Increasing Accessibility for Game-based Learning Experiences: Developing Brain Computer Interface Controls for Minecraft

Anonymous Submission

Abstract

The usage of games for learning has been proliferating for a long period of time. Video games specifically have become a cultural mainstay among youth and are also effectively used in classrooms and numerous informal learning environments for a variety of creative and engaging learning experiences. Minecraft is one of the most popular examples of a commercial videogame played by youth through personal interest and also leveraged in a variety of educational settings. Yet there is a dearth of work in making video game experiences like Minecraft easily accessible to youth and players with severe mobility challenges. In this work we describe the creation of a Brain Computer Interface (BCI) using Electro Encephalo Graphy (EEG) data to enable Minecraft gameplay, as a steppingstone for the creation of more BCIs across a variety of commercial and educational video games, for increasing access to these learning experiences as well as researcher and educator impact through the gameplay interaction data that these interfaces afford.

Introduction

Game based learning has been on the uptake across schooling for decades now and has a vast variety of research explaining and describing the kinds of learning that are supported by them. Facilitating and analyzing the learning process itself has largely relied on qualitative methods (Steinkuehler & Squire 2014), and log-based data (Owen & Baker 2019). Over the past ten years, researchers have become increasingly interested in using multimodal data to study game-based learning environments (Emerson et al. 2020; Gomes et al. 2009). In this paper, we are specifically interested in the affordances of leveraging multimodal data to support gameplay and player learning in Minecraft. In line with Worsley et al. (2018), researchers have begun to explore ways to utilize naturalistic interfaces to support increased accessibility and more complex analyses of learning. Some of the existing work with Minecraft uses natural language processing and computer vision, which require a certain level of physical movement (Worsley et al. 2021). While these technologies are still continuously growing, the next level of technology for consumers with such physical impairments or lack of voice sound must be further developed. This motivates using participant thinking as a possible direction for increasing accessibility. To enable this, the field of Electroencephalography (EEG) has been proposed with research. Electroencephalography is a signal that entails meaningful information from the brain activity that can be measured by electrodes measuring the voltages. However, the signals are so complex and non-stationary that efficient preprocessing methods must be applied to the raw EEG data for it to be accurately classified. Furthermore, to perform a real-time classification, the proposed deep neural network would require enough training to output the user's intention onto the game screen of Minecraft within a reasonable time frame. In this paper, we describe our effort to perform the real-time classification of the raw EEG input data from the user's gameplay, along with the classification results leading to certain movements in the Minecraft gameplay. The next section of this paper proceeds with related work, followed by the methodologies we adapted and proposed for our task, and finishes with the results and the conclusion.

Related Work

Learning from Games and Minecraft

The rich work in learning through games highlights how games provide creative, contextualized, complex systems experiences to understand phenomena in deeper and more engaging ways than afforded through typical classroom instruction and textbook learning (Steinkuehler & Squire 2014). This of course goes along with the significant amount of time that youth spend in video games for recreational as well as educational purposes (Chan et al. 2022).

This creativity and novelty of experience and engagement is further enriched in open world games – where players can often choose goals of their own interest, and move towards those goals in a huge variety of ways (Conole 2012). Minecraft in particular, one of the most popular open world (sandbox) games across the world (data deep dive), has been used in hundreds, if not thousands, of educational contexts and learning experiences in its last 11 years of existence and popularity (Nebel et al. 2016). This has been conducted through the base game of Minecraft itself, being used to explore topics spanning history, physics, chemistry, and computer science. The developers at Minecraft themselves have been engaged in supporting its educational uses, also developing MinecraftEDU (Cózar-Gutiérrez & Sáez-López 2016), as well as specific coding interfaces which both provide exposure to coding and legitimately help change and improve different laborious aspects of the gameplay experience itself (Kutay & Oner 2022).

This makes a strong case for increasing the ways learners are able to engage with such games, not only for equitable participation in social recreational opportunities, but also for access to creative learning experiences.

Accessibility in Games

Gameplay log data has been used to gather deeper lenses into learner understanding in a variety of ways. This work most notably emerges from analytics work surrounding interactive tutoring systems (ITSs), using varying complexities of artificial intelligence methods to deepen research findings around efficient ways to learn different concepts, identify learner understandings in response to their accuracy and approach in solving different problems and tasks, and create frameworks for educator orchestration, support, feedback, and assessment in response to student work in these systems (Baker, 2016). This was of course enabled by the specific design of ITSs for usage in classrooms, hosting tasks legible to traditional classroom environments, educators, and disciplinary goals, along with the technological ability of recording and analyzing highly detailed interaction data from learners – something not possible in classically used offline tools and materials.

