The Journal of Open Source Software

pixelLOG: Logging of Online Gameplay for Cognitive Research

Zeyu Lu [1] and **Dennis L. Barbour** [1,2¶]

1 Department of Computer Science and Engineering, Washington University in St. Louis **2** Department of Biomedical Engineering, Washington University in St. Louis ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor:

Submitted: 11 February 2026
Published: unpublished

Summary

Traditional cognitive assessments often rely on isolated, output-focused measurements that may fail to capture the complexity of human cognition in naturalistic settings. We present **pixelLOG**, a high-performance data collection framework for Spigot-based Minecraft servers designed specifically for process-based cognitive research. Unlike existing frameworks tailored only for artificial intelligence agents, pixelLOG also enables human behavioral tracking in multiplayer/multiagent environments. Operating at configurable frequencies up to and exceeding 20 updates per second, the system captures comprehensive behavioral data through a hybrid approach of active state polling and passive event monitoring. By leveraging Spigot's extensible API, pixelLOG facilitates robust session isolation and produces structured JSON outputs integrable with standard analytical pipelines. This framework bridges the gap between decontextualized laboratory assessments and richer, more ecologically valid tasks, enabling high-resolution analysis of cognitive processes as they unfold in complex, virtual environments.

Statement of need

Cognitive assessment methodologies have historically been constrained by laboratory-like conditions and narrowly defined tasks. These assessments often emphasize the measurement of singular executive functions—such as working memory, inhibitory control, or attention—through static, outcome-based metrics (Miyake et al., 2000). While foundational, these conventional approaches frequently lack ecological validity (Hedge et al., 2018), failing to capture the adaptive nature of human cognitive processes in multifaceted, real-world environments. Recent evidence suggests that many established tasks may primarily measure information uptake speed rather than distinct cognitive constructs (Löffler et al., 2024), limiting our ability to observe how individuals integrate strategies and respond dynamically to changing scenarios.

In response, the field is shifting towards assessment frameworks that provide process-oriented, fine-grained behavioral data. Minecraft—an open-ended sandbox environment—has emerged as a powerful platform for simulating complex tasks requiring navigation, resource management, and problem-solving. However, a critical gap exists in the tooling available for this platform. While tools exist for AI training, there is a lack of specialized infrastructure for human cognitive research that requires high-frequency, reliable, and unobtrusive data logging in multiplayer contexts.

We introduce pixelLOG (Logging of Online Gameplay) to address this need. pixelLOG is a plugin-based logging framework integrating with the Spigot modification layer. It enables fine-grained data acquisition at frequencies exceeding 20 Hz, capturing both continuous states (e.g., location, gaze direction) and discrete events (e.g., block placement, combat). This architecture supports the precise mapping of individual cognitive trajectories onto environmental cues, facilitating a process-based examination of how individuals engage with complex tasks.

The Journal of Open Source Software

State of the field

A variety of platforms and research tools, such as Microsoft's Project Malmo (Perez-Liebana et al., 2019) and the MineDojo framework (Fan et al., 2022), have emerged to facilitate experimentation, data collection, and reinforcement learning (RL) research within Minecraft (Hafner et al., 2025; Qin et al., 2024; G. Wang et al., 2023; Z. Wang et al., 2025). These frameworks provide standardized interfaces for agent interaction and observation in controlled environments. While these agent-centric RL studies collect in-game data as observation space, they typically lack the high-precision, high-frequency, and configurable data collection pipelines necessary for human cognitive research. Moreover, these experimental platforms often operate in isolation, with limited extensibility and customization capabilities.

In contrast, pixelLOG is specifically designed for researchers investigating human cognitive processes in dynamic virtual environments. By leveraging Spigot's event-driven architecture and implementing a custom plugin-based solution, the system delivers high-frequency polling, granular event capturing, and robust per-player data isolation. Unlike existing solutions that may impose constraints on data granularity or system extensibility, pixelLOG's modular architecture provides the flexibility and performance required for comprehensive cognitive research, while maintaining compatibility with standard Minecraft server environments. Additionally, while primarily designed for humans, pixelLOG can also provide richer behavioral telemetry for artificial agents than many existing solutions.

Software design

pixelLOG is designed as a modular, extensible framework operating within the Spigot Minecraft server environment (compatible with version 1.20.4 and adaptable to others). As shown in Figure 1, the system architecture comprises distinct layers for player management, data acquisition, and structured output generation.

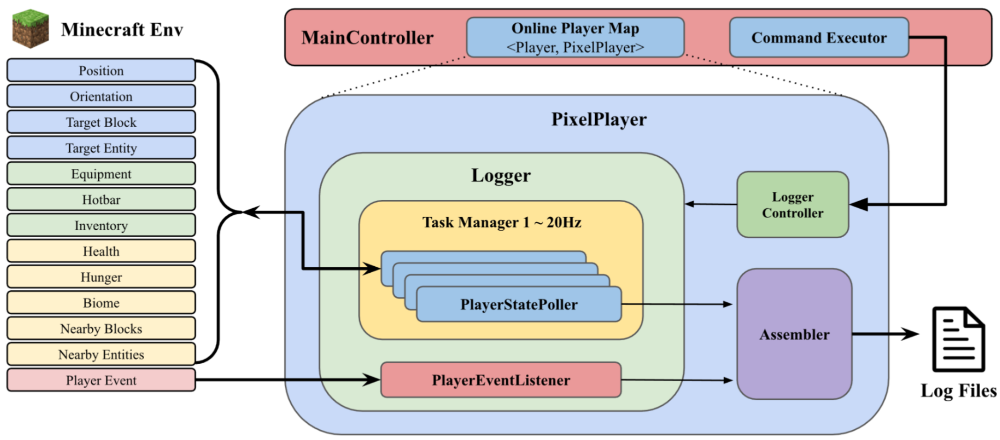

Figure 1: pixelLOG consists of several key components, each designed with specific responsibilities to ensure efficient data collection and processing.

