mediation whereas we only observed a partial mediation; it may be that our larger sample size gave us more power to detect a significant, direct effect of age on model-based learning, even after accounting for the indirect effect. Alternatively, the MaRs-IB may have less ex-

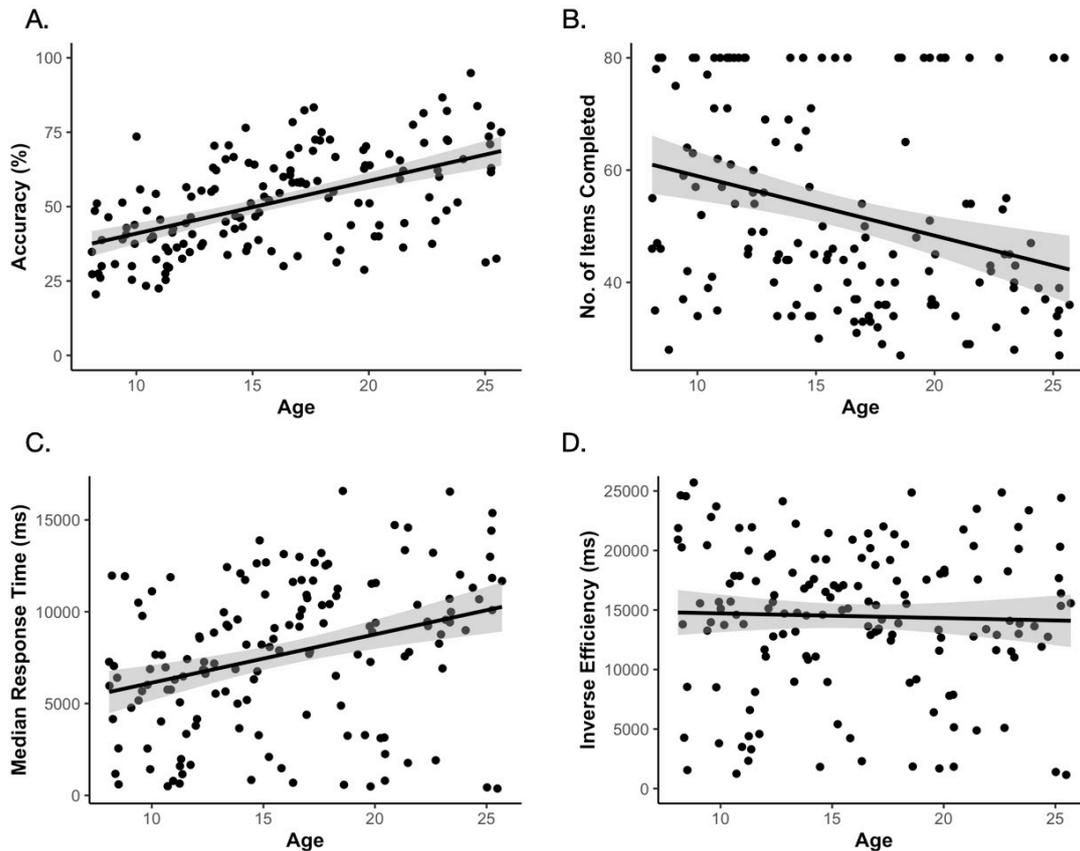


Figure 8. (A) MaRs-IB accuracy increased with increasing age. (B) Older participants tended to complete fewer puzzles within the 8-minute task. (C) Older participants also tended to make correct selections more slowly than younger participants. (D) Inverse efficiency did not vary as a function of age. The lines on the scatter plots represent the best-fitting linear regression line  $\pm$  1 standard error.

planatory power than the MR subtest of the WASI.

### Discussion

In an online administration of the two-step task, we successfully replicated the previous finding that model-based learning increases from childhood to adulthood (Decker et al., 2016; Potter et al., 2017). Across all three datasets, this effect emerged both when we considered only the influence of the previous trial on participant choices through our regression analyses and when we considered participants' full learning histories through our computational model. Beyond replicating this key finding, across all our analyses of the two-step task data, the pattern of results we observed in our online dataset closely mirrored those that we previously observed in the two in-lab samples. In addition to successfully replicating the age-related increase in model-based learning, we also replicated the pattern of reaction-time slowing after rare transitions, and the relation between this slowing effect and the recruitment of a model-based learning strategy. Taken together, our findings suggest that it is possible to examine age-related change in learning and decision-making strategies with data collected via online task administration, even through data-hungry computational methods that necessitate lengthy tasks.

