



# A Bayesian Analysis of Adolescent STEM Interest Using Minecraft

Matthew Gadbury<sup>(✉)</sup> and H. Chad Lane

University of Illinois Urbana-Champaign, Champaign, USA  
{gadbury2,hclane}@illinois.edu

**Abstract.** Minecraft continues to be a popular digital game throughout the world, and the ways in which adolescents play can provide insight into their existing interests. Through informal summer camps using Minecraft to expose middle school students to concepts in astronomy and earth science, we collected self-reports of STEM and Minecraft interest, as well as behavioral log data through player in-game interactions. Finding relationships between in-game behaviors and individual interest can provide insight into how educational experiences in digital games might be designed to support learner interests and competencies in STEM. Bayesian model averaging of data across camps was implemented to address the relatively small sample size of the data. Results revealed the important role of existing interest and knowledge for developing and sustaining interest.

**Keywords:** Digital games · informal learning · interest · Bayesian analysis

## 1 Introduction and Background

One conception of interest is that of a psychological state marked by heightened attention and focus, as well as a motivational force pushing an individual to explore and find additional information about an object or domain [3]. Interests can be activated or “triggered” by events in the environment, such as novelty, autonomy, personal relevance, or community [8], which in turn can lead to individuals developing greater competence and interest in the domain [5]. Adolescence tends to be a period when students are beginning to form identities, and this is also a time when interest in Science, Technology, Engineering, & Math (STEM) falters for many [7]. Utilizing technology to trigger and maintain interest in STEM for adolescents shows some evidence of being effective for developing interest.

Technology has been identified as both a trigger of interest learner interest [8] and a conduit by which other facets of interest development can be applied through formal or informal educational spaces. Minecraft remains one of the most played digital games in the world. Given its popularity, Minecraft stands as a low barrier educational tool for many students, in that students should be able to engage content without spending much time learning controls and how to navigate the environment. Research conducted using Minecraft has shown promise for increasing motivation to learn STEM content [1]. For these reasons, we use Minecraft to explore the following research question: To what

extent does a STEM-focused Minecraft summer camp influence adolescent interest in STEM and what is the role of Minecraft interest?

## 2 Methods

### 2.1 Participants

Participants ( $n = 96$ ) included middle school and early high school students ages 11–15 taking part in week-long summer camp programs implemented by our research team. A total of 7 separate camps across 3 separate sites (3 in the Western United States, 2 in the Midwestern U.S., and 2 in the Eastern U.S.) were held during the summer of 2022. Due to illness and absences amounting to more than half the camp days in a week, data of 22 participants were removed from the final count, resulting in ( $n = 74$ , 32% female). Breakdown of race/ethnicity showed: 29% White, 24% Black, 13% Hispanic, 3% Asian, 2% Native American, 8% Other and 21% Prefer Not to Answer (PNA). Written or online consent was obtained from at least one parent or guardian of each participant prior to participation.

### 2.2 Materials and Procedure

Each participant was provided with a laptop, mouse, and access to Minecraft: Java Edition. Most of each camp was spent playing custom-built worlds introducing themes in Astronomy and Earth Science in Minecraft. Worlds were inspired by “What if” questions posed by astronomers, such as “What if Earth had no moon?”. In total, our server hosts 2 orientation worlds, 5 hypothetical “What if” worlds, and 5 known exoplanets.

Each camp consisted of 5, 3-h meetings spanning one week. A pre-survey on STEM interest and Minecraft interest, validated by our lab but unpublished, was administered before the start of each camp. On the first day of each camp, participants were led through the 2 orientation worlds. The following 2 days had participants explore the rest of the curated worlds and make in-game observations inferring what is happening and why. Additionally, after exploring each world for 25–30 min, a set of 3–4 open-ended self-explanations were pushed out to each participant. The final 2 days of each camp had participants build habitats on Mars, considering known challenges to human survival on Mars. Participants also completed post-surveys on STEM and Minecraft interest, which were identical to the pre-survey.

### 2.3 Analysis

Assessing interest in STEM and Minecraft, models used either STEM interest or Minecraft interest as the dependent variable. The STEM survey consisted of 20 Likert-type questions asking students how interested they are in STEM activities, with 1 = “Not at all interested” to 5 = “extremely interested”. An example of a STEM survey question is, “How interested are you in using numbers to confirm ideas and solve problems?”. Scores were aggregated across questions. The same approach was taken with the Minecraft survey, which consisted of 20 Likert-type questions asking students how

interested in STEM-related Minecraft activities they are. An example of a Minecraft interest question is, “How interested are you in learning what colors each biome has and why?”. Demographic data, such as age and race/ethnicity, are also included in the final models.

### **Self-explanations (SEs)**

Participants completed a total of 22 open-ended self-explanations. An example of a self-explanation prompt is “How would you define habitability?” or “What might be a good way to generate energy on Earth without a moon and why?”. Answers were scored based on correctness and from 0 to 3. All self-explanations from the summer were scored by two graduate students. The aggregate percent agreement was 73%. All disagreements were resolved through discussion. After finalizing scores, an average score across answers was calculated for each participant.

