# CogSimulator: A Model for Simulating User Cognition & Behavior with Minimal Data for Tailored Cognitive Enhancement

**Weizhen Bian[1,4], Yubo Zhou[2], Yuanhang Luo[3], Ming Mo[5], Siyan Liu[4], YiKai Gong[2], Renjie Wan[2], Ziyuan Luo[2], Aobo Wang[4†1]**

[1] **Department of Computer Science, The Hong Kong University of Science and Technology**

[2] **Department of Computer Science, Hong Kong Baptist University**

[3] **Department of Applied Mathematics, The Hong Kong Polytechnic University**

[4] **Department of Systems Science, National University of Singapore**

[5] **Department of Methodology, The London School of Economics and Political Science**

## Abstract

The interplay between cognition and gaming, notably through educational games enhancing cognitive skills, has garnered significant attention in recent years. This research introduces the CogSimulator, a novel algorithm for simulating user cognition in small-group settings with minimal data, as the educational game Wordle exemplifies. The CogSimulator employs Wasserstein-1 distance and coordinates search optimization for hyperparameter tuning, enabling precise few-shot predictions in new game scenarios. Comparative experiments with the Wordle dataset illustrate that our model surpasses most conventional machine learning models in mean Wasserstein-1 distance, mean squared error, and mean accuracy, showcasing its efficacy in cognitive enhancement through tailored game design.

**Keywords:** Artificial Intelligence, Education, Group Behaviour, Skill acquisition and learning, Computational Modeling.

## Introduction

In recent years, the application of artificial intelligence has expanded across various fields such as music (Bian et al., 2023), gaming (Yin et al., 2023), and healthcare (Shaheen, 2021), demonstrating its broad impact and potential. Concurrently, the relationship between cognition and games has become a hot topic, particularly in research focused on using games to assess cognitive abilities and explore whether games can enhance cognitive skills (Boot, 2015). At the same time, the relationship between cognitive level and education is also the focus of attention, where cognitive ability is considered a crucial predictor of education and socioeconomic achievement, especially regarding strategies and methods in student learning, individual differences, etc (Yen et al., 2004; Van Hove et al., 2019).

In terms of using games to help measure cognitive ability, games can test different aspects of cognitive function, including aspects such as memory and attention, which can assist in diagnosing disease (Wiley et al., 2021). At the same time, games are thought to be more effective in assessing cognitive abilities, even in improving fairness and user experience (Leutner et al., 2023). Therefore, games are closely linked to cognitive ability, and in the framework of the importance of cognitive ability in education, the better use of games to contribute to the development of cognitive ability is a significant agenda. However, given the heterogeneous attributes of participants, a fixed game algorithm may not enhance the cognitive abilities of all individuals consistently (Manzano-León et al., 2021). Conventional approaches, encompassing behavior trees (Hecker, 2011) and data-driven methodologies (Kim & Ruipérez-Valiente, 2020), attempt to tailor game modes to diverse cognitive profiles of users yet necessitate substantial data acquisition. Technologies related to artificial intelligence, including Transfer Learning and Convolutional Neural Networks (CNN), are also challenged by over-fitting issues, consequently diminishing the precision in predicting user cognitive patterns (Zhao, 2017).

This work developed the CogSimulator model, which intends to capture and simulate the user's cognitive level based on a small amount of data. Further, using less data to simulate user groups, the game can implement a more targeted educational game for specific users. This paper uses the game "Wordle", a word game, to demonstrate how to evaluate and improve cognitive ability with a limited number of users, making the game a more acceptable form for teenagers in learning (Amory, 2010). It is important to highlight that due to the CogSimulator's efficient use of data, encompassing common word frequency metrics pervasive across numerous word-based games, it is well-suited for application to newly developed games or those aimed at niche user groups. By simulating individual player behaviors, the model offers game designers a valuable tool for tailoring game difficulty to match specific cognitive profiles.

## Related Work

Cognition includes the mental processes of acquiring knowledge and understanding through thought, experience, and the senses and plays a fundamental role in human development and interaction with the environment (Sommerville, 2020). This area, crucial for critical thinking, problem-solving skills, and the effective processing and interpretation of information, has attracted increasing attention in contemporary research, particularly at the intersection of cognition and play (Majuri et al., 2018; Sailer & Homner, 2020). Human research on educational games dates back to 1981 when Malone investigated in his seminal paper how to use the captivating effects

[1†] Corresponding author. Email: aobo.wang@nus.edu.sg



In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* 

of computer games to make learning fun and interesting (Malone, 1981). In the mid-1980s, research examined the link between video game play and cognitive performance, and this correlation was gradually confirmed (Dale et al., 2020). However, even using AI-based game-based educational technologies, such as AI applications for learning new skills or knowledge, may only improve learners' cognition if they motivate long-term use (Laine & Lindberg, 2020). Therefore, the challenge is producing a game design that suits the needs and preferences of the players to ensure the gamers' cognitive enhancement.

