Title
Distinct Developmental Trajectories In The Cognitive Components Of Complex Planning

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

Authors
Ma, Ili
Phaneuf, Camille V
van Opheusden, Bas
et al.

Publication Date
2022

Peer reviewed

Distinct Developmental Trajectories In The Cognitive Components Of Complex Planning

Ili Ma (i.ma@fsw.leidenuniv.nl)
Department of Developmental and Educational Psychology, Wassenaarseweg, 2333 AK Leiden
The Netherlands

Camille Phaneuf (cphaneuf@g.harvard.edu)
Department of Psychology, Harvard University, Cambridge, Massachusetts
USA

Bas van Opheusden (svo@princeton.edu)
Department of Computer Science, Princeton University, Princeton, 08540
USA

Wei Ji Ma (weijima@nyu.edu)
Center for Neural Science and Department of Psychology, New York University, 2-6 Washington Place
New York, NY 10003 USA

Catherine Hartley (cate@nyu.edu)
Department of Psychology, New York University, 2-6 Washington Place
New York, NY 10003 USA

Abstract

This study aimed to characterize the developmental trajectories of different cognitive component processes underlying planning decisions. Participants (ages 8-25 years) completed a planning task called Four-in-a-row. We used computational modeling to distinguish between three cognitive component processes of planning: planning depth, heuristic quality, and attentional oversights, each of which three contributed to better playing strength, but differed in their developmental trajectories. Specifically, from early to mid-adolescence, heuristic quality rapidly improved and contributed to better playing strength. From mid to late-adolescence, planning depth increased and supported better playing strength. Fewer attentional oversights were associated with better playing strength and this relation did not show age differences. Together, these results reveal sequential development of the cognitive component processes underlying planning, with early refinement of heuristic strategies, and gradual increases into young adulthood in the number of considered future actions, states, and outcomes. These findings provide a more complete account of the development of planning and its component processes.

Keywords: Development; Planning; Decision-making; Reinforcement learning

Introduction

Planning is a form of decision-making that involves the mental simulation of potential futures. Everyday life frequently requires planning for achieving short- and long-term goals, from selecting the most efficient commute home to making sound investment choices. Planning is a complex construct that has several cognitive component processes. To evaluate a potential sequence of actions, one must identify and attend to relevant features of the current state, and mentally simulate potential actions and their future consequences within working memory (Sutton & Barto, 2018). While a large developmental literature has demonstrated that planning improves from childhood well into early adulthood (Albert & Steinberg, 2011; Korkman, Kemp, & Kirk, 2001; McCormack & Atance, 2011), prior studies have not yet characterized the development of the component processes that underlie planning within a single task.

In recent years, there has been a growing focus on characterizing the computations underlying planning using simple decision-making tasks that are amenable to detailed mathematical modeling (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Huys et al., 2012). The paradigmatic example is the Two-Step Task (Daw et al., 2011), a reinforcement learning task which dissociates model-based and model-free decision-making (Drummond & Niv, 2020), and has been used in developmental studies of planning (Decker, Otto, Daw, & Hartley, 2016; Potter, Bryce, & Hartley, 2017; Vaghi et al., 2020). However, a shortcoming of the two-step task is that it only requires thinking two steps ahead, making the resulting decision tree small enough to exhaustively explore. The simplicity of this task (as well as other widely used tasks such as the Tower of London) limits their capacity to evoke the complex strategies characteristic of real-world planning, in which decision trees are typically too large for exhaustive search to be feasible (van Opheusden & Ma, 2019). In planning tasks with large decision trees, the values of intermediate states are often uncertain and have to be evaluated through heuristic means. (Sutton & Barto, 2018). Additionally, tasks with a small number of distinct states or that only allow for a few actions in a given state, will have a limited ability to dissociate component processes of planning such as the ability to evaluate states, think ahead, and attend to state features. As a result, there is a need for a sufficiently rich planning paradigm in which age-related changes in these

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

component processes can be examined. The current study uses a child-friendly game — the Four-in-a-row task (van Opheusden, Galbiati, Kuperwajs, Bnaya, & Ma, 2021) — and a corresponding computational modeling framework to formally quantify the development of complex planning and its cognitive component processes.

