AN EXPLORATORY ANALYSIS OF THE RELATIONSHIP BETWEEN INTEREST, KNOWLEDGE, AND ENGAGEMENT OF ADOLESCENTS PLAYING MINECRAFT IN STEM-FOCUSED SUMMER CAMPS

BY

MATTHEW GADBURY

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2024

Urbana, Illinois

Doctoral Committee:

Associate Professor H Chad Lane, Chair
Professor Jennifer Cromley
Associate Professor Mike Tissenbaum
Associate Professor Luc Paquette

ABSTRACT

The psychological construct of interest has been demonstrated to be a powerful tool for motivating learners to meaningfully engage with specific content or activities. When learners interact with a digital learning environment, their existing interests will likely influence the ways in which they engage, and the ways they engage will affect their level of interest in and knowledge of the topic. This study takes adolescent engagement patterns in a digital STEM learning environment, a customized version of *Minecraft*, and connects engagement patterns to the degree of interest in STEM learners enter and leave the experience with, as well as performance on knowledge assessments. Engagement is defined across three dimensions: affective, behavioral, and cognitive, and log data from the *Minecraft* server is extracted and engineered to be positioned within one of the dimensions of engagement. Time-series clustering is then used to extract patterns of behavior over time that differentiate learners. Emergent patterns revealed 5 profiles of engagement, including high performers, balanced performers, exploration only, point of interest only, and low performers. Scores on STEM, Astronomy, and *Minecraft* interest surveys were also divided by level of interest and correlated with time-series clusters. Bayesian models were used to incorporate additional factors into models that predict post-test individual interest. Knowledge and interest appeared to be deeply intertwined for these participants engaging in these *Minecraft* tasks, emphasizing the importance of knowledge for productive engagement, and increasing interest in a domain.

ACKNOWLEDGEMENTS

I want to thank all who have supported me through this process, including those who paved the way and will continue to build upon this work in the future. I was fortunate to work with Sherry Yi on this research my very first semester of grad school. Working with Sherry taught me a lot about how find meaning in data and organize it in efficient ways. She also pushed me to find conferences to present my work at, and this has proven to be very beneficial throughout my time as a grad student. I would also like to thank Meghna Paul and Kayla Zimmer for all of their hard work supporting my research, which included coding and scoring, collecting data, and even learning how to do some programming. They both approached the research with a strong desire to learn and excel in analysis, and I am extremely grateful for them. Evan Shipley and Sam Hum were also instrumental in collecting data and helping with analysis. I feel very lucky to have had them as lab mates, and I deeply appreciate their contributions.

Of course, I must thank all members of my committee. Luc Paquette was the very first instructor I had after starting my PhD program. The research methods class taught by him really eased me into the program and provided invaluable knowledge for writing and exploring methods. I have had the opportunity to lend a hand in work Luc is continuing WHIMC, and I hope to continue collaborating on projects examining learning and motivational processes through logs. Luc has always offered excellent advice on how to engineer features and think through how features align with methods and research questions. I only wish I had tapped into his expertise more throughout grad school.

I met Mike Tissenbaum at the admitted student info day in 2019, and I was struck by his enthusiasm and energy. At the time, I did not know much about his research, but taking a class with Mike made me instantly realize he would be an outstanding mentor. I always appreciate

Mike's honesty in terms of what makes good research, as well as what matters for succeeding in academia. His support has been amazing, and as with Luc, I really wish I had sought his advice and expertise more.

The whole journey really started with Jennifer Cromley (Dr. C). I took a class on learning theories with her in 2017, and I found it fascinating. I also found her teaching style refreshing and something I have tried to emulate since. Teaching with her for two semesters are experiences I will always cherish and never forget. I cannot overstate the impact Jennifer has had on my life. She has mentored and guided me as a scholar (I took almost every class she has offered!), showing me how to conduct rigorous research and appropriately apply statistical and other analytical techniques. I do not think I have ever met somebody who has impressed me more than Jennifer and inspired me to reach further and try harder. I have always endeavored to reach Jennifer's high standards, and I think this is a major reason why the quality of my work has increased over the years. I cannot thank her enough. If I do continue as a scholar, I will use everything I have learned from Jennifer to hopefully come closer to being as exceptional as she is. You are amazing!

Of course, H Chad Lane is the one who took me on and has been a stalwart support from the beginning. I met Chad prior to applying for to the PhD program, and he was so enthusiastic about me applying. Conversations with Chad were what really made me feel that leaving a job I actually enjoyed to pursue a PhD was the correct decision. He was right! Every interaction with Chad has been positive. I would come to him with untethered thoughts and ideas, and I left every meeting feeling uplifted and having more direction. He also afforded me the autonomy to really explore and determine what I want my course of research to be. I truly appreciate how Chad has been a mentor but also treated me as a colleague in this space. He put a lot of trust in me to collect data, to teach and lead class when he was away, and to represent our work at conferences. Thanks so much for giving me the chance and investing in me throughout these years. You have helped me grow and understand how to navigate this challenging landscape and feel more comfortable being in it.

I also want to thank my high school social studies teacher and coach, Keith Pogue. Not knowing it at the time, but it was him that taught me how to be disciplined and self-regulate, how to think critically, and how to write. I am not sure there are skills more important than these and he provided all three. I reflect a lot about what I learned from him, and I am grateful for his contributions to my growth.

I have thanked several people already, but my greatest thanks goes to my wife, Wang Shuo. She pushed me to pursue this degree when I had second thoughts, and if it was not for her I would not be writing this right now. She has worked tremendously hard to support me and our family throughout this process, and I am fortunate to have her as my partner. She keeps me in check but also helps me unlock potential I cannot see. I owe her a lot! I would not be here without her.

Finally, I must thank my mom, Lila, for emphasizing the importance of school throughout my formative years. She pushed me to do well, and in turn it ended up instilling in me a love for learning and a desire to know more. She is an avid reader and always encouraged me to read as much as I could. I often reflect on how important this has been throughout my life. She raised me and my siblings mostly on her own, and the fact that I am writing this now is a testament to the incredible work she has done.

This research is based upon work funded by the National Science Foundation, Grants 1713609 and 1906873. Opinions, findings, conclusions, or recommendations made in this document are those of me, the author, and do not reflect the views of the National Science Foundation.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

My goal in this dissertation study is to investigate affective and motivational states from user log data using a popular digital game, with the aim of developing a model of engagement closely linked to pre-existing individual interest and knowledge. The game in question, *Minecraft*, has become one of the most popular games in the world. As of January, 2024, *Minecraft*, has over 166 million monthly users, and 21.21% of daily user traffic in *Minecraft* originates from the U.S. (Woodward, 2024). Citing global numbers, 53% of *Minecraft* players are 21 years of age or younger, in other words close to 88 million youth actively play *Minecraft* (Woodward, 2024). Undoubtedly, *Minecraft* has achieved a great deal of popularity and name recognition around the world, especially with adolescents. Part of the popularity is due to accessibility of *Minecraft* on tablets, smartphones, consoles, and personal computers. Given the ubiquity of the game, as well as the identified opportunities to conduct research learning and motivation (Lane & Yi, 2017), *Minecraft* deserves attention regarding effectiveness for learning in formal and informal learning environments. The data used in this study comes from an NSF-funded project titled, “What-if Hypothetical Implementation in Minecraft” or “WHIMC” (Grants 1713609 and 1906873). The WHIMC project uses *Minecraft* as a vehicle for understanding how interest in a context can influence motivation to explore and learn more about STEM content, primarily Astronomy and Earth Science. Learners are presented with a variety of “what-if” scenarios, such as “What if Earth had no moon?” or “What if Earth orbited a colder sun?” (Comins, 1993). These questions pose novel scenarios that even seasoned *Minecraft* players are unlikely to have experienced in-game previously.

Novelty has been identified as trigger of interest, particularly in early phases of interest development (Renninger et al., 2019), and the scenarios presented in our customized version of

Minecraft have been shown to attract attention of players and trigger interest through novelty of the hypothetical environments (Yi, 2021; Yi, Gadbury, & Lane, 2021). The way children explore and uncover phenomenological relationships is learner-centered and meant to be intrinsically motivating. Possible actions, such as taking measurements and making observations is scaffolded by instructors, but the frequency and direction of use is determined by the individual player. Content is embedded throughout each world, providing myriad ways to explore and uncover information. An advantage of this approach is the potential for autonomy in engagement (across dimensions of behavior, affect, and cognitions) to stave off declines in interest and engagement. Additionally, behavioral educational data has historically come from constrained environments, such as in the context of a class, administering structured exams/questionnaires, or running learners through a lab study. Using open-ended data from games possesses a lot of opportunity for drawing novel inferences about behavior and learning. Markers drawn from fine-grained behavioral data may provide chances to intervene when interest is flagging or provide encouragement when an interest is peaking. Learning in context can be powerful when the context is meaningful and provides many opportunities for active engagement and interaction, and technology-rich learning environments afford capturing data to uncover such complex, contextually-driven relationships (Lajoie & Poitras, 2024). In this way, exploring the evidence for how to best use digital games for learning and motivation, while considering context can provide valuable information into learners' states and how they persist or change over time.

An impetus for such research stems from the trend of adolescents' declining interest in continuing to take STEM courses and pursue STEM-related career paths, termed the "leaky pipeline" (Maltese & Tai, 2011). Declines in STEM interest are especially prevalent among underrepresented groups in STEM (Burt & Johnson, 2018; Calabrese Barton & Tan, 2019;

Daniels & Robnett, 2021). Research in interest has posited that developmental factors begin occluding STEM interest, as adolescents' concern grows more for their personal identity and relations with others (Renninger, 2009). Adolescents may drop a previously dominant STEM interest if it negatively impacts their relationships with friends or presents them as an outsider in school settings. Additionally, STEM courses become more challenging as learners enter high school, causing many students to shift interest towards areas they feel more competent at understanding. Appraisals of competence and novelty have been shown as key ingredients for understanding how an individual's interest in a topic develops or not (Silvia, 2005). If a learner feels incapable of comprehending new information they may disengage and feel less interested in the content. Therefore, reducing barriers of participation are an important step towards maintaining existing STEM interest or triggering situational interest in STEM during a period of life where interests are highly malleable (Nye et al., 2021). As a game many adolescents have some familiarity with, *Minecraft* can serve as a familiar platform for engaging STEM content in informal learning environments, such as after-school programs or summer camps. My research centers around the use of *Minecraft* in informal learning settings and how the game can be used to trigger interest in STEM.

Considerable effort has been placed into identifying the ways learners engage in digital environments and how engagement translates into learning outcomes or achievement (Kew & Tasir, 2022; Martin & Borup, 2022; Wiedbusch et al., 2023). Some studies seek to predict when a learner begins disengaging from a learning module or course and what supports are necessary to retain the student's engagement in the learning environment (Bernacki et al., 2020). Through personalized supports, students who previously showed signs of disengagement may persist and maintain high engagement levels in the domain after receiving supports (Bernacki et al., 2021).

Educational research has focused previously on the factors that suggest disengagement in digital learning environments (Gobert et al., 2015), and researchers are calling for theoretically grounded approaches to understanding engagement in digital learning environments (Martin & Borup, 2022), including how and why learners attain varying levels of engagement (Dewan et al., 2019; Kew & Tasir, 2022). Current thinking suggests methods for improving performance and promoting engagement should not rely heavily on demographic data that is unalterable, but instead on the behavioral factors that can be nudged (Arizmendi et al., 2022). By providing an operationalization in *Minecraft* for the theory-driven dimensions cognitive, behavioral, and affective engagement (Fredricks, 2011; Fredricks et al., 2004; Sinatra et al., 2015), I aim provide a clearer picture of which activities (i.e. conversing with an in-game character, generating explanations, taking a measurement) in a digital game environment map to patterns of high engagement and knowledge construction, and seek means to promote engagement and interest. Furthermore, vast amounts of digital game research have focused on features of the game that engender engagement, ignoring the context and situation of the learners (Steinkuehler & Squire, 2024). My aim is to consider both the features of the game and the individual, vis-à-vis interest and knowledge, to uncover how these factors relate to productivity and engagement with the learning environment.

Interest is closely tied to engagement, as individuals interested in a domain or activity are by necessity engaged, yet those engaged may not always be interested (Renninger & Hidi, 2016). Individuals interested in a topic will exhibit interest through behaviors of engaging and re-engaging, cognitive effort to expand understanding, and through affective investment in the topic (Krapp & Prenzel, 2011). Given the relationship between learner interest in a domain and engagement, opportunity exists to tie patterns as detected by analyzing log files to student interest and thereby use digital learning environments to develop existing interest and/or promote new, related interests. In this dissertation, I investigate the following research questions to address the disconnect between a learner's interest in a domain or activity, their knowledge accumulation, and the dimensions of engagement they portray in a digital learning environment:

RQ: To what extent can interest be monitored and tracked in a digital, game-based STEM learning environment? How can the psychological and motivational construct of interest be monitored and track changes in STEM interest and knowledge across a diverse audience of participants in an informal, digital learning environment?

1.1 How can engagement be theoretically grounded, operationalized, and measured by analyzing interaction data in an open-ended, digital STEM learning environment?

a) What log data features effectively reveal affective, cognitive, and behavioral engagement?

b) Which combinations of features effectively capture engagement in this context across time?

1.2 In what ways can data about individual interest be triangulated to understand relationships between pre-existing interest, knowledge, and engagement?

a) How can content-related v. context-related interest be effectively distinguished through logs?

b) How can states of interest, as well as changes in interest, be captured through triangulation of data from log files, self-reports, and knowledge assessments?

c) What variables comprise the best predictive models of knowledge assessment scores, and interest in STEM, Astronomy, and *Minecraft*?

In this research, I seek to demonstrate that *Minecraft* can reveal behaviors and strategies of highly engaged learners compared to relatively disengaged learners, and that these behavioral patterns can be tied to knowledge and interest. For example, a highly engaged learner may exhibit a high frequency of using in-game tools to measure temperature or radiation on any given hypothetical world, a high frequency of typing findings out for other players to see, and ensuring typed findings are of high quality. Creating composites of engagement can then inform when levels of interest may be dipping or peaking, and contextual aspects of the learning environment can provide additional insight. Ultimately, my goal is to minimize reliance on self-reports to measure interest because trace data from digital games, or at least *Minecraft*, can be engineered to provide accurate measures of shifting interest. Additionally, working towards computational models of interest development will provide bounds around the reaches and limits of the theory.

CHAPTER 2: LITERATURE REVIEW

2.1 Why Interest Matters

When individuals act upon their interests, they do so in a manner characterized by effort, focused attention, and a desire to expand their knowledge. The characteristics of interest make it a variable that is beneficial for learning (Renninger & Su, 2012), as interest supports persistence and conscientiousness (Renninger & Hidi, 2020a; Thoman et al., 2011). Through developing interest, individuals can deepen learning and will voluntarily re-engage content or activities they find interesting. Because interest can develop over time (Renninger & Hidi, 2022), interest interventions can be used to great effect in learning environments to attract learners' attention and make content personally relevant (Harackiewicz et al., 2016; Renninger & Hidi, 2021). Interventions might take the form of utility-value exercises that ask learners to link content to their personal life (Canning et al., 2018a; Durik et al., 2015; Hulleman et al., 2017), utilizing novel activities or techniques in instruction (Glowinski & Bayrhuber, 2011; Hunsu et al., 2017; Laine et al., 2017), or allowing learners autonomy and freedom to explore environments (Flum & Kaplan, 2006; Tan et al., 2021). As both a state and motivational variable, interest matters when asking learners to actively participate in an activity or lesson. Through careful consideration of learner interests, instructors and researchers can encourage persistence and effort through promoting interest-enhancing activities, or in other words, "bootstrapping" interest. Inherently interesting topics or activities can serve as springboards to engagement. The notion of developing interest rests on the theory that interest can change ("state-based interest") and is not a fixed trait of an individual.

2.2 State-based Interest v. Trait-based Interest

Fixed, or “trait-based”, interests are seen as “trait-like preferences for activities, contexts in which activities occur, or outcomes associated with preferred activities that motivate goal-oriented behaviors” (Rounds & Su, 2014, p. 98), primarily measured through self-reports (i.e. *RIASEC*, *Strong, O-Net)* and used to determine the academic major or occupation most suitable for a given individual. Trait-based interests predict a fit between an individual’s interest and an environment, career or academic (Hanna & Rounds, 2020; Rounds & Su, 2014). A review done by Hanna & Rounds (2020) found that there is a 51% hit rate between career interest inventories and career choices, such that trait-based interest is a reliable predictor of the domain of work an individual will select into. Not only are trait-based interests linked to accurate predictions of career membership, they also can influence academic performance among college students, including GPA and attendance measures (Nye et al., 2021). Further, trait-based interest has been found to be a strong predictor for setting and pursuit of life goals, such as social, economic, aesthetic, political, and religious goals (Stoll et al., 2020). Trait-based interest does not favor a focus on specific domains of interest (Hidi & Renninger, 2006) or pose a relationship between a person and object necessarily (Krapp, 2005), rather a person has a stable characteristic of how they approach activities or search for information.

Inventories have been developed to categorize individuals based on their interest profile, stating that this profile influences choices. The RIASEC interest inventory (Holland, 1997) positions learner interest as a series of rankings among six distinct interest categories: *realistic, investigative, artistic, social, enterprising, or conventional*. A realistic person likes to work with their hands, an investigative person likes to solve problems, an artistic individual shows preference for creative expression in activities, a social person prefers environments that allow interactions with others, an enterprising individual prefers management or sales and the ability to persuade, and an individual with conventional interest likes well-structured environments that have clear rules. These categories can apply to individuals and to environments, therefore if an individual is able to align their interest with an environment possessing similar characteristics (*person-environment fit*) then productive performance and career success should follow (Nye et al., 2012; Su, 2020), even if measured as early as adolescence (Hoff et al., 2021). Research on state-based interest reinforces this point; the more interested an individual is in their studies or hobbies the more likely they are to persist and excel (Quinlan & Renninger, 2022), and evidence points to the possibilities of a person's interest being activated by either activity type (i.e. *realistic*) or content (*i.e. astronomy*; Blankenburg et al., 2016). However, the frameworks diverge when considering the trait-based account of interest stability and resistance to development (Renninger & Hidi, 2020b).

Where trait-based theories envision interest as a predisposition as how one makes decisions based on contexts, state-based interest is predominantly treated as both an affective state and motivational variable (Ainley, 2006; Hidi & Renninger, 2006; Izard, 1977). Interest has been researched as a knowledge emotion, which differs from other related emotions, such as happiness (Silvia, 2008). Interest as an emotion motivates a person to explore, whereas happiness motivates a person to revisit pleasurable experiences. The affective state of an emotion leads a person to appraise the novelty-complexity of a situation, as well as the ability to comprehend or cope with integration of new information (Silvia, 2005). If a person is presented with information about how the moon impacts life on Earth, they will appraise if the information is novel and almost simultaneously how complex the information is. They will then determine if the complexity exceeds their current capability to understand it. If the material is deemed learnable then the person will approach the information with heightened attention and concentration, characteristics of an individual who is experiencing an affective state of interest (Renninger & Hidi, 2016). Therefore, there must always be an object to activate an individual's state-based interest. The focus of this paper is on the role of state-based interests for learning and motivation, and the term "interest" from hereafter refers to "state-based" interests.

2.3 Defining State-based Interest

State-based interest operates under the belief that an individual's interest exists between the person and an object or domain, and both personal and environmental factors can lead to an interest either waning or gaining in strength over time (Hidi & Renninger, 2006; Krapp, 2005; Krapp & Prenzel, 2011; Schiefele, 2009). As Renninger & Hidi (2021, p. 4) state,

> From preschool through adulthood there may well be "off-ramps" as well as "on-ramps" in the development of interest such as science, and other people (e.g. educators, including parents) can play a supporting role.

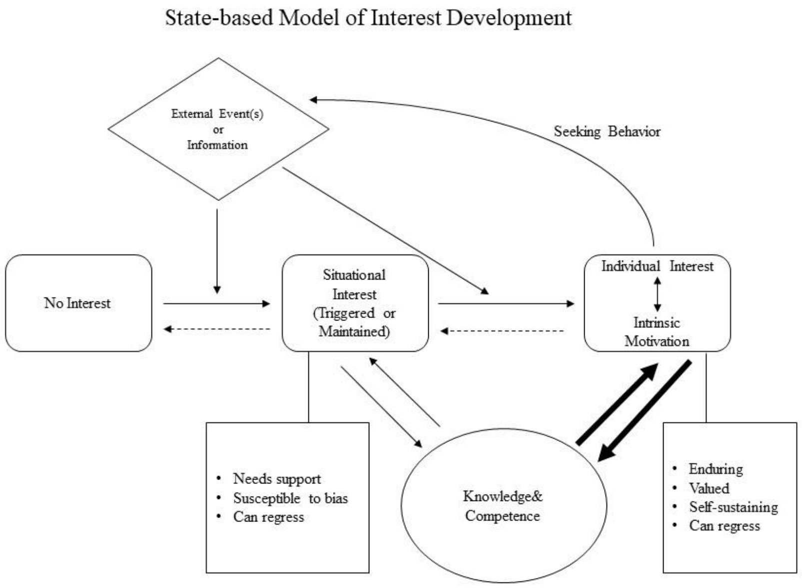
The malleability of interest in the state-based model poses opportunity for an individual to pursue any domain they wish, however this comes with the caveat that factors outside their control may influence their ability to reach certain interest-related goals, such as a targeted career (Hecht et al., 2019; Kier & Blanchard, 2021; Niese et al., 2019). Individuals can still feel threatened by stereotypes and or negative emotions that put them off a path they had initially thought they could pursue (Wang et al., 2018). On the other hand, many triggers exist for promoting the development of interest, and these are often aspects of the environment that can be imbued in activities and objects (Schank, 1979; Schiefele, 1991), characteristics of teachers or mentors (Linnenbrink-Garcia et al., 2013; Fauth, 2018), value a person finds in the activity (Canning et al., 2018b; Hulleman et al., 2017), or the amount of novelty one might expect to experience (Renninger et al., 2019; Renninger & Bachrach, 2015). These attention-grabbing aspects, or triggers, of constructed activities or inherent components of the environment typically associate with earlier phases of interest development, termed situational interest, but they can also be experienced by individuals in later phases of interest.

According to Hidi and Renninger's (2006) 4-phase model of interest development, situational interest represents new or rekindled phases of interest where environmental factors play a major role in ensuring the continuation of the interest. For example, the first time a child sees a group of older children playing basketball they may initially be interested in trying to shoot a basketball. The child's continued interest relies on opportunities to interact with basketball, whether through playing or watching (Renninger & Hidi, 2016). Experiencing of interest through environmental factors, such as playing basketball or reading interesting text (Krapp, 2002; Schiefele, 1991) elicits positive affect, which can push the individual to persist longer on a task that requires high effort (Thoman et al., 2011). Triggering situational interest can be a powerful motivator to further explore a domain or activity (Flum & Kaplan, 2006), thereby increasing knowledge (Hidi & Renninger, 2020; Rotgans & Schmidt, 2017) and positive feelings toward the domain (Silvia, 2008).

Interest has been shown to be important for promoting self-regulated learning in individuals (Schiefele, 2009), as interest closely relates to the three motivational aspects of Self-Determination Theory: autonomy, competence, and relatedness (Ryan et al., 2021). Self-Determination Theory states that autonomy, competence, and relatedness are psychological needs that must be met for an individual to feel motivated and develop (Ryan & Deci, 2000). Additionally, Renninger, Bachrach, & Hidi (2019) identified several more triggers eliciting positive emotions and continued engagement: ownership, hands-on experiences, technology, challenge, personal relevance, affect, and instructional conversation. Continued opportunities for engagement, alongside powerful triggers for interest development can direct an individual's development of interest for an object, domain, or activity towards a more self-sustained, intrinsically motivating individual interest (see Figure 1). The darkened arrows depict the growing importance of knowledge and competence for motivating the deepening of individual interest.

Figure 1

State-based model of how interest develops and characteristics of phases (created by me)



The 4-phase model of interest development positions individual interest as an enduring motivation to voluntarily seek out opportunities to engage an object, domain, or activity, as well as seek feedback and input from experts regarding directions forward (Hidi & Renninger, 2006). Individual interests are seen as intrinsically motivating (Hidi, 2016; Murayama, 2022), where the brain's natural inclination towards gaining rewards boosts the interested individual to continue pursuing knowledge to activate the reward circuitry (Hidi, 2006, 2016; Hidi & Renninger, 2019).

Creating a bank of stored knowledge is essential for interest development and constitutes a sort of ‘homeostasis’ point learners enter an experience with (Geerling et al., 2020). Interest, along with other affective variables, will fluctuate around this point and may raise or lower the point, depending on interactions between value, knowledge, and aspects of the environment. The development of interest relies on multiple interacting dynamic forces, and adequately capturing such complexity proves challenging.

Interest theory posits a complex system of intrapersonal and interactions that motivate behavior (Hilpert & Marchand, 2018). In interest research, literature on measurement consistently points toward the need to measure multiple points of interest to understand how the phenomenon takes place at an individual level (Krapp & Prenzel, 2011; Renninger & Hidi, 2011). A statement of liking “math” does little to inform a researcher about how interested an adolescent is in the domain of mathematics. Does the adolescent stay engaged during math classes, do they gleefully complete math homework, do they participate in voluntary activities related to math, do they ask questions to further their knowledge in math? These are all important questions to ask and measure to understand what comprises an interest profile of an individual. How interested a person is in a context or domain will influence the way in which they engage content, and in turn their engagement will reciprocate and push interest in a certain direction (Ainley, 2002; Arnone et al., 2011; Harackiewicz et al., 2016; Renninger & Su, 2012). My research contributes to this by showing that how a learner regulates their behavior in a digital game environment does matter for how much their interest increases and how they perform on knowledge assessments. I was able capture important engagement dimensions and tie them to the content and context interest of the individual, as well as the construction and display of acquired knowledge.

2.4 Prior Knowledge, Competence, and Knowledge Acquisition

Stored knowledge is seen as an important factor in individual interest, where the knowledge one has accumulated towards a domain or activity leads the individual to engage in more directed information-seeking behavior, as opposed to random behavior (Renninger & Su, 2012). Knowledge can also be considered the driver of interest, as gaining knowledge may precede one's interest towards a domain (Rotgans & Schmidt, 2017). While disagreements exist between directionality, researcher agree knowledge is an integral part of how situational interest develops into individual interest and lead to establishment of one's metacognitive awareness of how to further promote the growth of one's interest (Renninger & Su, 2012). Others argue that stored knowledge is not a necessary part of individual interest, but rather the associations between the interest object and individual feelings and valuing of the object matters more (Schiefele, 2009). Naturally, environmental and situational factors will impact one's interest experiences, and the inclusion or absence of triggers will determine whether or not one finds an object interesting and is able to find value in it.

Interest strengthens as one develops knowledge (Fastrich & Murayama, 2020), and the accumulation of further knowledge supporting one's interest serves as a self-boosting motivational effect, and thus the individual continues searching for additional information to further knowledge related to their interest (Murayama, 2022). People with high interest in a domain are likely to have more knowledge about the domain (Tobias, 1994), however others have argued that stored knowledge is not necessarily a facet of only well-developed individual interest (Schiefele, 2009). It is possible that a person could develop a strong feeling of interest for a domain without being an expert in it. Regardless, knowledge is still a part of the interest development process because knowledge provides the individual with the means to formulate questions and determine what knowledge to gather and where to look (Hidi & Renninger, 2020; Krapp & Prenzel, 2011; Renninger & Hidi, 2021). By presenting learners with challenging information that prompts them to explore category boundaries and reduce uncertainty they can extend their understanding towards their interest and subsequently increase competence (Ainley, 2002). Increasing competence can provide valuable information for learners to metacognitively evaluate their own interest phase and abilities (Renninger & Su, 2012), which aligns with Silvia's (2006) appraisal process of identifying novelty and ability to cope with new information. Examining the relationship between knowledge and interest reveals three important components: intrinsic reward of knowledge accumulation, competence, and exploration.

