Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Achievement, Behaviors, and STEM Interest of Frustrated and Bored Learners Using Minecraft

Maricel A. ESCLAMADO[ab*], Maria Mercedes T. RODRIGO[a] & Jenilyn A. CASANO[a]

[a]*Ateneo Laboratory for the Learning Sciences, Ateneo de Manila University, Philippines*
[b]*University of Science and Technology of Southern Philippines, Philippines*
*maricel.esclamado@obf.ateneo.edu

Abstract: In this paper, we analyze the relationship between behavioral and achievement outcomes and self-reported frustration and boredom levels of students using the What-If Hypothetical Implementations using Minecraft (WHIMC). We examine in-game data, out-of-game assessment data, self-reported frustration and boredom, and results of the STEM interest questionnaire (SIQ) from 175 Grade 8 learners from a school in the Philippines. We find that learners who are frustrated tend to disengage from the game. Learners who are bored tend to do less well on post-game assessments. While students' STEM interest generally increased, SIQ scores of frustrated and bored learners tended to remain unchanged. Among low-performing learners who did not express frustration or boredom, SIQ scores increased, implying that the absence of these states was beneficial to this category of learners. This study contributes to what is known about how affect relates to outcomes such as test scores, in-game behaviors, and attitudes. It also raises questions about how games can best be used to pave the way towards greater student engagement in STEM.

Keywords: Minecraft, WHIMC, Frustration, Boredom, Behaviors, STEM interest, Philippines.

1. Introduction

Games used for education can act as rich primers for engaged learning and were reported effective in improving learning outcomes (Zhonggen, 2019). A popular type of game used in education is open-ended games. Open-ended games like Minecraft provide players autonomy on how to do the tasks in the learning environment, thus there is no prescribed learning sequence and analysis of behavior becomes significantly important (Käser & Schwartz, 2020). Minecraft is an open-ended, sandbox-type video game first launched in 2009 (Bitner, 2021). It is open-ended in that it gives players freedom in exploring the environment and players can also mine resources and craft or build objects, often in collaboration with other players. With the open-ended nature of these games, it is expected that the players will have positive and negative emotional states. Frustration and boredom are some of the most frequent emotional states in open-ended learning and are associated with reduced learning outcomes (Sabourin & Lester 2014).

Frustration is defined as an emotion that results from interference that prevents achievement of a goal (Rajendran, Iyer, Murthy, Wilson, & Sheard, 2013). Boredom, on the other hand, occurs when individuals experience low arousal and dissatisfaction or disinterest in response to low arousal (Vogel-Walcutt, Fiorella, Carper, & Shaltz, 2012). Previous work has found that frustration, boredom, and other negative emotions are associated with negative learning and behavioral outcomes (Tze et al., 2015; Liu et al., 2013; Sabourin & Lester, 2014). For example, frustrated students tend to have a history of incorrect actions and help requests. Bored students tend to work slowly (Baker et al, 2012), have more idle time which is also an indicator of off-task behavior (Pardos et al., 2014), give up on a task (Andres et al., 2015), or engage in non-learning behaviors such as gaming the system (Baker et al., 2010).

In this paper, we analyze the relationship between frustration and boredom and learner in-game behaviors of students in Minecraft, an open-ended, less-structured game. We also investigate the relationship between learners' self-reported frustration and boredom with their knowledge assessments and STEM interest. This study sought to answer the following research questions:

RQ1: How do the behaviors of the learners who self-reported frustration and boredom differ from those who did not?

RQ2: What is the relationship between learners' self-reported frustration and boredom and their assessment outcomes?

RQ3: How did the STEM interest of the learners who self-reported frustration and boredom change after using WHIMC?

RQ4: How did the STEM interest of high- and low-performing learners who self-reported frustration and boredom change after using WHIMC?

2. WHIMC

What-If Hypothetical Implementations in Minecraft (WHIMC) is a set of simulations that learners can explore in order to learn more about science, technology, engineering, and mathematics (STEM). It is composed of a Rocket Launch Facility on Earth, Lunar Base LeGuin, a Space Station, alternate versions of Earth, exoplanets, and others. The Rocket Launch Facility is modeled after that of NASA. Here, learners can visit mission control and Mars rover test sites as well as talk to simulated NASA scientists. Lunar Base LeGuin is a tutorial level in which learners practice using in-game science tools to measure and record temperature, oxygen, pressure, and wind speed. The Space Station includes an unaltered version of Earth in which learners practice making different kinds of observations. The Space Station is also the jump-off hub from which learners travel to the different worlds.

The alternate versions of Earth present learners with opportunities to observe the planet under altered conditions. Although the worlds are fictional, they are created in consultation with scientists: They accurately depict conditions on Earth under these circumstances. For example, what if Earth had no moon (See Figure 1)? There would be no seasons, days would be shorter, and winds would be stronger. What if Earth was tilted by 90 degrees, that is, with the North Pole pointed towards the sun? There would be very different day/night cycles. What if Earth had a slightly colder sun? Water might only be able to exist in liquid form in a limited strip of green, and this is where we would all be forced to live.