Methods

Participants and procedure

Participants were 8 to 25 years old ($n = 156$), uniformly distributed across age and gender. Participants were fluent in English and reported no color blindness, learning disability, or neurodevelopmental or psychiatric disorders. Participants were instructed that they would receive a \$15 Amazon gift card plus a performance-based bonus for participation; in reality, all participants were compensated with a \$17 gift card to ensure that bonuses were not unethically biased toward older participants. Parental permission and child assent were obtained prior to participation. The study took place online. Participants completed the Four-in-a-row Task, followed by age-appropriate assessments of individual differences in fluid reasoning, daily life impulsive choices and future orientation using the Matrix Reasoning Item Bank (MaRs-IB (Chierchia et al., 2019; Nussenbaum, Scheuplein, Phaneuf, Evans, & Hartley, 2020)), Barratt Impulsiveness Scale - Brief (BIS-Brief (Steinberg, Sharp, Stanford, & Tharp, 2013)), and the Future Orientation Scale (FOS (Steinberg et al., 2009)), respectively. Parents were allowed to assist children 11 years or younger with reading BIS-Brief or FOS questions. This study was approved by the Institutional Review Board at New York University (IRB-FY2016-1194).

Four-in-a-row

The Four-in-a-row Task is a variant of Tic-Tac-Toe. The player and computer opponent alternate placing tokens on a 4-by-9 board (Fig 1A), and the first to complete four in a row (horizontally, vertically, or diagonally) wins. If no player wins before the board fills, the game is drawn. This board size allows for many winning opportunities, while being sufficiently small to be experimentally and computationally tractable. The task was programmed in JavaScript and played in a web browser. The computer opponents were calibrated to human play to create a wide range of difficulty levels, which were then used in a staircasing algorithm to approximate a 2 win to 1 loss ratio (see below). The task began with brief written instructions followed by a comprehension check to ensure that participants understood all three winning Four-in-a row orientations (i.e., diagonal, horizontal, and vertical). This quiz took the form of three multiple choice questions, in which participants were shown final game states and asked to indicate the winner of the game. If the response was incorrect, the instructions were repeated. After the quiz, the main task started, consisting of 35 games. Simulations demonstrated that 35 games were sufficient to reliably estimate the computational model parameters. The task has approximately 1.2×10^{16} non-terminal states — a state space complexity that far exceeds tasks commonly used in cognitive science (van Opheusden & Ma, 2019). Yet, this task has proven to be amenable to rigorous computational modeling (van Opheusden et al., 2021).

The computer opponents' algorithm was similar to our main model (see below). We created 200 computer opponents that all used the same algorithm but with different parameters. We started by fitting the model on all participants in laboratory experiments from (van Opheusden et al., 2021), resulting in 1650 agents. We then ran an all-versus-all tournament between these 1650 agents and ranked their performance using the Elo system (see below). Finally, we selected 200 agents such that their Elo ratings uniformly covered an interval ranging from slightly weaker than the worst human players to slightly stronger than the best. We fine-tuned this interval in pilot experiments. We divided the set of 200 agents into 20 categories, with 10 agents per category. We matched participants with computer opponents using a two-win-to-one-loss staircase, starting at category 2. That is, when a participant won two consecutive games, the category of their opponent in the next game increased by one. When a participant lost, the category decreased by one, and after a draw the category was kept keep the same. An opponent was randomly selected from the 10 agents within that category.

We estimated the depth of planning from a participant's moves using the computational model in (van Opheusden et al., 2021). This model combines a heuristic value function with best-first search. The value function $V(s)$ assigns values to board states s. We use a weighted linear sum of features

$$V(s) = \sum_{i=0}^{4} w_i f_i(s), \tag{1}$$

where f_i denote the features and w_i their weights. We use 5 features: center, connected 2-in-a-row, unconnected 2-in-a-row, 3-in-a-row and 4-in-a-row. The center feature measures how close to the center of the board a player's pieces are distributed, and the other 4 features count how often the corresponding patterns occur on the board (Fig. 1B).

The evaluation function guides the construction of a decision tree (Fig. 1D), using a best-first search algorithm that focuses computational resources by exploring promising branches of the decision tree first. More precisely, given a partially constructed decision tree, the algorithm decides which node to consider next by exploring the principal variation, that is, the sequence that results if both players choose the highest-value moves in the current tree. The algorithm expands the final node in this principal variation, evaluates candidate moves using the value function defined above, backpropagates the result according to the minimax rule, and continues to the next iteration. After each iteration, the model has a probability γ to terminate search, and make the move that is best according to its current tree.