Many academic studies have shown that positive motivation is crucial in enhancing human engagement, which in turn helps promote cognitive improvement (Teixeira et al., 2012; Ryan & Deci, 2000; Dweck, 2006). In this context, it can be broken down participation motivation into two categories: intrinsic participation motivation (stemming from the game's intrinsic design elements) and extrinsic participation motivation (stemming from the game's reward and punishment system) (Laine & Lindberg, 2020). Although the persistence of intrinsically motivated engagement is remarkable, intrinsic motivation depends on various factors, including player type, specific educational needs, and personal interests (Manzano-León et al., 2021). Therefore, if the game mode cannot change with different individuals, the degree of cognitive enhancement may vary widely between individuals. This variability poses a significant challenge to developing fixed game mechanics that can universally meet the diverse needs of all players (Manzano-León et al., 2021).

Traditional approaches to creating adaptive and responsive gaming environments, designed to cater to players' individual needs, have primarily relied on rule-based systems such as finite state machines and behavior trees (Park et al., 2023). These systems provide an easy way to build simple agents that provide different feedback based on user actions (Hecker, 2011). However, while this approach effectively creates a baseline interactive experience, it cannot dynamically adapt to individual players' subtle and changing preferences or capabilities. An alternative solution lies in data-driven approaches (Kim & Ruipérez-Valiente, 2020). This approach entails collecting extensive gameplay data, such as average completion times, to assess game difficulty and subsequently recommend games of varying difficulty levels to different players. However, this approach relies on substantial data accumulation for accurate predictions, making it challenging to apply effectively in scenarios requiring rapid adaptation to new tasks. As for artificial intelligence technology, most of the gamification research that appeared in the past decade from 2010 to 2020 failed to provide a structured overview of game elements such as NPCs working with artificial intelligence (Funk et al., 2020). Until 2023, Generative Agents realized the simulation of human behavior based on large language models (LLMs), even in zero-shot scenarios (Park et al., 2023). However, the need for extensive resources poses a significant obstacle, especially for small educational games in their embryonic stages (Jozefowicz et al., 2016). For these emerging games, the high threshold of data and computing requirements makes it challenging to fully exploit the potential of AI-driven interactive elements. In contrast, models such as CNN may cause overfitting problems due to insufficient data volume. Therefore, while advances in artificial intelligence technology offer promising avenues for enhancing the realism and engagement of game environments, their applicability still needs to be improved in resource-constrained settings.

This work will elucidate and test the model using Wordle as an illustrative case. Wordle is a word-guessing game that epitomizes its players' diverse needs and individual characteristics, reflecting the unique responses and strategies each person brings to the game. The goal of Wordle is to guess a five-letter word within six attempts. After each guess, the color gives feedbacks as: Green (the correct letter in the right spot), Yellow (The correct letter but in the wrong position) or Grey (the letter outside the word) (Match, 2022). Unaided players guess words mainly through word recall, largely limited by their vocabulary. Consequently, the CogSimulator was developed to emulate user cognition utilizing a limited dataset of game records, thereby facilitating the design of game difficulties optimized for cognitive enhancement.

## CogSimulator Model

### Model Overview

We analyze that traditional machine learning models often face severe overfitting due to insufficient sample sizes, limiting their effectiveness. In contrast, our model excels in explaining word difficulty and result distribution, offering a more robust solution as detailed in Table 1. Attributes such as word frequency, prevalent across various word-based games, provide a solid foundation for applying our model to different gaming scenarios. To address these challenges, we have developed a novel sampling simulator that better reflects the gameplay dynamics of the broader population, as illustrated in Figure 1.

