Title
A joint analysis of dropout and learning functions in human decision-making with massive online data

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

Authors
Kuperwajs, Ionatan
Ma, Wei Ji

Publication Date
2022

Peer reviewed

A joint analysis of dropout and learning functions in human decision-making with massive online data

Ionatan Kuperwajs (ikuperwajs@nyu.edu)
Center for Neural Science, New York University
New York, NY, United States

Wei Ji Ma (weijima@nyu.edu)
Center for Neural Science and Department of Psychology, New York University
New York, NY, United States

Abstract

The introduction of large-scale data sets in psychology allows for more robust accounts of various cognitive mechanisms, one of which is human learning. However, these data sets provide participants with complete autonomy over their own participation in the task, and therefore require precisely studying the factors influencing dropout alongside learning. In this work, we present such a data set where $1,234,844$ participants play $10,874,547$ games of a challenging variant of tic-tac-toe. We establish that there is a correlation between task performance and total experience, and independently analyze participants' dropout behavior and learning trajectories. We find evidence for stopping patterns as a function of playing strength and investigate the processes underlying playing strength increases with experience using a set of metrics derived from a planning model. Finally, we develop a joint model to account for both dropout and learning functions which replicates our empirical findings.

Keywords: dropout; learning; decision-making; behavioral modeling

Introduction

In recent years, psychology has trended towards massive data sets collected through online experiments (Stafford & Dewar, 2014; Mitroff et al., 2015; Steyvers, Hawkins, Karayanidis, & Brown, 2019; Holdaway & Vul, 2021). The purpose of this methodological shift is to obtain rich data in participants' real-world environments compared to the traditional practice of isolating a single variable and controlling for sources of variation in a laboratory setting. In turn, this data can be used to study a wide array of cognitive mechanisms from learning to decision-making to planning (Schulz et al., 2019; Steyvers & Schafer, 2020; Kuperwajs, van Opheusden, & Ma, 2019), or even for leveraging methods from machine learning to construct computational models (Agrawal, Peterson, & Griffiths, 2020; Peterson, Bourgin, Agrawal, Reichman, & Griffiths, 2021). These data sets have additional value in that they can clarify whether results or models derived from constrained laboratory tasks generalize.

One challenge that these kinds of data sets pose is temporal: participants now have complete autonomy over when and for how long to engage in the given task. In practice, this means that any investigation into learning using these data sets must also consider dropout behavior. In other words, when individuals drop out for reasons that are related to their current or future performance, their learning functions are directly biased. Modeling time to event data has a long history across many fields, notably survival analysis in statistics. Accounting for dropout behavior can be viewed through this lens by defining the hazard function as the probability that participants will stop participating in the task and asking which factors increase or decrease the probability of survival. In cognitive science, the existing literature on learning in massive data sets often ignores this type of analysis, with a notable exception extrapolating group learning policies for age-related differences in dropout in a large-scale learning study (Steyvers & Benjamin, 2019). Previous findings on curiosity and boredom argue that people are intrinsically motivated to participate in challenging tasks that provide new information while avoiding excessively simple or difficult tasks (Schmidhuber, 2010; Geana, Wilson, Daw, & Cohen, 2016; Ten, Kaushik, Oudeyer, & Gottlieb, 2020), and these elements may directly contribute to shaping hazard functions in large-scale data sets.

Here, we present a massive data set to investigate both dropout and learning functions in a combinatorial game of intermediate complexity, which people play on their mobile devices. We begin by characterizing the endpoint of participants' learning trajectories, establishing that there is a correlation between final playing strength and overall experience in our task. Then, we characterize the task components that drive dropout behavior and learning trajectories. In analyzing dropout behavior, we determine that people are more likely to stop playing when they have high playing strength. To more precisely study the learning process, we fit a planning model to people's choices and derive a set of metrics which lead to population level playing strength increases with experience while simultaneously mapping to distinct cognitive mechanisms. Finally, we construct a joint model of participants' dropout and learning functions which replicates our prior empirical results.