Figure 1. Earth with No Moon

In each of these alternate Earths, learners explore the terrain, describe the environment, report observations about how life on Earth is affected by these circumstances, and possibly create habitats that will enable them to survive. By immersing learners in these activities, WHIMC hopes to generate interest in and excitement for STEM among participating learners.

3. Review of Related Literature

Game-based learning environments have become popular learning tools because of their potential to deliver knowledge while keeping the experience fun for the learners (All et al., 2016). Their ability to strike a balance between teaching/learning and gaming is a characteristic leveraged by educators to support the learning process (Prensky, as cited in Beserra et al., 2014). Learners are likely to exhibit a wide range of emotions when interacting with game-based learning environments and the role of affect in this context has been the focus of continuing research (Sawyer et al., 2017; Sabourin & Lester, 2014). Studies have shown interesting relationships between affect and learning. For example, confusion, which is considered a positive emotion, can benefit learning when it is resolved or partially resolved (Zhang et al., 2021). On the other hand, negative emotions such as frustration and boredom have the potential to encourage unproductive behaviors like gaming the system and can thus have detrimental effects to learning (D'Mello & Graesser, 2011; Andres et al., 2015).

Frustration and boredom have been found to be associated with off-task behavior (Baker et al., 2010). Off-task behaviors are characterized by learners disengaging from the learning environment or task (Baker, 2007). These behaviors vary among learning environments and the learning goals they are trying to achieve. For example, Sabourin et al. identified (1) moving a task-related object to an unrelated location and (2) spending too much time in a location irrelevant to a task within Crystal Island as off-task behaviors (Sabourin et al., 2011). Baker (2007) operationalized a learner's idle time as an indicator of off-task behavior in a Cognitive Tutor software. He presented a machine-learned model for detecting off-task behavior by considering a learner's idle time along with other features. Idle time was considered to be all actions taken beyond a set cut-off time to perform a certain task. In game-based learning environments, students disengaging from a task but not necessarily from the learning environment also indicates off-task behavior. The students take a break from the difficult learning task and focus on the game-based aspects of the learning environment. Prior work found that this behavior is negatively correlated with positive learning outcomes but may also aid students in regulating negative affect and returning to the learning task (Sabourin & Lester, 2014).

In this paper, we look at the in-game behaviors of students who self-reported frustration and boredom in Minecraft, an open-ended, less-structured game. We also investigate the relationship of learner self-reported frustration and boredom to their knowledge assessments and STEM interest.

4. Data Collection

The data analyzed for this paper consisted of in-game data, out-of-game assessment data, and results of the game experience questionnaire (GEQ) and STEM interest questionnaire (SIQ) from 175 Grade 8 learners, 13 to 15 years old, from a school in the Philippines. Partner teachers from the school developed two (2) learning modules that integrated the use of WHIMC in classroom lessons about biotic/abiotic components of the environment and adaptation.

WHIMC is instrumented to collect data about learner actions as they use the game. WHIMC logs learner positions in intervals of 6 seconds as they explore the game, or whenever the learner makes an observation or access a science tool. Details of the interaction logs included the timestamp, coordinates, the world explored, observations made, and the science tool used. Time sequences of learner explorations done in Minecraft WHIMC of each learner were generated based on the interaction logs. The time sequence indicates the different worlds explored for the whole duration of the module implementation and the time spent on each world. The total time spent is computed by adding all the time spent in each exploration in the time sequence.

Knowledge assessments were done after every module. The assessments consisted of questions about the module topic and short essays. Questions for the short essays were the following:

Module 1 – "How can human activities affect the balance of ecosystems?"

Module 2 – "As people have moved from place to place, they have often brought plants and animals with them. How might the introduction of a new species of plant and animal in an area have disastrous effects on organisms already in that area?" and "How is biodiversity important in an organism's survival?"

The GEQ assessed players' levels of immersion, challenge, and negative / positive affect among others. The GEQ instrument used in this study is an abridged version of the instrument developed by IJsselsteijn et al. (2013). Two (2) items from GEQ: "*I felt frustrated*" and "*I felt bored*", were used for the data on student self-reported frustration and boredom respectively.

The SIQ is an instrument that is used to describe and measure the propensity of pursuing (or not pursuing) a career in STEM (See Table 1). Questions in the SIQ assessed students' self-efficacy as well as how they regarded the relevance and usefulness of science.

Table 1. *Items in SIQ*

SCCT Constructs		SIQ Item
Self-Efficacy (SE)	SE1	I know I can do well in science.
	SE2	I think Science is challenging to learn.
Outcome Expectations (OE)	OE1	After I finish high school, I will use Science often.
	OE2	I believe that I can use Math and Science to solve problems in the future.
Interests (IN)	IN1	I enjoy Science activities.
	IN2	I enjoy solving Science and Math problems.
Choice Goals (CG)	CG1	Learning Science will help me get a good job.
	CG2	Knowing how to use Math and Science together will help me to invent useful things.
	CG3	Understanding engineering is not important for my career.
Choice Actions (CA)	CA	I try to get a good grade in science because I have an interest in science jobs.