Complimentarily, the creation and proliferation of a variety of educational games has contributed to numerous frameworks that have enabled researchers and educators to glean information about learners from the ways they play games (Owen & Baker, 2019). While educational games predate the commonality of access to personal digital devices for each learner (Bigelow, 1997), supporting learners to engage with games designed for learning individually enabled access to learner behavior data similar to that studied in ITSs. At the same time, a key difference in the nature of this data lies in how games that center playful experiences – especially games developed primarily for commercial gameplay rather than specifically for classroom use – enable a variety of behaviors that do not have direct maps to classroom learning assessments and need more complex analytical and interpretation methods (Squire, 2007).

This gap is a double-edged sword – open ended commercial games have avid player bases, foster rich engagement by youth (as well as adults) even on personal time, and can much better respond to players' personal interests, curiosities, and tendencies. At the same time, this flexibility of gameplay and player approach often makes it harder to use in classroom environments with specific learning goals. Minecraft, with its vast open endedness (from a player as well as an educator perspective), provides a particularly productive space to explore this continuum – from creating highly constrained worlds and curricular designs that engage learners to solve specific problems demonstrating their understanding and learning (Pusey & Pusey, 2016); to letting learners explore with great freedom and identifying connections between their play behaviors and out-of-game skills and competencies (Rahimi et al. 2022; Samsudin et al. 2022).

Additionally, the lenses into player and learner behavior that gameplay data afford are lately being further extended through the rising use of multimodal learning analytics (Emerson et al. 2020). With the proliferation of advanced video analysis, different sensors to detect emotional states, and advanced interaction paradigms spanning movement in real world space, researchers are able to complement in-game actions around task specific progress (e.g., successes and failures in solving different problems) with meta-information about the learning experience itself (e.g., frustration, disinterest, excitement, etc.) (Henderson et al. 2020).

The creative open-ended forms of engagement that games support are often enabled by more complex ways of interaction than commonly supported in classroom experiences. Being able to navigate information rich interfaces through using keyboard-mouse interactions, or multi-tactile game controllers involves a technological fluency and sensory-motor control that is not equally accessible for all learners, youth, and players (Aguado-Delgado et al. 2020).

There is a rich breadth of work in making gameplay more accessible in the games field. This work spans from including accessible gameplay modes (Brown & Anderson 2021) – most popularly visible in Mario Kart 8's auto drive mode which reduces the active manipulation required for players. This is considered a powerful example of accessibility design which not only benefits players with mobility challenges, but also increases gameplay engagement and inclusivity for players with different participatory barriers like unfamiliarity with controller fluency. Apart from work within the game design space itself, there is also a breadth of work on making accessible controllers (Maggiorini et al. 2017) specifically aimed to reduce the sensory-motor requirements of many different games.

A key next frontier of enabling easier access to gameplay, in the pathway to more inclusive recreation as well as more inclusive creative education experiences, is enabled by Brain Computer Interfaces (Kerous et al. 2018). Despite work on this for over a decade, much work around BCI games still relies more on simplified game design than enabling full engagement with pre-existing commercial games and their associated player cultures and communities (Vasiljevic & de Miranda 2020). This gap has continued due to a spectrum of factors including limitations of computer processing capabilities and BCI data sophistication being capable enough for smooth and sophisticated gameplay.

Consequently, in this work we build on prior calls for creating multimodal interfaces to help the proliferation of inclusive learning tools and experiences (Worsley et al. 2018) by demonstrate a framework for analyzing EEG data that enables real time gameplay with Minecraft. Our analysis so far enables interpretation of EEG data towards 3 broad categories of gameplay actions inside Minecraft. In addition to the description above of the importance and potential value of this work for inclusive gameplay experience as well as educational experiences, we build on our analysis by discussing how this can be integrated with learning game data analytics frameworks for deeper lenses into styles and levels of learner sensemaking and understanding.