Task and data set

Our task is a variant of tic-tac-toe, in which two players alternate placing tokens on a 4-by-9 board (Figure 1A). The objective is to get four tokens in a row horizontally, vertically, or diagonally. The game, which we call 4-in-a-row, has approximately $1.2 \cdot 10^{16}$ non-terminal states, and therefore is at a level of complexity which far exceeds tasks commonly used in psychology (van Opheusden & Ma, 2019). We believe that a task for studying various components of cognition should fulfill multiple criteria:

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

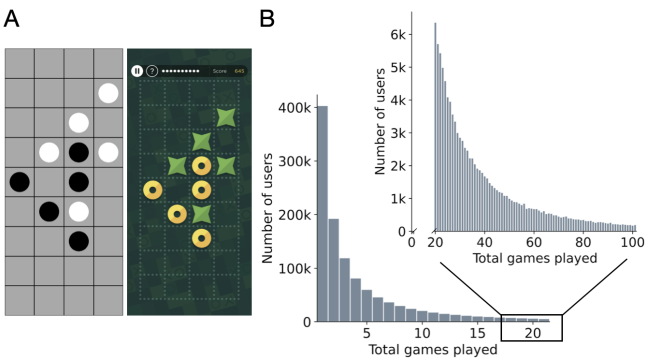

Figure 1: Task and data set. **(A)** Example board position in 4-in-a-row. The left is the laboratory version of the task, while the right is the gamified version used on the Peak platform. Two players, black and white or yellow circles and green stars, alternate placing pieces on the board, and the first player to connect four pieces in any orientation wins the game. **(B)** Histogram of the total number of games played by users in the data set. Note that the tail of the distribution, which consists of 20 to 100 total games played, still includes thousands of users.

1. The state space should be large enough so that participants continue to encounter states not previously experienced, and the task remains challenging.
2. The task should be novel, so that all participants are in the steep part of their learning curves.
3. The task should have simple rules, so that improvements are not due to participants learning the rules, but learning strategies.
4. The task should be engaging, so that participants remain motivated for many sessions.
5. The task should be amenable to computational modeling.

4-in-a-row satisfies these requirements by balancing complexity and computational tractability, making it an ideal candidate for studying expertise, planning, risk-taking, or dropout and learning, the latter of which we do in this paper.

Additionally, we partnered with Peak, a mobile app company, to implement a visually enriched version of 4-in-a-row on their platform (`https://www.peak.net`), which users play at their leisure in their daily environment. We are currently collecting data at a rate of approximately 1.5 million games per month, and here we analyze a subset consisting of 10,874,547 games from 1,234,844 unique users collected between September 2018 and April 2019. These users each play a wide range of games, but the data set includes thousands of users who have played upwards of 20 or even 100 games (Figure 1B). In this version of the task, users always play first against an AI agent implementing a planning algorithm, with parameters adapted from fits on previously collected human-vs-human games (van Opheusden et al., 2021).