At the start of the data collection, the students were asked to answer the SIQ. They then explored WHIMC, following the instructions in the teacher-prepared learning modules. After the exploration, the learners answered the knowledge assessment, the GEQ, and the SIQ again.

5. Analysis and Results

5.1 How do the behaviors of the learners who self-reported frustration and boredom differ from those who did not?

We started the analysis by first dividing the learners into two groups, the high-performing group and the low-performing group, based on the knowledge assessment. The high-performing learners were the learners who scored equal to or above the mean and the low-performing learners were the learners who scored below the mean score.

We then focused our study of in-game behavior on idle time, time exploring other worlds, and task completion. Idle time was defined as disengaging from the learning environment. It was operationalized as the time during which the learner stayed at a location point, with no change in distance traveled and no other activity, i.e. accessing science tools or making an observation.

We calculated the cumulative distance traveled by each learner in one-minute intervals. We also calculated the slopes between one-minute intervals and then calculated the average slope for each learner to determine how the learner explored the worlds. Lower values of the average slope mean there is a small change in the cumulative distance. Figure 2 shows that most of the learners with frustration have smaller slopes in the change of cumulative distance per minute and there are times that the slope is zero which indicates that the learner did not move from its location.

The GEQ used Likert scale thus self-reported frustration and boredom were treated as ordinal data and the Spearman rho assessment was used in the analysis. The average slope of change of distance per minute was negatively correlated with frustration for all learners and the result was marginally significant ($r = -0.136$, $p = 0.1$). This implies that frustrated students tend to move around the WHIMC world less. This tendency is most common among low-performing students as we found a significant positive correlation between these students' idle time and frustration ($r = 0.287$, $p = 0.04$). The average slope was not significantly correlated with frustration for both high- and low-performing learners. Idle time was also not significantly correlated with frustration for all learners and high-performing learners.

Both average slope and idle time were not significantly correlated with boredom for all learners, as well as for both high- and low-performing learners.

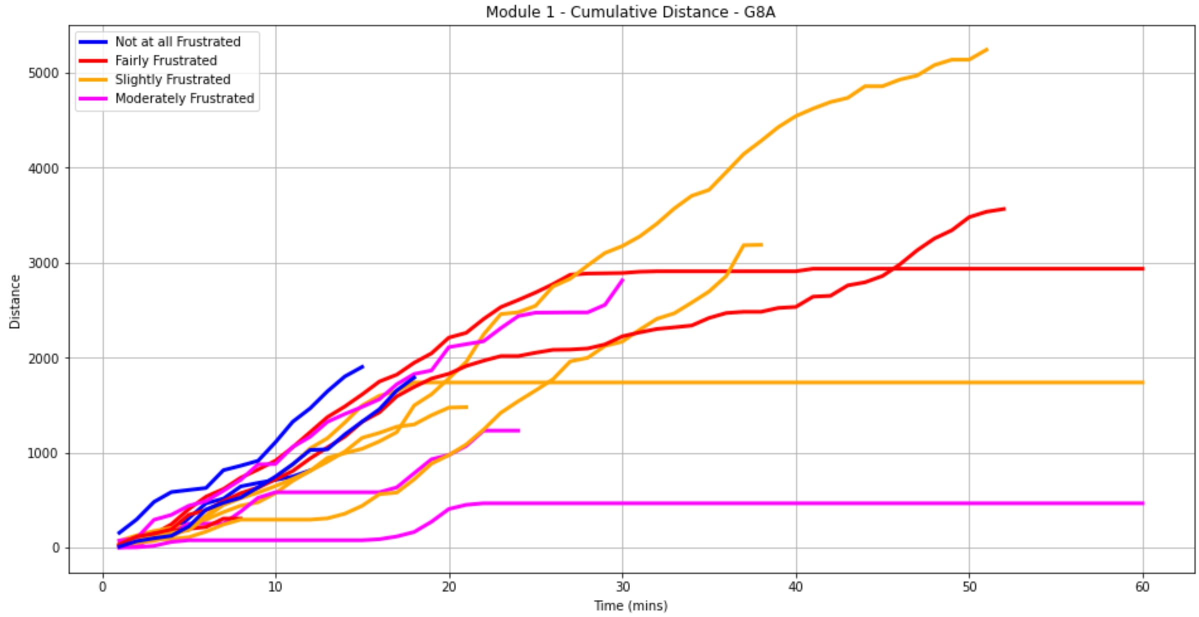


Figure 2. Cumulative Distance per minute interval of Learners and level of frustration in Module 1

Time exploring other worlds indicates the time that the learner visited worlds that were not part of the learning task. The time exploring other worlds was not significantly correlated with frustration or boredom for all learners, as well as for all high and low-performing learners.