The Cogsimulator operates through a process analogous to Markov Chain Monte Carlo (MCMC) as described by Geyer (Geyer, 2011), where parameters progressively converge to a steady state distribution and achieve detailed balance. The model incorporates hyperparameters that capture players' cognitive processes at each stage of their guessing attempts, alongside the stochastic variations observed in each trial. Unlike traditional optimization methods that require derivatives of the cost function, our approach uses Coordinate Search Optimization (Frandi & Papini, 2014), which optimizes parameters coordinate-wise within the hyperparameter space in each iteration, allowing the model

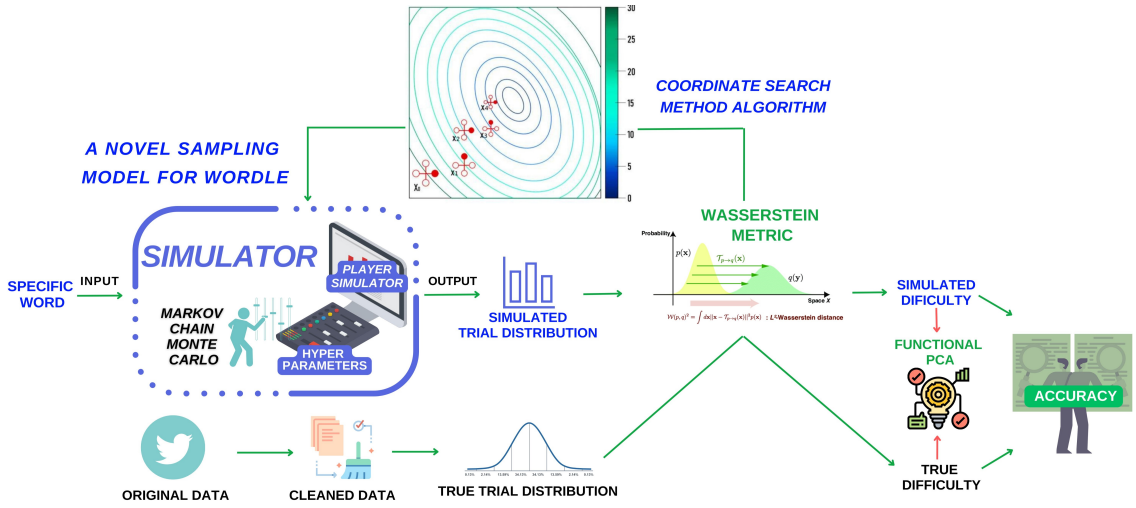


Figure 1: An overview of the CogSimulator model. It starts by aggregating Wordle data from Twitter and using it to form a true trial distribution for training. The simulator, via Markov Chain Monte Carlo, adjusts hyperparameters to match this distribution. Parameter optimization is performed with a Coordinate Search Method Algorithm, and model accuracy is ensured using the Wasserstein-1 metric to compare simulated outputs to actual trials. Functional PCA relates simulated and true word difficulties, enabling precise predictions for new word trials based on players' cognitive patterns.

to align with the actual trial distributions observed in data eventually. To measure the deviation between our training set and the outputs generated by the simulation, we employ the Wasserstein-1 distance (Liao et al., 2022), whose convex properties aid in steering the algorithm towards the global optimum. This setup ensures our model accurately estimates word difficulty and adeptly classifies the distribution of players' guessing attempts. Once optimized, the sampling simulator effectively replicates samples that reflect the average performance of the simulated player population.

Further refining our approach, we estimate the difficulty of newly introduced words by comparing their Wasserstein distance with that of the most straightforward word identified in the training dataset. This comparison uses the generated Wasserstein Metric to provide a relative measure of word complexity. Consequently, when a new word is introduced to the model, it predicts not only the difficulty of the word but also the expected number of guesses by players. This prediction is based on the collective cognitive profile of the player group derived from the training data, thereby aiding in selecting words that are optimally challenging yet engaging for players.

## Simulator

As the core component of CogSimulator, the simulator uses a coordinate search method algorithm to tune hyperparameters to match human performance in player simulators automatically. The simulator takes a single word as input and can generate a distribution of the number of times a user guessed the word, containing the percentage of attempts for each of the seven trial types. Other factors will be incorporated into the neural network alongside word guessing frequency distribution. This integration will occur through a process of 5-fold cross-validation to ensure a better fit for various real-world factors, thereby yielding the most accurate predictions. To determine whether the player selected a specific word in a step, given a qualified dictionary that allows him to choose, algorithms 1 and 2 output the words sampled in a round.