The model contains additional components which improve its ability to match human data. First, we include a pruning rule: when expanding a node, we prune all candidate moves whose value differs from the best by more than a threshold ω. To account for variability in people's choices, we add three sources of noise. First, before constructing the decision tree, we randomly drop features at specific locations and orientations with a probability δ; these features are omitted during the calculation of $V(s)$. This mechanism is intended to account for random attentional oversights. Second, during tree search, we add Gaussian noise to $V(s)$ at each node. Finally, we also include a generic lapse rate λ. For details on the computational model, see (van Opheusden et al., 2021).

Model-derived metrics There are four model-derived metrics: Playing strength (Elo ratings), planning depth, heuristic quality, and feature drop rate. To estimate a participant's playing strength from games against computer opponents, we used the Elo system (Elo, 1978), implemented using the publicly available Bayeselo algorithm (https://www.remi-coulom.fr/Bayesian-Elo/). This algorithm treats the problem as a Bayesian parameter estimation problem, with a model that specifies the probability of a win/loss as a logistic function of the rating difference of the players. The algorithm takes as input a database of game results and estimates all ratings. Constant offsets to all players' ratings do not affect the predictions of the algorithm. Individual players' Elo ratings are sufficiently precise that they can be used as a metric for playing strength. *Planning depth* was estimated as follows: After fitting the parameters of our main model for a given participant, we ran the model forward in generative mode. In each position, we generated 100 simulations using the fitted parameters. In each simulation, we recorded the depth of the principal variation (sequence of moves considered best); we then averaged across simulations and positions. The result is the planning depth estimate for that participant. *Heuristic quality* was defined as the correlation between $V(s,w)$ and the objective value (1 for wins, -1 for losses, 0 for draws), across a pre-generated set of observed game states s. The heuristic quality only depends on the feature weights in the model, and not on the parameters of the tree search algorithm. The *feature drop rate* is simply the parameter δ in the model.

Analyses

Age-related differences in model-derived metrics and relation with Elo To test the relation between age and the model derived metrics, we used robust regression models with bisquare weighting for outliers. We tested both linear and polynomial age effects and report the results for the best fitting age model. We also used robust regression models to predict Elo ratings using planning depth, heuristic quality, feature drop rate, and their interaction with age.

Model stability To assess parameter estimate stability, we follow the methods described in van Opheusden et al. (2021) by calculating the correlation between parameter estimates of independent fits. While the parameter estimates between the two independent fits should be correlated, they cannot be identical because the model fitting algorithm provides stochastic estimates of the log-likelihood and therefore returned stochastic parameter estimates (van Opheusden, Acerbi and Ma, 2020).

Results

Descriptive statistics

With age, participants took longer on average to make a move (Spearman rho correlation between decision time and age: $\rho = 0.234$, $p = 0.003$) and made more moves per game ($\rho = 0.200$, $p = 0.012$). On average, participants completed the task in 16.17 minutes (*sd* = 5.54). Task completion time correlated with age such that older participants took longer ($\rho = 0.277$, $p = 4.55 \cdot 10^{-4}$). Opponent category stabilized over games for all ages, which suggests that the staircasing algorithm converged within the 35 games. Most participants answered all comprehension questions correctly on the first attempt (n = 138; n = 16 made one mistake and n = 2 made more than one mistake). Age did not significantly correlate with the number of correct answers on the comprehension questions ($\rho = 0.136$, $p = 0.089$), suggesting that younger participants also comprehended the task well.