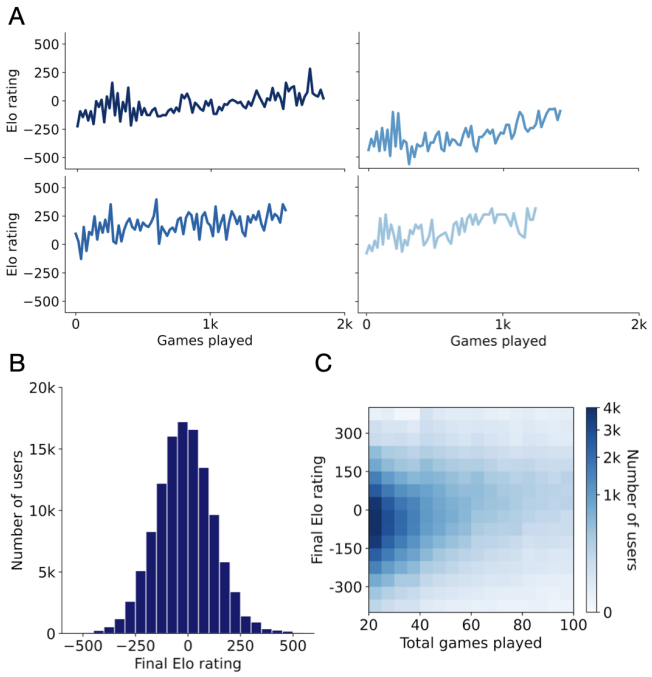

Figure 2: The relationship between task performance and experience. **(A)** Individual trajectories for 4 experienced users who played at least 1,000 games. **(B)** Histogram of the final Elo ratings for the 115,968 users who played at least 20 games in the data set. **(C)** Two-dimensional histogram of the final Elo rating and total number of games played for users in the data set. The visualization is limited to 104,681 users who had a final Elo rating between −400 and 400 and played at least 20 and less than 100 total games.

AI agent playing strength is modulated by changing model parameters based on game outcomes.

Results

Characterizing the endpoint of learning

In order to motivate the rest of our findings, we first investigate the relationship between task performance and total experience. We measure users' task performance using Elo ratings (Elo, 1978), a standard method for calculating the relative skill levels of players in zero-sum games such as chess. Ratings calculated for relatively few games can be statistically unreliable, so we exclude players with less than 20 total games played from our analysis and group each user's experience into blocks of 20 games. This results in 115,968 unique users, and we use a common baseline to compute Elo ratings across all experimental data. Figure 2A shows the full individual learning trajectories of a few experienced users, and Figure 2B the distribution of users' Elo ratings in the last full block of gameplay. In Figure 2C, we correlate this final playing strength with the total number of games played by each user ($\rho = 0.270$). Due to the size of the data set, our p-values are below the minimum representable float ($2 \cdot 10^{-308}$) unless

reported otherwise. This result illustrates that, at the endpoint of their learning trajectories, users are more likely to have higher task performance if they have accumulated more total experience with the task. In the following sections, we break down the factors of dropout and learning that contribute to this relationship between task performance and experience.

Dropout behavior

We hypothesize that dropout behavior in this game is strongly influenced by two factors: current playing strength and current number of games played. In order to test this hypothesis, we examine the effect of each user's Elo rating in a given window of 20 games on the probability that they stop playing in the next block of 20 games, and further bin these probabilities by number of games that the user has played so far (Figure 3). We find that as users play more games they have a lower stopping probability (logistic regression: $\beta = -0.020 \pm 0.058 \cdot 10^{-3}$), and their stopping probability increases with higher Elo ratings (logistic regression: $\beta = 0.631 \cdot 10^{-3} \pm 0.010 \cdot 10^{-3}$). Logistic regression models have also been used extensively in survival analysis to characterize hazard functions (Cox, 1972), albeit in our application we are using the number of games played as a proxy for time. This finding suggests that as users gain more experience, they are more likely to continue playing when the game is still challenging and they have lower Elo ratings. Conversely, users tend to quit when their playing strength is high. This aligns with the literature on intrinsic motivation in humans: if users devalue the task due to lack of a challenge or information gain, they will stop participating (Schmidhuber, 2010; Geana et al., 2016; Ten et al., 2020). However, we do not find evidence for the inverse trend that users find the task too difficult and are therefore more likely to stop playing with lower Elo ratings. One possible explanation for this is that even our strongest AI agents did not make the task difficult enough for users with high playing strengths.