Individuals who have established interests will want to find additional information to continue developing their interest, and in this regard, knowledge serves as an intrinsically motivating force for information pursuit (Murayama, 2022). The acquisition of information related to one's interests results in a pleasurable feeling associated with enjoyment that a person desires to experience repeatedly (Hidi, 2016), and does not depend on material gains of any type (Ten et al., 2021). Ainley & Ainley (2011) analyzed data from PISA assessment for four countries and found strong predictive relationships between learner enjoyment of science and interest in learning more about science topics. Students with higher interest get more enjoyment out of extending knowledge; the more information one gathers the more interest develops, to a point (Fastrich & Murayama, 2020). High school students participating in an out-of-school lab program at a university were surveyed about interest in science and found that prior interest showed highest correlation with interest in the lab experiments (Glowinski & Bayrhuber, 2011). In general, individuals who have more knowledge, exposure to, and ways to access a topic will be more prone to interest developing through new knowledge and experiences. Comparing

STEM curricula delivered in three distinct formats (design-based, inquiry, and traditional) to 8th grade students, Peterson (2020) found that the more connections made between what learners were studying and possible career paths led to increases in interest for those topics (*i.e.* studying Chemistry and discussing possible careers in Chemistry). Working with middle school students, Renninger et al., (2014) found that the more students reflected on acquired knowledge through procedural and conceptual prompts the more their interest in STEM increased. My research has also shown that students with high knowledge persist at a high level in their engagement, aiding in their ability to grow their knowledge.

Knowledge becomes more relevant to the individual through increasing understanding, and they develop stronger representations of the domain and activities associated with it. Having that domain knowledge and engaging in activities then becomes rewarding and enjoyable. Despite Carmen et al., (2017) seeing no change in interest in climate science through a short classroom intervention for 7th grade students, those who had existing interest in science broadly showed greater interest in the topic. Those with existing interest have a store of knowledge they can contribute to and develop further expertise in. If a consistency of individual interest and knowledge can be established, then situational factors should also be consistent for individuals in how they acquire new knowledge. Using a problem-based learning module on energy supply for early high school students, Knogler et al., (2015) found that situational interest in one section of the module only predicted the experience of situational interest in another module for learners with higher individual interest in the topic. Individuals with higher interest and more knowledge are amenable to continued engagement and situational factors that promote further engagement and knowledge-building. Prior knowledge and interest constitute an essential piece of how an individual will experience a situation aimed at influencing interest and should always be taken into account (Tobias, 1994). Knowledge levels and interest phases are not always consciously available to the learner (Renninger & Bachrach, 2015; Renninger & Su, 2012), yet learners will make competency judgments when assessing their interest in and willingness to engage a domain.

When interest researchers look at student interest in a domain they often think about how interest in the domain can be leveraged to help the student pursue their interest long-term, possibly into college and then a career (Harackiewicz et al., 2008; Hulleman et al., 2017; Potvin & Hasni, 2014; Renninger & Hidi, 2022; Saw & Agger, 2021), or how interests will predict selection into and performance in college or a career (Hanna & Rounds, 2020; Hoff et al., 2022; Nye et al., 2012; Rounds & Su, 2014). A learner will likely not continue to develop an interest into a long-term pursuit if they do not feel competent in the interest domain. Surveys and interviews with undergraduates showed students with more interest in their majors wanted to pursue their major domain into a career, and a big part of career planning was based on perceived competence in the domain (Quinlan & Renninger, 2022). Competence, stemming from knowledge and comfort in the domain, is a critical factor in determining future goals and intensity of interest experiences (Pugh et al., 2010). Students in varying phases of interest will need different supports, and providing the proper supports through personally relevant environments can offset a lack of prior interest (Glowinski & Bayrhuber, 2011; Habig & Gupta, 2021). In interviews with first-year college students in Physics or Biology, students spent most time describing interest in relation to competence and enjoyment of their major (Ruff, 2016). Students were more likely to see themselves in a future career related to their interest if they felt competent. My research draws upon this by investigating if a familiar context, *Minecraft*, instills feelings of competence in content because players feel competent in the environment. While most learners increased interest, no clear link between context (*Minecraft*) and interest development could be established.

When students do not feel competent in a domain, because of feelings induced by others or uncontrollable events, they will disengage and start to distance themselves from the domain (Hecht et al., 2019; Renninger, 2009; Wang et al., 2018). Observing interest practices of individuals over long periods, Azevedo (2018) noticed engagement tapered with an interest when the individual did not have time or financial resources to support it. This could lead to missed opportunities and decreasing competence while one is not developing further knowledge in the domain and others are doing so. Additionally, racialized experiences influence competence by leading learners of diverse backgrounds to question their abilities, regardless of how much knowledge they have in the domain (Kier & Blanchard, 2021; Master et al., 2021; Niese et al., 2019). Learners develop awareness of their knowledge in the interest domain over time, as well as knowledge of how they fit in the domain, which can have huge implications for determining their trajectory in the interest domain going forward. Interest develops out of previous experiences and can lead to learners creating expectations in the domain (Ainley & Ainley, 2011), therefore the valence of experiences, positive and negative, will influence one’s competence and desire for future knowledge pursuits in relation to interest. Knowledge pursuits can be understood as individual exploration patterns in information-rich environments, and these patterns will differ by phase of interest, interest in available topics, and value-orientations.

Exploratory behavior and interest share a close bond. Interest develops through individuals exploring environments for new information, and the information an individual possesses will determine how they explore the environment; interest is the motivator for learning and exploration (Silvia, 2008). Exploration is a deliberate action centered around gathering information in relation to the self (Flum & Kaplan, 2006; Renninger & Hidi, 2021). How we interpret a theory of interest should extrapolate how values and interests for specific content inform exploratory behavior (Krapp, 1994). However, not all exploratory behavior can be explained by interest, as there are cases where individuals explore and gather information for assignments (Baram-Tsabari, 2015), try to resolve an information-gap stemming from curiosity (Donnellan et al., 2022; S. E. Hidi & Renninger, 2020; Loewenstein, 2004; Shin & Kim, 2019), or seek stimulating experiences or information to alleviate boredom (Geana & Daw, 2016; Tanaka & Murayama, 2014). As long as an individual is interested in a domain, and their interest has been activated they will be engaged (Hidi & Renninger, 2006), and engagement can be understood in terms of how they are exploring or gathering information in regard to their interest (Baram-Tsabari, 2015). In this regard, I was able to draw connections between information-seeking behavior, knowledge, and interest, where effortful information gathering contributed to increases in interest and high performance on knowledge assessments.

Having opportunities to exercise one’s interest through self-directed exploration is important for developing interest further (Barron, 2006). Given a full choice condition on what information categories to select about a country, adults showed increased interest compared with a group that could not select categories (Fastrich & Murayama, 2020). Exploration, relating to autonomy and choice, has a marked effect on what individuals explore and how much interest or enjoyment they obtain as a result. In terms of physical health, exploration intention or what drives a learner to explore, was positively correlated with cadence, heart rate, and oxygen consumption, while undergrads played a bicycle exergame (Pasco & Roure, 2022). Having freedom to explore and concentrate on what they found interesting, students playing the exergame exerted more energy on the bike. It would not be too much of a stretch to consider how a seasoned runner might explore different shoes, cadences, routes, and distances because they have a strong interest in running. Through exploration this person would gain more knowledge about themselves, and the surrounding environment that will inform how they explore in the future. Individuals can be motivated to explore within their interests both by content and activity type (Blankenburg et al., 2016), which fuses trait-based and state-based interest to show that both types of interest do collaborate to determine how an individual will explore.

Interest relates to learning through the exploration of content an individual finds both relevant to the self and capable of being organized into existing structures of knowledge. Exploratory action may differ between interest, curiosity, and other motivators, however in any form exploration serves a strategic purpose (Flum & Kaplan, 2006). Exploration is how curiosity is resolved and interest develops because the acquisition of knowledge is the outcome and what motivates the behavior. Many educational studies regarding interest measure how likely a person is to reengage content over time, and this is used as a behavioral indicator of interest development. Additionally, the degree of "liking" is used to assess whether interest has developed. Only a few of the studies referenced in this review have examined the fine-grained exploratory behavior of learners and how it relates to their interests. On the other hand, cognitive psychology has spent a great deal of effort understanding how individuals make decisions, given multiple choices, to explore and how these choices impact learning and retention. Taking an approach to combine cognitive psychology with motivation research, I have been able to show how several forms of exploratory behavior in a learning environment can be measured and applied to static measures of interest and knowledge. This result provides insight into how what we know about exploratory behaviors and performance at a broader scale remains consistent in digitally rendered, content-rich worlds.

2.5 Measurement of Interest

Self-reports

The use of surveys is very common in measuring interest, however Renninger & Hidi (2011) have discussed at length the limitations of only relying on survey data to measure a person's interest in a domain. They advocate using multiple methods of measurement, or in other words triangulation, to see if interest can be linked across multiple measures, such as surveys, field notes, log data, interviews, or case studies. One reason for this is individuals in early phases of interest (*situational*) may not always be aware of their interests, which can lead to self-reports done through surveys being inaccurate (Renninger & Hidi, 2016). Surveys frequently ask "liking" questions, such as "How much do you like doing Math problems?", and these are used as approximations of interest. However, liking and enjoyment are only one component of interest, and it is important to capture other behavioral and cognitive components (Krapp & Prenzel, 2011). Interest does not stop at 'liking', and through my research I hope to show the learning strategies adolescents use then they are interested in subject and the degree of effort they put forth, connecting these dimensions to better understand how interest influence behavior. To establish this, though, self-reports will still be used to capture an affective dimension of interest.

One frequently used survey instrument is the *situational interest scale* (*SIS*, Linnenbrink-Garcia et al., 2010), with three factors predicting situational interest: triggered situational interest, affective maintained situational interest, and value-related maintained situational interest. One important distinction of this scale is that it separates out how students may react to the way content is delivered (*affective maintained*) versus reactions to the content itself (*value-related maintained*). The scale is also shown to generalize across domains and age groups. Another instrument, the *individual interest questionnaire* (Rotgans, 2015), measures individual interest across diverse disciplines and has shown predictive validity for affect towards a domain and on-task behaviors. Most survey tools have been adapted to match the research and environment of each study.

Several studies aimed to include constructs that related to interest as an improved estimation of what aspects matter most for energizing motivation and performance, and how those constructs interact with interest. To see if personalization of material and learner choice differently motivated students differently, Hogheim & Reber (2015) used and adapted the *SIS* by adding items related to perceived value (*i.e. "What I learned is valuable to me"*) and task effort (*i.e. "I tried hard when working on the problem"*). If students find personalized examples meaningful then this should correspond with their liking and enjoyment of the material, and likewise if they value the examples and then they should put forth more effort (Harackiewicz et al., 2016; Renninger & Su, 2012). Another study by O'Keefe & Linnenbrink-Garcia (2014) adapted the *SIS* (Linnenbrink-Garcia et al., 2010) to gauge affect-related interest towards the specific task (*i.e. "I really enjoy working on problems like this"*) and questions from other measures to capture achievement goals (*i.e. "I wanted to avoid performing poorly like others"*) more broadly. The study showed that if learners have high interest towards their goals, and the goals are perceived as personally important, then the tasks of interest are not considered effortful. These studies measured interest for the task at hand, rooting the measurements in value and what might be considered individual interest.

Measuring individual interest, Knogler et al., (2015) adapted another individual interest scale and based items on Krapp's (2002) individual interest conceptualization (*i.e.* "*I like studying the topic of energy supply"*). In this study, situational interest was also assessed using an SI scale validated in classroom contexts (Lewalter & Geyer, 2009) and adapted to the study (*i.e.*

“When you think of the previous module’s sessions, to what extent did the sessions capture your attention?”). While not entirely a different construct, the authors sought insight on the relationship between situational and individual interest in problem-based learning and energy systems. A history and continued tendency for adapting scales to match study aims can be seen, and multiple constructs are often brought together to examine divergence of interest from those constructs (Silvia, 2005). These studies focused primarily on using self-reports, and the self-reports relied on measuring value, effort, and affect to build a representation of interest, which does not fully capture the interest profile of an individual and how they behave towards realizing their interests. A handful of other studies have elicited several forms of measurement to obtain a bigger picture of how interest is experienced and displayed by the individual through behavioral and cognitive measures.

Combining Methods

Variables, such as knowledge, task persistence, or choice, are better measured by assessments or log data that do not force the participant to make judgments regarding their feelings towards domains and activities. Attempting to understand the relationship between knowledge and interest, Rotgans & Schmidt (2017) used the *IIQ* to measure students’ individual interest in science and combined this with a *Concept Recall Test*, where students had to recall as many concepts and ideas as possible about a designated topic. Through combining self-reports with knowledge assessments correlations between depth of interest in the subject and knowledge can be discerned, whereas self-report can only provide how learners perceive their own level of knowledge. Other studies have combined interest scales with student GPA and persistence in college major (Canning et al., 2018b; Harackiewicz et al., 2008; Hulleman et al., 2017), or persistence on a particular task (O’Keefe & Linnenbrink-Garcia, 2014; Thoman et al., 2011, 2020). These studies predict that students who either enter a scenario with higher interest, or experience events leading to heightened interest, perform better and exhaust less effort in a task. These studies still place a heavy burden on surveys capturing the individual's interest, however the added measures provide reinforcement of interest as a powerful antecedent for engagement and performance. Other research approaches have used multiple measures of interest itself through surveys, interviews, field notes, and log data.

By combining measures of interest from multiple instruments, a strong representation of behavior, cognitive, and affective dimensions of interest can be evinced (Krapp & Prenzel, 2011). A study looking at interest in science and research among high school students enrolled in an informal museum-based science program used repeated surveys at three time points (pre-, mid-, and post-intervention) that included open-ended questions, as well as post-intervention interviews to better understand if and how interest development occurred (Habig & Gupta, 2021). Surveys asked questions about interest in scientific research, self-perceived competence in science, and scientific engagement outside the intervention. Interviewees were grouped based on determined phase of interest (Hidi & Renninger, 2006) from surveys and asked questions about engagement outside the intervention. Surveys and interviews were combined to determine phase of interest and assess change in interest over time. This study combined self-reports with interviews, but some studies forgo the use of self-reports altogether in favor of observational data. Observational data removes the confounding of individuals either misremembering their interest experiences, misunderstanding questions, or inauthentically answering questions.

Observational data provides the advantage of capturing behaviors linked to interest in authentic learning environments. Relying on field observations and students' daily writings in a science notebook, Renninger et al., (2014), recorded middle school age students' reactions to events in an informal, inquiry-based biology summer program. Students were also interviewed at three time points (pre-, post-, and 5 weeks post-intervention). Observations were coded to reflect excitement and engagement with material, notebooks were analyzed for value and knowledge based on conceptual prompts, and interviews assessed understanding of science and how it can be personally useful. This approach avoids any unreliability associated with self-reports; however, the approach relies on the interpretation of interest from the researcher's perspective. Another observation-based study from Azevedo (2017) followed high school students and amateur hobbyists, collecting video data and field observations to create ethnographic cases studies of how individuals experience interest. Longitudinal collection of qualitative data allows for researchers to see how individuals behave and reengage their interest across experiences and over time, without having to rely on self-reports from surveys or interviews. Again, what constitutes interest behavior is determined by the researcher and should be grounded by theory. In a related vein, research examining interest using and in relation to technology offers a lot of opportunity for pooling behavioral, cognitive, and affective data across time to provide a unique perspective on interest and its development.

2.6 Measuring Interest with Technology

Technology as an object of interest

Technology has been identified as a trigger of situational interest in and of itself (Renninger et al., 2019), and it stands also as a means for collecting data and formulated novel measurements of engagement that can contribute towards understanding of interest. Technology, such as video games, has been used to measure learning and skill development for decades (Clark et al., 2016; Gopher et al., 1989; Mayer & Johnson, 2010; O'Neil et al., 2005). Research has also been conducted on the relationship between technologies, such as digital games, and motivation to learn (Mayer, 2018; Wouters et al., 2013; Wouters & van Oostendorp, 2017; Zheng & Gardner, 2017), but not a lot of research has looked into measuring interest with or through technology, as in how does the way the learner use the technology reflect existing interest.

Basic studies have utilized technology as a trigger for enhancing an interest experience, where a technology is used as part of an intervention, but no behavioral data is collected from technology use. For example, using VR as a pre-training environment for an inquiry-based climate change workshop in a lab, Peterson (2020) found the VR group showed significantly higher knowledge transfer than a non-VR group, yet both groups experienced significant increases in climate change theory interest. A study placing 7th grade students in a classroom utilizing a mobile learning app versus a traditional classroom found students in the digital group's overall interest in science and math remained at an average level, while interest technology and collaboration remained high (Laine et al., 2017). Over the course of the second semester the digital group's interest remained stable but the control group's interest dropped to below-average levels. Through interviews, students in the digital group expressed appreciation for the hands-on aspect of the technology, ease of use, autonomy, and collaborative features, and they consider those features as being interest-generating. These studies highlight how technology may be used for promoting interest development among adolescents, however how digital environments might be adapted to individuals' interests remains unclear. My research adapted a familiar environment that is already interesting to millions of adolescents across the country, and leveraging this did prove to bolster interest in Astronomy for most of the individuals. In a way, this approach is akin to personalization of STEM content, since learners could navigate the content through an environment that is personally meaningful to them.

Personalization

Personalization stands out as one of the greatest opportunities for targeting interest development in digital learning environments. Looking at log data in relation to interest, Cordova and Lepper (1996), conducted a study using personalization of textual components and individual choice for non-integral learning content in a math-based video game. Through self-reports they measured how interested elementary students were in the game and how much they enjoyed it. These measures were analyzed by condition (choice v. no choice; personalization v. no personalization) and engagement in the digital environment, such as taking on more challenges, use of more complex operations, and employment of strategies. Results showed children in the personalization condition exerted more effort, performed better on assessment, and found math more interesting than the control group. A non-game study using math problems on a computer, Hogheim & Reber (2015) measured middle school students' triggered and maintained situational interest throughout a learning module. Situational interest was enhanced by personalized problems for those with low individual interest and perceived competence in math, but this did not hold for students with high individual interest. Other research has found similar results, where additional measures to make content fun or personalized results in negative feelings toward the intervention for individuals with high interest in the content (Wang & Adesope, 2016). Students with high individual interests may feel distracted or uninterested in any extraneous details that obscure meaningful content.

Looking at how quality of personalization might matter, Bernacki & Walkington (2018) used an intelligent algebra tutor to personalize word problems for high school students. Students' out-of-school interests were assessed by self-reports prior to using the intelligent tutor, and targeted module questions were altered to reflect interests. Results showed the personalized exercises did contribute to increases in situational interest, as well as individual interest, contrary to results from Hogheim & Reber (2015). One possibility is Bernacki & Walkington (2018) used more in-depth personalization by tailoring problems instead of merely inserting names and cities into math problems. It is also possible technological advances allow more meaningful tailoring of content to individuals' interests and desires. If personalization is done in earnest and meaningfully connects to personal interests then interest can be enhanced for those with higher individual interest in a topic. The personalization studies described here used pre-surveys to determine how to personalize content, but continued research using trace data could personalize based on interactions in the digital learning environment.

Digital Games

As with other forms of technology described, digital games offer a great deal of autonomy for the player, progressively revealed information and abilities reflecting novelty, and increased feelings of competence and expertise with time (Boot et al., 2017; Granic et al., 2014; Green & Bavelier, 2012). A bike-based exercise game was able to link situational interest, in the form of challenge and enjoyment, with increased physical activity in undergraduates, compared to a control group (Pasco & Roure, 2022). Interest plays a role in exertion of attention and energy, in both physical and cognitive activities. *Crystal Island*, a digital game focused on collecting evidence to determine the biological origins of a disease sickening residents of an island, when used as part of an 8th grade biology curriculum produced a significant correlation between situational interest, in-game goals completed, and microbiology post-test scores (Rowe et al., 2011). A study embedding digital games into a formal curriculum, Rodriguez-Aflecht et al., (2018) had 5th grade students play a math game for two 30-minute sessions per week for 6 weeks and found maintained situational interest for only half the participants, with higher individual interest at the beginning predicting higher maintained situational interest. When situational interest was not maintained, students showed strong decreases in later individual math interest and worse performance on a post-assessment than those with higher initial interest. The game started as novel, triggering interest, however as appeal wore off the game more closely resembled homework. Novel effects pose a concern when utilizing newer technologies, however there are means to address these effects.

Interest develops over time through consistently having access to content and experiences that continue to increase knowledge in the target domain. When using digital environments, and especially games, it is important to have a variety of ways to interact with the game so that there are ample challenges or available interactions (Steinkuehler & Squire, 2024). Digital games, as with other forms of technology, offer expansive opportunities for exploration and information-seeking in exciting, low-stakes environments, and this data can be used to further create environments that are relevant and interesting to individuals. I focused heavily on how decisions to use available resources for learning in *Minecraft,* in a meaningful, productive way, contributed to growth in interest and knowledge. Exploration of resources available and how to consistently implement them in a self-regulated way is critical in open-ended, exploration-based learning environments. Underlying this research is the assumption interest does indeed play a role in how individuals interact with digital learning environments and that interest can be altered for the purpose of promoting engagement and persistence.

Multimodal Approaches

Classroom studies have started looking at the unique affordances of process data collected from hypermedia (Azevedo & Taub, 2020), and how interest can be measured in congruence with physiological measures of arousal (*e.g. skin conductance*) (Tan, Gillies, Jamaludin, 2021). Multimodal analytics has the potential to more fully understand interest and the processes under which it develops. In a study using the educational game *Crystal Island,* Emerson et al., (2020) used combinations of behavioral gameplay logs, facial recognition data, and eye gaze patterns to classify low, medium, and high interest groups after play. They used several dimensions of gameplay features (*e.g.* total time spent playing, movement to new locations, conversation with characters, reading a book, completing an in-game achievement, among others) along with an interest enjoyment subscale to encode actions into a performance metric. Results showed that a multimodal model incorporating the gameplay metrics along with eye gaze produced the best predictive model for interest, better than a singular modality and inclusion of all modalities. This study adds weight to the notion that interest can be derived from patterns of learner engagement in video games, however this study remained agnostic about theoretically grounding engagement measures, which is a needed component of understanding engagement in a digital learning environment (Martin & Borup, 2022). While lacking in as many measurement instruments, I did theoretically ground engagement and tie it to interest through various log data features. Whether through interaction with technology or measurement by technology, opportunities exist to measure interest and uncover relationships between variables, such as physical activity or exploration, that may be valuable to developing impactful environments for interest growth. Through incorporating other affective variables, interactions between interest, demographics, and triggers can be better understood.

Interest is considered a unique, affective state, and finding how it relates to and diverges from other affective states can be useful for constraining the theory. Using a dynamic systems analysis, Geerling et al., (2022) divided undergraduates in an elective computer science course into a utility value condition and control conditions. Student interactions with the learning management system (LMS) prompted automatic questionnaires asking students to rate their interest and confusion regarding the subject. Prompts were pushed based on events occurring, such as students making changes to code. Results showed interactions between interest, confusion, gender, and UV conditions, in that female students in the UV condition saw increases in interest cooccurring with increased confusion, whereas the control saw decreased interest with increased confusion. This dynamical systems approach to understanding interest sheds light on the transitory nature of interest across course activities and highlights important relationships to other affective variables and course content. Such fine-grained data provides evidence for what activities might spike interest or lead to its diminishment, and how interest might be measured in relation to engagement. Technology use lends itself to collecting longitudinal data on interest experiences without researchers needing to frequently be present to collect observations. Still, attention must be paid to how engagement is being operationalized in these environments and in what ways interest converges and diverges with engagement. Through my research I operationalized engagement through theoretical dimensions and applied the dimensions to features generated from player logs in *Minecraft*; in this way, interactions between affect, cognition, and behavior were used to inform how interest relates to patterns of engagement across the theoretical dimensions. In essence, all dimensions of engagement are important to get the most out of an open-ended *Minecraft* learning environment.

2.7 Defining Engagement in Digital Environments

Interest and engagement are intertwined concepts; it is possible for a learner to be engaged with content but uninterested, yet it is impossible to be interested in content and disengaged (Renninger & Hidi, 2016). Interest and engagement are both multidimensional constructs that include aspects of cognition and affect but are not fully captured by any single dimension (Krapp

& Prenzel, 2011). Interest has been mentioned as one of the affective aspects of engagement that drives behavioral engagement (Martin & Borup, 2022). A learner already interested in exoplanets will exert more effort and focus when this topic is taught in class or comes up in discussion. With interest, self-related information processing is an important facet of how individual interest develops and what makes specific content interesting or not for an individual (Harackiewicz et al., 2016; Renninger & Hidi, 2021). With the relation to the self and the potential to tailor information and make it more personalized digital technologies offer a lot of opportunities to develop interest through engagement (Bernacki et al., 2021; Bernacki & Walkington, 2018; Emerson et al., 2020; Høgheim & Reber, 2015). Interest has been pointed out as a complex intrapersonal system that informs how affect and motivation direct behaviors (Hilpert & Marchand, 2018), and undoubtedly these behaviors constitute diverse forms of engagement through the manifestation of interest. I have discussed how interest is defined, measured, and some means by which it can develop, and thus tying interest to engagement dimensions has provided insight into how interest drives engagement and how engagement in turn reverberates with interest.

Defining and measuring engagement has been a long-debated, important issue in education research. An entire special issue of *Educational Psychologist* was devoted to unpacking research surrounding engagement, and Azevedo (2015) sought unifying definitions of engagement that might constrain engagement research and broaden its reach, emphasizing that triangulation of multiple data sources is necessary to generate inferences from the complex nature of engagement. While the disparate lines of research surrounding engagement pose problems for defining the construct, many reviews and studies rely on core dimensions that have been identified: affective, behavioral, and cognitive (Fredricks, 2011; Fredricks et al., 2004), with some adding additional dimensions, such as agentic (Sinatra et al., 2015), or reducing dimensionality to delineate salient dimensions in specific learning environments, such as blended learning (Halverson & Graham, 2019). With technology mediating such a large part of individual learning currently, let alone all aspects of life, understanding engagement with and through technology can root our understanding of the affordances technology offers for learning and how interests might drive technologically mediated learning. I will explore how engagement has been measured through digital learning environments, and when the dimensions of cognitive, emotional, and behavioral engagement have been adequately defined and measured.

Behavioral Engagement

In a review of student analytics in digital learning environments, Kew & Tasir (2022) point out that most engagement research relies on behavioral dimension data (37% of studies identified), learning or cognitive level (25% of studies), and very few identify emotional engagement (8% of studies). This insight is also shared by Vytasek, Patzak, & Winne (2023) in their review of student analytics and engagement. Behavioral data can be intuitively drawn from logs that capture how a student behaves in a digital learning environment (*i.e.* using a tool, taking a measurement, submitting an assignment), which entails following task objectives and actively involving oneself through effort, attention, and participation (Fredricks et al., 2004). Going further, Sinatra et al., (2015) places information-seeking, persistence, and resiliency among the characteristics of a behaviorally engaged individual. Martin & Borup (2022) describe behavioral engagement as the "physical behaviors and energy that students demonstrate when completing learning activities", and they place importance on both the individual and the affordances of the technology.

Investigating self-directed learning in online environments, Al Mamun et al., (2022) measured persistence among undergrads who voluntary participated in a chemistry module, supplemental to their introductory chemistry course. Persistence was measured through combining task completion, time-on-task, and task completion. Results showed that students who had more experience with simulations performed better on post-assessments administered through one-on-one interviews. Using persistence as a multi-feature behavioral measure provided insight into how the behaviors of individuals contributed to performance, however effective strategy use was attributed to prior knowledge and not the quality of the simulation. If a technology lacks sufficient opportunities to invest effort then the technology itself is not engaging and may not lead to productive engagement. Behavioral data only scratches the surface of the complex interactions learners enact in digital environments, and determining how learners obtain high or low levels of engagement throughout the process is needed (Arizmendi et al., 2022; Kew & Tasir, 2022). What is the process a learner is using to gain requisite knowledge to provide inferential answers to conceptual prompts? How do the technology and content enable this, and how does the learner feel while interacting? Using time-on-task does nothing to inform stakeholders if learners are acquiring knowledge or what they are doing with it once they interact. Extracting strategy use and how information is integrated can be assessed on the cognitive level.

