A Comparative Assessment of US and PH Learner Traversals and In-Game Observations within Minecraft

Jonathan DL. Casano[1[0000-0001-8443-0857]] and Maria Mercedes T. Rodrigo[1[0000-0001-7881-7756]]

[1] Ateneo de Manila University, Philippines
jcasano@ateneo.edu

Abstract. What-If Hypothetical Implementations using Minecraft (WHIMC) is a set of Minecraft worlds that learners can explore to learn more about science, mathematics, engineering, and technology. Because Minecraft is open-ended by design, assessing whether students are learning is always a challenge. In addition to using methods such as post-tests and self-reports, how can we use in-game data to measure learning? In this paper, we analyze and compare American (US) and Filipino (PH) learner traversals and in-game observations against canonical answers from experts to determine the extent to which students achieved the desired learning outcomes and the ways in which outcomes varied. We grouped students into high- and low-performing categories based on an out-of-game post test. We found that high-performers tended to make more observations than low-performers. Observations of high-performers tended to align with canonical answers. Many PH students tended to be low-performers and tended not to make in-game observations. In contrast, even low-performing US students tended to record observations actively.

Keywords: Minecraft, WHIMC, Comparison, Philippines, United States

1 Introduction

The What-If Hypothetical Implementations using Minecraft (WHIMC) project is a set of Minecraft worlds that learners can explore as supplementary activities to learn more about science, mathematics, engineering, and technology. It logs both the ways in which learners traverse these words and the observations that learners make during their explorations.

When engaging learners in open-ended environments such as Minecraft, assessment is always a challenge: How can we determine whether students are learning and what they have learned? Educators use a variety of well-worn assessment methods, e.g. pre- and post-tests [7] self-reports [1, 4] and teacher or researcher observations [1]. In recent years, more and more researchers are examining game logs, sometimes in conjunction with out-of-game assessments, to measure learning.

In this paper, we analyze game logs of American (US) and Filipino (PH) learner traversals and observations in order to assess whether learners achieved expected learning outcomes and to determine ways in which outcomes varied. We first organize the

observations into word clouds, taking into account the context in which the observations were made. We then compare canonical answers from experts against answers from (1) All students, (2) High-performing and low-performing learners and (3) US and PH learners

2 Review of Related Literature

To guide our analysis, we drew on prior work regarding WHIMC, word clouds, and map exploration archetypes of [6]. As WHIMC logs all student activity and all observations that students make, we organized student observations made in WHIMC into contextualized word clouds and compare student answers with canonical answers from experts.

Word clouds or tag clouds "are visual presentations of a set of words, typically a set of tags, in which attributes of the text such as size, weight or color can be used to represent features (e.g., frequency) of the associated terms" [3]. By displaying text data in graphical form, word clouds help readers surmise the gist of a text quickly [5]. However, word clouds are only useful for quick looks and cannot replace careful analysis of student responses. It was therefore necessary to group observations by location before creating the word clouds, a process that will be discussed further in the section on the Word Cloud Visualization Tool.

Finally, to interpret the traversal patterns of the students, we referred to the four map exploration archetypes of [6] and attempted to operationalize some of these archetypes in the logs that we analyzed.

3 Data Collection

The demographic details of the respondents are shown in Table 1. High- and low-performers were determined using out-of-game assessments [2]. The observations data of these groups were considered in the analysis.

Table 1. Summary of demographic details of respondent groups.

Group	Inclusive Dates	Total	Year Level / Subject	Age range	# Observations
US 1	6/28 - 7/2, 2021	10	Grade 5 Science	11 - 14 y/o	110
US 2	7/12 - 7/28, 2021	11	Grade 5 Science	11 - 14 y/o	53
PH 1	8/27 - 9/15, 2021	16	Grade 6 Science	11 - 14 y/o	291
PH 2	9/29 - 10/2, 2021	24	Grade 5 Science	10 - 13 y/o	102

4 Results and Analysis

The observations were organized into word clouds using a custom visualization program written in Python.

Using the custom visualization tool, eighteen (18) comparison attempts were made across 6 WHIMC maps and 3 performance categories (high-, average- and low-performers) for a total of 108 comparison attempts. Upon analyzing the 44 successful attempts, we find that:

High-performing students make more observations. When there were insufficient observations for the K-means clustering, the visualization was not generated. Hence, we see that making more observations across all maps visited is characteristic of high-performing students. Low-performing students fewer than average number of observations or no observations at all.