Task completion refers to whether the missions per module were completed or not. Task completion of module 1 was marginally negatively correlated with frustration ($r = -0.185$, $p = 0.1$) among all learners and among high-performing learners ($r = -0.245$, $p = 0.1$). Task completion was not significantly correlated with boredom for all learners, as well as for all high- and low-performing learners.

5.2 What is the relationship between learners' self-reported frustration and boredom and their assessment outcomes?

We correlated the self-reported frustration and boredom with learners' assessment scores to determine the relationship between students' affective states and how they relate to their assessment outcomes. There was no significant correlation between frustration and assessment scores, but we found a significant negative correlation ($r = -0.17$, $p = 0.025$) between boredom and assessment scores. This suggests that bored students tend to learn less from playing with the game.

5.3 How did the STEM interest of the learners who self-reported frustration and boredom change after using WHIMC?

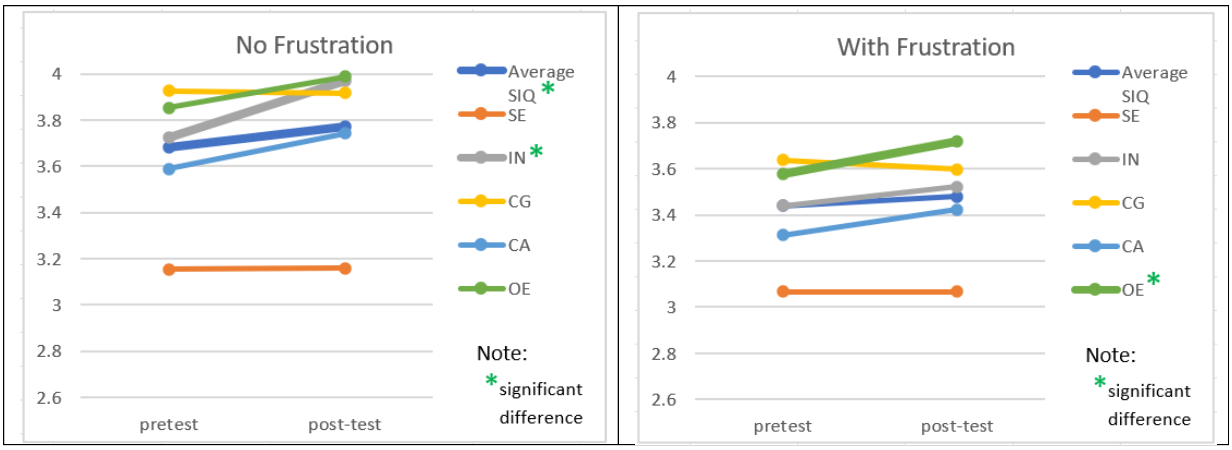
We then compared the pre-test and post-test results of the SIQ to determine how the STEM interest of the learners who self-reported frustration and boredom changed after they played with WHIMC. For self-reported frustration, we separated the learners into two groups: the first group was those learners who answered, "not at all" on the GEQ item "I felt frustrated" and the second group was those learners who have certain levels of frustration and answered "slightly", "moderately", "fairly", and "extremely". For self-reported boredom, another set of comparisons was also done. The learners were also separated into two groups: the first group was those learners who answered, "not at all" on the GEQ item "I felt bored" and the second group was those learners who have certain levels of boredom and answered "slightly", "moderately", "fairly", and "extremely".

Figure 3 shows that, on average, learners had a pretest SIQ of 3.56 and a post-test SIQ of 3.62. The difference was significant ($t(174)=2.13$, $p=0.03$). We also find significant differences in the Interest category (t(174)=3.26, p<0.01) and Outcome Expectations category (t(174)=2.69, p=0.01). This implies that learners had increased interest in STEM after playing WHIMC.

Figure 3. Comparison between pretest and post-test of SIQ of all learners

We now examine the differences in pretest and post-test SIQ among learners who expressed or did not express frustration. Figure 4a shows that, on average, learners who self-reported no frustration had a pretest SIQ of 3.68 and a post-test SIQ of 3.77. The difference was significant (t(84)=2.14, p=0.04). In the dimension of Interest in SIQ, the students had a pretest SIQ of 3.72 and a post-test SIQ of 3.97 and the difference between pretest and post-test was also significant (t(84)=3.92, p<0.01). This implies that students who self-reported no frustration had increased interest in STEM after playing WHIMC.