Cognitive Engagement

Empirically, there is ample evidence that the dimensions of engagement do overlap, especially between the cognitive and behavioral dimensions (Halverson & Graham, 2019; Martin & Borup, 2022). Still, distinct definitions of cognitive engagement point towards deliberate behaviors in service of processing information towards a goal, essentially organized pattern behavior (Vytasek et al., 2020). Cognitive engagement positions learners as going beyond the requirements of the task, seeking challenge, and investing in learning through metacognitive monitoring (Fredricks, 2011; Fredricks et al., 2004; Sinatra et al., 2015). Cognitive engagement moves beyond behavioral engagement, in that the learner is not just spending time on performing a task, such as writing a discussion forum post, but is doing so in service of gaining understanding of a topic and integrating it into their stored knowledge. Again, emphasizing the relationship between the individual and the affordances of the technology, Martin & Borup (2022) see cognitive engagement as productive engagement towards reaching a goal through interacting in the learning environment. Behavioral engagement is defined by effort, but this does not necessarily entail productivity, which is what sets the cognitive and behavioral dimensions apart.

Measuring cognitive engagement requires knowing the task and analyzing how the learner is approaching the task. In a study using personalization strategies to promote engagement through learning in a digital game, Cordova & Lepper (1996) saw higher levels of task involvement, such as taking on more challenges, using complex operations, and strategizing among elementary-age learners in the personalization and choice condition. These metrics were constructed through extracting the number of spaces the learner advanced in a turn, use of hints, difficulty level selected, how many times a complex math operation was used, and the percentage of problems solved correctly. In another study, Kang & Liu (2020) contrasted the relationship between at-risk students' strategy use in an open-ended, STEM-based digital game and non-at-risk students' strategy use. Log data collected was cognitive tools used, actions, notes taken, and timestamps. Mean posttest score was significantly different between students at-risk of dropping out of school, as defined by the Texas Education Agency, and those not-at-risk of dropping out. Sequential pattern mining showed non-at-risk students interacted with all in-game tools more often and became better problem solvers after a couple play sessions, while the at-risk students kept repeating the same actions. At-risk students were behaviorally engaging the game, however their strategies for learning were not as effective as non-at-risk students who exhibited deeper cognitive engagement. Goals will differ by the task assigned in the digital environment and should be considered when examining the cognitive engagement dimension of engagement. Additionally, affect will vary for individuals and across digital learning environments (Loderer et al., 2020), and should also be considered when assessing engagement.

Affective Engagement

Interest and curiosity are often identified among the emotions that constitute affective engagement (Martin & Borup, 2022; Sinatra et al., 2015; Vytasek et al., 2020). This reinforces the notion that if a student is interested then they will be engaged in the content or activity. However, interest itself is a complex phenomenon and contains many intertwined facets that work towards making a person interested, and these facets of interest possess commonalities with affective engagement. Other affective components of engagement are identifying with the environment or community involved, feelings of belonginess, and valuing of content (Fredricks et al., 2004). Sinatra et al., (2015) adds the importance of promoting positive affect in learning environments and invokes Eccles & Wigfield’s (2002) expectancy-value theory as being central to a learner’s emotional engagement. Learners will be motivated to engage if they see themselves succeeding in the task, where the cost of participation is outweighed by the benefit. The learner will approach any digital environment with pre-conceived feelings towards the environment (Martin & Borup, 2022), and these feelings can shift throughout the experience if the learner is unable to extract value from the environment (Hulleman et al., 2017). Summarizing affective engagement, Martin & Borup (2022) define it as emotional responses to activities in the learning environment (*e.g.* curiosity, confusion, boredom), as well as the emotional energy expended while completing task goals. Affective states have a profound impact on motivation to engage, as they can be activating or deactivating based on valence (Pekrun & Perry, 2014), and emotions can vary across multiple dimensions of the digital environment (*e.g.* topic, aesthetics, social, epistemic, among others) (Loderer, Pekrun, & Plass, 2019). Measuring attitudes and interest towards a technology (Clark et al., 2016; Laine et al., 2017; Plass et al., 2013; Potvin & Hasni, 2014), as well as experience (Al Mamun et al., 2022) is frequently done, however, capturing the moment-to-moment changes in affective engagement in digital environments has proven more challenging (Vytasek et al., 2020).

Engagement in Log Data

Game-based learning environments tend to offer a lot of autonomy for players to explore the different ways they can interact with the environment, sometimes productively and sometimes not. Performance and understanding can be tracked through player log data in educational games by looking at behavioral data and learning outcomes, or in other words correlating dynamic processes in the context with static measures external to the context. In a study assessing learning of qualitative physics using *Newton's Playground*, an open-ended physics game, Shute and colleagues (2013) collected log files from 167 late middle and early high school students as they played each level of the game, creating a set of features including: time spent on the level, restarts, objects used, whether the problem was solved, and the trajectory of the target object. Using log data, they operationalized engagement as the number of levels attempted and found students who played more levels (the high engagement group) showed significant gains between pre- and post-test, compared to non-significant gains for the low-engagement group.

Additionally, high pre-test scores correlated highly with gold trophies for each level, a measure of content mastery.

Log data from digital environments provides insight into the process of how learning gains emerge, as well as how individual differences in knowledge and motivation impact behavior seen in logs. Using the same game, Andres et al., 2014 explored how the affective state, confusion, related to in-game performance for 48 middle school students in the Philippines. Collapsed vectors of student in-game interactions, including achievement, were correlated with human observations of confusion, and findings showed an inverse relationship between confusion and achievement, and a positive correlation between lower achievement and confusion. Logs offer great versatility for tying learners' affective engagement to other dimensions of engagement, as well as performance-related measures, such as knowledge.

Given the overlap of engagement dimensions, some measures captured through log data have been positioned as proxies for affective engagement, such as patterns of click-stream data (D'Mello, 2017). Still, affective measures tend to require multimodal physiological data collection, such as facial recognition or eye-tracking (Wiedbusch et al., 2023) or rely on self-reports administered during or after the assigned task. Revisiting Geerling et al., 2020, undergraduates in an introductory computer science elective course were assigned to either a utility-value (UV) condition or control. Pop-up questionnaires asking students to rate their confusion and interest, respectively, appeared following pre-determined activities, such as changing code in an exercise problem. The questionnaires were tied to engagement so if a student did not actively participate in the course then they would have fewer or no questionnaires appear. Interestingly, students in the control condition answered more questionnaires than the UV condition, but similar patterns of engagement revealed differences in confusion and interest between groups, such that students in the UV condition showed increases in interest when they were confused, and the control did not. This study effectively tied self-reports of interest and confusion to engagement and showed that interest and confusion fluctuate depending on the level of engagement but tend to return to a more stable level between engagement episodes. A clear indication of triggered interest through logs and external measures does seem to provide solid evidence for the role of interest while learning through digital environments.

Also discussed earlier, Emerson et al., (2020) used an interest self-report and knowledge assessment following 180 minutes of playing *Crystal Island*, a biology-based digital game, and tied learning and affect to gameplay features. Gameplay logs and facial recognition provided the best predictive model for learning, and gameplay and eye gaze provided the best predictive model for predicting interest. Features of the best model included examining books in-game, having conversations with non-playable characters (NPCs), moving more, reading posters in-game, and longer gaze with food, the lab, and NPCS. Students who submitted the diagnosis worksheet more often tended to have lower interest. This study provides strong links for tying interest to gameplay behavior but does not ground how each gameplay feature indicates engagement, which leaves the authors guessing why certain features constituted interest and/or learning. Regardless, literature clearly mandates the need for multimodal, multisource measures of engagement to draw inferences about learning and motivation. My approach uses self-reports, logs, and interviews to draw clearer connections between interest and engagement, but attention is also paid to context and if each form of measurement is contributing to what model best predicts interest. Dynamical approaches to unpacking affective states and engagement can draw clearer pictures of how affect can be reflected differently, depending on motivational and engagement factors, and how these factors fluctuate.

Log data allows researchers to track patterns of behavior over time, treating behavior as a dynamic variable as opposed to a static measurement (Snow, Allen, & McNamara, 2015). Learner attitudes and behaviors may shift over time, and log data can reveal when and where these shifts happen. Creating an approach to capture the relationship between thought and action through logs, D'Mello & colleagues (2017) proposed the "Advanced, Analytic, and Automated" (AAA) approach to measuring engagement through logs. AAA aims to examine fine-grained time specific data for learners in digital environments through non-intrusive measures. AAA also affords a dynamical approach to considering log data, where a researcher might discern interest at quick, successive time intervals and correlate shift in interest with other mental states, such as mind-wandering. These mental states operate through a complex, dynamical relationship where they interact and influence each other (Marchand & Hilpert, 2024). Logs can be used to detect when players are engaged, disengaged, or disengaged with the goals of the task. Players may appear active, but engagement indicators from logs can tell a different story. Data collected by Gobert et al., (2015) analyzed data from 144 middle school students navigating the Inq-ITS digital learning environment. The environment aimed to promote scientific inquiry and analysis found students who exhibited stretches of activity or "clips" where they ran numerous simulations or took very quick actions to be active but disengaged (Gobert et al., 2015). While students may appear active in their activity, they may not be attending to the task(s) asked of them. Being able to track changes in states is critical for knowing when a player is being productive or if they have mastered a skill. Knowing when engagement and interest are waning across many learners can be difficult, and log data makes understanding and possibly intervening at these times feasible.

Allowing students to remain unproductive while interacting with digital environments can leave them unable to grasp the content or can generate misunderstandings, as well as foster feelings discontent with the contextual element. Analyzing logs of 3,546 participants interacting with a digital, circuitry-focused, touch tabletop exhibit, *Oztoc*, Tissenbaum (2020) developed a model of tracking productive and unproductive states of visitors. A "MakeCircuitCreate" event was coded for each participant, determining complexity, completeness, and uniqueness of the circuit. A Hidden Markov Model (HMM) effectively captured transitions and provided insight into the mechanisms that might lead from unproductive to productive states. Learners have the autonomy to explore options, as well as the ability to see the work of others, which supported these transitions and promoted collaboration. This further draws out the complexity of digital learning environments and how resources available to students, whether that be tools or the actions of others, can prompt learners to transition to productive states.

Using a complex systems approach, Dever et al., (2023) explored how pedagogical agents in an intelligent tutoring system can help scaffold self-regulated learning behaviors for 117 undergraduate students. Learners were divided into either a "Prompt and Feedback" condition, where the agent helped guide SRL strategies, or a control condition, where no agent was present. Time series of learner SRL strategy deployment was captured via log data, as well as whether the strategy was prompted or unprompted. Results showed that SRL strategies should be considered through non-sequential analyses, in that repetitive strategies should be demarcated from novel strategy deployment. In turn, learners scaffolded by agents had higher learning gains and greater frequencies of strategy usage, as well as lower recurrence rate of strategies. This suggests the prompted learners used more novel strategies in their learning. Through exploration of logs, productive patterns of behavior can be extracted, and methods for nudging learners towards productive patterns, through pedagogical agents or seeing the work of others, can defend against learner disengaging from task goals or the context altogether. Regardless, logs must be analyzed with other measures to build models of engagement and to infer the learner's state.

Several researchers have suggested including as many multimodal means as possible for capturing engagement, such eye-tracking, logs, and facial recognition (Azevedo & Taub, 2020; Wiedbusch et al., 2023), however others have cautioned against the "kitchen sink" approach, showing that the best predictive models usually do not include all means of measurement (Emerson et al., 2020, 2022). Including manually extracted metrics from self-reports and observations still proves an important part of detecting and understanding engagement (Dewan et al., 2019). All efforts extracting features from digital learning environments and combining them with other measures aims at understanding complex processes and how they can be influenced to aid learners in need through appropriate scaffolds and support (Marchand & Hilpert, 2024). Research on engagement with LMS' and games has been covered in this review, however research examining engagement and interest in widely popular contexts that learners may already have formed interests around remains relatively unexplored. Many of the digital games researched are researcher-designed (*i.e. Crystal Island, FactorReactor*), and learners are unlikely to have prior interest or much knowledge about such games. As such, K-12 age learners invest massive amounts of time into playing or learning about popular games, such as *Minecraft*, and extracting how they learn from play in a familiar context is a prominent question. Does their prior contextual knowledge and interest matter when learning, or does content interest and knowledge matter more? Through my research I expanded upon engagement and interest research by linking them through longitudinal gameplay data, self-reports, and knowledge assessments. This was done using popular media (*Minecraft*) that many learners frequently engage with outside of the classroom setting and likely have preconceived notions regarding it.

2.8 Minecraft, Learning, and Cognition

Minecraft is one of the best-selling games of all time (Yaden, 2021). *Minecraft* has two primary modes: survival and creative. In survival mode players must defend themselves from enemies by building shelters, crafting weapons, creating means of production for food, and designing transportation systems. Creative mode allows the player to freely explore the various biomes and build structures at their leisure. Because of the popularity of the game, as well as the autonomy and openness afforded to players, *Minecraft* is fertile ground for conducting research into educational outcomes (Lane & Yi, 2017). A lot of research has been done looking at possibilities for using *Minecraft* in the classroom and how the game can be adapted to incorporate concepts of Earth Science (Pusey & Pusey, 2015), urban planning (Magnussen & Elming, 2015), storytelling and creativity (Cipollone, 2015), cultural heritage (de Andrade & Poplin, 2020), computational thinking (Towhidnejad et al., 2014), math (Bos et al., 2021), and engineering (Dezuanni et al., 2015). Furthermore, a literature review exploring the various uses of *Minecraft* in education has pointed to *Minecraft* use resulting in higher interest, enjoyment, and persistence related to academic content for young learners (Baek et al., 2020).

Motivation and interest are important factors for engaging learners in formal and informal educational settings (Renninger & Hidi, 2016), and utilizing a medium that has appeal and personal value to a large swath of K-12 learners might hold unique motivational value. Previous work with WHIMC data has shown that regardless of individual experience with *Minecraft*, participants at varying levels of gameplay experience and interest were able to develop interest in STEM topics and successfully engage with STEM content (Lane et al., 2022). Evidence exists linking *Minecraft* to increasing motivation and pointing individuals toward targeted content, but efficacy of the medium should also be tied to evidence of learning (Mayer, 2014). While the literature is rife with examples of *Minecraft* being incorporated into both formal and informal education (*MinecraftEdu* was created for this purpose), the literature on learning and cognitive change resulting from *Minecraft* play is still relatively sparse.

Creativity is a popular cognitive variable to consider when researching *Minecraft* because of the autonomy offered the player to manipulate the environment in myriad ways (Baek et al., 2020). Research from Checa-Romero & Gomez (2017), had 85 adolescent middle school students participate in an 8-week workshop where students were instructed to build the house of their dreams and produce a narrative video about their creation. A pre- and post-test using the CREA Test was implemented to measure changes in student creativity, with results showing significant mean differences between times 1 and 2 for students and a moderate effect size (Cohen's d = .45). The CREA Test asks participants to generate questions based on three illustrations (Corbalan et al., 2003). Defining creativity can be a substantial challenge and understanding how it applies across situations is not straightforward, but the implication is important for recognizing the potential of the game to induce cognitive change.

Rich *Minecraft* environments allow players to explore different biomes (jungle, tundra, caves, tropical, coastal, etc.) and build structures of any scope, including functioning computers and musical compositions (Lim, 2020). Given the cognitive enhancements associated with gameplay in 3 dimensional environments, *Minecraft* can feasibly lead to greater spatial abilities and improvements in memory. Clemenson et al., (2019) used *Minecraft* to understand how exploration and building in a 3D environment impacts player performance on an unrelated memory task, the Mnemonic Similarity Task (MST). The researchers specifically looked at change (subtracting old from new score) in the Lure Discrepancy Index (LDI) of the MST to see how specific types of gameplay affect the cognitive ability to separate patterns and identify lures. Understanding recognition and familiarity as a graded scale (Wixted, 2007), this study might indicate that exploration and building in a 3D environment could increase a learner's ability to recognize images as being different, essentially reducing familiarity. Results from Clemenson et al., (2019) showed that players in a directed build condition, free exploration condition, and explore and build condition all showed significant change on the LDI between pre- and post-test. The only condition that did not improve was the free building condition. While three conditions improved on the LDI, none of the conditions showed overall improvement on the MST between pre- and post-test. Still, increased LDI from playing a game for a total of 7.5 hours over two weeks shows exploring and building in *Minecraft* can be used to affect cognitive change.

Minecraft research has mostly focused on learning, changes in cognitive variables, or collaborative behaviors. Some research has looked at *Minecraft* as an environment that can be motivating for learning (Baek et al., 2022), however studies have continued to rely on self-reports instead log data available from learner play patterns. I utilized the data available from logs of learner play patterns to assess how interest influences engagement over several distinct scenarios, evidencing deeper connections between contextual interest in *Minecraft*, interest in Astronomy content, and engagement. Through my research I provide insight into how learners with relatively low interest engage compared to those with higher interest in Astronomy and *Minecraft.* Furthermore, I examined if the learning environment was able to trigger their interest, and if learners less interested in *Minecraft* were able to engage content despite lack of familiarity with navigating the medium. Lack of interest in *Minecraft* did not appear as a roadblock to meaningful engagement, as stronger interest in content drove higher levels of engagement.

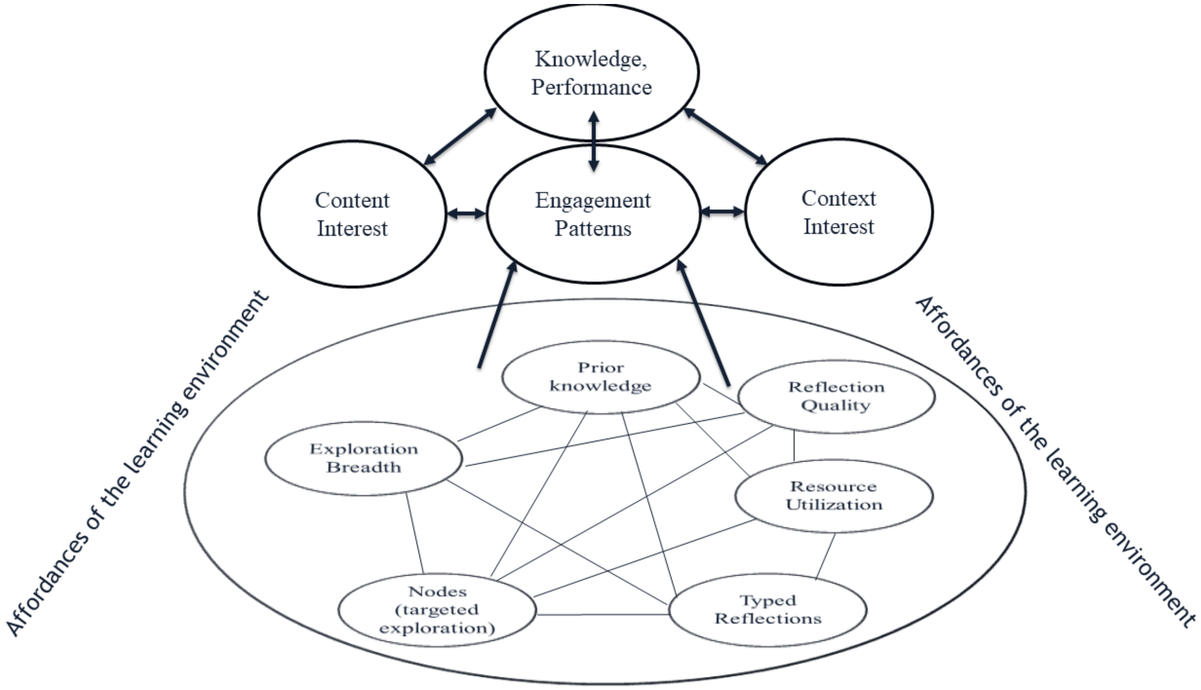
CHAPTER 3: RELEVANCE & SCOPE OF STUDY

My dissertation builds upon the work done with LMS technologies by extending time series clustering (Zhang et al., 2022) and categorization of interest levels (Emerson et al., 2020; Gadbury & Lane, 2022) to the ways adolescents engage with STEM content in open-world, digital environments. Furthermore, the digital environment is one that is massively popular (*Minecraft*) around the world (Woodward, 2024), and players that interact with the environment are going to have preconceived notions regarding the game. My research aims to disentangle contextual (*Minecraft*) engagement and interest from conceptual (STEM, Astronomy) engagement and interest through analyses that include both game-external and game-internal assessments of engagement, interest, and knowledge. Much research using learning analytics has been described as agnostic in linking engagement to log data patterns (Martin & Borup, 2022; Vytasek et al., 2020), and with this in mind, my research grounds each log feature in a dimension of engagement. Through grounding engagement, my research pays special mind to the effort put forth by the individual, the affordances of the digital learning environment, prior knowledge, utilization of learning tools available, and how all of this produces emergent patterns of engagement that affect and may be affected by interest, leading to knowledge construction (see Figure 2). My aim is to uncover emergent features of engagement and interest as learners explore content-rich environments primarily under their own volition, using the tools available to them. If the environment is physically narrow or does not include measurements for several tools then we cannot consider exploration breadth or tool use to be high counts, for example.

Patterns of behavior were extracted from the game log data and analyzed in a way that accounts for between subjects-differences (how different interest profiles realize specific patterns of engagement) and within-subjects differences (how persistent high v. low interest individuals differ in their behavior over time). This analysis treats the learner and the environment as interactive and consistently changing entities, in line with a complex systems approach to research (Hilpert & Marchand, 2018; Marchand & Hilpert, 2020). Complex systems research methodology is a nascent way of understanding contextualized experiences of learners as they navigate learning environments (Marchand & Hilpert, 2024). Given the complexity of interactions and lack of research into the role interest plays in wayfinding and learning in digital game environments, an exploratory approach is used in my research.

Figure 2

Depiction of the complex relationship between engagement, knowledge, and interest



Note: Affordances of the learning environment encapsulates possible activities (i.e. resource utilization) that might contribute to learning; the learning environment affords the opportunity for learners to engage in these activities. The large, bottom circle is the learning environments, and all smaller, internal circles capture an engagement activity. Activity in the digital learning environment incorporates prior knowledge, and all activities influence each other. Interactivity in the environment leads to emergent behavior patterns that are influenced by knowledge, and interest, and influence knowledge and performative outcomes. Adapted from Hilpert & Marchand (2018).

An exploratory approach in this case allows me to focus on the individual and not entirely on variables being manipulated. Interest develops differently for each individual, based on their learning characteristics, as well as the environment (Renninger & Su, 2012), and my focus is on the state of the learner and how that influences their individual engagement. My research does not draw assumptions about how manipulated variables will influence play behavior and conceptual understanding in this context, but rather how the variability in the individual lends itself to patterns of behavior that might then indicate interest and knowledge. Through this approach I have flexibility to understand the emergent patterns of individuals in the data without needing to hypothesis how a slight change in the environment will impact interest. The variability comes from the unique dispositions and knowledge of each individual. Relationships might be uncovered, such as a particular form of engagement (*e.g.* a learner who only likes to explore far and wide) and performance on an assessment or an effect drawn from the site the data was collected.

Theories of interest and engagement are well-developed, yet how they manifest in settings, such as open-ended games, when many variables are in play is deserving of further research (Renninger & Hidi, 2022). An exploratory approach raises questions, such as what influence does interest have on the directions individuals take in their moment-to-moment learning, or how does interest in the mode of delivery interact with interest in the domain. The context, *Minecraft*, is one in which familiarity and expertise varies from learner to learner, and context interest could have an additive or conflicting impact on engagement and domain interest. With so many intersecting and dynamic variables, an exploratory approach is warranted to draw out what connections to theory exist, setting up future research for verification of the hypotheses generated from an exploratory approach (Swedberg, 2020).

An exploratory approach affords researchers the chance to challenge or support existing theoretical assumptions in related but previously unexamined contexts. For example, interest theorists state there is a strong relationship between interest and knowledge (Renninger & Hidi, 2021; Rotgans & Schmidt, 2017; Tobias, 1994), and this can be considered a *working hypothesis* that guides further inquiry into an area that may already have a long history of inquiry (Casula et al., 2021). A *working hypothesis* entails a provisional state of the hypothesis that can be nullified or further supported by findings. In this sense, my work falls somewhere between exploration and explanation, asking questions of "what" that do not necessitate determining causation but that generate new avenues for testing and confirmation. This bottom-up approach does not assume effectiveness of any intervention, or that particular learners will engage in specific ways. All variables involved may impact each other in a complex way that should first be uncovered before being experimentally tested.

I address questions stemming from, "how do prior experiences and affect (*i.e.* knowledge, *Minecraft* interest and play experience, identity) influence learners' engagement patterns and interest development through game-based, STEM summer camps?" through engineered features from data, time series clustering, and Bayesian models aimed at predicting interest. The patterns of engagement across several time points are clustered to provide an assessment of which behaviors pertain to high performance and interest. Each time series is standardized based on individual performance on the first world explored, *Lunar Crater*; in this way individual fluctuations of engagement can be captured across worlds and comparisons can be made between highly interested individuals' emergent patterns of behavior. Do different patterns result in similar outcomes in terms of knowledge and interest? Time series clustering has been conducted on LMS logs (Zhang et al., 2022) but not digital game log data. Time-series clusters can be compared with knowledge and interest scores to see if those with high interest also exhibit effective patterns of behavior for building knowledge and developing their own interest. As a variable of focus, interest has rarely been examined through analysis of logs, with some exception (Emerson et al., 2020; Geerling et al., 2022), however the authors did not theoretically ground logs with engagement, and the game used was not one that is widely known around the world.

This study explores what dimensions of engagement, as well as demographic data, best predict contextual and conceptual interest using Bayesian Model Averaging, a method of analysis that holds great potential for social science research but is underutilized (Hinne & Gronau, 2020). Ultimately, if interest behaviors can be understood through engagement logs, then adaptive systems can be tailored more closely to individual interest, promoting value, knowledge creation, and personal relevance to the learner. If specific patterns of engagement are predictive of interest at the end of a week-long learning experience, then evidence from highly effective patterns can be used to inform the experience and promote higher engagement for others who may not be utilizing all the resources available to them. Researchers can then test interest interventions based on how interest ties to engagement to determine what strategies work and what strategies do not. Furthermore, extracting interest from logs can reduce the need for self-reports of interest, which can often misrepresent an individual's scope of interest (Renninger & Hidi, 2011). If interest does indeed possess robust motivational power, then it should be possible to examine how interest influences the decisions learners make in a variety of learning environments.

CHAPTER 4: METHODS

4.1 Data Source

Analysis was conducted on an IRB approved dataset. All data was collected as part of the WHIMC project from June – August 2022. The dataset consists of log files, 1-on-1 interviews, typed self-explanations answered in-game, and self-reports from learners participating in summer camps that met for 3 to 6 hours daily for the duration of a week (See Table 1).

Table 1

Summary of data collected and analyzed from summer camps

Data Source	In-game	KR	Time-stamped	Location Tracked	Text Input	Pre/Process/Post	Summary
In-game Logs:							
Location	X		X	X		Process	Recorded every 3 seconds
Proximity to POIs/NPCs	X		X	X		Process	Interaction recorded when within 10 block radius of POI/NPC
Science Tool Measurements	X		X	X	X	Process	Meant to be used repeatedly
Typed Observations	X	X	X	X	X	Process	Frequency & quality both matter
Semi-structured interviews		X				Post	1-on-1 interviews asking about STEM interest, STEM identity, and Astronomy knowledge questions
Self-explanations	X	X			X	Process	Completed after exploration phase, explain why phenomena occur/appear
Self-Reports						Pre & Post	1 - 5 Likert-type scale (1 = not interested, 5 = extremely interested) measuring STEM interest, Astronomy interest, and Minecraft interest

Note: POI = point of interest, NPC = non-playable character, KR = Knowledge/Reasoning Component

Data has been collected from WHIMC summer camps since 2018, however 2022 was the first year where camps were run in multiple locations and the sample size is considerably larger than in previous years. This allowed for greater diversity regarding participant gender and race as well. WHIMC is and has always been an evolving project, with new content and activities introduced each year. Therefore, only data from 2022 is used in my dissertation because tasks, instructors, and measurement instruments were held constant across sites, yet differ from previous and subsequent years. Some camps in 2022 were voluntary, where learners and their parents decided to enroll in the camp, while some of the camps were incorporated into summer programming at community centers and required learners' participation. This is accounted for in Bayesian models, as contextual factors may have an impact on engagement and interest in the content and learning environment. Still, the amount of time playing *Minecraft* remained roughly constant across all camps that I have included in my analyses.