Age-related differences in model-derived metrics

We found that performance on the planning task improved with age, as shown by an age-related improvement in Elo ratings, especially during early adolescence (linear age effect $B = 58.46$, $p = 8.55 \cdot 10^{-8}$, 95% CI [37.93; 78.99], quadratic age effect B = -30.77, $p = 6.06 \cdot 10^{-3}$, 95%CI[-52.62; -8.92]) (see Fig.2A). As a result of the five-fold cross validation method, each subject had five estimates of planning depth, heuristic quality, and feature drop rate. The average of each metric per subject was therefore used as a dependent variable in the regression models. We found that planning depth monotonically improved with age (B = 0.83, $p = 1.45 \cdot 10^{-6}$, 95%CI[0.50; 1.16]), while heuristic quality showed a polynomial age effect, suggesting that age-related improvements in heuristic quality emerged most strongly between childhood and early adolescence (linear B = 0.03, $p = 5.37 \cdot 10^{-7}$, 95% CI[0.02; 0.04]; quadratic: B = -0.02, $p = 2.32 \cdot 10^{-3}$, 95% CI[-0.03; -0.01]). Feature drop rate did not significantly change with age (B = 0.00, p = 0.884, 95% CI[-0.02; 0.02]).

Relation between Elo and model-derived metrics

We found that planning depth and heuristic quality indeed showed age-dependent effects on Elo ratings (Fig.2B). Specifically, planning depth became more predictive of Elo ratings with age (interaction between planning depth and age B = 35.17, $p = 1.19 \cdot 10^{-4}$, 95% CI[17.59, 52.75]), while heuristic quality became less predictive of Elo ratings with age (interaction between age heuristic quality and age B = -38.82, $p = 1.48 \cdot 10^{-5}$, 95% CI[-55.94, -21.70]). These findings suggest developmental changes in the contribution of these component processes to solving planning problems, with greater reliance on heuristic quality at younger ages, and

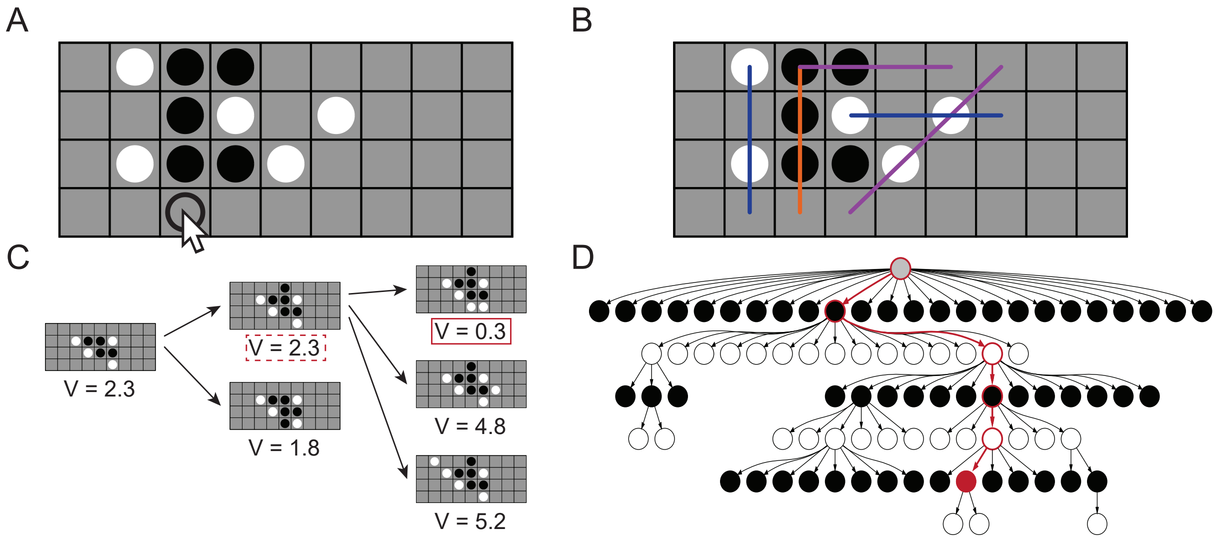

Figure 1: Four-in-a-row planning task and computational model re-printed from van Opheusden et al. (2021) with permission. A. Example board position in the Four-in-a-row game. Two players, black and white, alternate placing pieces on the board. The first player to achieve 4-in-a-row wins the game. In this position, black is about to win by moving on the 3rd square in the bottom row (open circle, mouse cursor). B. Features used in the heuristic function. Features with identical colors are constrained to have identical weights. The model also includes a central tendency feature and a 4-in-a-row feature. C. Illustration of the algorithm. Illustration of the heuristic search algorithm. Black is to move (left image). After considering two candidate moves for black and evaluating the resulting positions using V(s), the highest-value move (V = 2.3) is selected on the second iteration and expands this node with three candidate moves for white. The algorithm backpropagates the lowest value (V = 0.3), as this is the worst possible move for black. That value is compared against its alternatives in each intermediate node of the tree, to decide in which direction to expand the tree in the algorithm's next iteration. D. Actual decision tree for one simulation in one position for one participant, after parameter fitting. Red: principal variation, the sequence of most promising moves for both players. In this example, the depth of planning is 5. Different branches are evaluated to different depths.