### **Observations**

In-game observations were collected for each participant. Participants were encouraged repeatedly throughout the camp to make scientific observations about what they saw in the game and why they think it is the case. An example observation participant might make is, “Because there are two moons, more light is reflected to Earth”. Only the frequency of in-game observations made is used in this paper, which is considered a measure of behavioral engagement with the game and content.

### **Exploration**

Location data was captured for each participant every 3 s. Exploration was calculated as the number of squares a participant crossed into on a 10x10 grid overlay of each map. Exploration was averaged across all maps visited to provide a mean exploration measurement for each participant.

### **Bayesian Analysis**

Given the small sample size of the data, a Bayesian analysis is a more robust approach at analyzing the data than using frequentist statistics [6]. A non-informative prior is used in this approach, however the posterior probability distribution estimated from this research can be used for future interest research as the starting prior. We can establish a stronger connection between the predictors and dependent variable by estimating the probability of contribution, and this is accomplished using Bayesian Model Averaging (BMA). BMA is a way to estimate parameters that averages the predictions of different models being considered, and each model is given a weight based on its probability [4].

Correlations were run first to examine the relationship between the variables. All variables are then centered to avoid multicollinearity. All analyses were conducted in R using the “BMA” package.

## **3 Results**

A paired samples t-test was first conducted to see if there were any changes in STEM interest and Minecraft interest before and after the summer camp. The results from the pre-survey ( $M = 73.82$ ,  $SD = 15.9$ ) and post-survey ( $M = 74.39$ ,  $SD = 14.77$ ) of STEM

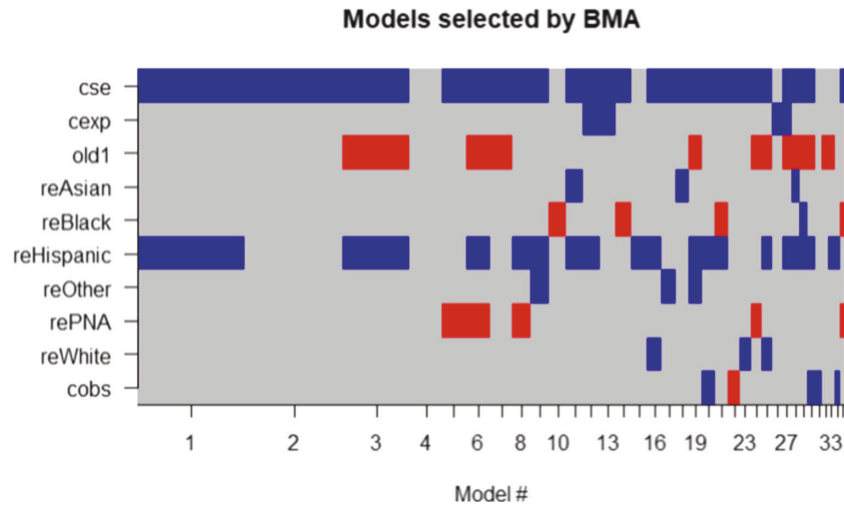
interest showed there was no change in STEM interest as a result of the camp  $t(61) = 0.11$ ,  $p = 0.914$ . Looking at Minecraft interest, the difference between the pre-survey ( $M = 70.93$ ,  $SD = 14.68$ ) and the post-survey ( $M = 68.13$ ,  $SD = 14.71$ ) was also non-significant  $t(60) = 1.45$ ,  $p = 0.152$ .

Table 1 shows correlations between analyzed variables. All continuous variables were centered before using Bayesian analysis.

**Table 1.** Correlation table of measured variables

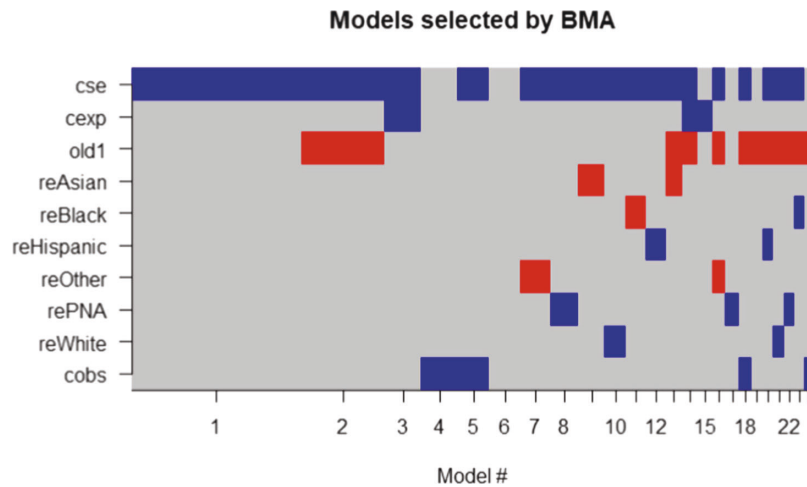
|      | SInt | MInt | SE   | Obs  | Exp |
|------|------|------|------|------|-----|
| SInt | --   |      |      |      |     |
| MInt | 0.61 | --   |      |      |     |
| SE   | 0.29 | 0.31 | --   |      |     |
| Obs  | 0.12 | 0.24 | 0.48 | --   |     |
| Exp  | 0.16 | 0.19 | 0.23 | 0.28 | --  |

Note: SInt is STEM Interest, MInt is Minecraft Interest, SE is self-explanation score, Obs is observation frequency, and Exp is exploration metric.