4.2 Participants

All participants were enrolled in public middle or high schools in the United States, and the age ranged from 11 to 15 years (see Table 2). Participants spread across three locations: (1) BR, FIS, SD all from the western United States, (2) UN1 and UN2 from the Midwest United States, and MA and MP from the eastern United States. See Table 3 for sample sizes and demographic breakdowns at each location.

Table 2
Breakdown of age by gender

	10	11	12	13	14	15	Total
Female	8% (2)	32% (8)	4% (1)	28% (7)	24% (6)	4% (1)	33% (25)
Male	22% (11)	20% (10)	31% (16)	16% (8)	12% (6)	0% (0)	67% (51)
Total	17% (13)	24% (18)	22% (17)	20% (15)	16% (12)	1% (1)	100% (76)

Participants were recruited through ongoing relationships with school and community partners, as well as through targeted marketing of camps and programs. Aside from being in the 11 to 15 years age range, no prerequisite was required for participation. Participants either voluntarily enrolled in a community center summer program or were required to attend as part of their community center's summer programming. Informed written consent was obtained from at least one parent/guardian of each participant through REDCap, an online consent form, and assent was obtained from each participant on the first day of the respective program or camp. Consent was obtained for all participants.

The participants in this study are important because they represent a diversity of backgrounds, in terms of socioeconomic status, experience with *Minecraft*, interest in STEM, and desire to be part of the summer camp. If all participants were already interested in Astronomy and *Minecraft* then changes in interest might be undetectable due to ceiling effects. However, if this summer camp can trigger interest even for those who had the experience programmed in by their community center (or parents) then it lends credence to the power of *Minecraft* to make learning feel less effortful, more invigorating, and interest triggering.

Table 3

Sample size and demographic breakdown for WHIMC summer camps 2022

Site	Sample size (*n*)	% Female	Race/Ethnicity
BR	10	20	33% White, 17% Hispanic, 50% PNA*
SDP	18	6	61% White, 6% Asian, 33% PNA
SK	10	30	70% Hispanic, 10% PNA, 10% Asian, 10% Black
UB	9	0	100% Black
UG	9	100	100% Black
MA	10	50	64% White, 9% Hispanic, 27% PNA
MP	10	40	30% White, 10% American Indian, 60% PNA
Total	76	32	37% White, 28% PNA, 24% Black, 13% Hispanic, 2% Asian, 1% American Indian

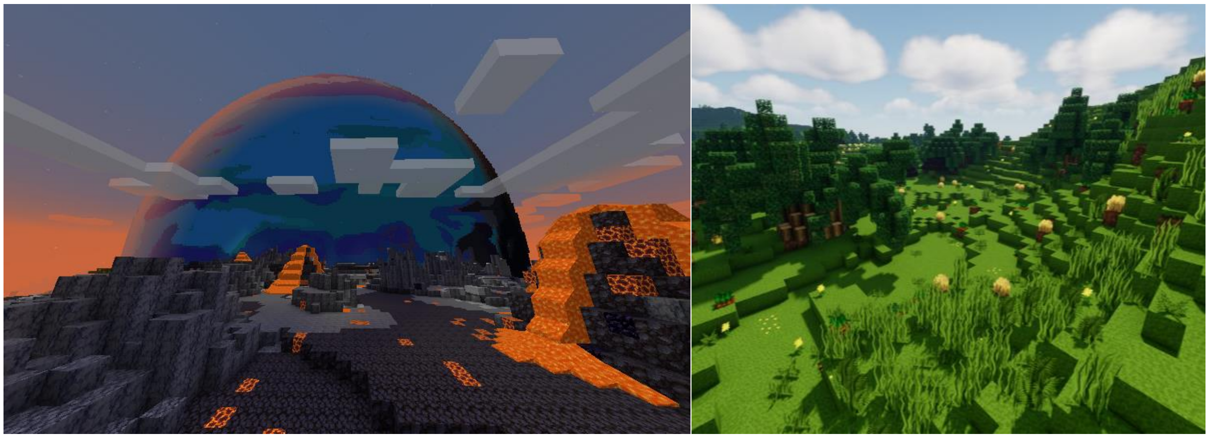
Note: *PNA stands for “Prefer Not to Answer”.

4.3 Materials

Participants were all provided with a laptop, mouse, and an individual account to play *Minecraft: Java Edition*. Participants used the same account for each session of their respective after school program or summer camp. All maps explored by participants were created by our lab and represent simulations of “What if” questions, such as “What if Earth was a moon to a larger planet?”, as well as known exoplanets (*e.g.* Kepler 186-f). The “What-if” worlds are all based on hypothetical scenarios described in *What if the Moon Didn't Exist? Voyages to Earth that Might have been* (Comins, 1993). Design of worlds was done in consultation with the author and feature extreme conditions, such as high winds, widespread volcanic activity, freezing temperatures, or low gravity, which can all be seen or measured using science tools. Examples of and detailed descriptions of each world can be found online at the WHIMC homepage: (https://whimc.education.illinois.edu).

Minecraft: Java Edition is an open-world sandbox style digital game, where players can explore vast worlds with varying terrain. See Figure 3 for examples from the WHIMC server. Players can interact with every block in the game, collect resources, and build any type of structure. There are different modes of the game, creative and survival, with creative mode providing players with unlimited resources to build, and survival requiring extraction of resources and battling against enemies. Our project relies mostly on creative mode to focus on exploration and building without the threat of players' characters falling, taking damage, or being

Figure 3
Examples of “What-if” WHIMC Minecraft environments (Left: Earth as a moon, Right:



attacked by in-game enemies (which are all neutral in creative mode). The “What if” scenarios and exoplanets were created using (1) specialized world authoring tools to create unique terrains, (2) datapacks that allowed for altering of basic world features like skies and liquid textures, and (3) “plug-ins”, or coded adjustments to the game that allow for customized tools and robust data collection.

Many other activities augmented *Minecraft* play in our camps, such as drawing exoplanets, using inflatable planets to represent relationships between celestial bodies, and discussion sessions to answer questions and breakdown what was seen on each world. These activities were designed to enrich the learning experiences and provide breaks from computer time, and certainly impacted the knowledge and interest a participant leaves the camp experience with; however, insufficient data regarding participation and engagement in these activities limits the inferences that can be made, and thus activities outside of gameplay, surveys, and interview transcripts are not considered in this dissertation. Participants in the full-day camps had more augmented activities, such as planetarium shows, and outdoor recesses.

4.4 Measures

Self-report surveys (Interest)

Participants completed a pre- and post-survey measuring STEM and STEM-related *Minecraft* interest, hereafter referred to as MC interest. The STEM and MC interest surveys consisted of a total of 40 (20 STEM interest and 20 MC interest questions), 5-point Likert-type items asking participants how much an activity interested them, ranging from “1 = Not Interested” to “5 = Extremely Interested”. A sample STEM item reads, “How interested are you in determining the amount of force needed to break a wooden board?”, and a sample Minecraft question reads, “How interested are you in crafting foods and making/using dyes?”. The survey was developed by our research lab to pinpoint learners' interest in various STEM domains perceived as being relevant to Minecraft play.

Previous unpublished analysis showed a significant positive correlation was found between participant-rated STEM interest and Minecraft STEM interest ($r = .882$, $p < .001$). Cronbach's Alpha was used to measure reliability across all item blocks, resulting in acceptable consistency for all blocks (α between .8 and .9 for all blocks). Validity was established by conducting a confirmatory factor analysis (CFA) on $n = 705$ responses. The final accepted model consisted of a bifactor model with an orthogonal 'General Interest' as the general factor, 'STEM interest' and 'STEM-related Minecraft interest' as correlated latent factors, and all composite CIP categories loading onto the general factor and their respective interest factor ($\chi 2 = 231.12$, $df = 87$, $p <.001$, CFI = .98, TLI = .98, RMSEA = .048, SRMR = .02). Items with absolute value factor loadings <.3 were removed, resulting in the shortened survey used in my research. The pre-survey also included demographic questions. Because the focus of the summer camps was on Astronomy concepts, an additional 5 Likert-type questions were added to gauge interest in Astronomy specifically. The Astronomy questions were in the same format as the STEM and MC interest questions, and examples were, "Please indicate how interested you are in learning more about exoplanets", and "Please indicate how interested you are working in a space-related career someday". Because the questions were not validated with the STEM and MC interest surveys, they will be analyzed separately. Additionally, the Astronomy questions represent situational interest in the domain being examined by my research, as opposed to more enduring, individual interest represented by the STEM and MC surveys. See Appendix A for a complete list of survey questions.

Log Files (Engagement and Interest)

Log data of participants' exploration patterns, science tool use, and observations were captured on the server. A plugin running in the background recorded student location data every three seconds, providing an output of coordinates traversed by each participant. Every map in *Minecraft* consists of "X", "Y", and "Z" coordinates. The "X" coordinate describes the "East/West" position in each map, the "Y" coordinate describes the up/down position of the player in each map, and the "Z" coordinate describes the "South/North" position of the player in the world. While all three coordinates are captured for each participant, the "Y" coordinate was not included in analysis, as terrain is relatively flat and players do not have the ability to fly upwards. Because of the limitations placed on "up/down" mobility, "Y" coordinates do not offer additional information on top of "X" and "Z" coordinates. Coordinate data is accessed from the server using SQL commands. Each output record includes "X", "Y", and "Z" coordinates and time.

Programmed into each map on the server are a collection of measurements that change, depending on the location of the participant. The measurements are referred to as "Science Tools", and include measuring the following variables: oxygen, temperature, pressure, magnetic field, tilt, wind speed, day/night cycle, radiation, and gravity, among others. In total, there are 17 science tools players can use in each world. See Table 4 for a list of science tools and associated outputs. Through typing "/gravity" into the command bar in *Minecraft* the participant receives a measurement output for their specific location, and the output is compared to the known measurement on Earth. See Figure 4 for an example. Some of the variables are global measurements, such as tilt, while others change frequently throughout exploration, such as temperature. Every science tool measurement is recorded on the server. Science tool use is accessed through SQL commands and each output record includes time of measurement, location, measurement taken (*i.e.* /oxygen), and the output viewed by the player. When learners approach NPCs dialogue appears that will, in some cases, offer suggestions for using specific science tools or provide valuable information regarding the phenomena occurring in the environment being explored.

At any point while exploring a map the player may use the command "/observe" to type an observation into the command bar at the bottom of the screen. See Figure 5. Observations are open-ended and can reflect what a player sees directly in front of them, or a possible hypothesis as to why a phenomenon occurs. Observations are accessed from the server through SQL commands and each output record includes observation text, time, and location.

Table 4

Available science tool inputs, variability, and associated outputs

Science Tool Input	Type	Output
/gravity	Global	[value] times the Earth, or [value] m/s^2
/radiation	Global	[value] times typical Earth radiation exposure
/atmosphere	Global	(Elements vary by world)
/pressure	Varies	[value] kPa, [value] times the Earth's pressure
/airflow	Varies	[value] m/s
/humidity	Varies	[value]
/oxygen	Varies	[value]%
/temperature	Varies	[value]ºC/[value]ºF/[value]ºK
/year	Global	[value] days
/rotation	Global	[value] hours
/tides	Global	Tidal variance [value] greater/less than Earth's
/tectonic	Varies	Tectonic activity is (low/average/high)
/magnetic_field	Global	The magnetic field is [weak,strong], influenced by [moon/nearby planet]
/cosmicrays	Global	[value]MeV
/tilt	Global	[value]º
/altitude	Varies	[value]meters
/radius	Global	[value] times the Earth's radius

Note: *Type indicates if the output varies by location on world or remains static (Global) on given world*

Figure 4

Examples of science tool use on "What if" WHIMC server worlds (Left: Earth with Colder Sun, Right: Earth with Two Moons)

/temperature

The ambient temperature is -103.1°C (-180.7 °F) (155.0 °K)

/tides

The tidal variance is About 8 times greater than the range of tides on Earth, currently high tide

Figure 5

Examples of observations made on "What if" WHIMC server worlds (Left: Tilted Earth, Right: Colder Sun)

is melting because the sun is rising after 3 months o

AIEW - March 2 2024, 9:15 AM CST

Expires March 7 2024, 9:15 AM CST

[Observations] Your observation has been placed:
"the ice is melting because the sun is rising after 3 months of dark"
(Completed) Observe why it is so hot on Mynoa!

the sun appears pink in the sky

SAIEW - March 2 2024, 9:10 AM CST

Expires March 7 2024, 9:10 AM CST

[Observations] Your observation has been placed:
"the sun appears pink in the sky"
(Completed) Observe why it is so hot on Mynoa!

Participants answered open-ended self-explanations after exploring each world for 20 – 30 minutes and before moving on to the next world. Questions were developed by the research team and reflect concepts that students encounter throughout exploration of each world. For example, students are prompted to measure wind on a map that has extreme global wind speeds. A related conceptual self-explanation is "Based on what you saw, what is a noticeable effect of high wind speeds on No Moon?". A student might respond by mentioning the short height but great width of the trees, or the use of wind energy to provide power. Another example is, "Could humans survive on Earth Tilted at 90º? Why?”, and a well-reasoned answer might mention difficulty surviving 3 – 4 months of no sunlight on the hemisphere facing away from the sun, and the challenge of migrating many humans to different hemispheres every year. See Appendix B for the full list of self-explanation prompts and possible correct answers. Each question can have multiple answers that can be considered correct, and the ability to answer it correctly requires time spent exploring the map, gathering information or suggestions for measurements from NPCs, and integrating measurements with visual observations. Self-explanations are pushed to each player in-game through a plugin called SurveyPlus. Answers are recorded on a database and accessed using SQL commands. Self-explanation records include question/prompt and response.

Interviews (Interest and Knowledge Assessment)

One-on-one interviews were conducted with each participant at the end of the camp week. Interviews were recorded on Sony ICD-UX570 audio recorders and uploaded to a Box folder and deleted from the device to protect participant confidentiality. Interviews were conducted by either me and other graduate assistants working on the project or Dr. Lane, project PI. Interviews consisted of questions regarding interest in STEM in formal and informal settings, *Minecraft* play habits, and a series of knowledge questions regarding content from the camp. For example, participants were asked, “Do you ever watch any YouTube videos or Google information about outer space in your free time?”, which captures a degree of voluntary re-engagement with STEM that indicates at least a triggered interest (Renninger & Hidi, 2020a). Full interview protocol can be found in Appendix C. All audio files were transcribed by a paid online service, TEMI, and checked for accuracy by an undergraduate or graduate student working on the project.

4.5 Procedure

Procedures regarding *Minecraft* play and data collection were identical across locations. Participants were assigned seats by researchers or camp staff (see Figure 6 for examples of camp setup) and directed to complete the STEM & MC interest surveys by following a link to the survey in Qualtrics. Participants were given 10 minutes to complete the survey. Following surveys, ground rules and expectations were communicated to participants regarding behavior, and the purpose of the research project was explained. Participants were then logged into *Minecraft* and teleported to the orientation world, *Rocket Launch*, which provided an opportunity for participants new to *Minecraft* time to learn the controls, and more experienced players could gain a sense of how the WHIMC server works (*i.e.* interactions with non-playable characters, exploration) and what the bounds of actions are (*i.e.* no flying). Examples of how to take measurements and make observations were provided on-screen by researchers in *Lunar Crater*, the second world explored. Participants were then allowed to explore and use science tools to uncover relationships between physical phenomena and habitability concerns on the imagined worlds. Observations were encouraged throughout exploration. If a player vocally expressed an

Figure 6

Camp setup at two distinct sites from summer 2022

idea, excitement about an in-game phenomena, or a description of the world they were asked by camp staff to turn that into an in-game observation. At the start of every world following *Lunar Crater*, participants were asked to explore the worlds as fully as possible in the given time, use science tools to take measurements of temperature, gravity, oxygen, pressure, among others, and to write observations of what they see and understand about the worlds. Please see Appendix D for an example full camp schedule.

After completing approximately 25 minutes of exploration time on a given world, all in-game activity was paused by a research team member and learners were asked to complete self-explanations. The reason for pausing participants was to prevent continued exploration and in-game distractions while completing the questions. Once paused, prompts were pushed to players by a research team member. Each world has an accompanying set of self-explanations that ask participants to draw on what they observed to hypothesize and describe reasons for phenomena they saw in-game. Participants were given 5 minutes to complete each accompanying set of self-explanations, with each set consisting of 2 – 3 self-explanations. Prior to completing the first self-explanation, an example was provided on-screen, and a researcher showed participants how to answer using the in-game command bar. Participants were asked not to confer with peers to answer prompts, and camp staff did not provide any guidance answering the prompts, however staff did help with any technical issues that may have arisen.

The final two sessions of the camp or program, participants were tasked with building a habitat on a 1:1 generated map of Valles Marineris on Mars. Participants were placed into groups with others sitting at their table, or researchers assigned groups if needed. Participants were told to design a habitat for humans to survive and carry out research on Mars that might allow future generations of humans to live on Mars. Participants were told to use what they learned from exploring all the previous hypothetical worlds and exoplanets to inform how they respond to conditions they can measure on Mars. Participants had around 3 hours to work together and build their habitats. About 40 minutes prior to the end of the last session, groups were asked to provide a tour of their habitat on the screen in the front of the room and explain the problems they addressed and how they solved them. This data is not considered in this analysis, as it constitutes a different way of interacting with the server from the exploratory data presented here.

With 15 minutes left in the final session of the camp or after-school program, participants were asked to close *Minecraft* and click on the survey link provided in the browser. The link took participants to the same STEM & Minecraft interest survey they completed in Qualtrics at the beginning of the camp or program. After completing the survey, participants were released, and the camp concluded.

CHAPTER 5: DATA ANALYSIS

All data included are expected to capture an aspect of either engagement and/or interest development, whether that be triggering of situational interest through novelty or knowledge displays for those with established individual interest in the topic. See Table 5 for Guzdial chart of how data addresses research questions. Because of expected non-normality in the data and a relatively small sample size, all measures were analyzed using non-parametric and Bayesian approaches (van de Schoot & Miočević, 2020), unless assumptions were met using statistical tests, such as normality using the Shapiro Wilk test. Feature engineering of log data will first be discussed and how each feature relates to one of the dimensions of engagement (Fredricks et al., 2004; Martin & Borup, 2022). Times series clustering was used to determine patterns of behaviors that reflect differing levels of emergent engagement (Aghabozorgi et al., 2015). Coding and scoring schemes are detailed, including calculations of inter-rater reliability and internal reliability using Cronbach's alpha. All measures were correlated using the non-parametric Spearman Rank Correlation, and Bayesian models were used to determine which aspects of engagement best predict a learner's interest in STEM, Astronomy, Minecraft-related STEM activities, and Astronomy knowledge, respectively. In short, profiles of varying patterns of engagement were expected to be tied to gradations of interest in STEM, providing insight into how conceptual interest (STEM) and contextual interest (*Minecraft*) inform the ways adolescents approach digital STEM activities in informal settings.

5.1 Spacing, Grain Size of Measurement, and Dimension Mapping

Spacing is an important decision that must be determined before measuring engagement (Hilpert & Marchand, 2018; Sinatra et al., 2015). This proposal considers engagement across approximately 3 hours of gameplay but segmented by each unique world visited. In total, data from six worlds was included, as these worlds are thoroughly constructed for engagement and do not require changing mode of play. Exoplanet worlds were not included because less time was spent on each individual world, and they have fewer NPCs and Points of Interest (POIs) than the "What-if" worlds. Some of the exoplanet worlds also require switching to spectator mode to fly, which entails exploring these worlds in a vastly different way than the worlds being analyzed. By looking at engagement on each unique world, a time series of engagement with determined turning points can be derived (Zhang et al., 2022) and used to find clusters of individuals with similar patterns of engagement (Aghabozorgi et al., 2015). This approach also allows an analysis of how an individual persists in various measures of engagement (Al Mamun et al., 2022), and then profiles of persistence (consistent engagement) can potentially be tied to degree of interest (Emerson et al., 2020).

Interest is operationalized through multiple, static measures that capture a more enduring interest in content and context, and how that interest informs the ways in which participants engage the learning material present in the digital environment. Each engagement feature extracted consists of either a frequency or average by world. See Table 6 for a summary of engagement features derived from logs and their mappings to engagement dimensions. Undoubtedly, some log features can be argued to exist in multiple dimensions, and there does tend to be overlap between the engagement dimensions (Sinatra, Heddy, & Lombardi, 2015). Next, I present arguments for inclusion in a primary dimension for each feature, acknowledging there is overlap and some features can exist in multiple dimensions simultaneously.

Table 5

Guzdial chart summarizing how research questions will be answered

Research Questions	Participants	Instruments	Analysis
What log data features best reveal affective, cognitive, and behavioral engagement?	• Middle school students on summer break participating in a week-long summer camp. • Ages range from 10 - 15 years • Race/Ethnicity diverse between camps, but mostly homogenous within • Wide variety of experience with playing Minecraft	Log data captured from server (movement, use of resources, time, and text inputs). Outlined goals of the summer camp.	Theoretically grounded based on affordances of the digital learning environment and tasks prescribed by instructors.
Which combinations of features effectively capture engagement in this context?		Log data from server (targeted movement, breadth of movement, measurement tool use, text inputs). Self-reports of interest in STEM (broadly), Astronomy, and *Minecraft*.	Examine correlations between features to eliminate redundant features, and to gauge if features have relationships with interest and/or knowledge.
How can content-related v. context-related interest be effectively distinguished through logs?		Log data from server (targeted movement, breadth of movement, measurement tool use, text inputs). Self-reports of interest in STEM (broadly), Astronomy, and *Minecraft*.	Exploring links between different groups of performers, as revealed by time series clustering, descriptives, and scores on pre-interest surveys. Compare the frequency of high performers who fall above interest means with that of low performers who fall above the same means.
How can states of interest, as well as changes in interest and knowledge, be captured through the triangulation of log data, self-reports, and knowledge assessments?		Log data from server (targeted movement, breadth of movement, measurement tool use, text inputs). Self-reports of interest in STEM (broadly), Astronomy, and *Minecraft*. In-game conceptual self-explanations, and post-experience knowledge assessment.	Log data analyzed using time series clustering to reveal patterns of engagement. Changes in interest captured through self-reports can be tied to engagement patterns to see if certain patterns closely associate with interest changes. Knowledge assessments can be linked to shifts in interest and patterns of engagement.
What variables comprise the best predictive models of interest in STEM, Astronomy, and Minecraft?		Log data from server (targeted movement, breadth of movement, measurement tool use, text inputs). Self-reports of interest in STEM (broadly), Astronomy, and *Minecraft*. In-game conceptual self-explanations, and post-experience knowledge assessment. Camp location as categorical predictor	Explore what variables have the highest probability of predicting knowledge and interest to see if malleable factors, such as specific engagement features, can be encouraged to develop interest and knowledge. Also, provide further evidence for/against theory as to the complex relationship between engagement, knowledge, and interest.

Table 6

Mapping of logs and network metrics onto engagement dimensions

Engagement Dimension	Definition	Example	Log/Metric
Affective	Emotional energy the learner uses when completing activities in a digital learning environment	A curious learner will navigate to more spaces of a map to uncover information.	**Nodes**: Count of POIs or NPCs interacted with in LE. **Science tools**: Count of measurement tools used **Exploration Breadth:** number of designated coordinate blocks traversed by the player
Behavioral	Physical actions and energy that students use to complete learning activities	A learner may type several observations while exploring the environment and exhibit less time between actions	**Average action latency:** average amount of time between actions (observations and science tools) **Observations**: count of open-ended observations made
Cognitive	Productive or strategic involvement with learning activities in the environment	A learner seeks information from a source, puts information to use, and then reflects on that information through writing	**Observation quality %:** % of observations categorized as comparative, descriptive, or inferential **Self-explanation scores**: scores for reflective prompts after exploring a given world **Unique science tools:** count of unique science tools used in each world

5.2 Affective Engagement Logs

Exploration Breadth

Exploration breadth is considered one important metric for understanding how players move around a map, but incorporating more forms is integral for triangulation. Maps are divided into grids based on the maximum and minimum “x” and “z” coordinates visited by players on a map. Each map was then divided into 100 squares, as this provides a middle ground between fine- and coarse-grained tracking. The width of each square, a step in the “x” or “z” direction, was calculated by subtracting the minimum value of “x” or “z” from the maximum value of “x” or “z”, and that number was then divided by the number of squares. A step is then going from one block to another, and the sum of steps for a given map provides exploration breadth. (See Figure 7). A player with a higher exploration breadth score has explored more of the map than a player with a low score (*i.e.* the learner has taken more steps). For this analysis, exploration breadth is on the affective dimension of engagement. Exploration breadth entails a player is expending emotional energy to satisfy curiosity through exploring the map and seeking information (Arnone et al., 2011; Hardy et al., 2019; Hidi & Renninger, 2020). Arguments could be made for exploration breadth being on the behavioral dimension, but the affective state of curiosity makes a stronger argument for being the driving factor in exploring novel, “what-if” scenarios. Exploration breadth does not ascertain time spent investigating, rather it is a breadth measure to capture extent of exploration; how much of a map did a learner uncover, and how probable are they to be found in any specific location on the map.

Science Tool Use Frequency

Science tool use frequency is another exploration metric, as it pertains to active information-seeking behavior on the part of the player. A player who is curious about the temperature on the side of a world covered entirely in volcanoes will take a measurement to satisfy their curiosity

(Ainley, 2019), which can be considered a possible pre-cursor to situational interest (Hidi & Renninger, 2019; Murayama, 2022). Repeated use of science tools reflects affective engagement through acting upon one's curiosity to know more about the environment, comparing conditions in different areas of the map. Science tools can also be considered a form of cognitive engagement, as they do require comparisons of information and an active attempt at sensemaking in the hypothetical worlds.

Nodes

From network analysis research, nodes provide a picture of where participant are visiting in an environment (Hilpert, Green, & Bernacki, 2023). In the WHIMC scenario, nodes constitutes the number of points of interest (POIs) visited by a participant, which includes non-playable characters (NPCs) visited. NPCs provide information on how to use science tools and which science tools are most advantageous to use in each world. Nodes captures another affective form of engagement, where the player may be curious about what NPCs have to say, as well as what they might see in various POIs. The count of nodes is another reflection of exploratory behavior that complements exploration breadth and provides evidence of exploration targeted at obtaining information. A higher count of nodes may serve as a precursor towards gaining targeted knowledge and developing greater depth of conceptual understanding for each world. In this way, nodes can possibly be viewed as a form of cognitive engagement, however whether the learner attended to the information meaningfully cannot be discerned from nodes count alone.

5.3 Behavioral Engagement Logs

Observation Frequency

Observation frequency accounts for the total number of observations made by a player. While science tool used constituted a measure of exploration and information-seeking, observations are a means of consolidating information and recording what has been experienced, trying to piece together answers for some. Observations require more effort than using a science tool, as the learner must take information and create an explanation in writing, and in this regard it is also a cognitive exercise of attempting to make sense of what has been experienced. However, frequency should only be considered on the behavioral dimension of engagement because it does not account for the quality of observations, and participants may use observations for chatter and off-topic discussion. Observations require effort but a high number of observations does not necessarily entail productive engagement towards conceptual understanding.