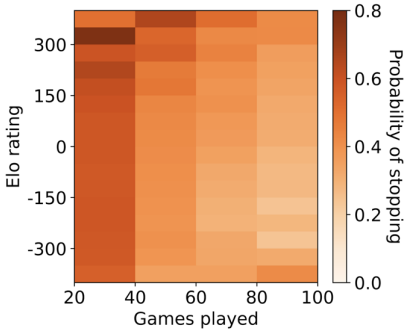

Figure 3: Dropout behavior is driven by recent playing strength and number of games played. Probability of stopping during the next block of 20 games as a function of the Elo rating and number of games played so far in the current block, again limited to the same ranges as in Figure 2C.

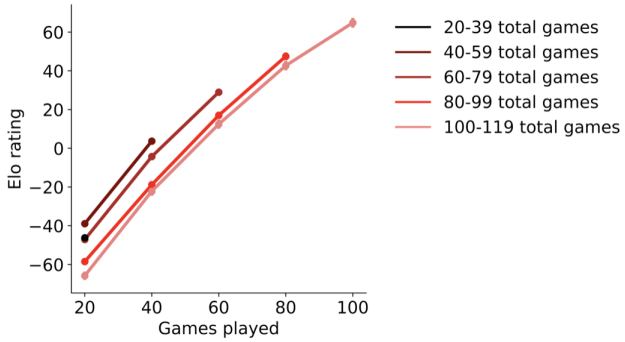

Figure 4: Playing strength increases during learning. Average user Elo rating as a function of current experience level conditioned on total number of games played. This is computed for the set of 107,769 users who played at least 20 and less than 120 total games.

Learning trajectories

To make more precise claims regarding the factors underlying users' learning trajectories, we first validate that a reliable increase in playing strength over time occurs at the population level. In Figure 4, we show that average Elo ratings increase as users gain more experience, and that this trend occurs irrespective of the total number of games each user ends up playing (linear regression: $\beta = 1.500 \pm 0.893 \cdot 10^{-2}$). We also find a reliable, albeit smaller, effect of initial Elo ratings on current playing strength (linear regression: $\beta = 0.664 \pm 0.160 \cdot 10^{-2}$). Note that the correlation from Figure 2 can be observed by connecting the endpoints of each learning curve. Importantly, we find no evidence for changes in the slopes of users' average learning trajectory when conditioned on either total number of games played or initial playing strength. This suggests that learning rates in this task are independent of these two factors, and primarily captured by current playing strength, current experience level, and individual differences.

To investigate which aspects of people's decision-making process underlie this performance increase, we fit a planning model to users' choices in the task and derive a set of metrics from the model's parameters (van Opheusden et al., 2021). This algorithm combines a heuristic function (Figure 5A), which is a weighted linear combination of board features (Campbell, Hoane Jr, & Hsu, 2002), with the construction of a decision tree via best-first search (Figure 5B). Best-first search iteratively expands nodes on the principal variation, or the sequence of actions that lead to the best outcome for both players given the current decision tree (Dechter & Pearl, 1985). To allow the model to capture variability in human play and make human-like mistakes, we add Gaussian noise to the heuristic function and include feature dropout. For each move the model makes, it randomly omits some features from the heuristic function before it performs search. Such feature omissions can be interpreted cognitively as lapses of selective attention (Treisman & Gelade, 1980). During search, the