gradually expansion of the decision tree size with age. In addition, lower feature drop rate was associated with better playing strength (main effect B = -38.19, $p = 9.60{\cdot}10^{-6}$, 95% CI[-54.65, -21.73]) and this relationship did not differ significantly with age (interaction effect B = -4.07, p = 0.615, 95% CI[-20.04, 11.90]), confirming that, across participants, fewer attentional oversights predicted better playing strength.

We performed post-hoc analyses to better understand the interaction effects between age and planning depth and age and heuristic quality on Elo ratings. We therefore created age groups to examine how each of these metrics contributed to Elo ratings within an age group. We applied Bonferroni-Holm corrections for multiple comparisons. We found that heuristic quality was significantly predictive of Elo ratings for the 8-12 year olds (B = 87.46, $p = 6.98{\cdot}10^{-5}$, 95% CI[46.96, 127.95]), but planning depth and feature drop rate were not (planning depth B = 10.23, p = 0.608, 95% CI[-29.61, 50.07]; feature drop B = -33.50, p = 0.110, 95% CI[-75.01, 8.03]). For the 13-17 year olds, heuristic quality remained a significant predictor of Elo ratings (B = 55.03, $p = 5.95{\cdot}10^{-4}$, 95% CI[24.97, 85.10]). Feature drop rate additionally became a significant predictor (B = -43.44, $p = 0.634{\cdot}10^{-3}$, 95% CI[-74.03, -12.86]), but not planning depth (B = -7.22, p = 0.622, 95% CI[-36.50, 22.05]). For 18-25 year olds, heuristic quality was no longer significantly predictive of Elo ratings (B = -11.97, p = 0.250, 95% CI[-32.64, 8.70]). Instead, for this age group planning depth was most predictive of Elo (B = 72.49, p= $1.61{\cdot}10^{-8}$, 95% CI[-51.06, 93.92]) followed by feature drop rate (B = -31.55, p = 0.005, 95% CI[-52.98, -10.12]). Together, these results further support an age-related shift in the use of the cognitive component processes that underlie planning decisions, with heuristic refinement during childhood to increasingly dominant reliance on expanding the decision tree towards young adulthood.

Model stability

All parameter estimates were positively correlated between independent fits in our entire sample as well as for each age group (Fig. 3A). Importantly, the correlations were not smaller for younger age groups than for adults, suggesting that the parameter estimates were stable across fits and age. In addition, the correlations between different model derived metrics were in expected directions, and not strong, suggesting that there was no strong trade-off between the metrics (Fig. 3B).

Individual differences

Age and fluid reasoning were positively correlated (Spearman's ρ = 0.301, $p = 1.80{\cdot}10^{-4}$). Robust mediation anal-