Figure 7

Example of how Exploration Breadth is calculated based on a player's path on a world

Note: Shaded grid boxes represent areas an individual visited at least once. Differences in line color represents changes in elevation.

Average Action Latency

Average action latency looks at the average amount of time between use of science tools and or observations for each world. Time-on-task is a measure frequently used in learning analytics as a measure of behavioral engagement (Martin & Borup, 2022), indicating the more time a learner spends on a task the more self-regulated behavior they are showing (Wiedbusch et al., 2023) and the more they are likely to learn (Kew & Tasir, 2022). The nature of the summer camps analyzed in this proposal necessitates equal amounts of time spent across worlds for each player, and therefore time-on-task is uniform for all players. Average action latency thus provides a variant time-on-task measure by extracting the amount of time players take between use of tools available to them for understanding the environment. It is likely that a lower mean action latency for each world negatively correlates with unique coordinates, as time is needed to type commands and observations. Average action latency constitutes engagement along the behavioral dimension, as it measures how frequently a player is utilizing a behavior to understand the world they are exploring. Some analytics research examining action latency in interactive web-environments posit low action latency means students are not taking time to reflect on their environment (Amershi & Conati, 2006). However, given limited time to explore, participants may have to make tradeoffs between putting effort into making observations or taking measurements with science tools and exploring a map broadly.

5.4 Cognitive Engagement Logs

Unique Science Tools Used

Unique science tools used is a measure how many of all possible science tools a player used per world. If a player took advantage of the full suite of science tools, then their score will max at 16. This also constitutes a form of cognitive engagement, in that the player is taking steps to increase knowledge and understanding of the world as far as they can (Fredricks, 2011; Fredricks et al., 2004). This can also be indicative of interest at a cognitive level, where the player desires to know as much about the worlds as possible due to an existing interest in planets and survival (Krapp & Prenzel, 2011). This builds upon frequency of science tools used and constitutes deeper, productive engagement by trying to piece together the combinations of unique variables might inform the conditions present on each "what-if" world or exoplanet.

Self-explanations

A scoring scheme based on the rubric provided by the expert was used to assess correctness. See Table 7 for scoring rubric. All scores are at a coarser grain size (Chi, 1994), where the entire answer is scored, as opposed to fragments or clauses. My approach borrows from Clark et al., (2016), in that self-explanations are scored based on abstraction, in a sense, however they are not scored based on level of abstraction in the question but based on level of abstraction and accuracy in the answer.

The scoring rubric was applied to all text responses by students. One researcher trained a second scorer by applying the scoring scheme to self-explanation responses from previous years. The second scorer was trained on a random sample of 40 responses. The scoring was done collaboratively to discuss how the scores apply and when exceptions might occur. Following training, 10% of all responses from the current sample were scored by both scorers (162 responses), resulting in 92% agreement and a reliability using Cohen's Kappa of $\kappa = .928$, indicating high agreement. An additional 10% of responses were then scored and combined with previous responses (325 responses in total), resulting in 91% agreement and an overall Kappa of $\kappa = .906$, maintaining strong agreement. Another 10% of responses were then scored and combined (487 responses in total), obtaining 90% agreement and $\kappa = .906$. A final 10% of responses were scored and combined with previous responses (649 responses in total), and 90% agreement was maintained and a final κ = .913, indicating strong agreement.

Table 7

Summary of SE scoring rubric and examples

Score	Description	Question	Example
"0"	Nonsense, "I don't know", or left blank.	How would you define habitability?	"I think bubblegum tastes better than spearmint"
"1"	Answer is brief (1 or 2 words), answer is topic adjacent but does not directly answer prompt	What would happen to polar ice caps if Earth's tilt was at 90° and why?	"Melt"
"2"	Answers only part of the question, provides some evidence or reasoning.	What would change if Earth had no moon and why?	"Earth would rotate more quickly without a moon"
"3"	Fully answers prompt, provides proper evidence and reasoning, attempts to explain the "why?"	How and why would tides change with two moons and what affect will this have on humans?	"If Earth had two moons tides would be higher because there is more gravitational pull from two moons. Also, people would need to move more inward from coasts.

After establishing reliability of the scoring scheme, the primary researcher scored the remaining responses. Once all responses were scored, Cronbach's Alpha was used to assess reliability of all self-explanation items, revealing α = .868, suggesting strong reliability of assessment items. Self-explanations regarding exoplanets were subsequently dropped from the final analysis. Exoplanet exploration differed substantially from how other worlds were explored, in that learners only had 5 – 10 minutes to explore each world, and the worlds included fewer POIs and NPCs, as compared with the counterfactual worlds. In total, 1,245 scored responses were retained. Missing data constituted another 161 responses (12.9% of retained scored responses), and multiple imputation at the item level was used to impute missing scores (Eekhout et al., 2014). The MICE package in R was used to impute missing data with 5 total imputations and 50 iterations (Buuren & Groothuis-Oudshoorn, 2011). Predictive mean matching (pmm) was used as the imputation method. Imputed values from the first imputation were thus added to the dataset. The final analysis consisted of 1,406 total scores.

The total self-explanation score represents the dimension of cognitive engagement, as it requires active synthesis of what a player learned during exploration of a given world. If a player is talking with NPCs, exploring POI's, using science tools, and making observations then they are incorporating a variety of strategies to provide an answer to the self-explanations that conveys effortful reflection of what was experienced in the learning environment. These responses are mostly open-ended and providing a quality answer deserving of a score of "3" requires effort, building upon strategies implemented during exploration. Self-explanations, however, serve a dual role in my analysis. SEs are treated as a dependent variable and measure of knowledge, but they are also collected from the in-game environment and do require self-regulation and focused engagement to meaningfully complete.

Observation Quality Percentage

Observations are open-ended and made at the discretion of the player, such that quality varies from one observation to the next. A coding scheme has been created to categorize the variety of observations made by players, which places each observation into one of five categories: "off-topic", "factual", "descriptive", "comparative", and "inferential" (Yi, Gadbury, & Lane, 2020). See Table 8 for descriptions and examples of coded observations. All observations were coded using the proposed coding scheme by myself, and a second coder. A random 10% of all observations (162 observations) were used as training, resulting in Cohen's Kappa of $\kappa = .702$, indicating moderate agreement. A further 10% of observations were coded by both coders (320 observations total), and Kappa of $\kappa = .864$ was achieved, suggesting strong agreement. Another 10% of observations (480 observations total) were coded and assessed inter-rater agreement was

$\kappa = .878$. A final 10% of observations were coded by both coders (640 observations total) and final inter-rater agreement was assessed at $\kappa = .883$, strong agreement. In total, the second coder coded 40% of all observations, including the training set. Any disagreements in coding were resolved through discussion, with myself making final decisions if consensus on any individual observation codes was not attained. I coded all additional observations individually, an additional 960 observations.

Table 8
Codes for observations and categorization by quality

Quality	Category	Definition	Example
Low	Off-Topic	Technology-related, conversational.	"I'm stuck in a hole!"
	Factual	Stating objects without any elaboration.	"Pumpkin"
High	Descriptive	Related to color, temperature, quantity, and other physical attributes of the environment, like weight or size.	"The temperature is 465º F"
	Comparative	Comparing one natural phenomenal to another; expectations are violated.	"This Earth has way more volcanoes than colder sun earth"
	Inference	States explanation(s) for why observed phenomena occur	"Solar would be a great source of energy on the desert side of colder sun earth because it always faces the sun"

Observation quality % encapsulates a form of cognitive engagement, as it shows the player is actively trying to compare worlds, derive explanations, or qualitatively describe the environment. A disengaged player may have a high observation frequency, yet most of their observations may be off-topic or merely stating the presence of an object (low quality). Therefore, incorporating quality is important for establishing if observation frequency is a useful measure of engagement. However, many participants made fewer than 10 observations in total, including several who made no observations at all. Observation quality was therefore not included in any predictive models and only accounted for in qualitative exploration of engagement.

5.5 Survey Data

Survey scores were aggregated for STEM interest, MC interest, and Astronomy interest separately. This was done for both pre- and post-survey scores. Demographic information was included in pre-surveys. Less than 10% of data was missing from survey responses, and therefore the MICE package in R was used to fill in missing values at the item level (Buuren & Groothuis-Oudshoorn, 2011; Eekhout et al., 2014). Missing data was generated using 5 imputations, 50 iterations, and the 'pmm' method. Values from the first imputation were added to the final dataset of self-reports. To determine if there is any difference in survey scores between the beginning and end of the camp, a dependent samples t-test was calculated. Typically, differences in interest scores are not seen resulting from short, informal STEM experiences (Lewalter et al., 2021). Given there is a null result comparing interest scores, we can infer that enduring interest is not easily changed and does, to some degree, influence the ways in which a learners engage our *Minecraft* learning environment.

5.6 Interview Data

STEM Interest

Interviews have frequently been used to as one measure to identify interest among participants in a study by asking about habits and attitudes toward specific domains or activities (Renninger et al., 2014; Barron, 2006; Habig & Gupta, 2021; Renninger & Riley, 2012; Ruff, 2016; Tan et al., 2021). One-on-one interviews conducted on the last day of each summer camp asked students several questions regarding interest in STEM. Interviews were segmented by a graduate researcher to identify utterances that reflect an interest in STEM. For example, a student who says they watch YouTube videos about outer space or use search engines to find additional information about Mars reflects voluntary engagement with a domain of interest (Renninger & Hidi, 2016). Semi-structured interviews provide a qualitative form of data that allows further information regarding individuals' interest to be gleaned, where surveys only capture self-reflections at a particular point in time. Also, student attitudes towards surveys may lead them to straight-line or complete surveys inauthentically (Kim et al., 2019). Interviews can more directly capture affect of individuals as they express excitement about their interests and expand upon topics they find particularly interesting. Interview responses were used to provide specific cases and examples of STEM interest behaviors of individuals who fell into specific categories based on their engagement pattern over time.

Astronomy knowledge assessment

One-on-one interviews included ten questions aimed at assessing what participants learned throughout the camp experience. Answers were scored based on correctness against a rubric. The rubric includes broad, surface answers as black bullet points that constitute a score of "2". Answers that cut deeper and include evidence or inferences abstracted from the game environment's "what-if" scenarios garner a score of "3". The questions and rubric were created in consultation with an astrophysicist. If an answer fell outside of possible answers on the rubric, the graduate researcher checked with the expert to see if the answer is feasible. See Appendix E for complete answer key and scoring rubric summary.

The scoring rubric was applied to all answers provided by students. An undergraduate researcher and I separately scored 69% of all knowledge assessment responses. The second scorer was trained on a random set of knowledge assessment responses, totaling 60 responses. Following training, 12% of responses (84 total responses) were scored independently and compared, with % agreement totaling 84.5% and κ = .768. An additional 15% of responses (184 responses total) were scored independently and compared, reaching 80.1% agreement overall and κ = .767, indicating moderate agreement. Another 20% of responses (348 responses total) were scored, resulting in 82.2% agreement overall and a cumulative κ = .768. A final 18% of responses (125 total responses) were independently scored and compared, achieving 89% agreement overall and κ = .867, suggesting strong agreement. All disagreements were resolved through discussion, with myself serving as the final decision maker if consensus could not be reached. I applied the scores to another 171 responses.

In total there were 46 missing responses (7% missing rate overall) because a question was inadvertently skipped by the interviewer or the interview had to end early due to a participant needing to leave. Multiple imputation using the MICE package in R was used to impute missing values at the item level (Buuren & Groothuis-Oudshoorn, 2011; Eekhout et al., 2014). A total of 5 imputations were conducted over 50 iterations using the 'pmm' method. Iteration 1 was thus included in the final dataset. With a full set of scores, an assessment of internal reliability of question items was conducted using Cronbach's Alpha, resulting in α = .823, indicating strong reliability across items.

5.7 Time-series Clustering

As a means of understanding change in systems over time, especially across several dimensions, Hilpert & Marchand (2018) point to the use of non-linear analyses, such as non-linear time-series, to address questions about how elements in a system interact and change over time. Before running whole time-series clustering, four important components must be taken into consideration: representation, distance measures, prototypes, and algorithm (Aghabozorgi et al., 2015). Representation concerns the reduction in dimensionality of multiple time-series, where data adaptive methods are best used for non-uniform lengths, non-data adaptive for equal length time series, and model-based for stochastic processes. Distance measures concern the level of distance and are usually approached at either the shape-level or structure-level, with shape used mostly for short, local time series and structure for longer, global structures, such as annual meteorological data. Finding the representative prototype for each cluster is mostly done through extracting the medoid sequence or identifying the sequence minimizing the sum of squared distances to other clusters. Other options include averaging (best for equal length time-series) or local search, which combines medoid and averaging approaches. Finally, the algorithm must be selected, and algorithms are grouped based on hierarchical, partitioning (includes *k*-means and fuzzy *c*-means), model-based, density-based, grid-based, and multi-step. Choice of algorithm is made based on computational power, complexity of the algorithm, and whether time-series are of equal length or not.

Time-series clustering of educational data is sparse, however one study using time-series clustering to extract patterns from educational trace data, Zhang et al., (2022) recorded undergraduate student behavioral interactions in a learning management system (LMS) aggregated weekly over 16 weeks. Time-series clustering was used to group students based on similar patterns of behavior, with distinct clusters showing significantly higher or lower behavioral frequencies across accessing content, checking grades, in-person interactions, and assessment performance. Analysis consisted of comparing passing rates for students represented by the various clusters, with balanced clusters (clusters showing persistently engaged behaviors) having the highest course pass rate. The authors used unsupervised deep learning models to perform dimension reduction, with the aim of maintaining as much information as possible while removing noise. They then used shape-based similarity measures to group patterns of behavior and a long short-term memory (LSTM) algorithm to find the clusters. The authors also divided the semester by turning points, which were identified by the deep learning model, so that behaviors could be viewed in moving windows. A student who displayed balanced behavior in the first segment may not for the rest of the semester. Determining turning points is an important part of understanding how interest develops over time, and where the points in an experience are that contribute to substantial change.

Participants in WHIMC camps visited all the same worlds in the same order, and therefore time-series are of equal length, however the overall number of time-series at the individual level is still relatively small (*n* = 76), and the time-series itself is short at 6 time points. Cluster analysis was conducted across each data feature (observations, exploration, tools, and nodes) representing the three dimensions of engagement proposed. Given these conditions, time-series was represented through a non-adaptive approach, Piecewise Aggregate Approximation (PAA). Distance for short-time series is best analyzed using Euclidean distance or Dynamic Time Warping (DTW, Aghabozorgi et al., 2015), and in this case DTW was used as it has shown better accuracy than Euclidean distance. DTW is a similarity measure method that takes separate individuals and compares their time series shapes, clustering times series of similar shape. Shape averaging was used to find the prototype for each cluster, as the time series are uniform in length.

The Partitioning Around Medoids (PAM) algorithm was used to determine a medoid that is the result of finding a representative time series with least distance between all the time series. However, the number of clusters must be determined before using PAM. Prior to running any clustering algorithms, all values for participants were normalized. The algorithm was run on a range of clusters from a minimum of two to a maximum of six, determined by the length of the time series. Cluster Validity Indices (CVI) were calculated for the range of time series, providing a set of seven internal model statistics that can be used to determine the optimal number of clusters for the data (Sobolewska, n.d.). See Appendix F for complete statistics for each of the five features analyzed through time series analysis. Some of the statistics should be maximized, such as the Silhouette index, while others (*e.g.* Davies-Bouldin index) should be minimized. To determine the ideal number of clusters, I looked at all statistics and chose the number of clusters (*k*) represented by the majority of maximized or minimized values (Sobolewska, n.d.).

By using time series clustering, engagement patterns of individuals across all worlds can be identified and these patterns can be used to categorize individuals into emergent forms of engagement. For example, a cluster of students might display a consistent upward trajectory in amount of map explored, and this engagement pattern, along with time series clusters on other variables can be used to determine an engagement profile for each student. The number of students per identified pattern can then be analyzed based on knowledge score, STEM interest, Astronomy interest, and MC interest, such as what percentile of interest do the students in a particular pattern lie. Features of the pattern can then be extracted to draw inferences about the way interest motivates behavior in an open-ended, digital STEM-focused environment. Cluster membership can also be used as categorical variable in Bayesian models. Theoretically grounding engagement informs what dimensions of engagement are most salient regarding interest-related behaviors (*i.e.* is a highly interested student engaging more cognitively, and at what intervals?). Additionally, transitions can be analyzed to see if students become disengaged at critical points in the camp experience. Transitions worth considering are worlds explored by day, where the metrics for the first world explored each day may be higher than subsequent worlds that day, or simply world-by-world. Looking at engagement on individual worlds provides insight into the design of the experience and if there are specific worlds that elicit higher engagement than others. Engagement is compared across days to see if persistent individual behaviors exist from day-to-day.

Analysis was conducted in R using the “dtwclust” package. The “dtwclust” package was specially designed for implementation of the dynamic time warping algorithm, but the package is flexible and can provide comparisons between various representations, distance measures, prototypes, and algorithms of time-series clustering (Sardá-Espinosa, n.d.).

5.8 Bayesian Model Averaging

Given the small sample size of the data, a Bayesian analysis is a more robust approach for analyzing the data than using frequentist statistics (König & van de Schoot, 2018; van de Schoot & Miočević, 2020). A non-informative prior is used in this approach, however the posterior probability distribution estimated from this research can be used as a prior distribution for continued interest research using *Minecraft*. We can establish a stronger connection between the predictors and dependent variables by estimating the probability of contribution to every possible model, 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 (Hinne & Gronau, 2020). The prior distribution, or our belief that the dependent variable is caused by any of the predictors, is updated with every piece of new information. Computationally, BMA is represented as follows:

$$p(t)=\sum_i p(t|H_i)\, p(H_i)$$

In our case t is the probability of either interest or knowledge, depending on the model, that is being estimated by all prediction models simultaneously. The prior model possibility is represented by *p*(H_i) before any data has been observed. The posterior *p(t)* updates based on successes or failures in the data to accurately predict the dependent variable. Variables with greater influence on the DV will have greater frequency of presence in models, and confidence intervals can be examined to determine if variables should be included in a final model. The final model will provide an indication of what factors, including clustered behaviors, best represent an individual's interest and knowledge in STEM broadly, *Minecraft*, and Astronomy. All variables were centered to avoid multicollinearity. If highly correlated variables represent similar constructs, one of the variables will be dropped. Model fit statistics were examined after running BMA. All analyses were conducted in R using the "BMA" package.

5.9 Descriptive Statistics

Following the creation of features, descriptive statistics were calculated for all continuous variables. Given the non-normal structure of several variables and the relatively small sample size, the conservative Spearman Rank Correlation was used for the continuous variables. Descriptive statistics were run following the creation of engagement patterns from time series clustering, which includes the number of learners above interest self-report means and knowledge scores by engagement pattern. Categorization of learners across qualitative variables by their respective engagement patterns were also determined, which includes the number of learners who identify as a "science person", the number of learners who stated their favorite school subject is a STEM subject, and the percentage of learners who exhibited high quality observations (*i.e.* comparative, inferential, or descriptive). Given the exploratory nature of this analysis, descriptive statistics are used extensively to uncover relationships between variables and suggest further avenues for research.

CHAPTER 6: RESULTS

The first RQ sought to understand engagement and how it can be understood and represented through log data from the *Minecraft* summer camp experience. The data analysis section laid out the features to be calculated and theoretically grounded them to the dimensions of engagement. For predictive models, the aggregates or sums of the features and other variables are used as predictors, and are the data used in the descriptive statistics provided. See Table 9 for means, correlations, and standard deviations of log features, interest scores, and knowledge scores. Spearman correlations were used to conservatively calculate correlations among all data since most of the log data proved to be non-normally distributed.

Table 9

Spearman Correlations for Interest, Knowledge, and Engagement

	STEM Int	MC Int	ASTR Int	KQs	Exp	SEs	Obs	Utools	Tools
STEM Int	--								
MC Int	.51*	--							
ASTR Int	.47*	.30*	--						
KQs	.24*	.25*	.35*	--					
Exp	-.14	0	.12	.27*	--				
SEs	.18	.28*	.25*	.3*	.21	--			
Obs	.14	.13	.09	.44*	.25*	.54*	--		
Utools	.01	0	.02	.4*	.38*	.31*	.5*	--	
Tools	-.05	-.02	0	.41*	.41*	.25*	.49*	.96*	--
Nodes	.08	.04	.07	.4*	.5*	.3*	.51*	.7*	.65*
mean	58.42	70.96	16.03	18.32	184.24	30.55	21.87	13.34	15.39
sd	13.52	14.26	3.67	4.31	34.99	9.12	18.97	14.18	18.5

Note: STEM Int = STEM interest, MC Int = Minecraft Interest, ASTR Int = Astronomy Interest, KQs = Knowledge Question Scores, Exp = Exploration Breadth, SEs = self-explanation scores, Obs = Observation Frequency, Utools = Unique Tool Use Frequency, and Tools = Tool Use Frequency, * = $p < .05$

1.1a) What log features best reveal affective, cognitive, and behavioral engagement?

In the methods section, eight features were proposed to be included in the analysis, however I determined that Action Latency too close a resembled outcomes from frequencies of observations and science tools and ultimately did not provide any additional information. Also, Unique Science Tools exhibited a near ceiling correlation with Science Tool Frequency (see Table 9), as well as near identical correlations between Tool Use Frequency and other log features. Unique Science Tool Use was thus determined to be a redundant feature. Observation quality is considered in combination with qualitative data, as many learners did not make observations or there were too few (*i.e.* < 5) to provide a meaningful measure of quality. Self-explanations scores are also included, representing cognitive engagement, however they are used as a dependent variable and separate from log data due to the need to manually score responses. Therefore, the log data retained representative of engagement across the camp experience is Exploration Breadth (*affective)*, Nodes (*affective/behavioral*), Observation Frequency (*behavioral*), and Science Tool Use Frequency (*affective/cognitive*).

All derived log features positively and significantly correlated with each other, indicating that individuals who explored more, both broadly (Exp) and in a directed fashion (Nodes) also took more science tool measurements and made more observations. These features significantly and positively correlated with knowledge assessments (*i.e.* KQs and SEs), apart from exploration and aggregated SE score. None of the log features significantly correlated with any of the measures of interest, and in turn no determination can be made regarding the relationship between interest and engagement from correlations. The correlations represent overall frequencies or aggregate scores, and do not pertain to engagement patterns and how they change or emerge over time. Time series clustering is therefore implemented to see if the derived features evince persistence patterns over time.

1.1b) How can combinations of features effectively capture engagement in this context across time?

Time series clustering was used to determine engagement across four features derived from the log data. The optimal number of clusters for each feature were determined by CVI statistics across the range of possible clusters determined by the number of worlds analyzed, $k = 2:6$. The final number of clusters ("k") for each feature, as determined by the provided CVI, are as follows: Exploration ($k = 2$), Observations ($k = 2$), Tool Use ($k = 2$), and Nodes ($k = 2$). See *Figures 8 - 11* for the visual representations of clusters, medoids, and codes assigned to each cluster. Worlds are ordered chronologically from the first world explored in each summer camp to the last world explored, such that 1 = "Lunar Crater", 2 = "No Moon", 3 = "Colder Sun", 4 = "Tilted Earth", 5 = "Earth as a Moon", and 6 = "Two Moons". The codes assigned to each cluster were based on emergent features of the clusters, primarily dictated by how the medoid line moved relative to the grand mean.

Borrowing from the work of Zhang et al., (2022) who labeled time series clusters based on emergent patterns, High (H) performers are those included in clusters whose medoid remained around .5 standard deviations or higher above the mean for most of the worlds, and Low (L) performers are those whose medoid remained close to .5 SD or below for most of the worlds. Time series analysis resulted in two clusters for each of the features of interest, mostly dividing individuals into categories of either high engagement or low engagement, except for observations. In terms of observations, there appears to be an upward trajectory of the medoid that suggests increasing use over time. However, overall observation use is still consistently higher than the "L" group and could be considered another high engagement cluster. For this analysis observations were considered as “H” for the purposes of coding and classification. The upward trend in observations could speak to increasing familiarity with and comfort using the feature. At the very least, these clusters do capture a dichotomy of engagement across features characteristic of how learners can engage with content on the *Minecraft* server.

Based on the coding of the clusters, engagement profiles can be created by aggregating codes into resulting groups. In total, five codes emerged from the data: “Balanced High” (BH), “Balanced Low” “BL”, “Balanced” (B), “Explorers” (EXP), and “Point of Interest Explorers” (POI). Table 6 lists the five emergent codes and the number of participants categorized into each engagement pattern. BH learners are those who were categorized into “H” on a majority of clusters (> 2), BL learners are those who were categorized into “L” on a majority of clusters (> 2), B learners are those who had an even mix of “H” and “L” categorizations on clusters, and looking at the combinations most B learners ($n = 9$) clustered high in exploration and nodes, second most ($n = 6$) clustered high in nodes and observations, and only a single person clustered high on nodes and tools. EXP learners are those who clustered into “H” on exploration but “L” on the remaining features, and POI learners are those who clustered into “H” on nodes but “L” on the remaining features. EXP and POI are considered separately because EXP captures the breadth of exploration but does not indicate interactions with NPCs or points of interest, whereas POI explorers showed targeted exploration but not necessarily breadth. Quality of exploration in *Minecraft* environments has been shown to be a significant predictor of memory and learning (Clemenson et al., 2019), and therefore a distinction between exploration types in our *Minecraft* environment reasserts this notion.