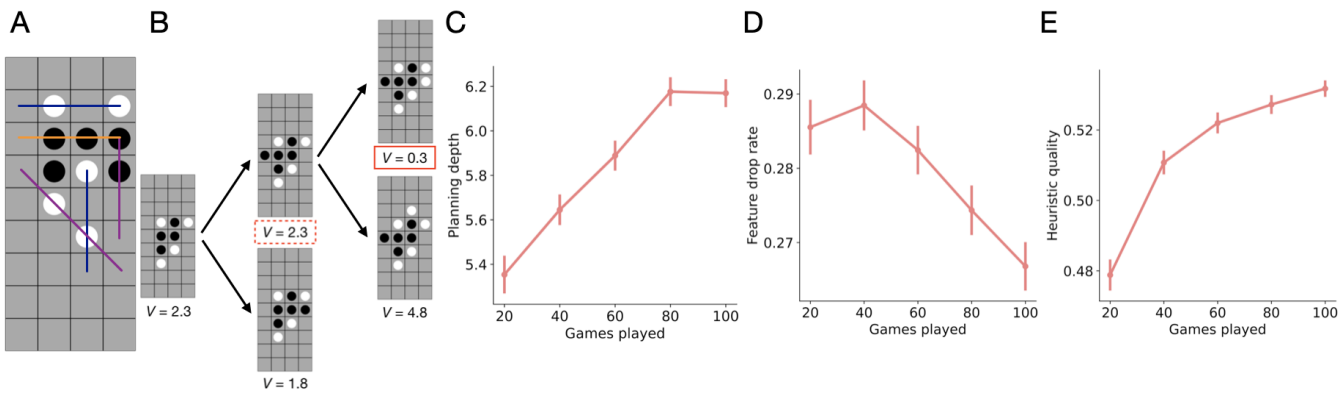

Figure 5: Using a planning model to investigate components of learning. **(A)** Features used in the heuristic function. Features with identical colors are constrained to have the same weights. **(B)** Illustration of the best-first search algorithm. In the root position, black is to move. After expanding the root node with two candidate moves for black and evaluating the resulting positions using the heuristic function, the algorithm selects the highest value node ($V = 2.3$) on the second iteration and expands it with two candidate moves for white. The algorithm evaluates the resulting positions, and backpropagates the lowest value ($V = 0.3$), since white is the opponent. That value will be compared against alternatives in each intermediate node of the tree to decide in which direction to expand the tree on the next iteration. **(C)** Average depth to which users plan as they gain experience, as estimated by the planning model. **(D)** Same as (C) for feature drop rate. **(E)** Same as (C) for heuristic quality.

model also prunes the decision tree by removing branches with low heuristic value (Huys et al., 2012).

Due to computational constraints, we analyze data from 1,000 pseudo-randomly selected users who played at least 100 games each. We estimate the log probability of a user's move in a given board position with inverse binomial sampling (van Opheusden, Acerbi, & Ma, 2020), optimize the log-likelihood function with Bayesian adaptive direct search (Acerbi & Ma, 2017), and account for potential overfitting by reporting 5-fold cross-validated log-likelihoods. We convert the set of parameters inferred for each user in each block to three metrics: planning depth, feature drop rate, and heuristic quality. Planning depth roughly corresponds to the number of steps people think ahead, feature drop rate measures the frequency of attentional lapses, and heuristic quality measures the "correctness" of the feature weights. Figure 5C-D show that depth of planning increases with experience (linear regression: $\beta = 0.011 \pm 0.796 \cdot 10^{-3}$), while feature drop rate decreases (linear regression: $\beta = -0.258 \cdot 10^{-3} \pm 0.036 \cdot 10^{-3}$, $p = 4.49 \cdot 10^{-13}$). We verify that users' response time decreases across blocks, signifying that the planning depth increase is not a result of slower play. Figure 5E also shows a reliable increase in heuristic quality with experience (linear regression: $\beta = 0.611 \cdot 10^{-3} \pm 0.029 \cdot 10^{-3}$). However, the heuristic quality in the first 20 games of this data set is much lower than in previously collected laboratory data, suggesting that users have more opportunity to improve their feature weights rather than starting at ceiling. These results demonstrate that users' learning trajectories are underscored by deeper planning, fewer lapses of attention, and bounded improvement of feature weights.

Joint modeling of dropout and learning

Given our insights into the factors driving dropout and learning functions in this task, we constructed a computational model that replicates our previous findings (Figure 6A). The intuition for the model is that after each additional game played, users receive a new Elo rating which increases or decreases based on their individual learning rate. Based on that rating and the number of games they have played thus far, they decide whether to continue playing or drop out. In the former case, users again adjust their rating and make a decision on whether to stop, and in the latter case the model outputs their final playing strength and total number of games played. Additionally, we make the following assumptions: each user begins with their computed initial Elo rating after 20 games, has an underlying true learning rate which is consistent across gameplay, and gains a noisy amount of playing strength per game until they drop out. This results in three parameters of interest, which are the learning rate α, the variance of the noise on Elo rating increase per game σ^2_{noise}, and the stopping probability P_{stop}.