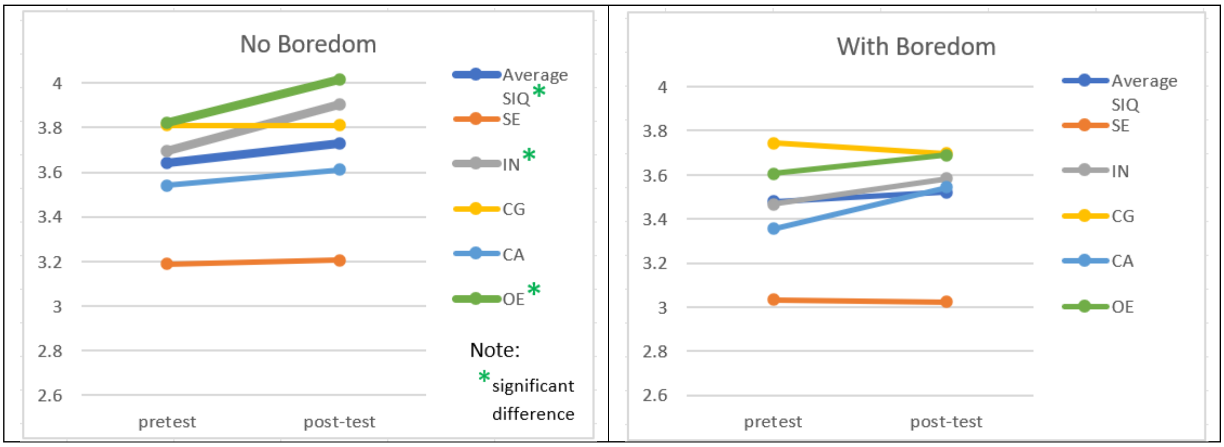
We compared the pretest and post-test SIQ scores of those who expressed frustration. Learners who self-reported frustration had a pretest SIQ of 3.44 and a post-test SIQ of 3.48 (see Figure 4b). We did not find a significant difference. However, when we examined student responses per category of SIQ, there was a significant increase in the category of Outcome Expectations (t(89)=2.07, p=0.04).



(a) (b)

Figure 4. Comparison between pretest and post-test SIQ of all learners who self-reported no frustration (a) and learners who self-reported frustration (b)

We also examined the differences in pretest and post-test SIQ among learners who expressed or did not express boredom. Learners who self-reported no boredom had a pretest SIQ of 3.64 and a post-test SIQ of 3.73 (see Figure 5a). The difference was significant (t(84)=2.6, p=0.01). We also find significant differences in the Interest category (t(84)=3.72, p<0.01) and Outcome Expectations category (t(84)=2.8, p=0.01). On the other hand, there was no significant difference in pretest and post-test SIQ among learners who expressed boredom (see Figure 5b). This implies that learners who did not express boredom had increased interest in STEM after playing WHIMC but those who expressed boredom had no significant change in STEM interest.



(a) (b)

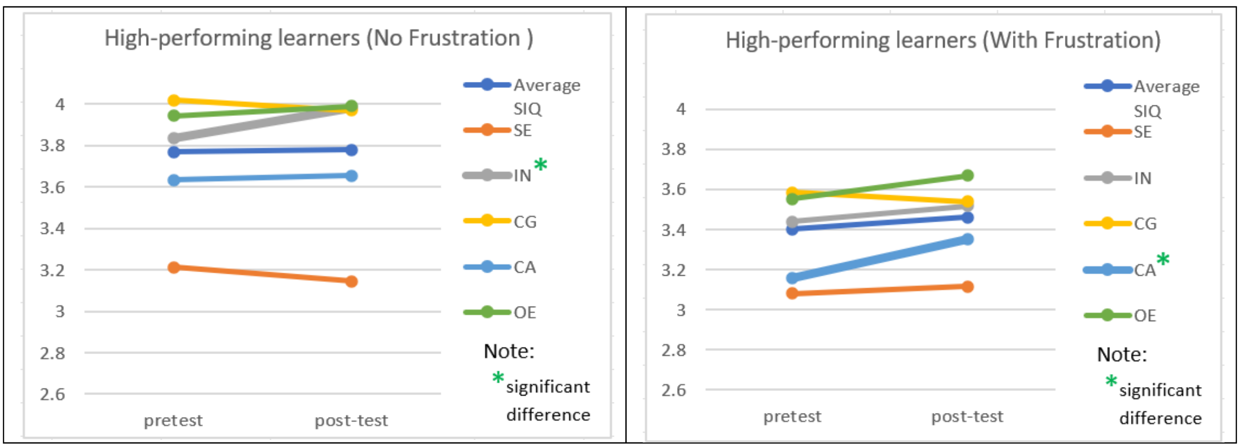
Figure 5. Comparison between pretest and post-test SIQ of all learners who self-reported no boredom (a) and learners who self-reported boredom (b)

5.4 How did the STEM interest of high- and low-performing learners who self-reported frustration and boredom change after using WHIMC?

Finally, we then tried to determine whether frustration affected high and low-performing learners differently. We found that high-performing learners who self-reported no frustration had a pretest SIQ of 3.77 and a post-test SIQ of 3.78 (see Figure 6a). The difference was not significant. However, there was a significant difference in the SIQ category of Interest ($t(51)=2.01$, $p=0.05$) with a pretest SIQ of 3.84 and a post-test SIQ of 3.98. On the other hand, the high-performing learners who self-reported frustration had a pretest SIQ of 3.4 and a posttest SIQ of 3.46 (see Figure 6b), in which the difference was not significant. However, there was a significant difference in the Choice Actions category ($t(56)=1.96$, $p=0.05$) with a pretest SIQ of 3.16 and a post-test SIQ of 3.35.