**Algorithm 1:** Choose A Word algorithm $\mathcal{CW}$

**Data:** Dictionary $\mathcal{D} \ni \{A_i\}^N$ with word frequency $\{p(A_i)\}^N$
**Given hyper-parameters:** $K, T$ ;
**Initialise:** Probability list of length $N$:
PL $= [0, 0, \ldots, 0]$;
**while** $i \leq N$ **do**
  **update** PL$[i] \leftarrow \max(0, p(A_i) \times T)$ ;
  Delete words from $\mathcal{D}$ according to Wordle clue rule (colored clue);
  **update** PL$[i] \leftarrow$ normalise(PL) ;
  **update** Chosenword $\leftarrow$ randomly choose a word from $\mathcal{D}$ with probability PL;
**end**
**Return** Chosenword

In simulating Wordle player choices, the *CogSimulator* posits that players are more inclined to guess words they encounter more frequently. However, the human capacity to recall words is finite, and the probability of recalling a specific word is not strictly proportional to its word frequency.

**Algorithm 2:** Single word Trial Simulation Algorithm

**Data:** Any target word $A^*$, dictionary $\mathcal{D} \ni \{A_i\}^N$ with word frequency $\{p(A_i)\}^N$
**Given hyper-parameters:** $K, T$ ;
**Initialise:** Count of steps: cnt = 1;
**Set:**Chosenword = $C\mathcal{W}(\mathcal{D}^*, \{p(A_i)\}^N, K, T)$;
**while** Chosenword $\neq A$ **do**
  **update** $\mathcal{D}^* \longleftarrow$ Delete words from $\mathcal{D}$ according to Wordle clue rule (colored clue);
  **update** Chosenword = $C(\mathcal{D}, \{p(A_i)\}^N, K, T)$ ;
  cnt = cnt+1 ;
**end**
**Return** cnt

Thus, we introduce two parameters to adjust for this. The first parameter, $K$, represents the cognitive limit, or the maximum number of words a person can typically remember at one time. The simulator only considers up to $K$ most frequent words as viable options for player guesses. The second parameter, $T$, represents a scaling factor for the frequency of the most common word, providing a baseline for comparison. To calculate the selection probability for each word, we multiply its frequency by $T$ and then normalize by dividing over the sum of the scaled frequencies for all $K$ words considered. This method ensures that while the likelihood of selecting highly frequent words is amplified, less common words maintain a non-zero probability of selection proportional to their relative frequency. This nuanced approach balances the natural human tendency to favor familiar words and the game's challenge to recall less frequent words.

On each trial, the simulator randomly samples words based on their probability of selection (El-Hadidy et al., 2021). After 1000 samples, a trial distribution of words is generated. At the same time, a coordinate search algorithm will be applied to automatically adjust hyperparameters to improve the fitting accuracy of the simulator. Finally, the simulator can generate seven probability distributions of word guesses containing the percentages for seven categories (1, 2, 3, 4, 5, 6, X). Finally, the model can predict the difficulty of a word for a specific cognitive user through distribution and Wasserstein Metric.

## Wasserstein Metric

To test the difficulty of a word for a particular cognitive group, we first obtain the distribution of word guessing times for this cognitive group's best record in past games. Then, by comparing the distance between the distribution of new input words and the distribution of the best record, the difficulty of the word for this cognitive group is judged. For this purpose, we propose using the Wasserstein-1 distance to evaluate the discrepancy between two trial distributions. Let $p, q$ be two probability distributions on compact spaces. Denote $\Pi(p,q)$ as the set of all distributions $\pi(\omega, \omega')$ on $\mathcal{X} \times \mathcal{X}'$ such that the marginals are $p(x)$ and $q(y)$ respectively. Then the Wasserstein-1 distance between $p$ and $q$ is

$$\begin{aligned} W_1(p,q) &= \inf_{\pi \in \Pi(p,q)} \int_{\mathcal{X}\times\mathcal{X}} \|\omega - \omega'\| d\pi(\omega, \omega') \\ &= \inf_{\pi \in \Pi(p,q)} \mathop{\mathbb{E}}_{(\omega,\omega')\sim\pi} [\|\omega - \omega'\|] . \end{aligned} \tag{1}$$

when 1 applies to discrete sample spaces, let us assume $\mathcal{X} = \{\omega_i\}_{i=1}^m$ and $\mathcal{X}' = \{\omega'_i\}_{i=1}^{m'}$. $p$ and $q$ are trial distributions of words $A$ and $A*$, respectively. The distance between $p$ and $q$ can be obtained by solving the following linear programming problem

$$\begin{aligned} \mathcal{W}(A, A*) = \inf_{\{\gamma_{i,j}\}, i, j} \Bigg\{ \sum_{i=1}^{m} \sum_{j=1}^{m'} \gamma_{i,j} |\omega_i - \omega'_j| : \\ \sum_{i=1}^{s} \gamma_{i,j} = q_j, \sum_{j=1}^{s'} \gamma_{i,j} = p_i, \gamma_{i,j} \geq 0 \Bigg\}. \end{aligned} \tag{2}$$