To fit this model to data, we decompose the problem into two parts. First, we take each user's learning trajectory, or Elo rating as a function of games played, and perform a linear regression per user. The learning rate parameter α for each user is then drawn once for each user from a normal distribution with its mean at the weighted average of the user slopes and its variance at the variance of the user slopes. The noise on playing strength increases is drawn from a normal distribution per iteration with its mean at 0 and its variance at the variance of the residuals. This gives us an update rule for

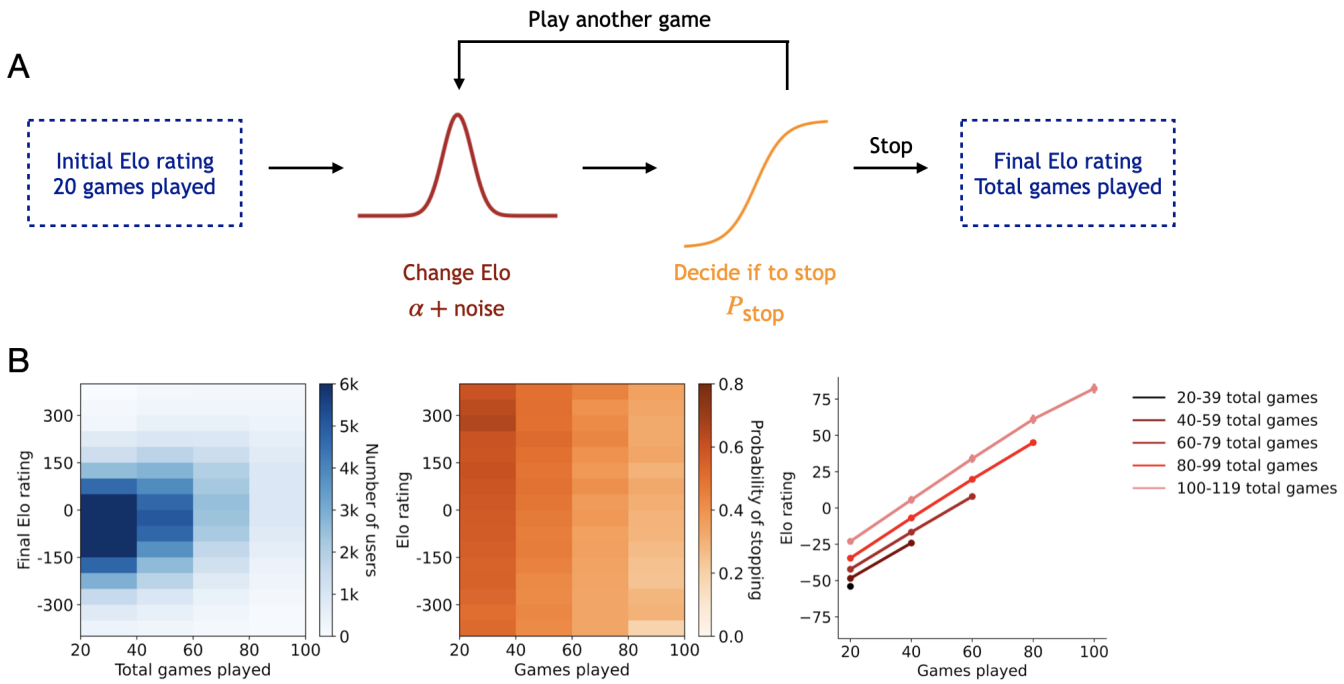

Figure 6: A joint model of dropout and learning. **(A)** Illustration of the model, which receives initial playing strengths for each user after 20 games played as input and returns a final playing strength and total number of games played. Each simulated game changes playing strengths by individual learning rates drawn from a normal distribution with added noise, and determines whether each user drops out based on probabilities from a logistic function. A linear regression on individual learning trajectories computes the mean and variance of the learning rates, and a logistic regression that takes into account current playing strengths and number of games played computes the coefficients for the logistic function. **(B)** The model replicates the trends found in the data, including the two-dimensional histogram of the final Elo rating and total number of games played from Figure 2C (left), the probability of stopping during the next block of 20 games as a function of the Elo rating and number of games played in the current block from Figure 3 (middle), and the average user Elo rating as a function of current experience level conditioned on total number of games played from Figure 4 (right). Results are simulated for the set of 115,968 users who played at least 20 games in the data set.

ratings r where η is standard normal noise:

$$r \leftarrow r + \alpha + \sigma_{\text{noise}}\eta. \tag{1}$$

Second, we utilize the coefficients from the logistic regression on users' dropout behavior to compute the stopping probability for each user in each simulated game based on current ratings and experience n.

$$P_{\text{stop}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 n + \beta_2 r)}}. \tag{2}$$

After we fit these parameters, we run the model forward to simulate our empirical results. As our Elo ratings are restricted to blocks of 20 games, we constrained our predictions to these same chunks. Figure 6B shows that the correlation between final playing strengths and total number of games played, the probability of dropout as a function of current playing strength and number of games played, and the population level learning trajectories conditioned on total experience can all be replicated reliably by the model.