Classifying EEG Data for brain Computer Interfaces

In order to classify EEG data, the first step is to preprocess the data by removing artifacts and extract meaningful information out from the raw data. The main kinds of artifacts are categorized into several types: Ocular artifacts: eye movement and blinks, Muscle artifacts: myogenic activity, Cardiac activity (Urigüen & Garcia-Zapirain 2015). As these movements come as signals in raw EEG data, filtering techniques are often applied such that artifact signals that are above or below a certain frequency are simply ignored in the collection. However, as they are not enough to denoise concise portions of the data, the next step of signal processing methods is required. Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) are common signal processing methods for denoising, and they are adapted as common techniques to generate features from the raw data, which have shown efficiency in denoising raw datasets of mental states (Bashivan et al. 2016) and eye movements (Ruffini et al. 2019). Depending on the feature generation methods, different types of models are also proposed. EEGNet is a deep learning architecture that has its core as Convolution Neural Network (CNN), that takes in input as the raw EEG data reformed as a N x N matrix (Lawhern et al. 2018). While using similar architecture, generating a spectrogram using Azimuthal Equidistant Projection (AEP), also known as Polar Projection as input has also proven high accuracy in classifying different memory tasks (Bashivan et al. 2015). While there is a certain level of intersection between these previous analyses, the difference in the measure of electrodes and the experimental setup leads each of them to be unique and also presents a challenge to easily adapt to our work at the same time. Even for a study of using a Muse Headband for data collection of Minecraft gameplay, the difference in intention of the research, analysis of the neural activity of children of age around 8 playing Minecraft in a public museum setting (Ravindran et al., 2019), makes our work a more unique subject. Our work not only focuses on the whole process of classification tasks with different Minecraft plays but also applying it to be done in real-time such that the players' imaginations are being the outputs to the screen in front of them.

Methods

Experiment Setup and Dataset

The Muse Headband was used for data collection. The four dry electrodes (TP9, AF7, AF8, TP10) are placed on the scalp of the participants, where TP refers to the temporal and parietal lobe, and AF refers to the prefrontal and frontal lobe of a brain. The Muse Headband is connected via Bluetooth using BlueMuse (Kowaleski, 2022) and is able to record the brain wave activity at a sampling frequency of 256 Hz using muse-lsl (Barachant et al., 2019). We collected EEG data from multiple individuals of different ages, typically between 11 to 14 and 20 to 26. Participants were to play three different tasks in the Minecraft game, 1) wandering in any direction in the map, 2) mining underground and collecting blocks, 3) building blocks of any type that participant desires (required to build N x N shapes for default choice) for each of 60 seconds. A Notch filter of 60 Hz is applied to the raw data to remove baseline noise at a certain level.

Feature Extraction

Feature extraction of the EEG signal is the core task in developing brain-computer interface (BCI) applications, as the raw data itself at a single time point does not convey enough information to be applied to the algorithms. Therefore, we adapted Jordan's short-time windowing technique of feature extraction method to our data (Ashford et al., 2019). EEG signal points are divided into a sequence of windows for the interval of one second, along with another window with an overlapping of 0.5 seconds in between the intervals, which makes a sequence of windows as [(0s − 1s), [0.5s − 1.5s), [1s − 2s), . . .]). From each 1-second time window, the two half-windows and quarter-windows are also generated by dividing the number of samples by 2 and 4 accordingly within the 1 second window period. From these divided periods, the following sets of statistical features are generated.

For the full-time window, each signal's sample means, standard deviation, sample skewness, sample kurtosis, minimum and maximum value, sample variance, sample covariances of all signal pairs, eigenvalues of the covariance matrix, upper triangular elements of the matrix logarithm of the covariance matrix, the magnitude of frequency components using Fast Fourier Transform (FFT), and frequency values of the ten most energetic components of the FFT. For the two half-windows, the delta of sample means, standard deviations, maximum, and minimum values between the first and second half-windows are generated. For the quarter windows, the sample means, all paired differences of sample means between the quarter windows, the maximum and minimum values, all paired differences of maximum and minimum values between the quarter windows for all signals are generated. In the end, a total of 989 features were generated, with corresponding labels to determine their states.

Classifying different states

We employed classifiers on the dataset to compare the performances and to proceed with the best performing model. Our dataset, which is feature-extracted, is in a 2-D shape with a 70/30 ratio of training and test split. Details of each of the classifiers are summarized below.