In this special case where the support of any trial distribution is the same, *i.e.* $m = m'$ with uniform weights, it can be easily shown that Wasserstein-1 distance has a nice closed and compact form $\frac{1}{m}\sum_{i=1}^{m} |\omega_{\eta(i)} - \omega'_{\vartheta(i)}|$ where $\eta$ is a sorting permutation of $\omega_i$ and $\vartheta$ is a sorting permutation $\omega'_j$. The difficulty of word $A$ is determined by calculating its Wasserstein-1 distance w.r.t trial distribution of the 'easiest' in the dataset, which we assume is the word "train". Therefore, we propose the Wasserstein-1 difficulty measure of the form.

$$\mathcal{W}^*(A) := \mathcal{W}(p_{\text{"train"}}, p_A). \tag{3}$$

## Coordinate Search Optimization

To obtain parameters to simulate the cognition of the target population, we used Coordinate Search Optimization. The algorithm is as follows:

**Algorithm 3:** Coordinate Search algorithm.

**Data:** Ground truth trial distribution $f_{A_i}$,data set word ${A_i}^{355}$, dictionary $\mathcal{D} \ni \{A_i\}^N$ with word frequency $\{p(A_i)\}^N$
**Initialise:** Hyper-parameters: $K_0, T_0$;
**while** $j \leq$ *maximum iteration (or not converge)* **do**
  **update** $K_j \longleftarrow \arg\min_{K^*} f(K^*, T_{j-1})$ ;
  **update** $T_j \longleftarrow \arg\min_{T^*} f(K_j, T^*)$ ;
**end**
**Return** cnt

By limiting the set of search directions to the axes of the input space, the coordinate search/descent technique is an alternative zero-order local approach that addresses the scaling issue seen in traditional local search. The theory is intuitive (Gürbüzbalaban et al., 2020): random search was created to simultaneously minimize the mean W-1 discrepancy about all

of its parameters, which takes the form.

$$f(T,K) = (\frac{1}{355}) \sum_{i=1}^{355} \mathcal{W}(f_{A_i}, f^*_{A_i}), \tag{4}$$

where $f_{A_1}$ and $f^*_{A_1}$ are trial distribution of target word $A_i$ from ground truth and simulated result, respectively. Note that the distribution $f^*_{A_i}$ is a realization from our Wordle simulator given hyper-parameter $K$ and $T$. A coordinate-wise algorithm reduces this function to one coordinate or weight at a time or, more generally, to a subset of coordinates or weights at a time while holding the other coordinates or weights constant. Despite this limitation, these algorithms are much more versatile than random search (in fact, they may be used to solve even medium-sized machine learning issues efficiently in practice) (Deshpande et al., 2018), even though they require additional steps to establish approximate minima and restrict the number of descent directions that can be discovered. These algorithms also act as predicates for a whole line of higher-order coordinate descent techniques, just like they did with the random search strategy.

## Experiment

### Dataset

The model employed in this study utilized Wordle results sourced from Wordle Stats over the full year from January 7, 2022, to December 31, 2022. This one year was chosen to capture longitudinal data, reflecting genuine player interactions throughout different seasons and stages of player development, thereby minimizing pre-selection bias and providing a comprehensive basis for understanding group behavior in word-guessing activities. The results include the distribution of the number of trials it took players to succeed, a critical measure of the game's difficulty, and player engagement. These details can be found at `https://shorturl.at/ade06`. Additionally, to construct a robust model, we integrated a dictionary database comprising five-word English terms from the Google Books Ngram Corpus from 1970 to 2019. This corpus provided the foundational data for calculating each word's usage frequency, a key determinant of user selection within our predictive framework. We define word frequency as the relative occurrence of a word in the corpus, which is a direct measure of its commonality and presumed familiarity to players, reflecting the likely cognitive effort required for players to guess the word correctly (Solovyev et al., 2019).