We also examined the “noisiness” of our online dataset, through qualitative examination of key quality metrics and through a quantitative, post-hoc power analysis. Though a small number of participants in our online study did not ap-

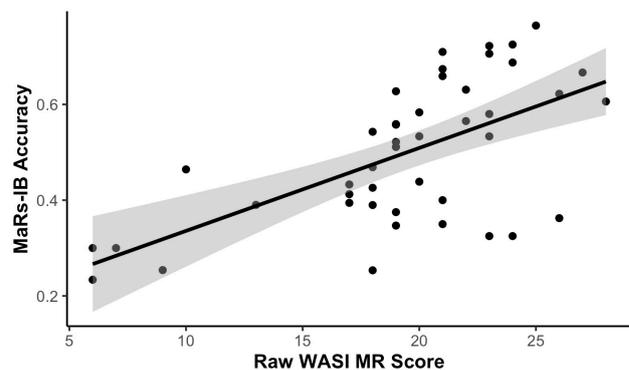


Figure 9. Participants' accuracy on the MaRs-IB correlated with their raw scores on the matrix reasoning portion of the WASI that they had previously completed as part of a separate, in-lab study. The line on the scatter plot represents the best-fitting linear regression line  $\pm$  1 standard error.

pear to be focused while completing the two-step task, as evidenced through extensive interaction with other browser windows and a large number of missed trials and fast responses (Figure 3), the majority of participants showed no obvious indicators of inattention. Our post-hoc power analysis did, however, suggest that the key reward x transi-

tion x age interaction effect on repeated first-stage choices was slightly weaker in our online sample relative to the Decker dataset, requiring 21 participants per age group to be reliably detected, as opposed to the 15 participants per age group Decker et al. required. While the effect was stronger in the online dataset than the Potter dataset, the Potter task involved 50 fewer trials per participant, making the comparison less apt.

One interpretation of our power analysis is that the online dataset was, as we initially anticipated, “noisier” than the comparable in-person dataset. However, it is not clear that the weaker three-way interaction was driven by an increase in random patterns of responses in the online dataset. Instead, the weaker age interaction may have been driven by an elevated use of model-based learning by child and adolescent participants in the online dataset (Figure 5; Table S1). Why would younger participants who completed the task online show a greater tendency toward model-based behavior? One possibility is that this qualitative difference across datasets emerged due to variance across samples and did not have to do with differences in task settings across studies. The patterns of repeated choices across age groups in the Potter dataset also show small, qualitative differences to the Decker dataset (Figure 5); it is not clear that the differences we observed in the online dataset are any greater than the difference between these two, in-person datasets.

Alternatively, however, it may be the case that administering the task online did shift decision-making strategies. Typically for our in-lab studies, families schedule their child’s appointment 1 - 2 weeks in advance, and children come to the lab in the afternoon, after a long day of school. When participating online, children can participate whenever they want, without any need to schedule the session in advance. We ran this study in June and July 2020, when schools were closed due to COVID-19 and most children and adolescents did not have strict schooling schedules to follow. It may be the case that children who participated online were actually less tired and better able to concentrate on the task. Another possibility is that parents interfered with their children’s choices during the task, potentially encouraging them to make more “model-based” decisions. We think this is unlikely, as previous research suggests rates of parental interference are low (<1% of trials), even for studies with young children (Rhodes et al., 2020). In addition, this version of the two-step task is designed such that model-free and model-based decisions, on average, yield equivalent reward, such that there would be no reason for parents to encourage a “model-based” strategy. However, future studies could measure parental interference more directly even without recording video, by having parents complete a survey at the end of the study in which they are asked if they helped their child in any way. Finally, though our online task instructions were identical to those used in the lab, we did add three comprehension questions to the end of them, two of which may have reinforced the importance of the task’s reward and transition structure. Recent research on decision strategies in the two-step task suggests the recruitment of model-based learning is highly sensitive to task framing, suggesting that even subtle differences in instructions may shift the mental models individuals use during learning, making their behavior appear more “model-based” (Feher da Silva & Hare, 2020). Our comprehension questions may have enabled participants to form better models of the task structure and increased their apparent use of model-based learning. A future study could remove, or manipulate, these comprehension questions to examine