Figure 8

Exploration Clusters and Emergent Patterns

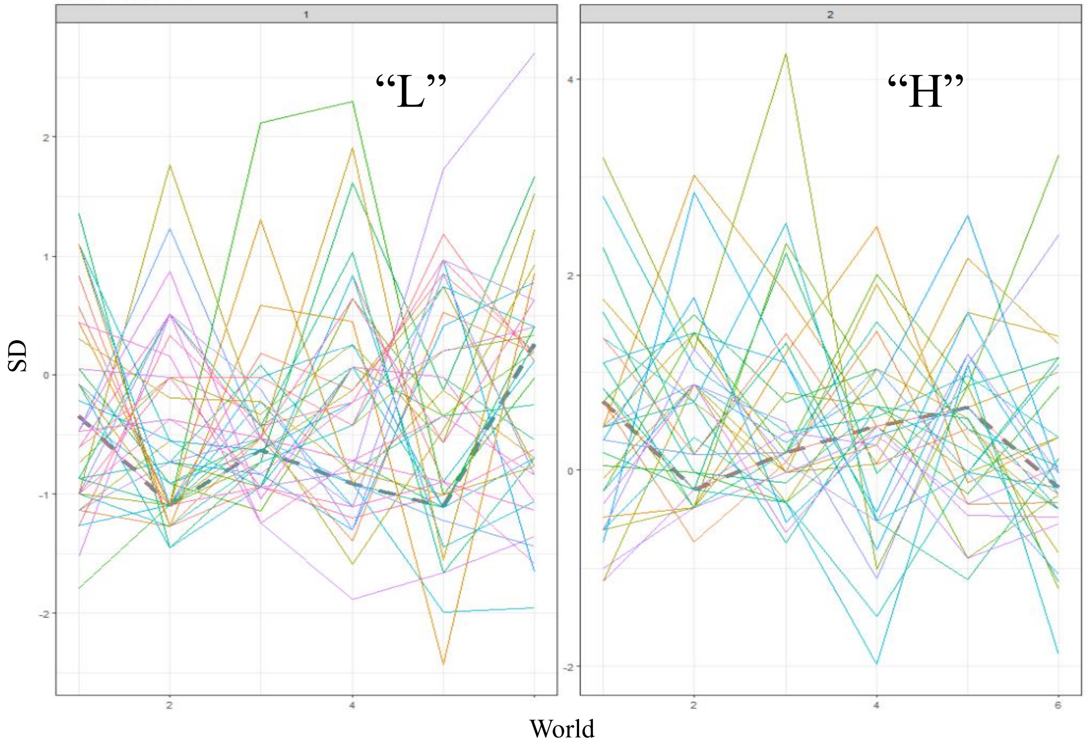


Figure 9

Observation Clusters and Emergent Patterns

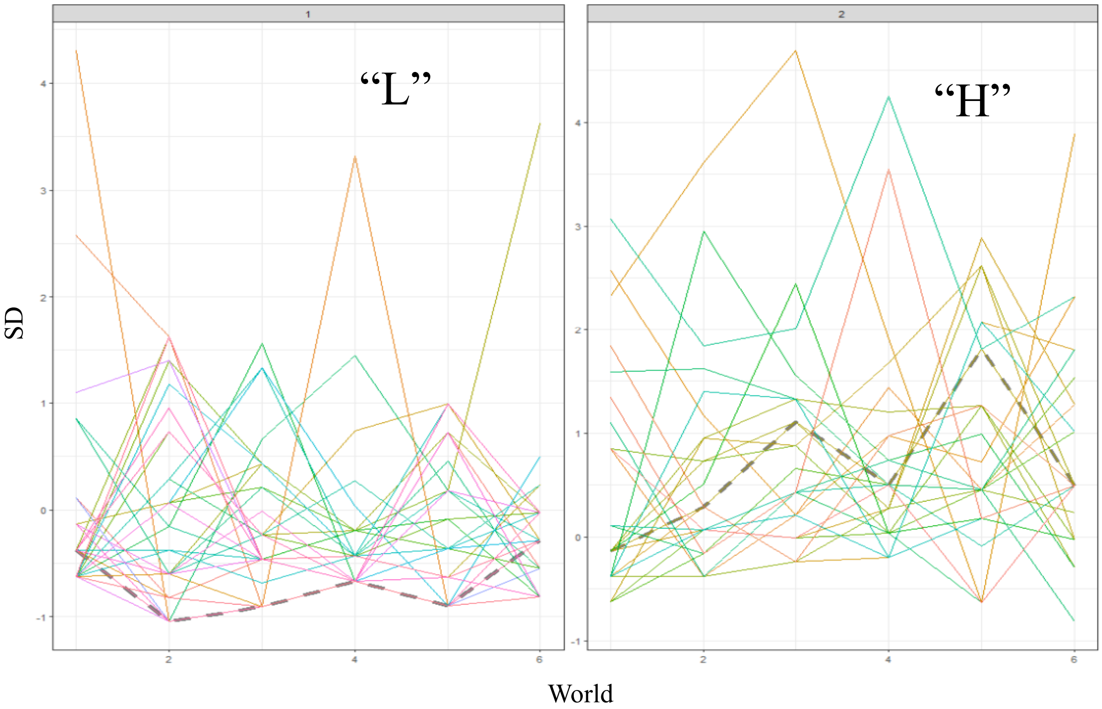


Figure 10

Science Tool Use Clusters and Emergent Patterns

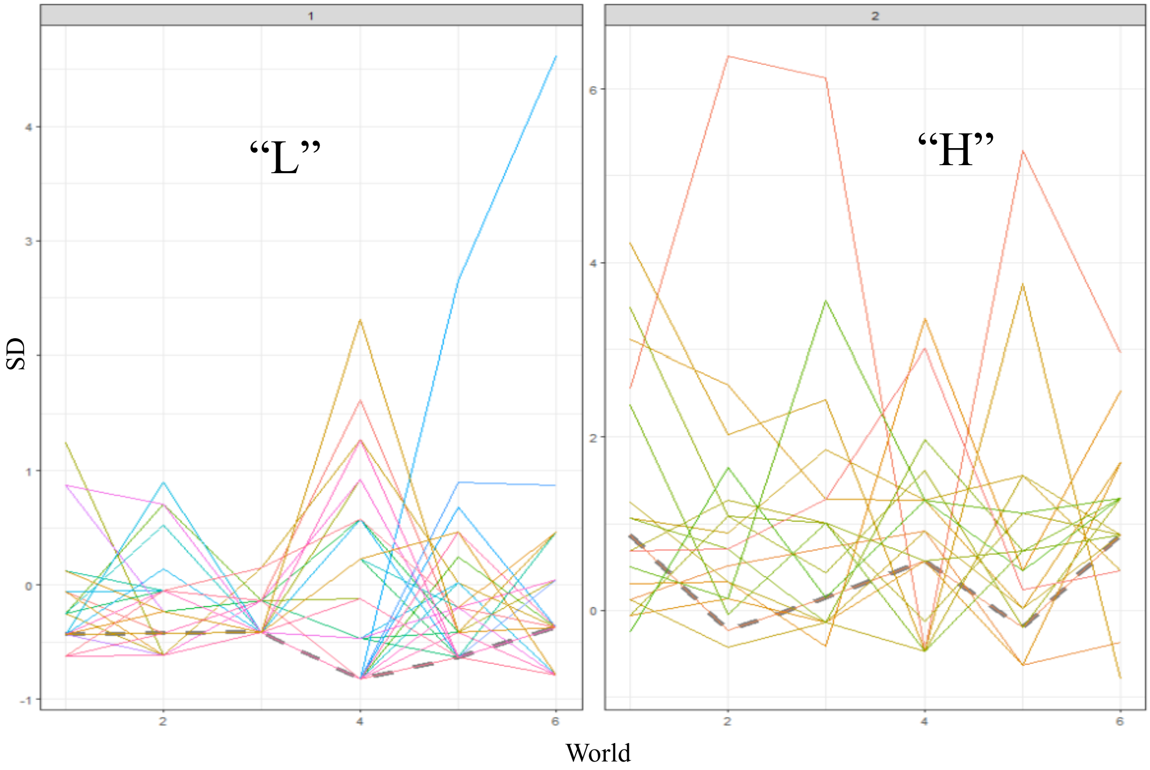
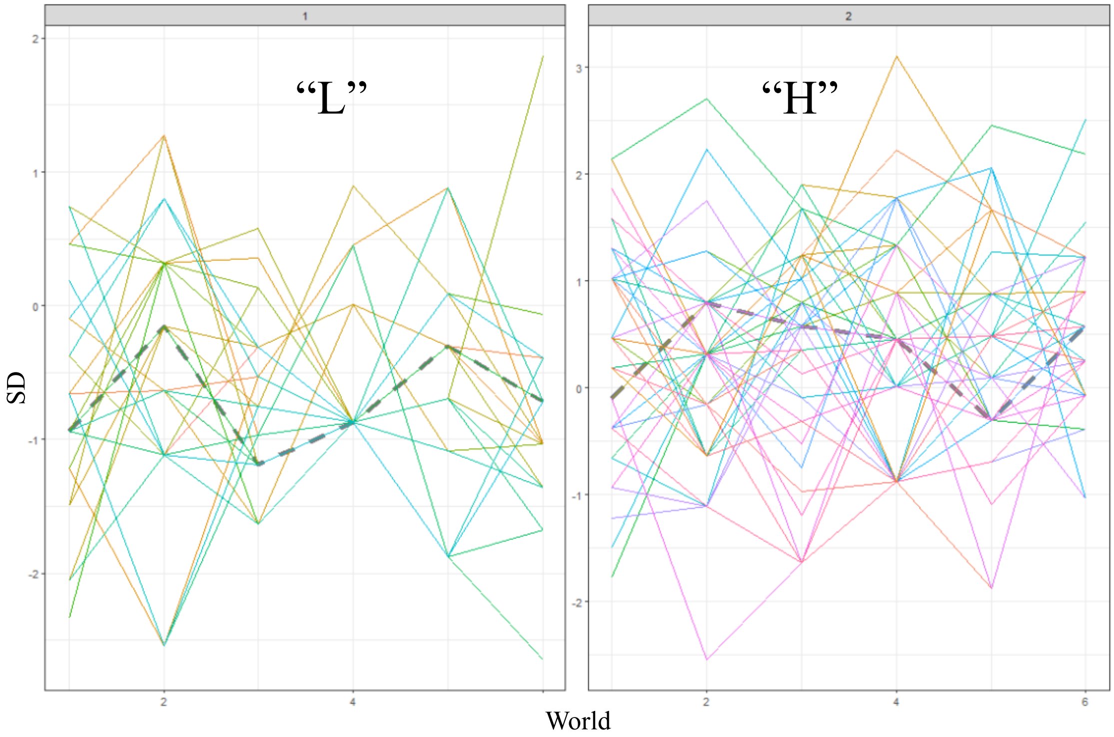


Figure 11

Nodes Clusters and Emergent Patterns



Most learners were categorized as BH engagers on the server, see Table 10, but only slightly more than those categorized as BL engagers. To determine whether the combination of features across clusters captured engagement, I looked at how many learners in each category fell above mean scores across the interest measures (STEM, *Minecraft*, and Astronomy self-reports at the start of camp) and knowledge assessments (SE's and knowledge questions from interviews). Since the focus of the camp is on Astronomy, I also included the Astronomy pre- and post-assessment categorization, as well as how many learners increased by at least one point in Astronomy interest. For the STEM interest measure, there does not appear to be any group highly associated with above average mean scores. Based on this descriptive analysis, STEM interest does not appear to be associated with higher engagement on our *Minecraft* server. *Minecraft* interest is most associated with BL learners, and this might indicate higher *Minecraft* serves as a distraction from task goals and learning objectives. BH learners do mostly sit above the mean for Astronomy interest at both pre- and post-assessment, and this could indicate that learners interested in Astronomy do find opportunities on the server to act upon their existing Astronomy interest. The major differences appear in the knowledge area, where BH learners overwhelmingly dominate in terms of scores above the means (SEs and KQs).

Table 10

Learners Above Mean in Interest and Knowledge & Changes in Astronomy Interest

Pattern (*n*)	STEM Int	MC Int	AstI Pre	AstI Post	AsInt Incr	KQs	SEs
B(16)	50%(8)	56.2%(9)	25%(4)	37.5%(6)	50%(8)	50%(8)	31.2%(5)
BH(22)	45.5%(10)	45.5%(10)	40.9%(9)	54.5%(12)	77%(17)	68.2%(15)	86.4%(19)
BL(20)	50%(10)	65%(13)	35%(7)	50%(10)	55%(11)	30%(6)	30%(6)
EXP(8)	75%(6)	50%(4)	62.5%(5)	50%(4)	20%(2)	37.5%(3)	50%(4)
POI(10)	70%(7)	30%(3)	50%(5)	50%(5)	50%(5)	50%(5)	40%(4)

Note: STEM Int = STEM Interest, MC Int = Minecraft Interest, ASTR Int = Astronomy Interest, KQs = knowledge questions score SEs = self-explanation score, B = Balanced Engagement, BH = Balanced High, BL = Balanced Low, EXP = High Exploration, POI = High Nodes

The combination of features does point to a relationship between engagement and knowledge, where BH learners perform better overall on the knowledge assessments throughout the summer camp and at the end. A clear line can be drawn between the BH learners, B learners, and the BL learners in this case, where only 30% of the BL learners and 31% of the B learners fell above the mean for self-explanations. The B and BL learners have different paths for arriving at similar knowledge outcomes, although B learners have a slight advantage in terms of KQs. The EXP and POI learners also have relatively low numbers of learners above the means for knowledge assessments, although a small advantage on KQs for POI learners exists. Since all the B learners fell into "H" clusters for nodes, they can be considered an extension of POI learners, as they incorporated one additional engagement feature. However, this additional engagement did not translate into higher knowledge scores or increases in interest. What can be concluded from this data is that the most effective pattern of engagement is that of high persistent engagement across all the features analyzed. A relationship between engagement and knowledge can be inferred, however directionality cannot be discerned form this data; it is possible BH learners enter the experience with higher knowledge, which in turn is reflected in scores on knowledge assessments. The combination of features does provide some insight into engagement and how it can be realized through various assessments used throughout the summer camp experience, but I remain cautious in how exploratory data is interpreted.

The second set of research questions digs deeper into how interest can be understood through engagement, possibly tying interest to aspects of engagement. Adequately analyzing interest requires triangulation of data sources (Renninger & Hidi, 2011; Krapp & Prenzel, 2011), and interviews, self-reports, log files, and conceptual understanding comprise the data used in this section.

1.2a) How can content-related v. context-related interest be effectively distinguished through logs?

A reasonable assumption from literature on game-based learning (Novak & Tassell, 2015) is that those most interested or experienced in the context, *Minecraft* in this case, should be able to immediately interact with advanced features, such as the *Minecraft* command bar, a facet of the game that goes beyond surface level engagement. Time spent in the orientation world should mitigate any gaps, though. Referring to Table 10, most learners in the BL group were above the mean for *Minecraft*-related STEM interest, evoking the notion that higher *Minecraft* interest or experience does not necessarily associate with high overall engagement. One might infer from this that higher *Minecraft* interest is a hindrance to the forms of engagement that align with task goals in this case. However, Table 9 does show a small, yet significant, relationship between *Minecraft* interest and performance on self-explanations. There are also small to moderate relationships between STEM, *Minecraft*, and Astronomy interest, based on the self-report survey. However, none of the engagement features of interest significantly correlated with *Minecraft* interest, except for self-explanations. Self-explanations are delivered following exploration of each world, and this could be a signal that interest in and knowledge of *Minecraft* do contribute to players' ability to navigate the environment and determine where relevant content is located. Counter to this, though, is the fact that *Minecraft* interest is not significantly correlated with interview knowledge questions at the end of camp. Additionally, all engagement patterns aside from BH seem to represent alternative patterns to reaching a similar end goal, being relatively lower scores on KQs and SEs.

Astronomy interest did significantly correlate with self-explanation performance and total number of observations, albeit the correlation is small. In terms of observation quality, see Table

11, 81% of learners in the BH engagement category had an observation quality measure over 50% (*n* = 18) compared to 45% (*n* = 9) of learners in the BL category. The BH group had only one learner with fewer than 5 observations, while the BL group had 9 learners who made five or fewer observations. Learners in the B group also showed high observation quality, as compared to other groups, yet this did not appear to have a tremendous influence on knowledge scores or levels of interest.

Table 11

Qualitative Displays of Interest in STEM

Pattern (*n*)	SPId	STEMSub	Obs Qual
B(16)	44%(7)	50%(8)	75%(12)
BH(22)	73%(16)	68%(15)	81%(18)
BL(20)	45%(9)	40%(8)	45%(9)
EXP(8)	13%(1)	50%(4)	25%(2)
POI(10)	50%(5)	50%(5)	60%(6)

SPId = identified as science person, STEMSub = named STEM subject as favorite, Obs Qual = Observation Quality, B = Balanced, BH = Balanced High, BL = Balanced Low, EXP = Exploration only, POI = High Nodes

Also, a small majority of learners in the BH group did fall above the mean for Astronomy interest but less than 50% fell above the mean for *Minecraft* interest. A small majority of learners in the BL category did fall above the mean for *Minecraft* interest, but fewer than 50% fell above the mean for Astronomy interest. Looking at Astronomy interest increases in Table 10, 77% of BH learners increased in Astronomy interest by at least 1 point, while 55% of BL learners did, relatively consistent with the B and POI groups. The EXP group had only one individual who increased in Astronomy interest. I cautiously suggest that content knowledge is more closely associated with productive engagement patterns on the server, and context interest less so, and in turn productive engagement patterns associate more closely with increases in content interest. General STEM interest poses no relationship with any of the engagement variables and is likely too broad of a measure to provide any meaningful data regarding engagement. In the end, knowledge positively and significantly correlates with all engagement features from log data, however neither context nor content interest do. Specific content interest may be an accurate predictor of how a learner will engage with *Minecraft* in an informal, STEM-based learning environment, however broad STEM interest and context knowledge do not seem to bear a strong relationship with engagement, as evidenced through log data.

1.2b) How can states of interest, as well as changes in interest, be captured through triangulation of log files, interviews, self-reports, and knowledge assessments?

Dependent samples t-tests were used to check for significant mean differences between pre- and post-interest assessments. Prior to running any t-tests, the Shapiro Wilk test of normality was applied to all interest assessments, revealing no violation of the normality assumption for any of the self-reports. No significant difference was found in STEM interest between pre- ($M = 58.42$, $SD = 13.53$) and post-assessment ($M = 59.03$, $SD = 11.86$), $t(75) = 0.456$, $p = 0.65$. Assessment of *Minecraft* interest between pre- ($M = 70.96$, $SD = 14.26$) and post- ($M = 68.45$, $SD = 14.12$), also revealed a non-significant change in interest scores, $t(75) = -1.7519$, $p = 0.08$. Astronomy interest assessed at pre- ($M = 16.04$, $SD = 3.67$) and post- ($M = 16.71$, $SD = 4.03$) did reveal a significant change between the beginning and end of the summer camp experience $t(75) = 2.12$, $p < .05$. Cohen's d was used to measure effect size, $d = 0.24$, showing a small effect. The summer camp experience did evoke a minor change in situational interest in Astronomy (the focus of the camp content), but none of the other measures of interest.

As evidenced in Table 10, most BH learners exhibited an increase in Astronomy interest, pointing towards a relationship between engagement patterns and changes in states of interest. This might indicate a triggering of situational interest, and this is supported by interest theory

(Hidi & Renninger, 2006). The fact that B, POI, and BL learners showed 50% to 55% increases in Astronomy interest between the beginning and end of the camp experience is a sign that the experience triggered situational interest in Astronomy for some of these learners. The EXP group only had two learners increase in Astronomy interest, which suggests a unidirectional focus on exploration depth does not translate to increasing knowledge and interest in content. More than half the participants in the camps showed increases in Astronomy interest ($n = 43$), and with an overall significant increase in Astronomy interest, general triggering of situational interest for most learners can be concluded.

Looking at Figures 8 – 11, there are specific time points (worlds) where steep increases in engagement do occur. For example, both "L" and "H" learners decline in exploration breadth from time 1 (T1) to T2, suggesting that the first world may induce a novelty effect (Rodriguez-Aflecht et al., 2018; Hunsu et al, 2017). The "L" exploration group does show a steep increase at T6, but this contrasts sharply with the "H" group. Based on Table 9 high exploration does not correlate with content interest. In terms of observations, which exhibited a significant relationship with Astronomy interest, the "H" group increased their number of observations from T1 to T3, showed a dip, then increased again from T4 to T5, before dipping at the end. The "L" observation group dipped from T1 to T2 and did not reach similar levels as T1 until the final world. The "H" groups' trajectory suggests becoming accustomed to making observations and this behavior persisting throughout the camp experience, whereas the "L" only made efforts at the beginning and end of the camp experience. Future studies should closely examine the features of the worlds and their specific triggering of interest.

Given the relationships between the BH learners, observation quality, and increases in astronomy interest, an indication of explaining phenomenon in a meaningful way might be associated with changes in interest, however this indication should only be taken lightly. Examining nodes, the “L” and “H” learners both increased in the POIs and NPCs visited from T1 to T2, but the “L” learners decreased while the “H” learners remained relatively constant from T2:T4. From T4:T5 the “L” learners increased their targeted exploration while the “H” learners decreased, but again switched from T5 to T6. The POI group (Nodes-focused engagement learners) only showed 50% of learners in the category increased interest and only 50% fell above the mean for interest, therefore no conclusions can be drawn from a focus on interactions with POIs and NPCs only, whereas as part of a pattern of overall high engagement, changes in content interest do appear to occur. The “L” learners show a relatively flat time series of tool use, whereas the “H” learners oscillate from world to world. The flow of the camp introduces and highly encourages use of science tools in the first world, and then allows learners to use them at their discretion for the rest of the worlds. New tools are introduced for each world, and it is possible the features of specific worlds, trigger interest in the use of certain measurement tools. Again, though, a consistent pattern of higher engagement, relative to the “L” groups is more indicative of changes in content interest.

Interest and identity do seem to be closely related in results, and this is also upheld by theory (Renninger, 2009). Students who express “science person identity” (SPI) comprise a majority of the BH group, whereas only one member of the EXP group saw themselves as a science person, see Table 11. Learners in the B & BL category were split almost evenly on SPI, and the POI group did evenly express SPI. Similarly, most of the BH learners mentioned a STEM class as their favorite class in school, and only half or few did so for the rest of the engagement pattern groups. Except for observation quality, the B and BL groups had similar numbers of learners increase in astronomy interest, express SPI, and mention a STEM subject as their favorite, and similar percentages were found for POI learners. These three categories show that similar results can be attained in different ways, and only learners that engage highly across most of the targeted features either benefit most or are predisposed to do so by their knowledge of or interest in the content. The EXP group showed the lowest number of learners who increased in STEM interest (Table 10), and knowledge scores for this group were also low compared to other groups. A breadth of exploration without interaction and availing oneself of the tools at their disposal appears to have no influence on interest or knowledge, possibly suggesting that effort and productive engagement are indeed needed to increase interest.

1.2c) What variables comprise the best predictive models of interest and knowledge?

A series of Bayesian multiple regressions were conducted using uniform priors and Bayesian Model Averaging (BMA) to find the set of predictors that best predict two knowledge dependent variables (KQs & SEs) and three dependent variables of interest (STEM interest, *Minecraft* interest, and post-Astronomy interest). The final selection of models can be seen in Table 11, and graphs of predictors averages across models can be seen in *Figures 7 – 11*. Full BMA statistical outputs for each dependent variable can be found in Appendix G, which includes probability of each predictor being included in the best model, and the posterior probabilities of each of the five best models. Posterior probability refers to the probability of the model as being the correct model. A combination of low BIC and high posterior probability indicates the most appropriate model unless there are theoretical considerations to choose an alternative model.

The first three models all look at post-interest scores and what predictors adequately inform us of where a learner's interest falls. The first model in Table 12, also see Figure 12, regresses STEM interest on the engagement variables (observations, tools, exploration, nodes), knowledge scores, engagement categories, and camp location. There are no predictors with a high enough probability to predict scores on the STEM interest assessment, reinforcing previous results of broad STEM interest having little to no impact on how learners perform in the camp experience or on knowledge assessments. Still, trends can be observed regarding what predictors dominate most models and their directionality. For example, Camp UG does seem to exhibit a positive relationship with STEM interest, as compared to the baseline camp, whereas no other camps do. At the same time, performance on knowledge assessments relates positively to STEM interest. Any trends in models should be used cautiously but can reinforce theory and provide directions for future analysis and research. The second model, *Minecraft* interest (Figure 13) regressed on the same predictors as STEM interest, shows that SE scores have a high probability of predicting where a learner's *Minecraft* interest might stand at the end of the camp experience. The amount of variance explained by the model (R^2 = 9.2%) is relatively low, but the posterior probability, which tells us the probability of the event taking place after accounting for new information, is in line with or relatively high compared to others (θ = .256). Based on the *Minecraft* interest model we can expect to see small increases in interest the better learners score on the self-explanations, and the probability the predictor is meaningful is quite high at 81.9%. Learners may be developing greater familiarity with the *Minecraft* environment as they play through the simulated counterfactual scenarios, or it is possible having greater *Minecraft* interest reduces some barriers to engagement that allow for deeper processing of content available.

Lastly, in terms of interest, Astronomy interest (Figure 14) is strongly predicted by performance score on summative knowledge questions. The model variance explained (R^2 = 11.5%) is slightly higher than *Minecraft* interest, but the posterior probability is slightly lower (θ = .23). The only meaningful predictor to result from the Astronomy interest model, *KQs*, showed a very high probability of inclusion (93.9%) and a small increase in Astronomy interest for each point increase. Once again, the relationship between interest and knowledge is on display, with learners who provide well-reasoned answers to knowledge questions being more likely to develop interest. Looking at contributions to knowledge provides a different story as to most influential factors.

Table 12

Bayesian Model Averaging Statistical Outputs

DV/Predictors	R^2/p!=0	BIC	θ	β
STEM Interest	0	0	0.255	--
Minecraft Interest	0.092	-2.997	0.256	--
SEs	81.9	--	--	0.4694
Astronomy Interest	0.115	-4.927	0.23	--
KQs	93.9	--	--	0.3165
Knowledge Questions	0.337	-18.298	0.113	--
Observations	55.9	--	--	0.069
Camp - UG	97.9	--	--	-3.395
Camp - UB	78.6	--	--	-4.456
Self-Explanations	0.447	-27.724	0.267	--
Engagement - BH	100	--	--	10.275
Camp - SDP	100	--	--	-9.312
Camp - UG	96.7	--	--	-7.809
Camp - UB	100	--	--	-9.226

Note: p!=0 denotes the probability of the including the predictor in the model. θ = the posterior probability of the model.

Knowledge question (*KQ* model) score, see Figure 15*,* was regressed on the engagement variables and background, resulting in a model with three predictors and a higher explanation of variance than any of the interest models (R^2 = 33.7%) but the lowest posterior probability (θ = .113). The *KQ* model had three predictors with probabilities above chance, including observations, Camp UG, and Camp UB. Camp UB and Camp UG included learners who predominantly qualified for free and reduced lunch at school, compared to the baseline group, and they showed large negative impacts for site on knowledge questions. The probability of inclusion in the final model is higher for UG (97.9%) than for UB (78.6%), and both are higher than the probability of inclusion for Observations (55.9%). The impact of making more open-ended observations is small for increasing scores, however it provides an opportunity to increase scores that can elevate students from underrepresented backgrounds to engage content and improve scores.

The other knowledge assessment model (*SE* model, Figure 16), regularly administered self-explanations, resulted in a strong explanation of variance ($R^2 = 11.5\%$), and a strong posterior probability ($\theta = .267$). Four variables emerged as highly probable for predicting SE scores: inclusion in BH engagement group (100%), participating in camp SDP (100%), participating in camp UG (96.7%), and participating in camp UB (100%). As with the *KQ* model, the *SE* model included two camps that both include mostly learners with free and reduced lunch (camp UG and camp UB), and they both have strong negative impacts on SE score, compared with the baseline group. The camp SDP group was included this time, too, but that site did not include free and reduced lunch learners, yet they still exhibited a large negative effect of site on overall SE score compared to the baseline group. Regardless of site, though, engagement patterns reflecting consistently high engagement across strongly predict higher scores on the self-explanations, than learners who were balanced (baseline group). SEs are considered a form of cognitive engagement, as they do require open-ended written responses in the game, and it is possible learners less engaged do not complete them in good faith or attempt to skip them altogether. Also, as has been evidenced by data presented so far, the relationship between knowledge, engagement, and interest remains strong. Directionality may not be inferred, but in accordance with theory there is an undeniable self-boosting relationship between the three components that propels learners to greater heights.

Figures 12 – 16 show combinations of predictors for top models averaged using BMA. Blue indicates positive predictors and red indicates negative predictors. The more horizontal space a model occupies, the higher the probability of that being the correct model is.