### Model Evaluation

**Wasserstein Metric** This algorithm uses Wasserstein-1 distance to evaluate the difference between two trial distributions. Here, the attempt distribution refers to the attempts the user requires to complete the task in the Wordle game. This model aims to determine the difficulty of the word A relative to the word A* in the Wordle game. This is done by calculating the Wasserstein-1 distance of word A relative to the overall attempt distribution. Figure 2 shows the difficulty distribution of 355 ground truth words. It shows that the proposed metric is sufficiently consistent, representative of the difficulty realized by records, and mediates between other quantifiers. Clearly, 'easier' target words exhibit a trial distribution shifted towards the left on the x-axis, and 'harder' target words ( yellow colored) have a right-shifted trail distribution.

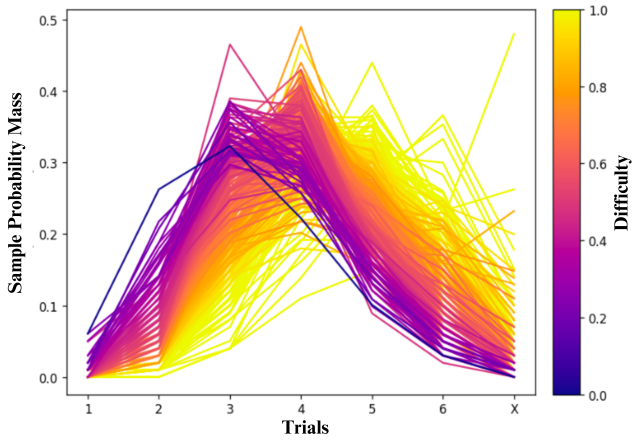


Figure 2: Distribution of word difficulties in the Wordle game. The x-axis represents the number of attempts, and the y-axis represents the frequency of each attempt number.

**Coordinate Search Optimization** To assess the efficacy of the Coordinate Search Optimization algorithm, it is essential to consider the broader statistical profile rather than just a single winning result. This algorithm's utility lies in its ability to iteratively explore and optimize a multi-dimensional space by adjusting one coordinate at a time. It is particularly suited for problems with a complex objective function or lack an analytical gradient. The evaluation of this optimization technique is predicated on its capacity to train generators that output a discrete target trial distribution accurately. We plan to invoke the same number of generative realizations for a robust evaluation as in fixed-length accurate data batches. This approach ensures a fair comparison between the model's output and the empirical ground truth. Consistency with the ground truth is then assessed by visualizing the distribution of densities using the same projection of functional PCA (Shang, 2014), as depicted in Figure 3 (a). Further, we solidify our statistical analysis by constructing an empirical 95% confidence interval for the optimization outcomes, with the results presented in Figure 3 (b) indicating that the sampling algorithm is robust, as evidenced by the slight variance observed. The predictions and summaries derived from this approach are systematically tabulated for further scrutiny.

**Simulator** The model is benchmarked with several other machine learning algorithms that comprehensively estimate trial distributions. Note that we are using the attributes to

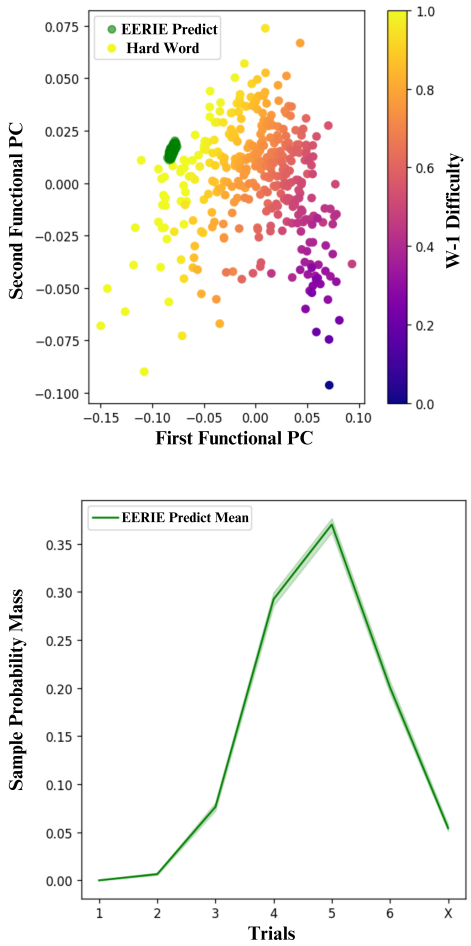


Figure 3: We obtain 200 simulated trial distributions of the word eerie. Displayed at the top, the visualization method projects these results into a 2-dimensional functional space. At the bottom, the predicted mean is highlighted with a green line, and a 95% empirical confidence interval is constructed around it.