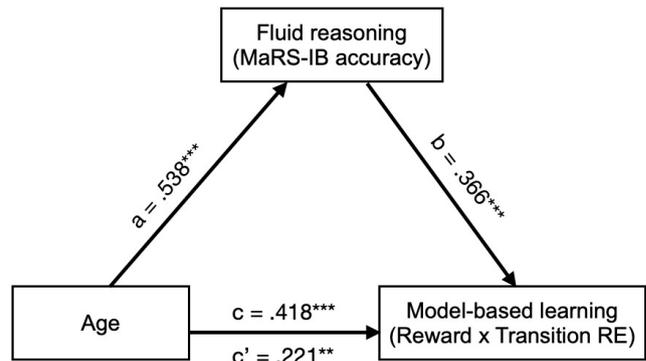


Figure 10. MaRs-IB accuracy partially mediated the relation between age and model-based learning. Path a shows the regression coefficient of the relation between age and MaRS-IB accuracy. Path b shows the regression coefficient of the relation between MaRS-IB accuracy and model-based learning, while controlling for age. Paths c and c’ show the regression coefficient of the relation between age and model-based learning without and while controlling for MaRS-IB accuracy, respectively. \*\* Denotes  $p < .01$ ; \*\*\* denotes  $p < .001$ .

their influence on learning.

Across datasets and analysis approaches, we observed mixed evidence for developmental change in model-free learning, though these results were consistent in the Decker and online datasets, suggesting these differences were not driven by the online task administration. In the Decker and online datasets, we observed a reward x age interaction effect on repeated first-stage choices in our regression analysis, which suggests that model-free learning may have also increased with increasing age. However, this effect did not emerge in our computational modeling results or in the Potter dataset. These mixed findings align with recent results that suggest that model-free learning may be difficult to capture in the two-step task (Feher da Silva & Hare, 2020). Other task designs may be better suited to characterize developmental change in habitual or reflexive decision-making (Miller et al., 2019), through both in-lab and online testing.

While our two-step task results appeared largely similar across in-lab and online task administration, our results from our abstract reasoning measure, the MaRs-IB, did not align as closely with those reported from a previous, in-person study (Chierchia et al., 2019). While we observed age-related improvements in reasoning accuracy, and, in a subset of participants, a correlation between MaRs-IB accuracy and MR scores on the well-validated WASI (Wechsler, 2011), participants in our online study seemed to respond to the puzzles more quickly than in the in-person study, completing more puzzles at the expense of task accuracy. One reason for this discrepancy may not have had to do with online administration, but rather be due to our overall task sequence: We had all participants complete the MaRs-IB after completing a lengthy (30 - 45 minutes) and somewhat cognitively demanding decision-making task. It is possible that they completed the MaRs-IB while fatigued and were thus unmotivated to spend the full amount of time on each puzzle. Future studies that involve two different experimental tasks could instead administer them via separate sessions to mitigate the influence of fatigue. Because participants do not need to travel to the lab to participate, online testing may facilitate data collection spread across multiple time-

points.

In addition, while we told participants that their performance on the two-step task would determine their monetary bonus, we did not incentivize the MaRs-IB. While reasoning tasks are typically not incentivized when administered in person, motivation can have strong effects on performance (Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011). Motivation may be particularly important to consider when an experimenter is *not* present during task administration because being observed may in and of itself offer motivation for performing well (Bond, 1982; Bond & Titus, 1983). Finally, even if participants were motivated to perform to the best of their ability in the MaRs-IB, the instructions do not provide participants with an explicit strategy for considering speed-accuracy trade-offs. Though we followed the example of Chierchia et al. (2019) and used task accuracy (number of puzzles correct / number of puzzles completed) as our primary index of reasoning ability, participants were never instructed to maximize accuracy. Participants instead may have tried to maximize the total number of puzzles they answered correctly, strategically guessing on puzzles they determined would require a long time to solve. During in-person studies, experimenters have the opportunity to observe how participants approach the task and, critically, to answer questions if participants are unclear as to what their goal is. Neither of these opportunities are available in online experiments, increasing the importance of highly detailed instructions that explicitly lay out what participants' goals should be and why they should be motivated to do well in their pursuit of them. Future online studies could add more explicit instructions and a clearer incentive structure to the MaRs-IB to determine if doing so would make data collected online more closely approximate data collected in person.