Observations of high-performing students matched the canonical observations from experts. Most of the observation matches occur in comparisons where the high-performing students were included. The matching observations were mostly from the high-performing group.

High-performing students tend to wander. A comparison between the performance categories within each data set was. It is noted that for each comparison, there are more wanderer clusters generated by the data coming from the high-performing group. Wandering around the Minecraft map and making observations beyond the prompts of the NPC seems characteristic of high-performing students.

Low-performing students from PH did not make any observations. None of the low-performing students from PH made any observations as opposed to the low-performers in the US data who all made observations. We speculate that active observation-making, might be an indicator of positive performance.

5 Contributions and Future Work

This work contributes to the literature in at least two ways. First, we contribute a tool that helps cluster and visualize word clouds and maps them onto the contexts from which the data was generated. This context can help with their interpretation. Second, we perform a cross-cultural comparison of data from two different populations, US and PH, using the same game-based learning environment, teasing out some of the similarities and differences between these populations.

In summary, high-performing students from the US and PH made the most in-game observations and that these observations tended to match the canonical observations from experts. US and PH high-performers tended to exhibit more wandering or moving around the map without a defined destination or purpose.Low-performers tended to make fewer observations in general. However, low-performers from the PH set did not contribute any observations while those from the US data set actively recorded their observations. The observations from low-performers from the US tended to match the canonical answers.

The differences between these groups may have been a function of differences in the way WHIMC sessions were introduced and conducted. US learners were asked to listen to 10-minute lectures that were specifically about the WHIMC worlds they were about to explore. The PH students were introduced to WHIMC, but preparatory lessons leading up to the explorations tended to be more generic, e.g. a lesson on making

observations or a lesson on ordinal numbers. WHIMC served as a backdrop or context against which students demonstrated knowledge of these concepts.

We would like to see how we can leverage on word embeddings libraries to merge or combine synonyms or related words, rather than treating them as unique and would like to further investigate the types of observations made by low-performing students. The data suggests that low-performing students who actively record their observations still manage to produce quality outcomes. However, more data is needed to validate this finding.

Finally, we hope to continue the cross-cultural comparisons we began in this study. WHIMC gives us a unique opportunity to capture and study the behaviors of US and PH learners within the same learning context. Investigating these differences may give us more insight about how to best use environments such as WHIMC to support learning goals in culturally appropriate ways.

Acknowledgement

The authors thank H Chad Lane and Jeff Ginger for their enthusiastic collaboration, Dominique Marie Antoinette B. Manahan, Mikael William Fuentes, and Ma. Rosario Madjos for their support, the Ateneo Laboratory for the Learning Sciences, the Ateneo de Manila University, and the Department of Science and Technology's Philippine Council for Industry, Energy, and Emerging Technology Research and Development for the grant entitled, "Nurturing Interest in STEM among Filipino learners using Minecraft."

References

1. Callaghan, N. (2016). Investigating the role of Minecraft in educational learning environments. *Educational Media International, 53*(4), 244-260.
2. Conati, C., & Kardan, S. (2013). Student modeling: Supporting personalized instruction, from problem solving to exploratory open ended activities. *Ai Magazine, 34*(3), 13-26.
3. Halvey, M. J., & Keane, M. T. (2007, May). An assessment of tag presentation techniques. In *Proceedings of the 16th international conference on World Wide Web* (pp. 1313-1314).
4. Melián Díaz, D., Saorín, J. L., Carbonell-Carrera, C., & de la Torre Cantero, J. (2020). Minecraft: three-dimensional construction workshop for improvement of creativity. *Technology, Pedagogy and Education, 29*(5), 665-678.
5. Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007, April). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 995-998).
6. Si, C., Pisan, Y., & Tan, C. T. (2016, February). Understanding players' map exploration styles. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-6).
7. Tangkui, R., & Keong, T. C. (2021). The effects of digital game-based learning using Minecraft towards pupils' achievement in fraction. *International Journal on E-Learning Practices (IJELP), 4*, 76-91.