Figure 12

Bayesian Model Averaging with STEM Interest as DV

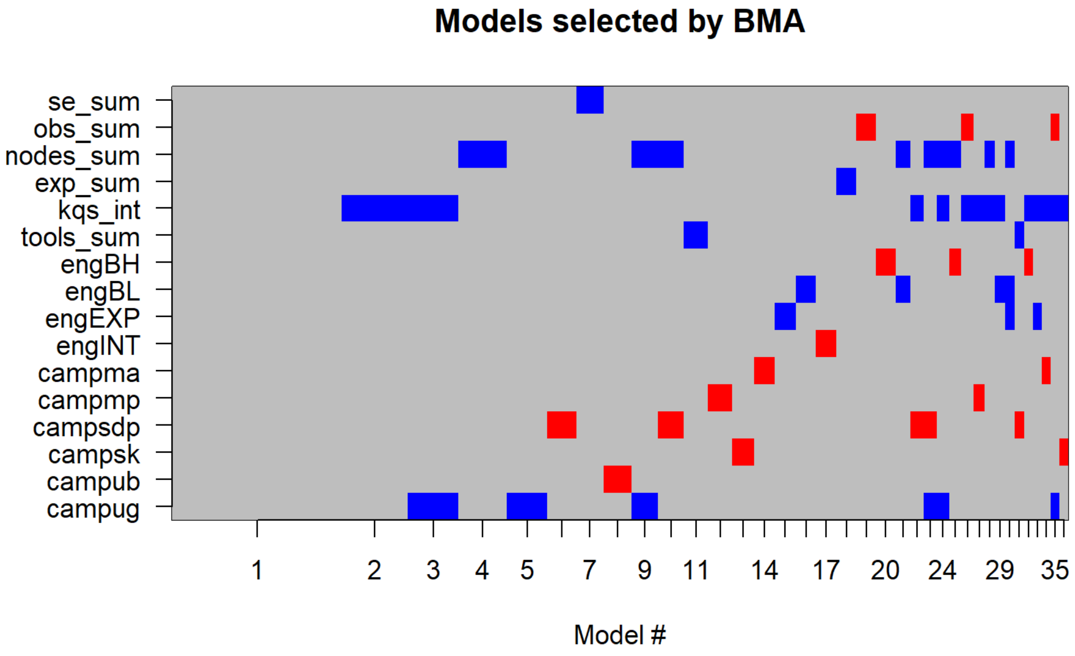


Figure 13

Bayesian Model Averaging with Minecraft Interest as DV

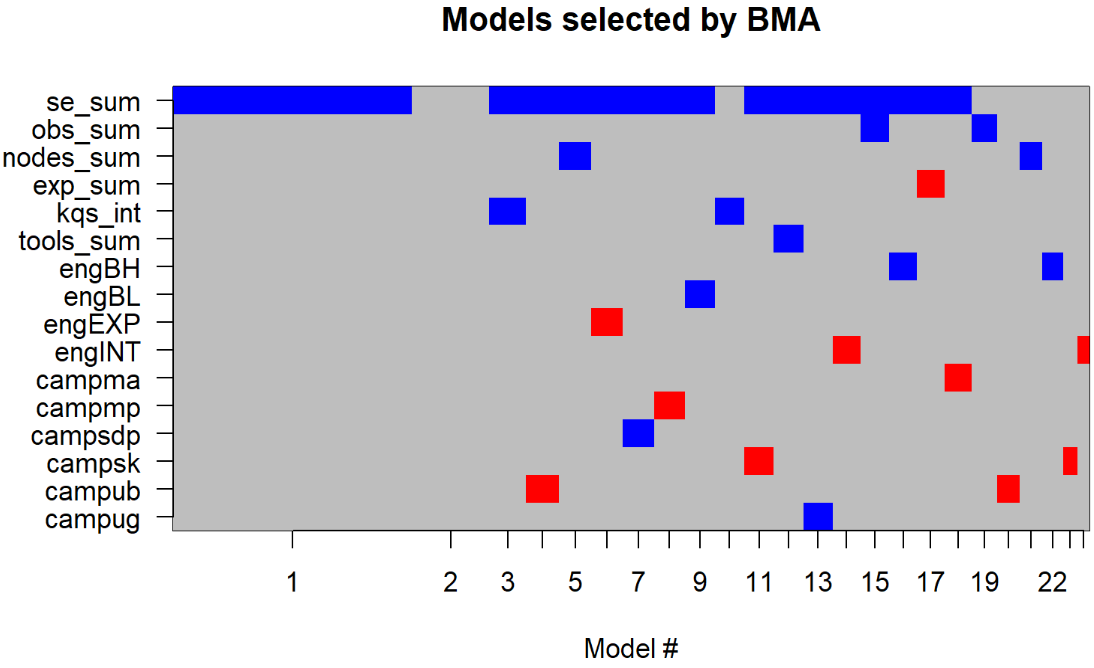


Figure 14

Bayesian Model Averaging with Astronomy Interest as DV

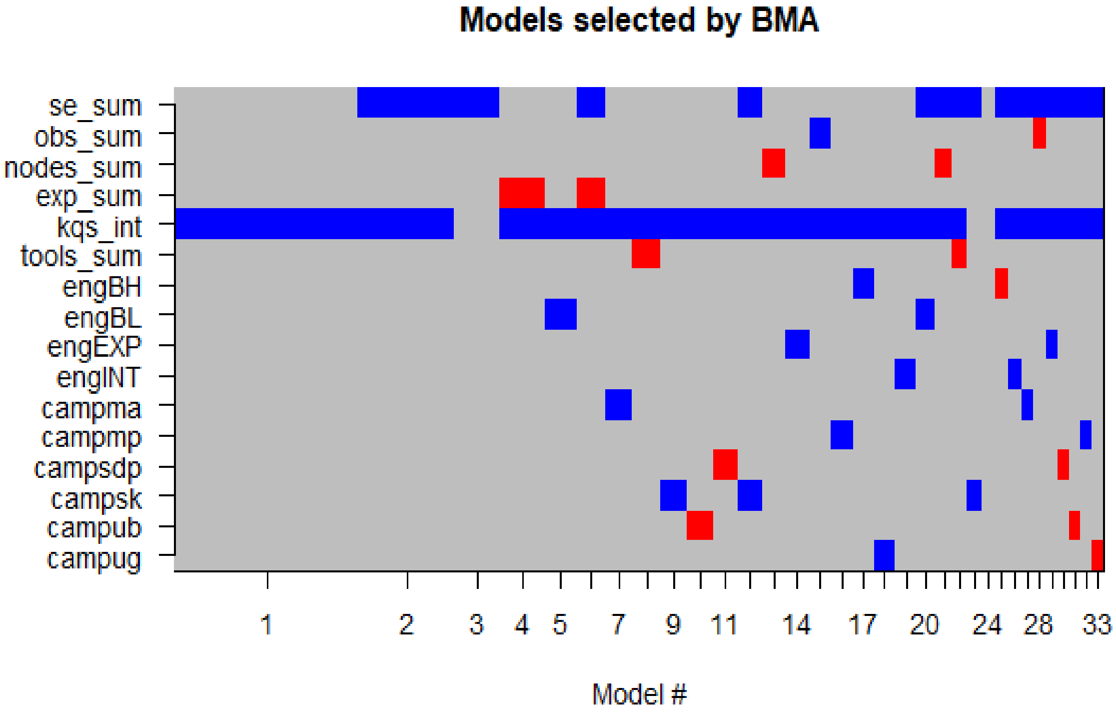


Figure 15

Bayesian Model Averaging with Knowledge Questions Score as DV

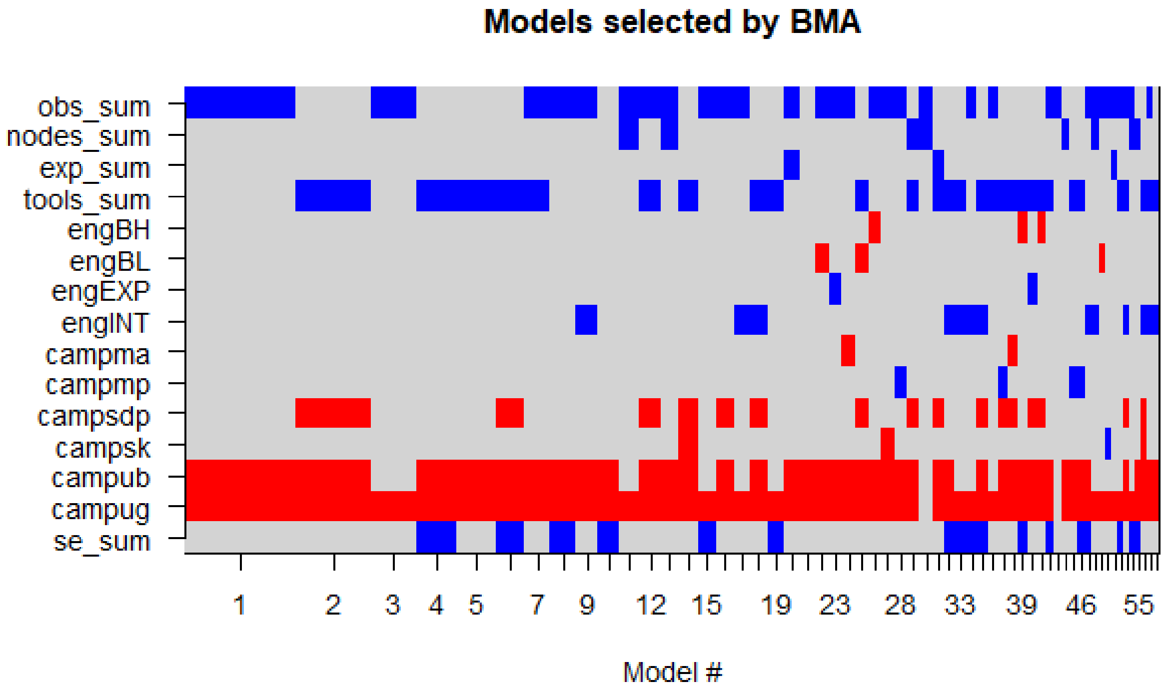
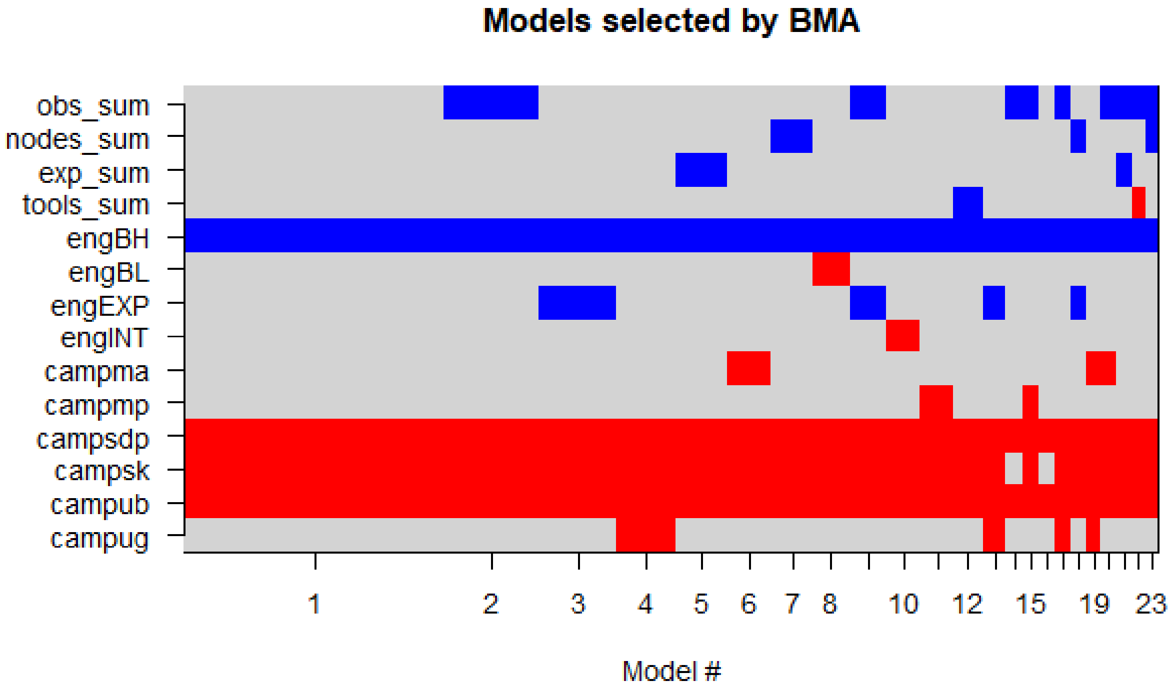


Figure 16

Bayesian Model Averaging with Self-Explanation scores as DV



CHAPTER 7: DISCUSSION

7.1 Engagement

The first set of research questions address what activities in the learning environment constitute engagement across three theoretical dimensions: affect, behavior, and cognition. Determining what constitutes “engaged” activities relies heavily on the affordances of the learning environment and the prompts given to learners (Martin & Borup, 2022), as well as the internal feelings of the individual and the time the learning is taking place. In essence, the combination of all these variables comprises a complex system of engagement, where the patterns of engagement emerge through the interaction of all the variables present (Marchand & Hilpert, 2024). Because of this complexity, an exploratory approach was taken to examine the data and better understand the patterns present as they relate to individual engagement.

The open-endedness of the *Minecraft* learning environment provides learners with a lot of autonomy in terms of wayfinding and where they decide to allocate their attention. The worlds are vast and have myriad areas to visit and thus a learner who explores more of the map might be considered curious, an affective state of information-seeking (Silvia, 2008). This feature (EXP) captures breadth of engagement, but this does not inform if the exploration is directed towards content present in the environment. In this way, another exploration metric was calculated using network analysis to determine how many POIs and/or NPCs a learner visited and how often. This form of exploration bridges a gap between affect and cognition, as it incorporates curiosity with goal-oriented wayfinding to uncover relevant information. These features capture engagement through movement, while the other features capture engagement through deliberate interaction with the game environment.

The task given to learners was to take measurements and make observations, integrating information to comprehend the phenomena they saw. Measurements were made using science tools (*e.g.* “/oxygen”), and the use of science tools captures an affective and cognitive measure of interest. A learner expresses curiosity about the environment and might use a science tool to fill the curiosity gap (Ainley, 2019), while also increasing knowledge about the world. The use of science tools can be seen as a strategic investigation of the environment, productively engaging with the resources available to meet the goals of the experience. The number of unique science tools used per world served as another form of engagement that digs deeper into cognitive engagement, hypothesizing using a diversity of tools is indicative of deeper interest to comprehend complex phenomena.

Observation frequencies were also tabulated to provide an indication of behavioral engagement, since the frequency itself does not provide insight into the quality of the observations, and historically learners have used observations to chat or distract others. Observation quality was determined by manually coding all observations based on the coding scheme described in the methods. Finally, average action latency was also calculated as the average amount of time between either tool use or observation placement per world. However, this feature was discarded due to the redundancy of information provided; smaller action latencies are associated with more tool use and more observation placement. The diversity of science tool use was also discarded because of how closely it correlates with overall tool use. Therefore, based on the affordances of the environment and the tasks assigned to learners the following engineered features were retained to best reveal the dimensions of engagement: exploration breadth (EXP), directed exploration (POIs), science tool frequency (STF), and observation frequency (OF). These features are the primary methods for interacting with our

Minecraft server and cover the three primary dimensions of engagement set forth by Fredericks et al., (2011).

Engagement is a requisite component of interest-driven behavior (Hidi & Renninger, 2006). All engagement features were found to significantly correlate with each other, affirming the positive relationship between the features. Those who make more observations are likely to use more tools, are likely to explore more, and are likely to visit more places of interest. These features were also found to correlate highly with knowledge assessments, evidencing a strong connection between engagement and knowledge, supported extensively by interest theory (Fastrich & Murayama, 2020; Tobias, 1994; Renningher & Hidi, 2021;, Rotgans & Schmidt, 2017).

Exploration was the only feature to not correlate significantly with performance on the open-ended self-explanations delivered at the end of each world. This indicates exploration that is not goal-oriented tends not improve knowledge construction, however EXP did significantly correlate with performance on the knowledge assessment at the end of the camp. Given this correlation, the fact that the EXP group exhibited mediocre performance on knowledge assessments could mean that breadth of exploration is important but only when it is done in tandem with other measures of engagement (*i.e.* observations). Extensive exploration may be effortful, but the short-term gains may not be realized, such as in the case of implementing "desirable difficulty" strategies in learning (*i.e.* distributed practice, interleaving; de Bruin et al., 2023; Zepeda, Martin, & Butler, 2020). This would explain why exploration does not correlate with SEs but does with the post-experience knowledge assessment. Time may be needed to consolidate information from such worlds for certain learners, yet this did not appear to the case for BH learners who engaged highly across all features and still did well on reflective self-explanation prompts at the end of each world. They likely have larger knowledge networks to plug new information into and process, as opposed to less engaged learners.

Correlations could only capture individual measures of engagement, while time series clustering provides insight into patterns of engagement over time. Time series analysis divided learners into groups of “High” v. “Low” performers on each of the engagement features and combined patterns of engagement could be drawn from these clusters. Coded engagement behaviors from the time series revealed 5 codes: persistently high engagers (BH), persistently low engagers (BL), equal high and low engagers (B), exploration only learners (EXP), targeted exploration only learners (POI). The largest category of learners to emerge was that of the BH learners, suggesting that most students maintained high engagement throughout, and that the server provided sufficient opportunities to remain engaged through tools, observations, and exploration. Having opportunities to engage is a major factor in developing interest (Renninger & Hidi, 2020), and having adequate prior knowledge to fold in novel information is also critical (Fryer & Ainley, 2019; Silvia, 2006; 2008). The second highest category was BL learners, and one reason for this is that some students did not voluntarily sign up for this camp, rather the camp was incorporated into their summer programming. Most of the students in the BL condition were in this situation. Other students may have heard the camp used *Minecraft* but did not realize the STEM bent of the camp. For some, *Minecraft* and digital games broadly can possess considerable amounts of seductive details (Mayer, 2019; Z. Wang & Adesope, 2016), where players may have had preferred ways of playing *Minecraft* that did not align with camp objectives, and they therefore are trying to align the experience more with their prior expectations or personal goals.

The next largest group to emerge was the balanced learners (B). These students unanimously visited NPCs and POIs at high frequencies and in addition either explored extensively or made observations. Their engagement patterns consisted of highs and lows, displaying features they preferred over others. Balancing nodes and exploration make sense, as they both capture a form of exploration, albeit one appears more purposeful than the other. However, learners may be visiting the NPCs but not reading the dialogue, and they may be visiting POIs but not taking time to consider the importance. Evidence of effort, through additional observations or tool use, seems essential here to perform highly on knowledge assessments. The more effort students put forth the more they gain from the learning environment (Bjork & Bjork, 2020; de Bruin et al., 2023), but it is still possible learners are drawing from their prior knowledge and interest to perform highly on knowledge assessments, creating continuities of interest (Azevedo, 2017; Canning et al., 2008; Harackiewicz et al., 2008). Both EXP and POI learner categories were small in comparison, and these learners had different ways of seeking information but engaged in no other way. Interestingly, no categories of "observer only" or "tool user only" emerged. This could be a result of the cognitive and (minimal) physical effort of using these features, and EXP and POI are more closely related to BL learners than BH learners in terms of knowledge assessment performance.

The open-endedness, and informal nature, of the experience allowed for such a diversity of engagement patterns to emerge. Learners were encouraged by camp staff to engage through use of science tools and observations, but ultimately learners' self-directed goals are what drove their engagement. Persistent high engagement in relatively unstructured tasks requires a degree of self-regulated learning behavior that many adolescents may have difficulty maintaining (de Bruin et al., 2023; Greene et al., 2024; Wiedbusch et al., 2023), especially if they lack control over the task goals (Hardy, Day, & Arthur, 2019; Zepeda et al., 2020) . This connects to theory of interest where learners who are able to self-relate information are more likely to maintain interest than those who struggle to find personal relevance in the content or context (Hulleman et al., 2017; Renninger et al., 2019; Renninger & Hidi, 2021). Of course, looking at the time series (Figures 8 - 11) shows there are times where BH learners' engagement dipped for certain features and BL learners' engagement increased. It is to be expected that individuals have limits to the length of time they can remain focused and engaged, but what separates them is the ability to reengage and maintain an overall persistent level of quality engagement (O'Brien & Toms, 2008). Less explicit is the need for examining engagement in digital environments across multiple dimensions and metrics. Looking at only one metric might indicate poor cognitive engagement, when the learner is engaging cognitively on a different metric (*i.e.* high-quality observations, as opposed to diversity of science tools). Engagement across metrics and across time is necessary for unpacking patterns indicating a leaner's overall engagement with the learning environment.

The worlds explored were spread across three days, and there might be excitement for some learners at the beginning of the day, but this may fade as they are tasked with comprehending the complicated relationships between variables on each planet. Interest has been shown to be a coping mechanism in frustrating learning situations, and if interest is not present then learners have more difficulty staying on task and self-regulating engagement (Thoman et al., 2011). Future research can explore the transitions in engagement or interest from world to world and day to day to capture the probabilities of moving from unproductive (uninterested?) to productive (interested?) behaviors or vice versa (Tissenbaum, 2020). Pinpointing when BL learners transition from engaged to disengaged provides opportunity to "bootstrap" interest in that area to promote continued engagement. For example, if a learner has stopped using science tools or making observations and started only exploring, they can be prompted to take a measurement in a specific area that might yield a different value from where they are standing. This allows them a chance to continue exploring but also working productively to uncover differences in the maps. Even simpler, nudges or directions to POIs could introduce novelty (*e.g.* traveling to a moon with volcanoes) that prompts learners to ask questions and wonder why the world appears as it does. The value of the experience comes from asking questions and inferring answers to those questions. Additional missions that ask learners to uncover a secret of the planet could also serve as challenges triggering their interest. The goal is to help all learners engage at the level of the BH group without feeling the effort to do so is overwhelming.

Looking at descriptive analyses capturing the number of learners above the mean for knowledge and interest measures by engagement category revealed explicit trends. The BH group substantially outperformed all other groups on knowledge assessments, while the BL group had the lowest percentage of learners above the mean for both KQs and SEs. A likely possibility is the BH group entered the experience with relatively high interest and knowledge in Astronomy, the targeted domain, while the BL group entered with lower knowledge and were not all attending by their own accord (*i.e.* lower Astronomy interest). On the SEs the BL group performed similarly to the B, EXP, and POI groups, in terms of number of learners above the mean. These learners achieved similar results through a variety of patterns, none of which were even as close to effective as persistent high engagement. This stresses the importance of utilizing the tools available to extract meaning from the learning environment to consolidate information and derive meaning from it. Additionally, SEs comprised another form of engagement, where learners had to critically reflect to draw inferences and reason through what was happening on each world.

The interview KQs were more varied in results. The BH group still outperformed all other groups by almost twice as many learners, however the B group did better relative to the other lower performing groups. In effect, the B learners were mostly engaging at the affective level (EXP & Nodes), which shows that curiosity may inhabit positive outcomes for learning. Selectively engaging with at least two features seems to suggest better performance than productively engaging with no features or just one. KQs are broader in their application than the SEs and therefore deep interaction with each world may not be necessary to perform well on those questions. Also, learners have the benefit of having answered SEs and participating in discussions following exploration that they do not have when answering SEs. Once again, though, a high degree of knowledge and interest entering the experience is a strong explanation for higher, persistent engagement and higher knowledge scores. Learner background and feelings of identity also contribute to this.

During interviews, learners were asked if they view themself as a “science person”. They were allowed to interpret this however they wanted, and once again the BH group overwhelmingly identified as “science people” compared to all other groups. The BH group also had close to twice as many learners declaring a STEM subject as their favorite subject in school, compared to all other groups. Links between interests and identity have been frequently posited in the literature (Hecht et al., 2019; Renninger, 2009; Wang et al., 2018; Perez et al., 2024), where an affective state of interest is integral to how one identifies themself with a domain. Identification as a “science person” and liking STEM subjects in school is associated with higher engagement and better knowledge scores in this context. Learners who identify as “science people" likely see this experience as an opportunity to create continuity with their existing interests (Azevedo, 2017). Still, there are several learners who identify as "science people" but fall into lesser engagement groups, possibly meaning their interests are more in line with other scientific disciplines (*e.g.* health, biology) not represented in this camp experience, or the experience proved to be a discontinuity with their interests (Azevedo, 2017).

The EXP group had only one learner who identified as a "science person", which suggests learners who lack STEM identity may be bored and looking for things to entertain them on the server. Learners who perceive a lack of useful information in the learning environment tend to exhibit increasing boredom, which correlates with greater exploration (Geana & Daw, 2016). Such learners either do not want to put the effort into using the tools or making observations because the content might not feel relevant or they are generally uninterested, and therefore they prefer to wander. Future research might consider the dimensions of disengagement, such as boredom, which would fall into the affective dimension. While the novelty of hypothetical worlds may be sufficient to trigger interest for some, there are still learners where the worlds are insufficient for promoting engagement and triggering interest. The EXP learners may be coming up empty-handed when the only content available is in a domain they do not identify with or have interest in engaging, in other words no utility value from the information available (Harackiewicz, Smith, & Prinski, 2016; Tanaka & Murayama, 2014). Based on previous camp observations, they may also be playing games on the server with friends, such as hide-and-seek, inflating exploration while ignoring all other engagement tools.

Bayesian models did not reveal any effect of engagement variables on interest, however the number of observations a learner makes did predict performance on KQs during interviews. The impact is small but shows that utilizing this engagement feature goes beyond any of the fixed demographic predictors. In other words, regardless of which camp a learner was part of they could increase their knowledge score by making more observations; engaging only on the behavior dimension does exhibit some progress. This result did not capture quality, rather increasing the number of observations (the pure effort of doing so) does seem to force learners to reflect on the content. Quality observations are an exercise of consolidating information and trying to make sense of the novel environment, what I would consider a greater effort than just considering observation frequency. More time spent effortfully engaging and trying to learn likely lead to better knowledge scores (de Bruin et al., 2023). For self-explanations, being in the BH group entailed a 10-point increase over the B group (baseline). Engaging highly across all engagement features indicates better performance world-by-world, which includes making more inferences and piecing more information together. It is also worth noting that self-explanations can be seen as an additional engagement measure, as some students do not want to complete them (as noted by staff observations) and may put in little effort. Therefore, the BH students add another dimension of high engagement here. Ultimately, a strong relationship between identify, engagement, and knowledge presents itself in the context of a STEM-focused summer camp using *Minecraft* for the delivery of content.

7.2 Interest

Looking towards interest, none of the engagement features significantly correlated with any form of interest (STEM, *Minecraft*, or Astronomy), while all forms of interest significantly correlated with each other. This result could indicate that interest has no bearing on how a student will engage in this context, but it does show that students in the camp tend to be interested in all targeted interest objects of the camp (content & context). One interest relationship uncovered is the significant correlations between *Minecraft* interest, Astronomy interest, and self-explanation scores. Having more interest in *Minecraft* might equip learners with more tools to navigate the digital environment, a readiness to learn and engage that those less interested in *Minecraft* lack. This ability to engage the worlds displays a tendency to perform well when asked to explain why the conditions of each world produce certain outcomes. Further, Astronomy interest acts as a base upon which the learner can gain and activate knowledge, a content readiness. STEM interest is likely too broad to relate to the experience, as learners may have high interest in other areas of STEM not represented in the experience, such as Biology, Chemistry, or Engineering. Yet, all forms of interest significantly correlated with interview knowledge questions as the end, and this again indicates the interplay between knowledge and interest, where those that know more tend to be more interested and therefore do better when being assessed (Murayama, 2022; Rodríguez-Aflecht et al., 2018; Tobias, 1994). Engagement and knowledge, as well as knowledge and interest, positively relate to each other, however no relationship between individual engagement features and interest can be concluded from correlations.

No changes in STEM interest or *Minecraft* interest were found because of participating in the camp, and this result is not surprising as these two measures of interest capture a more enduring, individual interest. Research has shown that individual interest does not significantly change from short duration, informal learning experiences (Lewalter et al., 2021). However, there was a significant change in Astronomy interest between the beginning and end of the camp experience across all learners, and this suggests the camp experience did, at the very least, trigger situational interest for most learners. The experience provided a novel experience for learners that was accessible regardless of prior knowledge and experience with *Minecraft* and prompted them to think about how different conceptualizations of Earth can lead to monumental changes in how humans survive. The angle of survivability or "life and death" draws upon theory on how to imbue experiences with universally interesting topics (Schank, 1979), which may have contributed to overall increases in Astronomy interest. Where most of the changes originate from also deserves attention. Still, the change in only Astronomy interest really points to using assessments targeted at the content and not broad measures of STEM interest. The experience focused on Astronomy, and it worked to develop interest in Astronomy. As research focusing on interest persists, my work highlights a need to be deliberate about measuring specific interests that align directly with content and the ability to engage the content.

A common thread has been established regarding the highly engaged learners and positive outcomes, and the story with Astronomy interest is no exception. The BH had only 40% of learners above the mean for Astronomy interest at the start of the camp but increased to 55% by the end. Even if a learner in the BH group did not increase to above the mean, 77% of learners in the BH group increased in Astronomy interest from pre- to post. As with knowledge and engagement measures, this is way more than any of the other groups, as a percentage or raw number. The BL group also saw great improvement for individuals above the mean from pre- to post, and overall, 55% of learners in this group increased. Despite the drastic contrast in engagement, even learners who do not productively engage can still have their interest triggered. Highly engaged individuals gain the most in terms of knowledge, but even less engaged individuals can still find something that increases their interest in the domain. The only group to not see at least 50% improvement was the EXP group, but this could be because most of the learners in this group entered with relatively high Astronomy interest. However, this was the only group to decrease in Astronomy interest from pre- to post. Possibly, they were unable to find any information pertinent to their own interests and therefore wandered aimlessly, ignoring areas and interactions that offered opportunities for productive engagement. Even within a domain there is considerable variability with possible targets of interest.