Scalable Multi-Player Data Management

To support concurrent data collection, pixelLOG utilizes a hierarchical architecture centered on a `MainController`. Upon a player's connection, the system instantiates a dedicated `PixelPlayer` object that encapsulates all participant-specific data collection processes. This isolation is critical for data integrity.

Early iterations utilizing centralized logging revealed bottlenecks under high load. Our implementation uses distributed, thread-safe queues for individual players. This ensures that the high-frequency data streams of one participant do not interfere with the collection stability of others, maintaining linear scalability with increasing player counts.

Adaptive Temporal Resolution (Hybrid Data Capture)

pixelLOG employs a hybrid data collection strategy to capture the full spectrum of cognitive behavior:

1. **Active State Polling (Continuous):** A `Logger` component coordinates `PlayerStatePollers` that operate at configurable frequencies (e.g., 20 Hz). These tasks capture rapidly evolving attributes such as player avatar position (x,y,z), velocity, and view orientation (pitch, yaw). Lower-frequency pollers simultaneously survey static environmental parameters, such as biome types or nearby entities, optimizing computational overhead.
2. **Event-Driven Monitoring (Discrete):** To capture episodic markers, the system implements `PlayerEventListeners`. These intercept specific game events via the Spigot event system, such as block interactions, inventory changes, or combat.

By fusing asynchronous event data with continuous polling trajectories, researchers can anchor moment-to-moment behavioral patterns within the context of meaningful actions.

Structured Data Integration

An `Assembler` component harmonizes the heterogeneous data streams into chronologically ordered, structured JSON output. This format was selected for its compatibility with standard data science toolchains (e.g., Python pandas, R). The output structure hierarchically organizes session metadata, high-frequency state logs, and discrete event logs, enabling straightforward ingestion for subsequent statistical modeling or machine learning analysis.

Research impact statement

This utility has been demonstrated in recent applications: the framework served as the data collection backbone for *pixelDOPA (Digital Online Psychometric Assessment)*, enabling the validation of immersive cognitive minigames against the NIH Toolbox (Marticorena, Lu, et al., 2025), and supported real-time data integration for *AMLEC*, a multidimensional Bayesian active machine learning study of working memory (Marticorena, Wissmann, et al., 2025).

Acknowledgements

This study was supported in part by the Washington University OVCR Seed Grant and Here and Next Programs, as well as NSF National Artificial Intelligence Research Resource award NAIRR240235.

AI usage disclosure

Generative AI tools (Google Gemini) were used to assist in the drafting, formatting, and refining of the text in this paper. The authors reviewed, edited, and validated all AI-assisted outputs and take full responsibility for the content.

References

Fan, L., Wang, G., Jiang, Y., Mandlekar, A., Yang, Y., Zhu, H., Tang, A., Huang, D.-A., Zhu, Y., & Anandkumar, A. (2022). *MineDojo: Building open-ended embodied agents with internet-scale knowledge*. https://doi.org/10.48550/arXiv.2206.08853

Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2025). Mastering diverse control tasks through world models. *Nature*, *640*(8059), 647–653. https://doi.org/10.1038/s41586-025-08744-2

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Löffler, C., Frischkorn, G. T., Hagemann, D., Sadus, K., & Schubert, A.-L. (2024). The common factor of executive functions measures nothing but speed of information uptake. *Psychological Research*, *88*(4), 1092–1114. https://doi.org/10.1007/s00426-023-01924-7

Marticorena, D. C. P., Lu, Z., Wissmann, C., Agarwal, Y., Garrison, D., Zempel, J. M., & Barbour, D. L. (2025). *Immersive virtual games: Winners for deep cognitive assessment*. https://doi.org/10.48550/arXiv.2502.10290

Marticorena, D. C. P., Wissmann, C., Lu, Z., & Barbour, D. L. (2025). *Multidimensional bayesian active machine learning of working memory task performance*. https://doi.org/10.48550/arXiv.2510.00375

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Perez-Liebana, D., Hofmann, K., Mohanty, S. P., Kuno, N., Kramer, A., Devlin, S., Gaina, R. D., & Ionita, D. (2019). *The multi-agent reinforcement learning in malmö (MARLÖ) competition*. https://doi.org/10.48550/arXiv.1901.08129

Qin, Y., Zhou, E., Liu, Q., Yin, Z., Sheng, L., Zhang, R., Qiao, Y., & Shao, J. (2024). MP5: A multi-modal open-ended embodied system in minecraft via active perception. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 16307–16316. https://doi.org/10.1109/CVPR52733.2024.01543

Wang, G., Xie, Y., Jiang, Y., Mandlekar, A., Xiao, C., Zhu, Y., Fan, L., & Anandkumar, A. (2023). *Voyager: An open-ended embodied agent with large language models*. https://doi.org/10.48550/arXiv.2305.16291

Wang, Z., Cai, S., Liu, A., Jin, Y., Hou, J., Zhang, B., Lin, H., He, Z., Zheng, Z., & Yang, Y. (2025). JARVIS-1: Open-world multi-task agents with memory-augmented multimodal language models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2024.3511593