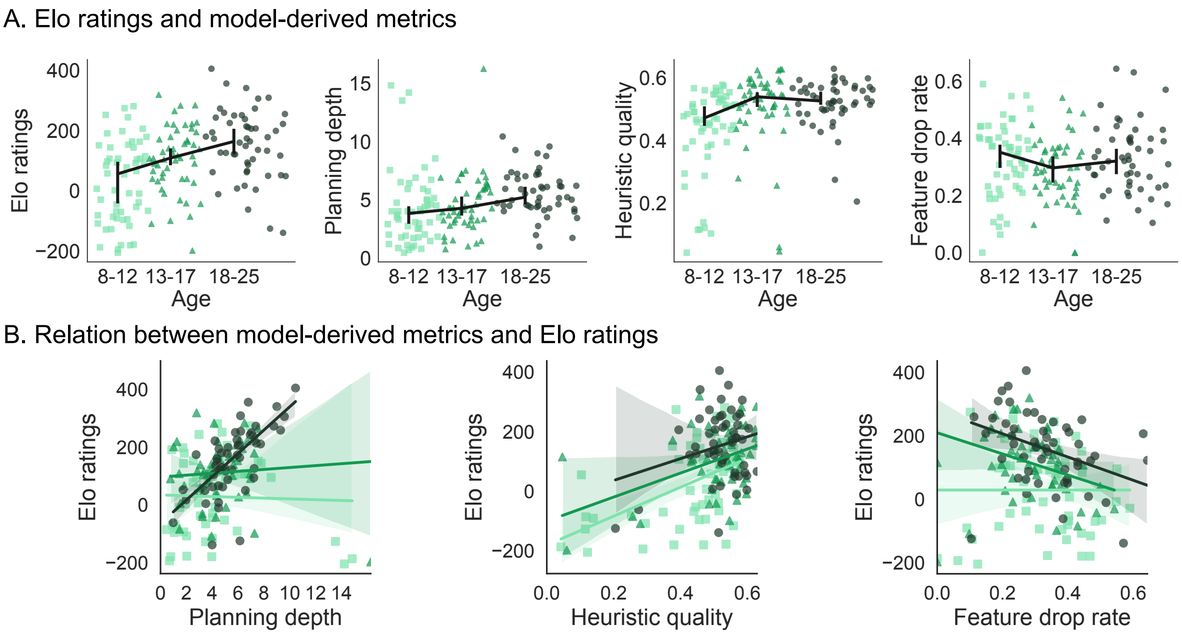

Figure 2: A. Elo ratings and Model-derived metrics per age bin (in years). Age was a continuous variable in the main analyses; bins are for visualization purposes. Line represents the median per group and the errorbars show the bootstrapped 95% confidence interval of the median. B. Elo rating as a function of the model-derived metrics per age bin. Lines show the robust regression fit and the bootstrapped 95% confidence interval. In all plots light green = children (8-12 years old), green = adolescents (13-17 years old), dark green = adults (18-25 years old)

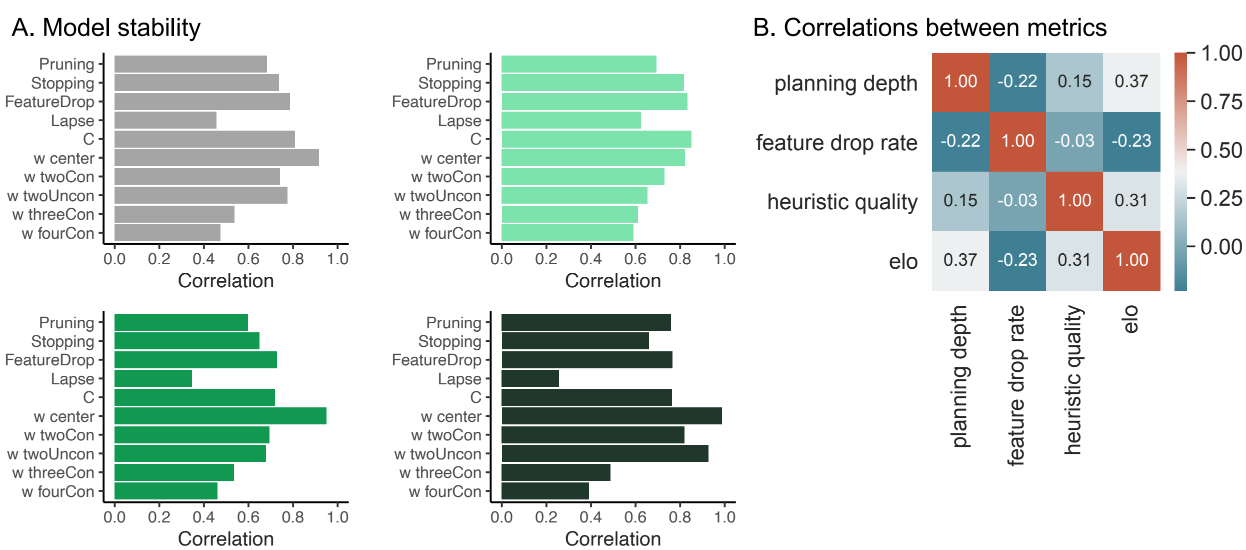

Figure 3: A. Spearman rho correlations between repeated fits for each parameter estimate. Grey = all subjects, light green = children 8-12 years old, green = adolescents 13-17 years old, dark green = adults 18-25 year old. B. correlations between metrics shows no strong parameter trade-off.