Finally, beyond replicating previous, in-lab developmental findings, one of our goals in conducting this study was to develop a pipeline that could be re-used for other online studies of learning and decision making in children and adolescents. As such, we were interested in determining the relative ease of recruiting and testing a large sample of participants, as well as how the diversity and representativeness of our online sample compared to previous, in-lab studies. In terms of the ease of participant recruitment, online testing surpassed our cautious expectations: We recruited and tested 151 child, adolescent, and adult participants in 5 weeks. For comparison, it normally takes our lab about 6 months to test that number of participants. Testing online also enabled participants to safely complete the experiment from the comfort of their homes (or any location of their choice), allowing data collection to proceed in the midst of a global pandemic. Further, as described in the supplement, a large proportion of participants (> 60%) who we invited to complete the study actually did so, and only 15 participants started but did not complete the experiment. The demographics of our participant population looked largely similar to that of a previous, in-lab study, though for this first foray into online testing, we contacted many families who had previously participated in in-person experiments in our lab. We expect that over time, as we advertise our studies more widely, our online participant pool will more closely resemble the national population, as opposed to the more racially diverse New York City population that we typically draw from for in-lab research. In addition,

in both our in-lab studies and our online study, we have undersampled Black adults. Moving forward, we intend to think critically about where and how we advertise research opportunities, to ensure that our participant population is more representative of the general population we hope to make inferences about.

Overall, we believe our results demonstrate that it is not only possible, but relatively straightforward, to collect high-quality value-based decision-making data from children, adolescents, and adults via online task administration. Our study adds to the growing literature examining online testing as a way to extend and replicate in-lab developmental findings, which have often relied on small and geographically constrained samples of participants (Rhodes et al., 2020; Sheskin et al., 2020). Critically, we find that online testing can be used not just to administer short experiments, but also lengthier tasks that require sustained focus and many, repeated decisions, even in children as young as 8 years old. That said, our findings also suggest that, as with in-lab testing, the collection of high-quality data likely requires limits to the length or number of tasks participants should be asked to complete within a single testing session and highlight the importance of clear instructions and incentive structures. Future work is needed to further delineate the possibilities and limitations of online testing, as well as how they compare to in-lab task administration.

## Acknowledgements

We thank the Niv Lab at Princeton for providing us with a version of the two-step task in jsPsych, Perri Katzman for recording audio instructions for the tasks, [ChildrenHelpingScience.com](https://www.ChildrenHelpingScience.com) for facilitating participant recruitment, and all the children and families that participated in the study.

## Funding

We thank our funders: The National Science Foundation (CAREER grant 1654393 to C.A.H.), the Jacobs Foundation (Early Career Fellowship to C.A.H.), the Department of Defense (NDSEG Fellowship to K.N.), and New York University (Katzell Summer Fellowship to K.N.).

## Author Contributions

K.N. and C.A.H. conceptualized the study. All authors contributed to the study design. K.N. and M.S. collected and analyzed the data under the supervision of C.A.H. K.N. drafted the manuscript. All authors provided critical feedback and revisions.

## Conflicts of Interest

The authors have no conflicts of interest to report.

## Data Availability

All data described in the manuscript are publicly available on the Open Science Framework: <https://osf.io/we89v/>



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## REFERENCES

- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Bolenz, F., & Eppinger, B. (2020). *Valence bias in metacontrol of decision making in adolescents and young adults*. Center for Open Science. <https://doi.org/10.31234/osf.io/5u9jq>
- Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology*, *42*(6), 1042–1050. <https://doi.org/10.1037/0022-3514.42.6.1042>
- Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94*(2), 265–292. <https://doi.org/10.1037/0033-2909.94.2.265>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2016). *Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data?* <https://psycnet.apa.org/record/2015-32022-009>
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*(10), 190232. <https://doi.org/10.1098/rsos.190232>
- Chuey, A., Lockhart, K., Sheskin, M., & Keil, F. (2020). Children and adults selectively generalize mechanistic knowledge. *Cognition*, *199*, 104231. <https://doi.org/10.1016/j.cognition.2020.104231>
- Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133. <https://doi.org/10.1016/j.cogpsych.2015.02.004>
- Cohen, A. O., Nussenbaum, K., Dorfman, H., Gershman, S. J., & Hartley, C. A. (in press). The rational use of causal inference to guide reinforcement learning changes with age. *NPJ Science of Learning*. <https://doi.org/10.31234/osf.io/j9zuc>
- Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. <https://doi.org/10.1016/j.intel.2014.01.004>
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS One*, *8*(3), e57410. <https://doi.org/10.1371/journal.pone.0057410>
- 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. <https://doi.org/10.1016/j.neuron.2011.02.027>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- de Leeuw, J. R., Coenen, A., Markant, D., Martin, J. B., McDonnell, J., Rich, A., & Gureckis, T. (2014). Online Experiments using jsPsych, psiTurk, and Amazon Mechanical Turk. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*. <https://escholarship.org/content/qt3266p2m8/qt3266p2m8.pdf>
- Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 310–320. <https://doi.org/10.3758/s13415-014-0332-5>
- 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. <https://doi.org/10.1177/0956797616639301>
- Dorfman, H. M., Bhui, R., Hughes, B. L., & Gershman, S. J. (2019). Causal Inference About Good and Bad Outcomes. *Psychological Science*, *30*(4), 516–525. <https://doi.org/10.1177/0956797619828724>
- Feher da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-020-0905-y>
- Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*(1). <https://doi.org/10.1038/s41467-020-16964-5>
- Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *ELife*, *5*. <https://doi.org/10.7554/elif.e11305>

- Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(3), 523–536. <https://doi.org/10.3758/s13415-015-0347-6>
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychological Science*, *28*(9), 1321–1333. <https://doi.org/10.1177/0956797617708288>
- Leshin, R., Leslie, S.-J., & Rhodes, M. (2020). *Does It Matter How We Speak About Social Kinds? A Large, Pre-Registered, Online Experimental Study of How Language Shapes the Development of Essentialist Beliefs*. Center for Open Science. <https://doi.org/10.31234/osf.io/nb6ys>
- Lourenco, S. F., & Tasimi, A. (2020). No Participant Left Behind: Conducting Science During COVID-19. *Trends in Cognitive Sciences*, *24*(8), 583–584. <https://doi.org/10.1016/j.tics.2020.05.003>
- Mathworks Inc. (2020). *MATLAB: R2020a*. Mathworks Inc.
- Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, *126*(2), 292–311. <https://doi.org/10.1037/rev0000120>
- Nussenbaum, K., Cohen, A. O., Davis, Z., Halpern, D., Gureckis, T. M., & Hartley, C. A. (2020). Causal information-seeking strategies change across childhood and adolescence. *Cognitive Science*, *44*, e12888. <https://doi.org/10.1111/cogs.12888>
- Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20941–20946. <https://doi.org/10.1073/pnas.1312011110>
- Potter, T. C. S., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, *25*, 272–280. <https://doi.org/10.1016/j.dcn.2016.10.005>
- R.Core Team. (2018). *R: A language and environment for statistical computing* (Version 3.5.1). <https://www.R-project.org>
- Rhodes, M., Rizzo, M., Foster-Hanson, E., Moty, K., Leshin, R., Wang, M. M., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*. <https://doi.org/10.1080/15248372.2020.1797751>
- Sauter, M., Draschkow, D., & Mack, W. (2020). Building, Hosting and Recruiting: A Brief Introduction to Running Behavioral Experiments Online. *Brain Sciences*, *10*(4), 251. <https://doi.org/10.3390/brainsci10040251>
- Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the Viability of Online Developmental Research, Results From Three Case Studies. *Open Mind*, *1*(1), 15–29. [https://doi.org/10.1162/opmi\\_a\\_00001](https://doi.org/10.1162/opmi_a_00001)
- Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online Developmental Science to Foster Innovation, Access, and Impact. *Trends in Cognitive Sciences*, *24*(9), 675–678. <https://doi.org/10.1016/j.tics.2020.06.004>
- Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *Afex: Analysis of factorial experiments* (Version .27-2). <https://CRAN.R-project.org/package=afex>
- Smid, C. R., Kool, W., Hauser, T. U., & Steinbeis, N. (2020). *Model-based decision-making and its metacontrol in childhood*. Center for Open Science. <https://doi.org/10.31234/osf.io/ervsb>
- van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex*, *22*(6), 1247–1255. <https://doi.org/10.1093/cercor/bhr198>
- Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence (WASI-II)*. NCS Pearson.

## SUPPLEMENTARY MATERIALS

### Author Response Letter

Download: [https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44038.pdf?auth\\_token=ubdhhxBCW9mUKq57aW7U](https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44038.pdf?auth_token=ubdhhxBCW9mUKq57aW7U)

---

### Peer Review Comments

Download: [https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44039.pdf?auth\\_token=ubdhhxBCW9mUKq57aW7U](https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44039.pdf?auth_token=ubdhhxBCW9mUKq57aW7U)

---

### Text S1. Supplementary information and analyses

Download: [https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44040.pdf?auth\\_token=ubdhhxBCW9mUKq57aW7U](https://collabra.scholasticahq.com/article/17213-moving-developmental-research-online-comparing-in-lab-and-web-based-studies-of-model-based-reinforcement-learning/attachment/44040.pdf?auth_token=ubdhhxBCW9mUKq57aW7U)

---