XGBoost. XGBoost has its algorithm foundation in gradient boosted trees, a supervised learning method to predict target value with high accuracy by ensembling weaker models. The goal is to optimize with reducing the loss of the following equation:

$$min\ \mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(\boldsymbol{x}_i)) + \Omega(f_t)$$

in which y_i is the real value (label) known from the training dataset at iteration $\boldsymbol{t}$.

K-Nearest Neighbor. K-Nearest Neighbor (KNN) is another type of supervised algorithm to determine how similar vectors are from each other by measuring the distances of the new data to all other classified data. Here we use Euclidean distance to measure the distance between vectors.

Gated Recurrent Unit (GRU). GRUs uses a hidden state to transfer information over. The update gate, z_t, remembers what t memories should be included by outputting the result between 0 and 1 by using a sigmoid function.

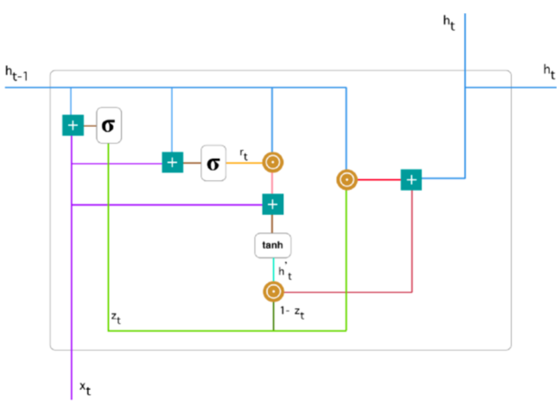

Fig. 1. Gated Recurrent Unit

It uses a hidden state to transfer information over. The update gate, zt, remembers what memories should be included by outputting the result between 0 and 1 by using a sigmoid function. The reset gate, rt, to determine what information to lose, which also uses a sigmoid function. The current memory unit, $h't$ passes on the information using hyperbolic tangent activation function, which will output between -1 and 1. For calculating the final result of the gate at current timestep t, we use the following equation:

$$h_t = z \odot h_{t-1} + (1 - z_t) \odot h'_t$$

Our GRU network consists of 256 neurons with helping layers to extend and flatten the dataset vector before and after the GRU layer to fit its required input shape. The model was trained in Google Colab where 12GB NVIDIA Tesla K80 GPU was provided free. Cross entropy function was used to minimize loss during the training step.

Results

We evaluated each model's performances by comparing the prediction results on the test set. Overview of the evaluation metrics for each model can be seen in Table 1.

Table 1. Evaluation metrics scores of each models on all datasets (70/30 split validation).

Model	Acc. (%)	Prec.	Rec.	F1
XGBoost	**93.15**	0.93	0.93	0.93
KNN	86.99	0.87	0.87	0.87
GRU	91.78	0.92	0.92	0.92

Figure 2 shows the confusion matrices of each model above for the Minecraft play states: wandering, building, and mining.

Conclusions and Future Work

In this work, models were trained and tested with a new form of the dataset of which several features were selected from Jordan's earlier short-window approach. The models each performed prediction in real-time using the muse-lsl package's built-in functions to return the electrodes values every second. Based on the classified results, commands are executed in Minecraft gameplay that matches the user's intentions. All in all, our work initiates further interactive applications to incorporating EEG signals in interacting with computers, with the purpose of supporting the physical impairments to access their opportunities to the current developments.

Our future work intends to introduce eye-gaze features into our classified results, developing our interface further into a multimodal interface, which would better understand the user's intentions. Challenges for us are to combine the results into one output without losing any important information from one. To accomplish this, a further dataset that takes into account the user's eye movement must be collected for accurate measure. Before that, converting each classified result into appropriate and precise Minecraft movements must also be carefully implemented.

As discussed in the introduction, this thread of work has powerful potential in not only increasing access for youth

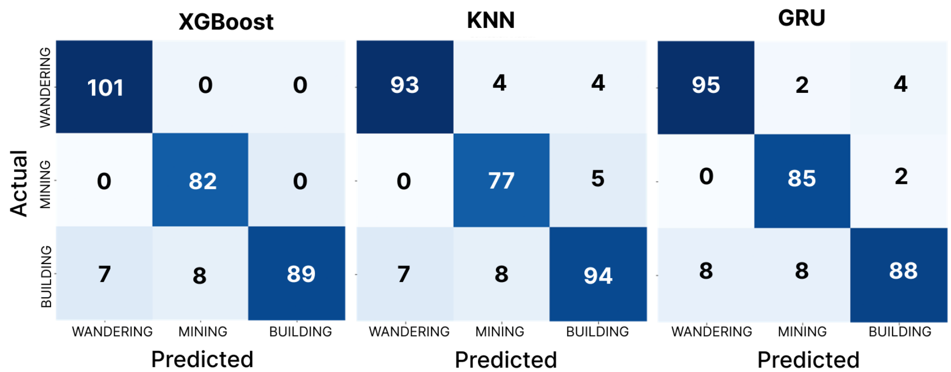

Figure 2. Confusion Matrices for unseen playing states across the 3 different methods we tested