feed as the input; in regression-type benchmarks, the machine learning algorithms output a 7 dimensional vector filled with floating number loadings, whereas in classification tasks, benchmarks output a category. We train these machine learning algorithms under the canonical parameter settings. Representative methods, namely, linear regression, decision tree regression, random forest regression, and Multiple layer perceptron regressor, are selected for the experiments, and the actual performance of the algorithms is compared by studying the training and validating performance among these algorithms. For the sake of fairness, hyperparameters and settings for each method are set by default.

The simulator has proven a successful method for predicting the difficulty and distribution of words in the simulator game. The model achieves an accuracy of 87%, outperforming other machine learning algorithms we have tried. The results demonstrate that the simulator is effective in predicting the distribution of future reports and the difficulty and distribution of specific words in Wordle.

Table 1: For distribution learner, We report mean squared loss for all models. Mean accuracy is obtained from 5-fold cross-validation; note that our model outperforms all benchmarks.

| Model | MSE | Mean Acc |
|---|---|---|
| CogSimulator | 0.512 | **87**% |
| Linear Regression | 1.206 | 19% |
| MLP Regression | 1.724 | 62% |
| Regression Tree | 0.758 | 60% |
| RandomForest Regression | 0.620 | 68.45% |

Overall, the simulator significantly improves the prediction of word difficulty and distribution in Wordle. Its superior performance compared to other machine learning algorithms highlights the effectiveness of the Convolutional Networks under the Wasserstein training approach. This work proposes a robust model for predicting the difficulty of educational games tailored to specific cognitive groups, ultimately enhancing overall game playability and engagement. The results of different models are shown in the table 1.

## Conclusion and Future Work

This study introduces the CogSimulator, designed to simulate cognitive distributions in contexts with limited sample sizes, exemplified by its application to Wordle. The model competes against machine learning metrics such as Wasserstein-1 distance, mean squared error, and mean accuracy. Leveraging the universal relevance of word frequency attributes, the CogSimulator shows promising generalization across word-based games, suggesting significant potential impacts on cognitive game development for niche user groups. Despite its strengths, the model tends to represent an "average" player, which may not reflect the diversity in player strategies and cognitive processes, particularly where data distributions are multimodal.

Future work will enhance this model by integrating clustering algorithms to detect and model unique player profiles, thereby better capturing the diversity of players. Plans also include integrating player feedback as a dynamic reward mechanism and developing an advanced parameter optimization method for dynamic loss functions to suit simulation tasks better and extend applicability to a broader range of educational games.

## References

Amory, A. (2010). Learning to Play Games or Playing Games to Learn? A Health Education Case Study with Soweto

Teenagers. *Australasian Journal of Educational Technology*, *26*(6).

Bian, W., Song, Y., Gu, N., Chan, T. Y., Lo, T. T., Li, T. S., ... Trillo, R. A. (2023). MoMusic: A Motion-Driven Human-AI Collaborative Music Composition and Performing System. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 37, pp. 16057–16062).

Boot, W. R. (2015). *Video Games as Tools to Achieve Insight into Cognitive Processes* (Vol. 6). Frontiers Media SA.

Dale, G., Joessel, A., Bavelier, D., & Green, C. S. (2020). A New Look at the Cognitive Neuroscience of Video Game Play. *Annals of the New York Academy of Sciences*, *1464*(1), 192–203.

Deshpande, I., Zhang, Z., & Schwing, A. G. (2018). Generative Modeling Using the Sliced Wasserstein Distance. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 3483–3491).

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. Random House.

El-Hadidy, M. A. A., Alfreedi, A. A., & Alzulaibani, A. A. (2021). Optimal Multiplicative Generalised Coordinated Search Technique to Find a D-dimensional Random Walker. *International Journal of Operational Research*, *42*(1), 1–33.

Frandi, E., & Papini, A. (2014). Coordinate Search Algorithms in Multilevel Optimization. *Optimization Methods and Software*, *29*(5), 1020–1041.

Funk, M., Dieber, B., Pichler, H., & Coeckelbergh, M. (2020). Gamification of Trust in HRI? In *Culturally sustainable social robotics* (pp. 632–642). IOS Press.