Figure 7a shows that low-performing learners who self-reported no frustration had a pretest SIQ of 3.55 and a post-test SIQ of 3.76 ($t(32)=2.8$, $p=0.01$). We also find significant differences in the Interest category ($t(32)=3.66$, $p<0.01$) which has a pretest SIQ of 3.55 and a post-test SIQ of 3.95, and in the Outcome Expectation category ($t(32)=2.06$, $p=0.05$) which has a pretest SIQ of 3.7 and a post-test SIQ of 3.99. On the other hand, low-performing learners who self-reported frustration had a pretest SIQ of 3.5 and a post-test SIQ of 3.53 (see Figure 7b), in which the difference was not significant.

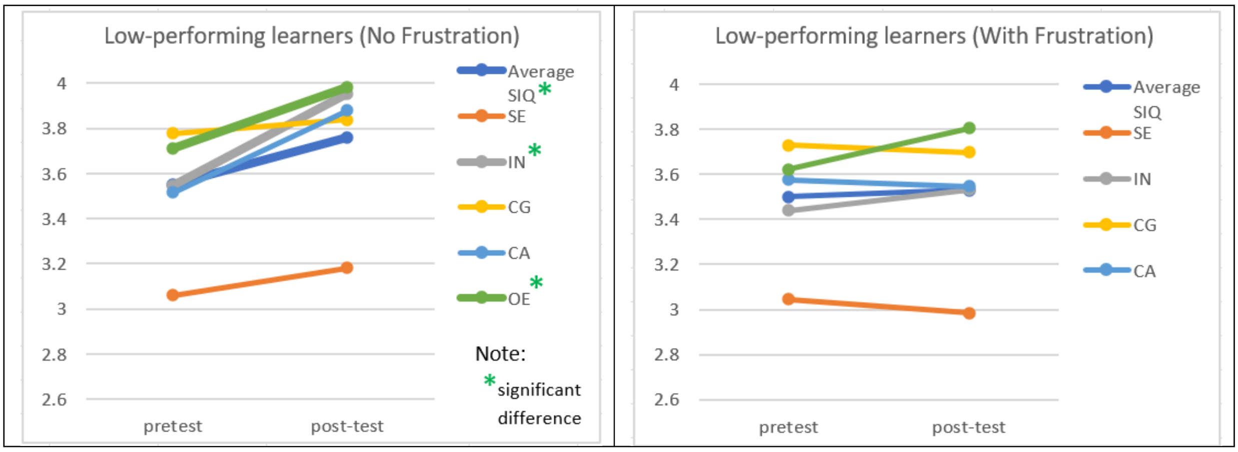
Based on the results, both the high-performing and low-performing learners who self-reported no frustration showed increases in some aspects of STEM interest.



(a) (b)

Figure 6. Comparison between pretest and post-test SIQ of high-performing learners

who self-reported no frustration (a) and those who self-reported frustration (b)



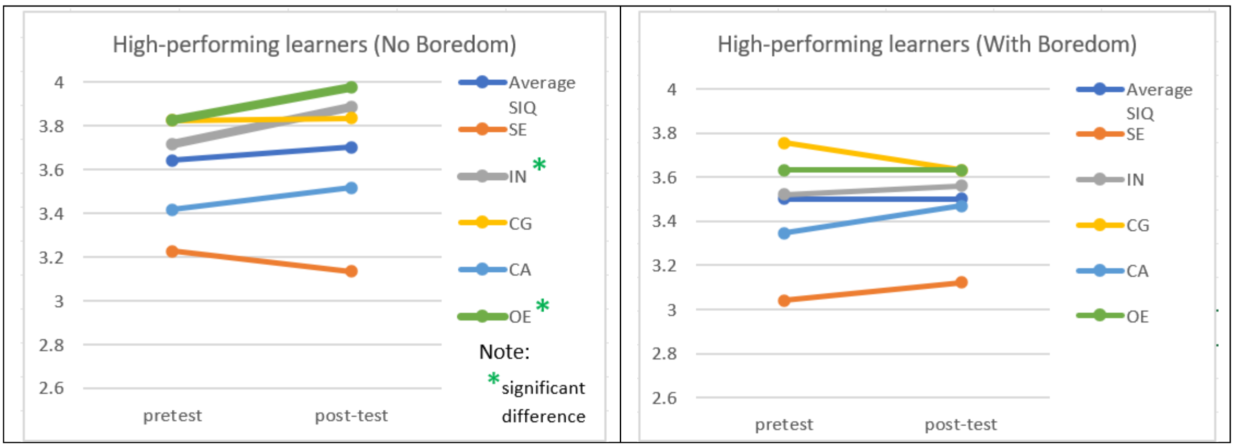
(a) (b)

Figure 7. Comparison between pretest and post-test SIQ of low-performing learners who self-reported no frustration (a) and those who self-reported frustration (b)

We also tried to determine whether boredom affected high and low-performing learners differently. The high-performing learners who self-reported no boredom had a pretest SIQ of 3.64 and a post-test SIQ of 3.7 (see Figure 8a). The difference was not significant. However, there was a significant difference in the Interest category (t(59)=2.95, p=0.05) with a pretest SIQ of 3.72 and a post-test SIQ of 3.88. On the other hand, the high-performing learners who expressed boredom had a pretest SIQ of 3.5 and a post-test SIQ of 3.5 (see Figure 8b), in which there was no significant difference. Neither did we find significant differences in the disaggregated SIQ categories.