and players with severe mobility challenges to recreational gameplay experiences but also to creative learning experiences which are extensively supported by Minecraft's design, popularity, and open-ended sandbox nature. Additionally, supporting (open-sourced) BCIs for gameplay control enables greater open access to tracking player-learner actions in ways that supports researchers to track, investigate, and understand different ways that these learners engage in different tasks.

Leaning on the field of learning game analytics work (Reardon et al. 2022), this data can be used to identify states of dis/engagement, un/productive exploration (Tissenbaum et al. 2016), and tinkering vs. refinement stages of understanding (Berland et al. 2013) for a few examples. Further, larger sets of these data can also be used to identify player styles (Heeter 2009), which in turn can inform the design of curricular materials and facilitation styles to respond to different patterns of learner interest and exploration in such environments. Work such as this enables easier and broader access to this data by expanding the player base, as well as widening the kinds of games (educational as well as commercial) where players' gameplay data can be accessed through greater ease than is always enabled by commercial game developers.

References

Aguado-Delgado, J., Gutierrez-Martinez, J.-M., Hilera, J. R., de-Marcos, L., & Otón, S. 2020. Accessibility in video games: A systematic review. *Universal Access in the Information Society*, *19*(1), 169–193.

Ashford, J., Bird, J. J., Campelo, F., & Faria, D. R. 2019. Classification of eeg signals based on image representation of statistical features. *UK Workshop on Computational Intelligence*, 449–460.

Baker, R. S. 2016. Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600–614.

Barachant, A., Morrison, D., Banville, H., Kowaleski, J., Shaked, U., Chevallier, S., & Tresols, J. J. T. 2019. Muselsl. *Epub Ahead of Print*.

Bashivan, P., Rish, I., & Heisig, S. 2016. Mental state recognition via wearable EEG. *ArXiv Preprint ArXiv:1602.00985*.

Bashivan, P., Rish, I., Yeasin, M., & Codella, N. 2015. Learning representations from EEG with deep recurrent-convolutional neural networks. *ArXiv Preprint ArXiv:1511.06448*.

Berland, M., Martin, T., Benton, T., Petrick Smith, C., & Davis, D. 2013. Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences*, *22*(4), 564–599.

Bigelow, B. 1997. On the road to cultural bias: A critique of The Oregon Trail CD-ROM. *Language Arts*, *74*(2), 84–93.

Brown, M., & Anderson, S. L. 2021. Designing for disability: Evaluating the state of accessibility design in video games. *Games and Culture*, *16*(6), 702–718.

Chan, G., Huo, Y., Kelly, S., Leung, J., Tisdale, C., & Gullo, M. 2022. The impact of eSports and online video gaming on lifestyle behaviours in youth: A systematic review. *Computers in Human Behavior*, *126*, 106974.

Conole, G. 2012. *Designing for learning in an open world* (Vol. 4). Springer Science & Business Media.

Cózar-Gutiérrez, R., & Sáez-López, J. M. 2016. Game-based learning and gamification in initial teacher training in the social sciences: An experiment with MinecraftEdu. *International Journal of Educational Technology in Higher Education*, *13*(1), 2. https://doi.org/10.1186/s41239-016-0003-4

Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. 2020. Multimodal learning analytics for game-based learning. *British Journal of Educational Technology*, *51*(5), 1505–1526.

Gomes, J., Yassine, M., Worsley, M., & Blikstein, P. 2013. Analysing engineering expertise of high school students using eye tracking and multimodal learning analytics. *Educational Data Mining 2013*.

Heeter, C. 2009. Play styles and learning. In *Handbook of research on effective electronic gaming in education* (pp. 826–846). IGI Global.