yses with fluid reasoning as a mediator were performed on model-derived metrics demonstrating evidence of age-related change (i.e., planning depth and heuristic quality) using 5000 bootstrapped replicates. While fluid reasoning partially mediated the relationship between age and heuristic quality (bootstrapped total effect 0.162, $p = 2.40 \cdot 10^{-4}$), the relationship between age and heuristic quality remained significant (bootstrapped direct effect 0.135, $p = 1.61 \cdot 10^{-3}$), confirming that heuristic quality improved with age. Fluid reasoning also partially mediated the relationship between age and planning depth (bootstrapped total effect 0.359, $p = 2.88{\cdot}10^{-6}$), again not fully accounting for the age-related changes seen in planning depth (bootstrapped direct effect 0.295, $p = 4.83{\cdot}10^{-05}$). Together, these results suggest that the age-related changes in heuristic quality and planning depth were not fully explained by individual variation in fluid reasoning, supporting the interpretation that they are unique measures.

As expected, self-reports showed an increase in daily life future orientation with age ($\rho = 0.374$, $p = 1.661{\cdot}10^{-6}$) and a decrease in impulsive choices, albeit to lesser extent ($\rho = -0.177$, $p = 0.027$). However, we did not find a relation between either measure and any task metrics (all $p \geq 0.156$) after controlling for age. This suggests that the Four-in-a-row model-derived metrics capture different cognitive processes than those indexed by these self-report measures.

Discussion

Measuring planning ability is notoriously complex as it relies on heuristics, forward reasoning, and attention (Ward and Morris, 2004). Here we used a novel task called "Four-in-a-row" and a computational model, previously used only in adult participants (van Opheusden & Ma, 2019)), to assess the developmental trajectories of these distinct cognitive components. Four-in-a-row has a larger state space than existing planning tasks, which makes it impossible to make a decision by reasoning about all the steps needed to reach a goal state, including reasoning backwards from a goal state. Thereby, Four-in-a-row is a planning task with increased ecological validity for real-world complex planning problems.

As expected, Four-in-a-row playing strength, as measured by Elo ratings, improved with age. By fitting a best-first-search computational model, we found distinct developmental trajectories of cognitive component processes that accounted for this improvement. Specifically, at younger ages (approximately early to mid-adolescence), we observed a rapid improvement in heuristic quality, which contributed to better playing strength. In contrast, planning ability showed stronger improvement and supported better playing strength at older ages (approximately mid to late-adolescence). Fewer attentional oversights were associated with better playing strength and did not show age differences. Together, these results suggest an order in which the use of cognitive component processes of planning develop into, during, and out of adolescence, starting by first refining the heuristic strategies, then gradually increasing the number possible future actions, states, and consequences considered towards young adulthood.

To further examine the uniqueness and novelty of the model-derived metrics, we assessed to what extent variation in planning depth and heuristic quality could be explained by individual differences in fluid reasoning (the capacity to apply logic to solve problems in new situations), future orientation, and impulsivity. We focused our analyses on planning depth and heuristic quality, as these showed age-related changes. As expected based on previous studies (Cattell, 1987; Ferrer & McArdle, 2004; Albert & Steinberg, 2011), fluid reasoning and future orientation increased with age, while impulsivity decreased. Neither planning depth nor heuristic quality was correlated with self-reported future orientation or impulsivity. Moreover, fluid reasoning only partially mediated the developmental changes in both planning depth and heuristic quality. These results complement a prior study in which fluid reasoning was found to be an important component process of model-based learning (Potter et al., 2017). Finding that fluid reasoning only partially mediated the relation between age and planning depth is consistent with the notion that planning crucially involves thinking ahead about future states, actions, and outcomes (i.e. planning depth), which is not part of fluid reasoning.