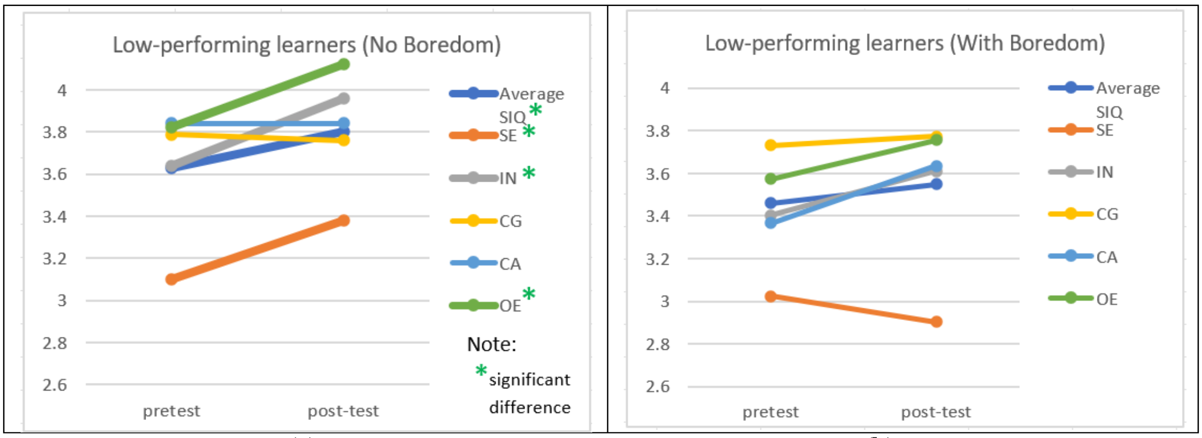
Low-performing learners who did not express boredom had a pretest SIQ of 3.63 and a post-test SIQ of 3.8 (see Figure 9a). The increase was significant (t(24)=2.1, p=0.05). We also found significant differences in the Interest category (t(24)=2.32, p=0.03) with a pretest SIQ of 3.64 and a post-test SIQ of 3.96, Self-Efficacy category (t(24)=2.58, p=0.02) with a pretest SIQ of 3.1 and a post-test SIQ 3.38, and Outcome Expectations category (t(24)=2.08, p=0.05). On the other hand, low-performing learners who expressed boredom had a pretest SIQ of 3.46 and a post-test SIQ of 3.55 (see Figure 9b), in which there was no significant difference.

Based on the results, high-performing learners who did not express boredom showed an increase in some aspects of STEM interest after playing WHIMC. Low-performing learners who did not express boredom had increased levels in the Self-Efficacy category as well as the aggregated SIQ. For those who self-reported boredom, there was no significant change in the STEM interest after using WHIMC.



(a) (b)

Figure 8. Comparison between pretest and post-test SIQ of high-performing learners who self-reported no boredom (a) and those who self-reported boredom (b)



(a) (b)

Figure 9. Comparison between pretest and post-test SIQ of low-performing learners who self-reported no boredom (a) and those who self-reported boredom (b)

6. Conclusion

In this paper, we presented an attempt to analyze the behavior of learners who self-reported frustration and boredom and determine the relationship between frustration and boredom to their in-game behaviors, knowledge assessment outcomes, and STEM interest.

We found that learners who are frustrated tend to move around the WHIMC world less. Frustrated learners also tended not to complete the missions or tasks. Low-performing frustrated learners tend to have more idle time, which is an indicator of off-task behavior.

When we examined the relationship between learners' frustration and boredom with their knowledge assessment outcomes, bored learners tended to have lower academic outcomes.

Learners who did not express frustration or boredom had increased interest in some categories of STEM interest after using WHIMC. Frustrated learners also had increased SIQ in the Outcome Expectations category while learners who expressed boredom had no significant change in SIQ. Frustrated learners tend to believe that after they finish high school, they will use Science often, and they also tend to believe that they can use Math and Science to solve problems in the future.

Both high- and low-performing learners who did not express frustration and boredom had increased interest in some categories of STEM. High-performing learners who expressed frustration had increased SIQ in the Choice Actions category. On the other hand, there was no significant difference in the pretest and post-test SIQ for both high- and low-performing bored learners. Note that low-performing learners who did not express frustration or boredom had increased SIQ scores, implying that the absence of these states was beneficial to this category of learners.