Discussion

In this paper, we leveraged a large-scale data set of human participants playing a two-player combinatorial game in order to interrogate their dropout and learning functions. We first established that playing strength and overall experience are correlated before characterizing the factors that underlie dropout and learning. For dropout, we found that current playing strength and experience level drive stopping probabilities. For learning, we demonstrated that playing strength increases with the number of games played, and that a planning model can attribute these improvements to increased planning depth, decreased feature dropping, and bounded increase in heuristic quality. Finally, we combined these components into a joint model of dropout and learning that is able to reproduce the patterns that we found in the data.

While our results and modeling provide a baseline for understanding dropout and learning in 4-in-a-row, there are additional factors we could analyze to more thoroughly explain human behavior. In terms of dropout, we primarily investigate user playing strengths rather than opponent playing

strengths. For example, it is feasible that stopping probabilities are influenced by which AI agent is being played against, or whether users are stuck in a staircase of AI playing strengths. If we were to find that opponent playing strengths contribute to dropout, either because they are too challenging or not challenging enough to play against, this could be easily incorporated into our model with an additional mechanism, such as that user Elo ratings must increase by a fixed amount over a set period of time or by changing the shape of the stopping function. Further, we could study the interval between games played by each user, as there may exist a complex interaction between these inter-game intervals, playing strength, and dropout. Given that some users play many games in a short period of time while others play games spaced apart over days or even weeks, replacing the number of games played with real-world time could strengthen the parallel between our work and survival analysis models. In terms of learning, we have assumed static learning rates thus far and can also consider variations such as nonlinear models for the number of games played or learning rates that vary with experience. Additionally, our results from the planning model show how experienced participants differ from novices, but do not shed light on how those differences emerge from their experience. In future work, we aim to model this learning process more explicitly by matching people's experience with their future actions. These results can then be integrated into a more sophisticated model of dropout and learning which accounts for additional components of each cognitive process.