Our findings contribute to the growing literature on the development of model-based decision-making and reinforcement learning (for review see (Drummond & Niv, 2020)). Model-based decisions are more prevalent when playing Four-in-a-row than model-free decisions, as participants often encounter novel states in this task. Of all the states our participants encountered, on average 88.7% (sd = 2.0%) were unique states that the participant did not observe before. It is particularly important to understand age-related change in model-based decision-making, as several recent studies suggest that it shows a more protracted development than model-free learning (Decker et al., 2016; Palminteri, Kilford, Coricelli, & Blakemore, 2016; Potter et al., 2017) (but see (Smid, Kool, Hauser, & Steinbeis, 2020). Our findings make novel contributions to this literature by examining planning in a large state-space and revealing distinct developmental trajectories of the component processes that underlie planning.

Taken together, the current study investigated the developmental trajectories of heuristic quality, planning depth, and feature drop rate in a single planning task. Established assessments of complex planning traditionally rely on courser outcome measures such as accuracy or decision times, which would be affected by developmental changes in any of the underlying cognitive components, and thus unable to dissociate their contributions. In the future, our approach may be useful for the study of psychiatric disorders, as planning deficits are prevalent in a wide range of disorders, including ADHD, OCD, and schizophrenia (Harrier & DeOrnellas, 2005; Kofman, Gidley Larson, & Mostofsky, 2008; Nigg, Blaskey, Huang-Pollock, & Rappley, 2002; Morris, Rushe, Woodruffe, & Murray, 1995). Further research is needed to identify how our model-derived metrics relate to other cognitive mechanisms. Planning depth for example, likely relies on working memory to remember the consequences of possible moves (Gilhooly, 2005). Nevertheless, our findings move the field of cognitive development towards a more complete account of the development of planning and its component processes.

References

Albert, D., & Steinberg, L. (2011). Age differences in strategic planning as indexed by the tower of london. *Child development*, *82*(5), 1501–1517.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Elsevier.

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.

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.

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.

Drummond, N., & Niv, Y. (2020). Model-based decision making and model-free learning. *Current Biology*, *30*(15), R860–R865.

Elo, A. (1978). The ratings of chess players. *Past and Present, Arco Pub*.

Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental psychology*, *40*(6), 935.

Gilhooly, K. (2005). Working memory and planning. *The cognitive psychology of planning*, 71–88.

Harrier, L. K., & DeOrnellas, K. (2005). Performance of children diagnosed with adhd on selected planning and reconstitution tests. *Applied Neuropsychology*, *12*(2), 106–119.

Huys, Q. J., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS computational biology*, *8*(3), e1002410.

Kofman, O., Gidley Larson, J., & Mostofsky, S. H. (2008). A novel task for examining strategic planning: Evidence for impairment in children with adhd. *Journal of Clinical and Experimental Neuropsychology*, *30*(3), 261–271.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the united states. *Developmental neuropsychology*, *20*(1), 331–354.

McCormack, T., & Atance, C. M. (2011). Planning in young children: A review and synthesis. *Developmental Review*, *31*(1), 1–31.

Morris, R. G., Rushe, T., Woodruffe, P., & Murray, R. (1995). Problem solving in schizophrenia: a specific deficit in planning ability. *Schizophrenia research*, *14*(3), 235–246.

Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and dsm-iv adhd subtypes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(1), 59–66.

Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology*, *6*(1).

Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S.-J. (2016). The computational development of reinforcement learning during adolescence. *PLoS computational biology*, *12*(6), e1004953.

Potter, T. C., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, *25*, 272–280.

Smid, C. R., Kool, W., Hauser, T. U., & Steinbeis, N. (2020). Model-based decision-making and its metacontrol in childhood.

Steinberg, L., Graham, S., O'brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child development*, *80*(1), 28–44.

Steinberg, L., Sharp, C., Stanford, M. S., & Tharp, A. T. (2013). New tricks for an old measure: The development of the barratt impulsiveness scale–brief (bis-brief). *Psychological assessment*, *25*(1), 216.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Vaghi, M. M., Moutoussis, M., Váša, F., Kievit, R. A., Hauser, T. U., Vértes, P. E., … others (2020). Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. *Proceedings of the National Academy of Sciences*, *117*(41), 25911–25922.

van Opheusden, B., Galbiati, G., Kuperwajs, I., Bnaya, Z., & Ma, W. J. (2021). Revealing the impact of expertise on human planning with a two-player board game. *PsyArXiv*.

van Opheusden, B., & Ma, W. J. (2019). Tasks for aligning human and machine planning. *Current Opinion in Behavioral Sciences*, *29*, 127–133.