Henderson, N., Rowe, J., Paquette, L., Baker, R. S., & Lester, J. 2020. Improving Affect Detection in Game-Based Learning with Multimodal Data Fusion. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial Intelligence in Education* (pp. 228–239). Springer International Publishing. https://doi.org/10.1007/978-3-030-52237-7_19

Kerous, B., Skola, F., & Liarokapis, F. 2018. EEG-based BCI and video games: A progress report. *Virtual Reality*, *22*(2), 119–135. https://doi.org/10.1007/s10055-017-0328-x

Kowaleski, J. 2022. *BlueMuse Project*.

Kutay, E., & Oner, D. 2022. Coding with Minecraft: The Development of Middle School Students' Computational Thinking. *ACM Transactions on Computing Education*, *22*(2), 21:1-21:19. https://doi.org/10.1145/3471573

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. 2018. EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, *15*(5), 056013.

Maggiorini, D., Granato, M., Ripamonti, L. A., Marras, M., & Gadia, D. 2017. Evolution of game controllers: Toward the support of gamers with physical disabilities. *International Conference on Computer-Human Interaction Research and Applications*, 66–89.

Nebel, S., Schneider, S., & Rey, G. D. 2016. Mining learning and crafting scientific experiments: A literature review on the use of minecraft in education and research. *Journal of Educational Technology & Society*, *19*(2), 355–366.

Owen, V. E., & Baker, R. S. 2019. Learning analytics for games. *Handbook of Game-Based Learning. Cambridge, MA: MIT*.

Pusey, M., & Pusey, G. 2016. Using Minecraft in the science classroom. *International Journal of Innovation in Science and Mathematics Education*, *23*(3).

Rahimi, S., Walker, J., Lin, L., & Shin, J. 2022. *In Pursuit of Creativity in Minecraft: A Mixed-Method Approach*.

Ravindran, A. S., Mobiny, A., Cruz-Garza, J. G., Paek, A., Kopteva, A., & Vidal, J. L. C. 2019. Assaying neural activity of children during video game play in public spaces: A deep learning approach. *Journal of Neural Engineering*, *16*(3), 036028.

Reardon, E., Kumar, V., & Revelle, G. (2022). Game Learning Analytics. In C. Lang, G. Siemens, & A. F. Wise (Eds.), *The Handbook of Learning Analytics* (2nd ed., pp. 152–162). SOLAR. https://doi.org/10.18608/hla22.015

Ruffini, G., Ibañez, D., Castellano, M., Dubreuil-Vall, L., Soria-Frisch, A., Postuma, R., Gagnon, J.-F., & Montplaisir, J. 2019. Deep learning with EEG spectrograms in rapid eye movement behavior disorder. *Frontiers in Neurology*, *10*, 806.

Samsudin, M. A., Ming, G. K., Ahmad, N. J., & Abrose, Y. 2022. *Levelling Up Primary School Students' 21st Century Skills Through Minecraft-Game-Based Learning* [Chapter]. Handbook of Research on Acquiring 21st Century Literacy Skills Through Game-Based Learning; IGI Global. https://doi.org/10.4018/978-1-7998-7271-9.ch038

Squire, K. 2007. *Open-ended video games: A model for developing learning for the interactive age*. MacArthur Foundation Digital Media and Learning Initiative.

Steinkuehler, C., & Squire, K. 2014. Videogames and learning. *Cambridge Handbook of the Learning Sciences*, 377–396.

Tissenbaum, M., Kumar, V., & Berland, M. 2016. Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models. *International Educational Data Mining Society*.

Urigüen, J. A., & Garcia-Zapirain, B. 2015. EEG artifact removal—State-of-the-art and guidelines. *Journal of Neural Engineering*, *12*(3), 031001.

Vasiljevic, G. A. M., & de Miranda, L. C. 2020. Brain–computer interface games based on consumer-grade EEG Devices: A systematic literature review. *International Journal of Human–Computer Interaction*, *36*(2), 105–142.

Worsley, M., Barel, D., Davison, L., Large, T., & Mwiti, T. 2018. Multimodal interfaces for inclusive learning. *International Conference on Artificial Intelligence in Education*, 389–393.

Worsley, M., Mendoza Tudares, K., Mwiti, T., Zhen, M., & Jiang, M. 2021. Multicraft: A Multimodal Interface for Supporting and Studying Learning in Minecraft. *International Conference on Human-Computer Interaction*, 113–131.