This study contributes to what is known about how affect relates to outcomes such as test scores, in-game behaviors, and attitudes regarding domains such as STEM. In general, students who expressed boredom or frustration tended to demonstrate minimal of engagement with STEM, as indicated by lower test scores and flat SIQ results. The study also contributes by raising the larger question about whether and how game-based interventions can be used to mitigate affective states such as frustration and boredom. How can educators best use games to pave the way towards greater student engagement in STEM? This is a research area that is open for further study.

Acknowledgements

The authors thank H Chad Lane and Jeff Ginger for their enthusiastic collaboration, Dominique Marie Antoinette B. Manahan, Jonathan D. Casano, and Ma. Rosario Madjos for their support, the Ateneo Laboratory for the Learning Sciences, the Ateneo de Manila University, the Department of Science and

Technology (DOST) as funding agency, and the Department of Science and Technology's Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) as monitoring agency for the grant entitled, "Nurturing Interest in STEM among Filipino learners using Minecraft."

References

All, A., Nuñez Castellar, E. P., & Van Looy, J. (2016). Assessing the effectiveness of digital game-based learning: Best practices. *Computers & Education*, *92-93*, 90–103. https://doi.org/10.1016/j.compedu.2015.10.007

Andres, J. M. L., Rodrigo, M. M. T., Baker, R. S., Paquette, L., Shute, V. J., & Ventura, M. (2015). Analyzing Student Action Sequences and Affect While Playing Physics Playground. *In EDM (Workshops).*

Baker, R. S. J. (2007). Modeling and understanding students' off-task behavior in intelligent tutoring systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*

Baker, R. S. J., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241. https://doi.org/10.1016/j.ijhcs.2009.12.003

Beserra, V., Nussbaum, M., Zeni, R., Rodriguez, W., & Wurman, G. (2014). Practising arithmetic using educational video games with an interpersonal computer. *Journal of Educational Technology & Society*, *17*(3), 343-358.

Bitner, J. (2021). What is Minecraft? Gaming. Digital Trends. Accessed from the Digital Trends website: https://www.digitaltrends.com/gaming/what-is-minecraft/

D'Mello, S., & Graesser, A. (2012). Dynamics of Affective States during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition and Emotion*, *25*(7), 1299–1308. https://doi.org/10.1080/02699931.2011.613668

IJsselsteijn, W. A., De Kort, Y. A., & Poels, K. (2013). The game experience questionnaire.

Käser, T., & Schwartz, D. L. (2020). Modeling and analyzing inquiry strategies in open-ended learning environments. *International Journal of Artificial Intelligence in Education*, 30(3), 504–535.

Richey, J. E., Andres-Bray, J. M., Mogessie, M., Scruggs, R., Andres, J. M. A. L., Star, J. R., Baker, R. S., & McLaren, B. M. (2019). More confusion and frustration, better learning: The impact of erroneous examples. *Computers & Education*, 139, 173–190. https://doi.org/10.1016/j.compedu.2019.05.012

Liu, Z., Pataranutaporn, V., Ocumpaugh, J., & Baker, R.S.J.d. (2013). Sequences of Frustration and Confusion, and Learning. *Proceedings of the 6th International Conference on Educational Data Mining*, 114-120.

Pardos, Z. A., Baker, R. S. J. D., San Pedro, M., Gowda, S. M., & Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *Journal of Learning Analytics*, *1*(1), 107–128. https://doi.org/10.18608/jla.2014.11.6

Rajendran, R., Iyer, S., Murthy, S., Wilson, C., & Sheard, J. (2013). A theory-driven approach to predict frustration in an ITS. *IEEE Transactions on Learning Technologies*, *6*(4), 378-388.

Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing*, *5*(1), 45–56. https://doi.org/10.1109/t-affc.2013.27

Sabourin, J., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. *In International Conference on Artificial Intelligence in Education* (pp. 534-536). Springer, Berlin, Heidelberg.

Sawyer, R., Smith, A., Rowe, J., Azevedo, R., & Lester, J. (2017). Enhancing student models in game-based learning with facial expression recognition. *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. https://doi.org/10.1145/3079628.3079686

Tze, V. M., Daniels, L. M., & Klassen, R. M. (2015). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review*, *28*(1), 119–144.

Vogel-Walcutt, J. J., Fiorella, L., Carper, T., & Schatz, S. (2012). The definition, assessment, and mitigation of state boredom within educational settings: A comprehensive review. *Educational Psychology Review, 24*(1), 89-111.

Zhang, Y., Paquette, L., Baker, R. S., Ocumpaugh, J., Bosch, N., Biswas, G., & Munshi, A. (2021). CAN strategic behaviour facilitate confusion resolution? the interplay between confusion and metacognitive strategies in Betty's brain. *Journal of Learning Analytics*, *8*(3), 28–44. https://doi.org/10.18608/jla.2021.7161

Zhonggen, Y. (2019). A meta-analysis of use of serious games in education over a decade. *International Journal of Computer Games Technology*, 2019, 1–8. https://doi.org/10.1155/2019/4797032

WHIMC. (n.d.) What-If Hypothetical Implementations using Minecraft. Available online: https://whimcproject.web.illinois.edu/