**Fig. 1.** Bayesian Model Averaging with STEM interest as dependent variable

BMA was implemented to understand how much each predictor contributes to the overall model, and which predictors are candidates for removal. See Fig. 1 for BMA results. Blue bars represent a positive predictor included in a model and red represents a negative predictor included in a model. What is clear from BMA is that self-explanations appear most frequently in models. The only other predictor that appears frequently is the positive effect of Hispanic on STEM interest. Interestingly, age negatively predicted STEM interest, where older learners showed less interest in STEM.



**Fig. 2.** Bayesian Model Averaging with Minecraft interest as dependent variable

As with STEM interest, Age is a negative predictor of Minecraft interest. Taken together, this might suggest that as learners age their interest in STEM decreases along with interest in Minecraft. Figure 2 shows the results of BMA. Again, self-explanations prove to be the most common predictor across models.

## 4 Discussion

No change in STEM or Minecraft interest between the beginning and end of the camp experience was observed. Interest in Minecraft and STEM showed moderate correlations with each other, which may point to inherent STEM activities and content built into Minecraft. Finding out what activities and camp aspects contribute most to STEM interest is needed to allay uncertainty as to how interest is triggered in the experience and what learners find most engaging.

Data collected reflects unique parts of the informal environment that can be measured using a digital game: observation frequency, exploration patterns, and in-game self-explanation responses. Given the novelty of the “What if” scenarios, we thought exploration would be quite extensive and would be reflective of participant information-seeking behavior for STEM content. However, exploratory behavior proved to be a very weak predictor of STEM interest, as did frequency of observations. The possibility exists that participants did not explore extensively but did take time to interact with non-playable characters (NPCs) who provide detailed information about each hypothetical. The only predictor that bore any substantial relationship to STEM interest was average score on the self-explanations, which was essentially a knowledge assessment for each world aimed at measuring cognitive interest. Higher knowledge in a domain is usually an indicator of higher interest [2].

Bayesian analysis is a powerful tool for exploring the triggering and development of interest. The data analyzed here can be used to inform a prior distribution for future analysis of camp data. Also, given the frequent small sample sizes we encounter when running summer camps and after school programs, Bayesian analysis can be used with relatively small sample sizes to estimate the posterior probability distribution. The power of BMA arises from the idea that no predictor ever dominates the entire model but also each predictor does contribute to some degree [4].

Future analyses will divide participants by high or low interest to see if learners with low interest experience any changes in interest, and if the magnitude of change is greater than those entering with high interest. In addition, analysis can benefit from precise measurements that capture situational interest from specific aspects of the environment, and that measure interest in domains dominating camp content.

## References

1. Baek, Y., Min, E., Yun, S.: Mining educational implications of minecraft. *Comput. Schools* **37**(1), 1–16 (2020). <https://doi.org/10.1080/07380569.2020.1719802>
2. Fastrich, G.M., Murayama, K.: Development of interest and role of choice during sequential knowledge acquisition. *16* (2020)
3. Hidi, S.E., Ann Renninger, K.: On educating, curiosity, and interest development. *Current Opinion Behav. Sci.* **35**, 99–103 (2020). <https://doi.org/10.1016/j.cobeha.2020.08.002>
4. Hinne, M., Gronau, Q.F., van den Bergh, D., Wagenmakers, E.-J.: A conceptual introduction to Bayesian model averaging. *Adv. Methods Pract. Psychol. Sci.* **3**(2), 200–215 (2020)
5. Hulleman, C.S., Thoman, D.B., Dicke, A.-L., Harackiewicz, J.M.: The promotion and development of interest: the importance of perceived values. In: O’Keefe, P.A., Harackiewicz, J.M. (eds.) *The Science of Interest*, pp. 189–208. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-55509-6\\_10](https://doi.org/10.1007/978-3-319-55509-6_10)
6. König, C., van de Schoot, R.: Bayesian statistics in educational research: a look at the current state of affairs. *Educ. Rev.* **70**(4), 486–509 (2018). <https://doi.org/10.1080/00131911.2017.1350636>
7. Maltese, A.V., Tai, R.H.: Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among U.S. students. *Sci. Educ.* **95**(5), 877–907 (2011). <https://doi.org/10.1002/sce.20441>
8. Ann Renninger, K., Bachrach, J.E., Hidi, S.E.: Triggering and maintaining interest in early phases of interest development. *Learn. Cult. Soc. Interact.* **23**(2019), 100260 (2019). <https://doi.org/10.1016/j.lcsi.2018.11.007>