Geyer, C. J. (2011). Introduction to Markov Chain Monte Carlo. *Handbook of Markov Chain Monte Carlo*, *20116022*, 45.

Gürbüzbalaban, M., Ozdaglar, A., Vanli, N. D., & Wright, S. J. (2020). Randomness and Permutations in Coordinate Descent Methods. *Mathematical Programming*, *181*, 349–376.

Hecker, C. (2011). My Liner Notes for Spore. *Retrieved September*, *20*, 2011.

Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., & Wu, Y. (2016). Exploring the Limits of Language Modeling. *arXiv preprint arXiv:1602.02410*.

Kim, Y. J., & Ruipérez-Valiente, J. A. (2020). Data-driven Game Design: The Case of Difficulty in Educational Games. In *Addressing Global Challenges and Quality Education: 15th European Conference on Technology Enhanced Learning, EC-TEL 2020, Heidelberg, Germany, September 14–18, 2020, Proceedings 15* (pp. 449–454).

Laine, T. H., & Lindberg, R. S. (2020). Designing Engaging Games for Education: A Systematic Literature Review on Game Motivators and Design Principles. *IEEE Transactions on Learning Technologies*, *13*(4), 804–821.

Leutner, F., Codreanu, S.-C., Brink, S., & Bitsakis, T. (2023). Game Based Assessments of Cognitive Ability in Recruitment: Validity, Fairness and Test-taking Experience. *Frontiers in Psychology*, *13*, 942662.

Liao, Q., Chen, J., Wang, Z., Bai, B., Jin, S., & Wu, H. (2022). Fast Sinkhorn I: An O (N) algorithm for the Wasserstein-1 metric. *arXiv preprint arXiv:2202.10042*.

Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of Education and Learning: A Review of Empirical Literature. In *Proceedings of the 2nd international gamifin conference, gamifin 2018.*

Malone, T. W. (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science*, *5*(4), 333–369.

Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M. A., Guerrero-Puerta, L., Aguilar-Parra, J. M., Trigueros, R., & Alias, A. (2021). Between Level Up and Game Over: A Systematic Literature Review of Gamification in Education. *Sustainability*, *13*(4), 2247.

Match, S. (2022). The New York Times Buys Wordle. *The New York Times*.

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative Agents: Interactive Simulacra of Human Behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology* (pp. 1–22).

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1), 68.

Sailer, M., & Homner, L. (2020). The Gamification of Learning: A Meta-analysis. *Educational Psychology Review*, *32*(1), 77–112.

Shaheen, M. Y. (2021). Applications of Artificial Intelligence (AI) in Healthcare: A Review. *ScienceOpen Preprints*.

Shang, H. L. (2014). A Survey of Functional Principal Component Analysis. *AStA Advances in Statistical Analysis*, *98*, 121–142.

Solovyev, V. D., Bochkarev, V. V., & Akhtyamova, S. S. (2019). Google Books Ngram: Problems of Representativeness and Data Reliability. In *International Conference on Data Analytics and Management in Data Intensive Domains* (pp. 147–162).

Sommerville, J. A. (2020). Social Cognition.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced Engagement in Internet Video Advertisements. *Journal of Marketing Research*, *49*(2), 144–159.

Van Hove, O., Van Muylem, A., Leduc, D., Legrand, A., Jansen, B., Feipel, V., ... Bonnechere, B. (2019). The Use of Cognitive Mobile Games to Assess Cognitive Function of Healthy Subjects Under Various Inspiratory Loads. *Medicine in Novel Technology and Devices*, *1*, 100005.

Wiley, K., Robinson, R., Mandryk, R. L., et al. (2021). The Making and Evaluation of Digital Games Used for The Assessment of Attention: Systematic Review. *JMIR Serious Games*, *9*(3), e26449.

Yen, C.-J., Konold, T. R., & McDermott, P. A. (2004). Does Learning Behavior Augment Cognitive Ability as an Indicator of Academic Achievement? *Journal of School Psychology*, *42*(2), 157–169.

Yin, Q.-Y., Yang, J., Huang, K.-Q., Zhao, M.-J., Ni, W.-C., Liang, B., ... Wang, L. (2023). AI in Human-Computer Gaming: Techniques, Challenges and Opportunities. *Machine Intelligence Research*, *20*(3), 299–317.

Zhao, W. (2017). Research on The Deep Learning of The Small Sample Data Based on Transfer Learning. In *Aip Conference Proceedings* (Vol. 1864).