Would our results generalize to other tasks or real-world environments? Given that our dropout results are consistent with the literature on intrinsic motivation, we suspect that some form of performance-mediated stopping occurs in most tasks for which participation is autonomous, and that the thresholds for dropout vary based on the nature and difficulty of the task. We also speculate that the same effects of experience on search and attention will exist in tasks for which planning ahead is beneficial, and that the accuracy of feature weights will depend on the amount of domain-specific knowledge that is required. Games like chess and Go contain many non-trivial features, and tasks with stochastic environments might involve distinct computational mechanisms altogether. If the factors underlying dropout and learning can be reliably identified, our joint modeling framework should be relatively straightforward to adapt to other tasks. More broadly, a general learning model that decides whether or not to continue engaging with a task can be combined with a more task-specific model that actually performs the task itself. While our two models are currently fairly distinct, bringing together these modular components is a more concrete step forward for understanding how humans cognitively navigate such complex decisions.

One of the main contributions of our work is to advocate for concepts from survival analysis and human studies on motivation to be integrated into the analysis of learning in massive cognitive science data sets. Our use of a hazard function and logistic regression to study dropout as well as intuition for factors that mediate dropout and learning, such as playing strength as a proxy for task difficulty, are derived from these fields. As the use of large-scale data sets where participants have autonomy over participation become more ubiquitous, we hope that it will become standard for accounts of human behavior to model motivation and learning simultaneously.

Acknowledgments

This work was supported by Graduate Research Fellowship number DGE183930 from the National Science Foundation, grant number IIS-1344256 from the National Science Foundation, and grant number R01MH118925 from the National Institutes of Health.

References

Acerbi, L., & Ma, W. J. (2017). Practical bayesian optimization for model fitting with bayesian adaptive direct search. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 1834–1844).

Agrawal, M., Peterson, J. C., & Griffiths, T. L. (2020). Scaling up psychology via scientific regret minimization. *Proceedings of the National Academy of Sciences*, *117*(16), 8825–8835.

Campbell, M., Hoane Jr, A. J., & Hsu, F.-h. (2002). Deep blue. *Artificial intelligence*, *134*(1-2), 57–83.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202.

Dechter, R., & Pearl, J. (1985). Generalized best-first search strategies and the optimality of a. *Journal of the ACM (JACM)*, *32*(3), 505–536.

Elo, A. E. (1978). *The rating of chessplayers, past and present*. Arco Pub.

Geana, A., Wilson, R., Daw, N. D., & Cohen, J. (2016). Boredom, information-seeking and exploration. In *Cogsci.*

Holdaway, C., & Vul, E. (2021). Risk-taking in adversarial games: What can 1 billion online chess games tell us?

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.

Kuperwajs, I., van Opheusden, B., & Ma, W. J. (2019). Prospective planning and retrospective learning in a large-scale combinatorial game. In *2019 conference on cognitive computational neuroscience. berlín, alemania* (pp. 13–16).

Mitroff, S. R., Biggs, A. T., Adamo, S. H., Dowd, E. W., Winkle, J., & Clark, K. (2015). What can 1 billion trials tell us about visual search? *Journal of experimental psychology: human perception and performance*, *41*(1), 1.

Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, *372*(6547), 1209–1214.

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, *2*(3), 230–247.

Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences*, *116*(28), 13903–13908.

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological science*, *25*(2), 511–518.

Steyvers, M., & Benjamin, A. S. (2019). The joint contribution of participation and performance to learning functions: Exploring the effects of age in large-scale data sets. *Behavior research methods*, *51*(4), 1531–1543.

Steyvers, M., Hawkins, G. E., Karayanidis, F., & Brown, S. D. (2019). A large-scale analysis of task switching practice effects across the lifespan. *Proceedings of the National Academy of Sciences*, *116*(36), 17735–17740.

Steyvers, M., & Schafer, R. J. (2020). Inferring latent learning factors in large-scale cognitive training data. *Nature Human Behaviour*, *4*(11), 1145–1155.

Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2020). Humans monitor learning progress in curiosity-driven exploration.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97–136.

van Opheusden, B., Acerbi, L., & Ma, W. J. (2020). Unbiased and efficient log-likelihood estimation with inverse binomial sampling. *PLoS computational biology*, *16*(12), e1008483.

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

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