An underlying reason for using *Minecraft* as the primary content delivery vehicle for this summer camp experience stems from adolescents' familiarity with and affinity for the game (Lane & Yi, 2017). If adolescents see they can engage STEM domains in an environment, they feel comfortable navigating and enjoy immersing themselves in then barriers to STEM content should be low. In other words, the content can be made more interesting through an environment that is personally relevant. Interest in *Minecraft* should cascade into developing interest in Astronomy and possibly STEM more broadly through the experience. When looking at *Minecraft* interest, though, the BL group had the most learners above the mean, while the BH group had the second lowest in terms of percentage. The BH group's interest in Astronomy and STEM was likely a greater motivating factor than interest in the context. If the domain content is the primary motivational variable and *Minecraft* is an added bonus, then engagement should be persistently high. The inverse may be true for the BL learners, where their motivation to participate in the camp was driven by *Minecraft* interest, but once they realized they had to play *Minecraft* in a way conducive to scientific inquiry they chose to disengage. Player v. player (PvP) is a popular way to play *Minecraft*, where individuals battle each other until one person's avatar is defeated in combat. Many learners in the camp expressed a desire for PvP that was not granted because it did not align with camp goals. If anything, high interest in *Minecraft* only seems to distract learners from engaging productively and achieving high performance on the knowledge measures. On the other hand, BL learners did still mostly increase STEM interest, and with the goal of triggering interest this can be considered a success. Work remains to investigate how both interests can be triggered, and knowledge increased for those entering the experience with less knowledge and interest.

Ultimately, none of the engagement variables or clusters predicted any form of interest through Bayesian Model Averaging, however knowledge assessments predicted *Minecraft* interest and Astronomy interest. At risk of sounding utterly redundant, knowledge and interest are intimately tied together, but chasing the source of knowledge may prove fruitful for determining where changes can be made to help learners engage and excel in developing conceptual understanding. While site of camp did not predict interest, it did predict performance on both knowledge assessments. Those in the UG and UB groups underperformed compared to the baseline group, BR. The BR group consisted of learners who voluntarily signed up for the camp, in consultation with parents, and who were of relatively high socioeconomic status. On the other hand, the UG and UB groups were learners who were required to participate in the camp and all qualified for free and reduced lunch. This contrast stresses the disparity in educational opportunities, where students of means are more likely to have access to more opportunities to participate in out-of-school STEM programs (Saw & Agger, 2021). The SDP group also scored lower than the baseline group, although they were all voluntarily enrolled and of relatively high socioeconomic status. It is unclear why they performed poorly on the SEs compared to their peers.

What is clear from the analysis is the importance of taking a learner-centered approach to discerning patterns of behavior that capture interest and engagement. Instead of focusing on variables and harnessing assumptions about what changes might lead to greater increases in interest, learners engaged on their own accord, and I was able to find patterns of behavior consistent with high engagement and increases in interest. Furthermore, grounding engagement features with theory helped to show that to get the most out of the experience learners should engage behaviorally, cognitively, and affectively. Some evidence showed purely relying on a form of behavioral engagement could increase performance on a knowledge assessment, but those that gained the most from the experience were the individuals who engaged across all dimensions consistently. Lessons can be gleaned from the highly engaged, and future iterations of this work can consider how to ensure engagement across all dimensions remains high, or at least all dimensions are reengaged after periods of disengagement.

CHAPTER 8: LIMITATIONS

Collecting data in informal learning environments is inherently messy, in addition the context being researched here has limited capacity for number of participants. We were limited by the number of laptops and available staff to ensure summer camps went smoothly, and in turn the relative sample size is small for generating inference. Experimental conditions comparing instructional techniques or value-added approaches to games research were not really feasible, as control and testing conditions could not randomly be assigned with such few learners. In several instances, a learner might be present for one day but absent the next day. Therefore, several instances of missing data had to be accounted for through multiple imputation. Multiple imputation has been shown to be the most effective method for handling up to 50% missing data for a measurement instrument, but it is unclear how effective it is when there are multiple instruments using imputation (Eekhout et al., 2014). For each instrument in this analysis the amount of missing data did not exceed 12%. One problem that arose was that missing data was greater for the learners identified as BL, as compared to the BH category. On one hand, this indicates BH learners were more persistent and put in effort to complete all assessments and surveys, but on the other hand this means BL data may not be fully representative.

While survey instruments were validated via a national campaign with over 1200 responses, the surveys used in this context were seen as a nuisance by some students and not completed with complete honesty. Except for complete straight lining, this can be difficult to detect (Kim et al., 2019). However, this is a potential challenge for interpreting interest self-report survey data from any group, where the possibility of not accurately assessing one's own interest is a perpetual concern (Renninger & Hidi, 2011). Given this constraint, saying an individual's interest has been triggered or developed can be fraught with uncertainty. Even in interviews, learners may answer STEM-related questions in a way consistent with high STEM interest, but they do not identify as a “STEM person”.

Overall, higher knowledge scores are associated with positive changes in content-interest, however the lack of a prior knowledge assessment makes this question difficult to resolve. A prior knowledge assessment would provide insight to increases in knowledge, determining if increases correlate with increases in content interest. Knowledge was measured throughout with self-explanations and at the end with interview knowledge questions, but without a sense of prior knowledge each learner entered the camp with no sense of change in knowledge can be discerned. Conclusions drawn regarding the link between knowledge and interest can be seen throughout the results, yet the influence of prior knowledge may be a major contributor to measured knowledge scores and it is unknown here. In predictive models, prior knowledge may render other predictors of interest or knowledge less probable. Future research that implements a pre-post knowledge assessment can make stronger claims about the role knowledge plays in changing states of interest in this context. The results presented comprise realistic hypotheses reinforced by interest literature, but continued data collection and analysis is needed to verify findings.

This study only captures interest from static time points, as opposed to repeatedly measuring interest while learners are engaged in the digital environment. While intrusive, pop-up assessments of interest could provide stronger links between engagement patterns and interest at given points. This would further establish links between the affordances of the digital environment and the engagement exhibited between the learner and environment. Perhaps, there are environments that are not interesting to the learner and unpacking why that environment is uninteresting can be used to understand the engagement patterns. Furthermore, unpacking what aspects of the environment are triggering or leading to the increase in Astronomy interest can provide valuable insight for highlighting specific features of the environment for who and when. This information is important for building environments that can cater to diverse conceptual interests. Though, the point of this research is to identify interest through engagement logs, moving away from the need for self-reports of interest.

Exploratory analysis sheds light on emergent patterns in the data that can be connected to theory, however causality cannot be inferred from such results. The analysis is largely descriptive, and cutoffs or thresholds can be argued to move in differing directions (*i.e.* learners above the mean v. in highest quartile). Care was taken to ensure clusters were valid through several statistics, however the small sample size and low number of time points does not provide as in-depth analysis as may be needed to determine meaningful patterns of engagement. Again, looking at the percentage of learners above the mean for a measurement instrument could possibly be happenstance. More participants are needed to implement robust statistical analysis, such as Structural Equation Modeling, to see if there are mediating effects of triggered interest on knowledge and engagement.

Finally, most of the students represented in this study enrolled in the camp because of their interest in *Minecraft*. This biases findings towards those already familiar with playing *Minecraft* and who have an established interest in the game. While there is a sample of learners represented who do not frequently play *Minecraft*, they are a small percentage of the total, and all participants had at least heard of *Minecraft*. An interesting direction for the future would be to include more participants who have never played *Minecraft*. This will provide insight into how *Minecraft* can be used with younger and older audiences, and if it is an effective medium for promoting interest and developing knowledge in STEM among non-players.

CHAPTER 9: CONTRIBUTIONS & FUTURE DIRECTIONS

One major contribution of this dissertation is the evidence provided for *Minecraft* as a learning environment where situational interest in a STEM domain can be triggered and developed over the course of a week playing the game. Additionally, knowledge in Astronomy can be cultivated using counterfactual examples of Earth. The learning environment is one where the learner is afforded autonomy to explore and engage as they choose. While tasks are prescribed to learners, they are not strictly enforced. Therefore, the patterns of engagement that emerged from the log data provided differences in self-regulated learning approaches, and these patterns were used to deduce how engagement, knowledge, and interest intersect. I showed that learners who engage across all dimensions (affective, behavioral, cognitive) tend to perform better on knowledge assessments, and they were also the ones that were most likely to increase their interest between the beginning and end of camp. Accounting for how the dimensions of engagement apply to activities in the learning environment is unexplored in the literature, and I have shown how important this can be for understanding how engagement impacts performance. Future research should continue to theoretically ground activity in digital environments with dimensions of engagement to understand at what level learners are engaging and whether the dimension(s) engaged is adequate or productive for learning in that environment.

The interest-knowledge connection uncovered in my dissertation reinforced what interest theory has shown to be the case in myriad conditions, however I was able to show that this holds even when popular digital games are used as the medium for instruction. While the *Minecraft* context in my research is specific and holds a lofted position in the cultural zeitgeist, the outcomes can be extrapolated to apply to other simulated learning environments that foster autonomy, novelty, and exploration of learning resources as their primary motivators and interest triggers. Learners will apply adaptive or maladaptive exploration strategies to meet task goals, and their prior interest and knowledge play a large role in how strategies and engagement unfold.

The *Minecraft* context did not appear to be a highly motivating force in this context for driving engagement or knowledge, except for *Minecraft* interest predicting self-explanation score performance. What really seemed to motivate learners to engage across all dimensions and perform highly on knowledge assessments was prior knowledge in the subject, identity as a science person, and existing interest in astronomy. The appeal of the content and the task goals seem to override the contextual appeal, causing players to lean into stalwart self-regulated learning strategies. Both a blessing and a curse, those who enter with high knowledge and interest are the ones likely to exit with higher knowledge and interest. Yet, overall interest increased for most learners, with reflective, information synthesizing observations serving as a possible means to overcome any social factors impacting a learner. Encouraging observations in a way that makes the effort seem effortless could elicit greater knowledge gains at the end of camp. Other categories of learners did still increase interest and perform well on knowledge assessments so there is some benefit to utilizing at least one dimension of engagement while learning. A question remains, “is there an optimal path for developing an interest?”, and I have come a little closer to answering this. There are many ways to engage an interest, but some ways are better than others. Of course, no interest outside of evolutionary imperatives is going to appeal to all individuals. A key takeaway is that no matter the context, learning is effortful, and those who engage across as many dimensions of engagement possible given the affordances of the environment will perform best on interest and knowledge measures.

I have hinted at many future directions in the limitations section. Importantly, one is the need for further data collection and inclusion of an instrument to measure prior knowledge. In addition to prior knowledge, more frequent self-reports of situational interest could provide stronger connections between behavioral patterns and interest. In turn, this could also help validate data mining methods looking at interest development. Personalized interventions can then be implemented to deepen moments of interest for learners across all phases of interest development. To reduce reliance on self-reports, though, we must first show they are a redundant form of information with what is being generated through log files and other process data collection instruments.

While this research accounts for the affordances of the digital learning environment when determining the features to extract, the affordances vary from world to world and this is not wholly captured. Each world has a different number of NPCs, points of interest, and science tools of interest. Additionally, some worlds are divided into separate "biomes", where the player travels through a portal to either another area of that planet or that planet at another time. These variations may complicate how players explore, take measurements, and observe. While the variations were explained as best as possible, capturing every possible confound is immensely challenging. Exhaustive analysis of all possible factors are not fully explored in this research but would be of interest to those wanting to conduct a finer grain analysis of interest and engagement.

More attention can also be paid to the events happening outside of *Minecraft* throughout the summer camp experience. Recording discussions, chatter between participants, impromptu curiosity questions, and other field observations will provide further insight into interest and engagement profiles of learners. Instructors' prompts and explicitly stated task goals should also be well-documented and kept as consistent as possible. For example, how science tools and observations are introduced can dramatically impact learners' interest in using them. Framed as tools for exploration and problem-solving will likely be more productive than setting quotas for using either feature. This was achieved with the presented research, but the camp experience changes from year to year, and generalizing across years can be difficult. Also, screen capture of *Minecraft* play can help situate some of the recorded behaviors through log data to see if a EXP student is aimlessly wandering, playing with a friend, or trying to resolve a knowledge gap. Other process data collection techniques, such as cognitive interviews or think-alouds, could potentially provide more insight, if they are carefully included into the design of the camp and aligned with the directions of the research questions.

Finally, collecting data from if and how learners play this customized version of *Minecraft* outside of the camp experience would be highly beneficial for understanding their interest in content. Learners have the affordance of choosing how they want to play *Minecraft* in their free time, and choosing our server would be a major indication of wanting to extract further information in service of interest. Voluntary reengagement is a sign that interest has reached a more advanced phase, and identifying such behavior speaks loudly about a learner's engagement through the camp experience. While our *Minecraft* server is available to the public, barriers in terms of access and data collection have precluded us from collecting data in this context. However, additional effort should be placed into making this a viable stream of data.

CHAPTER 10: CONCLUSION

Human interests can be wide and varied, and how much we know about a topic or environment impacts how we will engage with it. As interest theory has established, knowledge and interest are strongly connected and feed into each other to create a continuously evolving machinery promoting the development of both knowledge and interest. (Fastrich & Murayama, 2020; Glowinski & Bayrhuber, 2011; Murayama, 2022; Renninger & Hidi, 2021; Rotgans & Schmidt, 2017). The more knowledge a learner has about a topic or context the more avenues they possess to access related information and incorporate it into their knowledge base, hence more active engagement and 'boosting' of knowledge. Learning through *Minecraft* proves to be no different from how knowledge and interest interact in other formal and informal settings, such as in labs (Glowinski & Bayrhuber, 2011), studying climate science (Carmen, Jennifer; Zint, Michaela;, 2017), or learning about geography (Fastrich & Murayama, 2020).

The uniqueness of learning through *Minecraft* comes from the contextual factor of being a massively popular game that adolescents no doubt at least know by name. The contextual factor proved to be a minor contributor to knowledge building or other forms of interest, but not adequate to trigger and sustain the interest of all learners involved. Factors of knowledge and identity remained paramount in determining if learners increase interest and if they perform well on knowledge assessments centered around content learned primarily in-game. Overall, though, situational interest was triggered for most over the course of a week learning about astronomy concepts in *Minecraft*, and having this added opportunity to engage STEM can cascade into curiosity questions, sharing of new knowledge with friends and family, or a desire to engage other similar STEM-related experiences. Provided such out-of-school STEM activities are not leading to decreases in interest, then their value in promoting engagement, even if consistently low, is valuable.

The continued challenge lies in designing environments that provide ample opportunities for engagement and interest development, regardless of the knowledge and interest profile of the learner entering the experience. Students who do not engage meaningfully can still increase in interest, however their knowledge in the domain may remain comparatively low to those who engage at consistently high levels. If the acquisition of knowledge is enjoyable and feels effortless then a learning environment is working remarkably well. Engagement patterns firing on all cylinders in dimensionality (affect, behavior, cognition) provide the best outcomes, however relying on only two out of three can still have some impact. A digital environment for learning must be designed to consider all the dimensions of engagement by providing challenge, allowing autonomy, and instilling enjoyment. With this goal in mind, continued work can dig deeper into specific triggers in the environment that promote engagement, especially for those who might be less engaged. If all dimensions of engagement can be utilized, then it stands that a learner will likely increase both their knowledge and interest in the targeted domain. In turn, a learner may not be ready to identify themself as a "STEM person", but they have hopefully increased feelings of competence that open doors to future opportunities where they can more comfortably situate themself in a STEM domain. Informal STEM experiences, such as learning through *Minecraft*, are a viable option for a diversity of learners, and should continue to be expanded and tailored to the needs of diverse audiences. They should be considered key steppingstones in the development of interest over time, harnessing all we know about interest and making the experience as memorable and enjoyable as possible.

REFERENCES

Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems*, *53*, 16–38. https://doi.org/10.1016/j.is.2015.04.007

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review*, *18*, 391-405.

Ainley, M. (2019). Curiosity and interest: Emergence and divergence. *Educational Psychology Review*, *31*(4), 789–806. https://doi.org/10.1007/s10648-019-09495-z

Ainley, M. D. (2002). Interest in learning and classroom interactions. In D. Clarke (Ed.), *Perspectives on Practice and Meaning in Mathematics and Science Classrooms* (Vol. 25, pp. 105–130). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47228-7_6

Al Mamun, M. A., Lawrie, G., & Wright, T. (2022). Exploration of learner-content interactions and learning approaches: The role of guided inquiry in the self-directed online environments. *Computers & Education*, *178*, 104398. https://doi.org/10.1016/j.compedu.2021.104398

Amershi, S., & Conati, C. (2006). Automatic recognition of learner groups in exploratory learning environments. In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *Intelligent Tutoring Systems* (Vol. 4053, pp. 463–472). Springer Berlin Heidelberg. https://doi.org/10.1007/11774303_46

Andres, J. M. L., Rodrigo, M. M. T., Sugay, J. O., Baker, R. S., Paquette, L., Shute, V. J., ... & Small, M. (2014, November). An exploratory analysis of confusion among students using Newton’s playground. In *Proceedings of the 22nd international conference on computers in education* (pp. 65-70). Nara, Japan: Asia-Pacific Society for Computers in Education.

Arizmendi, C. J., Bernacki, M. L., Raković, M., Plumley, R. D., Urban, C. J., Panter, A. T., Greene, J. A., & Gates, K. M. (2022). Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01939-9

Arnone, M. P., Small, R. V., Chauncey, S. A., & McKenna, H. P. (2011). Curiosity, interest and engagement in technology-pervasive learning environments: A new research agenda. *Educational Technology Research and Development*, *59*(2), 181–198. https://doi.org/10.1007/s11423-011-9190-9

Azevedo, F. S. (2018). An inquiry into the structure of situational interests. *Science Education*, *102*(1), 108-127.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational psychologist*, *50*(1), 84-94.

Azevedo, R., & Taub, M. (2020). The challenge of measuring processes and outcomes while learning from multiple representations with advanced learning technologies. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.), *Handbook of Learning from Multiple Representations and Perspectives* (1st ed., pp. 532–553). Routledge. https://doi.org/10.4324/9780429443961-34

Baek, Y., Min, E., & Yun, S. (2020). Mining educational implications of minecraft. *Computers in the Schools*, *37*(1), 1–16. https://doi.org/10.1080/07380569.2020.1719802

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, *49*(4), 193–224. https://doi.org/10.1159/000094368

Baram-Tsabari, A. (2015). Promoting Information Seeking and Questioning in Science. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in Mathematics and Science Learning* (pp. 135–152). American Educational Research Association. https://doi.org/10.3102/978-0-935302-42-4_8

Bernacki, M. L., Chavez, M. M., & Uesbeck, P. M. (2020). Predicting achievement and providing support before STEM majors begin to fail. *Computers & Education*, *158*, 103999. https://doi.org/10.1016/j.compedu.2020.103999

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose(s)? *Educational Psychology Review*, *33*(4), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology*, *110*(6), 864–881. https://doi.org/10.1037/edu0000250

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied research in Memory and Cognition*, *9*(4), 475.

Blankenburg, J. S., Höffler, T. N., & Parchmann, I. (2016). Fostering today what is needed tomorrow: Investigating students' interest in science: *Science Education*, *100*(2), 364–391. https://doi.org/10.1002/sce.21204

Boot, W. R., Sumner, A., Towne, T. J., Rodriguez, P., & Anders Ericsson, K. (2017). Applying aspects of the expert performance approach to better understand the structure of skill and mechanisms of skill acquisition in video games. *Topics in Cognitive Science*, *9*(2), 413–436. https://doi.org/10.1111/tops.12230

Bos, B., Wilder, L., Cook, M., & O'Donnell, R. (2014). Learning mathematics through Minecraft. *Teaching Children Mathematics*, *21*(1), 56-59.

Burt, B. A., & Johnson, J. T. (2018). Origins of early STEM interest for Black male graduate students in engineering: A community cultural wealth perspective. *School Science and Mathematics*, *118*(6), 257–270. https://doi.org/10.1111/ssm.12294

Buuren, S. V., & Groothuis-Oudshoorn, K. (2011). **mice**: Multivariate Imputation by Chained Equations in *R*. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

Calabrese Barton, A., & Tan, E. (2019). Designing for rightful presence in STEM: The role of making present practices. *Journal of the Learning Sciences*, *28*(4–5), 616–658. https://doi.org/10.1080/10508406.2019.1591411

Canning, E. A., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Tibbetts, Y., & Hyde, J. S. (2018a). Improving performance and retention in introductory biology with a utility-value intervention. *Journal of Educational Psychology*, *110*(6), 834–849. https://doi.org/10.1037/edu0000244

Canning, E. A., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Tibbetts, Y., & Hyde, J. S. (2018b). Improving performance and retention in introductory biology with a utility-value intervention. *Journal of Educational Psychology*, *110*(6), 834–849. https://doi.org/10.1037/edu0000244

Carmen, Jennifer; Zint, Michaela;, I., Inez. (2017). *Assessing Student Interest and Desire to Learn More About Climate Change Effects on Forests in Middle School: An Intervention-Based Path Model*. Electronic Journal of Science Education.

Casula, M., Rangarajan, N., & Shields, P. (2021). The potential of working hypotheses for deductive exploratory research. *Quality & Quantity*, *55*(5), 1703–1725. https://doi.org/10.1007/s11135-020-01072-9

Cipollone, M., Schifter, C. C., & Moffat, R. A. (2014). Minecraft as a creative tool: A case study. *International Journal of Game-Based Learning (IJGBL)*, *4*(2), 1-14.

Checa-Romero, M., & Pascual Gómez, I. (2018). Minecraft and machinima in action: Development of creativity in the classroom. *Technology, Pedagogy and Education*, *27*(5), 625-637.

Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. https://doi.org/10.3102/0034654315582065

Clemenson, G. D., Henningfield, C. M., & Stark, C. E. L. (2019). Improving hippocampal memory through the experience of a rich Minecraft environment. *Frontiers in Behavioral Neuroscience*, *13*, 57. https://doi.org/10.3389/fnbeh.2019.00057

Corbalan, J., Martinez, F., Donolo, D., Alonso, C. Tejerina, M. & Liminana, M. R. (2003). CREA. *Creative Intelligence, a cognitive measure of creativity.* Madrid: TEA Editions.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*(4), 715–730. https://doi.org/10.1037/0022-0663.88.4.715

Daniels, E. A., & Robnett, R. D. (2021). The STEM pipeline: Do media and objectified body consciousness create an early exit for middle school girls? *The Journal of Early Adolescence*, *41*(7), 1099–1124. https://doi.org/10.1177/0272431620983442

de Andrade, B., Poplin, A., & Sousa de Sena, Í. (2020). Minecraft as a tool for engaging children in urban planning: A Case study in Tirol Town, Brazil. *ISPRS International Journal of Geo-Information*, *9*(3), 170.

De Bruin, A. B. H., Biwer, F., Hui, L., Onan, E., David, L., & Wiradhany, W. (2023). Worth the Effort: The Start and Stick to Desirable Difficulties (S2D2) Framework. *Educational Psychology Review*, *35*(2), 41. https://doi.org/10.1007/s10648-023-09766-w

Dever, D. A., Sonnenfeld, N. A., Wiedbusch, M. D., Schmorrow, S. G., Amon, M. J., & Azevedo, R. (2023). A complex systems approach to analyzing pedagogical agents' scaffolding of self-regulated learning within an intelligent tutoring system. *Metacognition and Learning*, *18*(3), 659–691. https://doi.org/10.1007/s11409-023-09346-x

Dewan, M. A. A., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: A review. *Smart Learning Environments*, *6*(1), 1. https://doi.org/10.1186/s40561-018-0080-z

Dezuanni, M., O'Mara, J., & Beavis, C. (2015). 'Redstone is like electricity': Children's performative representations in and around *Minecraft*. *E-Learning and Digital Media*, *12*(2), 147–163. https://doi.org/10.1177/2042753014568176

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational psychologist*, *52*(2), 104-123.

Donnellan, E., Aslan, S., Fastrich, G. M., & Murayama, K. (2022). How are curiosity and interest different? Naïve Bayes classification of people's beliefs. *Educational Psychology Review*, *34*(1), 73–105. https://doi.org/10.1007/s10648-021-09622-9

Durik, A. M., Shechter, O. G., Noh, M., Rozek, C. S., & Harackiewicz, J. M. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interest and performance. *Motivation and Emotion*, *39*(1), 104–118. https://doi.org/10.1007/s11031-014-9419-0

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eekhout, I., De Vet, H. C. W., Twisk, J. W. R., Brand, J. P. L., De Boer, M. R., & Heymans, M. W. (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *Journal of Clinical Epidemiology*, *67*(3), 335–342. https://doi.org/10.1016/j.jclinepi.2013.09.009

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.

Emerson, A., Min, W., Azevedo, R., & Lester, J. (2022). Early prediction of student knowledge in-game based learning with distribued represntations of assessment questions. *British Journal of Educational Technology*. DOI: 10.1111/bjet.13281

Fastrich, G. M., & Murayama, K. (2020). Development of interest and role of choice during sequential knowledge acquisition. *Aera Open*, *6*(2), 2332858420929981.

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2018). Exploring teacher popularity: Associations with teacher characteristics and student outcomes in primary school. *Social Psychology of Education*, *21*(5), 1225–1249. https://doi.org/10.1007/s11218-018-9462-x

Flum, H., & Kaplan, A. (2006). Exploratory orientation as an educational goal. *Educational Psychologist*, *41*(2), 99–110. https://doi.org/10.1207/s15326985ep4102_3

Fredricks, J. A. (2011). Engagement in school and out-of-School contexts: A multidimensional view of engagement. *Theory Into Practice*, *50*(4), 327–335. https://doi.org/10.1080/00405841.2011.607401

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Fryer, L. K., & Ainley, M. (2019). Supporting interest in a study domain: A longitudinal test of the interplay between interest, utility-value, and competence beliefs. *Learning and Instruction*, *60*, 252–262. https://doi.org/10.1016/j.learninstruc.2017.11.002

Gadbury, M. (accepted). Measuring changes in misconceptions among adolescents in an informal, Minecraft-focused STEM summer camp*s.* Short paper submission to the *International Society of the Learning Science* (ISLS) Annual Meeting 2024.

Gadbury, M., & Lane, H. C. (2022, July). Mining for STEM interest behaviors in Minecraft. In *International Conference on Artificial Intelligence in Education* (pp. 236-239). Cham: Springer International Publishing.

Geana, A., Wilson, R., Daw, N. D., & Cohen, J. (2016). Boredom, Information-Seeking and Exploration. In *CogSci* (1751 - 1756).

Geerling, D., Butner, J., Fraughton, T., Sinclair, S., Zachary, J., & Sansone, C. (2020). The dynamic association of interest and confusion: The potential for moderation by utility value and gender. *The Journal of Experimental Education*, *88*(3), 407–430. https://doi.org/10.1080/00220973.2018.1561403

Glowinski, I., & Bayrhuber, H. (2011). Student Labs on a University Campus as a Type of Out-of-School Learning Environment: Assessing the Potential to Promote Students' Interest in Science. *International Journal of Environmental and Science Education*, *6*(4), 371-392.

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist*, *50*(1), 43–57. https://doi.org/10.1080/00461520.2014.999919

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, *71*(1–3), 147–177. https://doi.org/10.1016/0001-6918(89)90007-3

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66–78. https://doi.org/10.1037/a0034857

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Greene, J. A., Bernacki, M. L., & Hadwin, A. F. (2024). Self-regulation. *Handbook of educational psychology*, 314-334.

Habig, B., & Gupta, P. (2021). Authentic STEM research, practices of science, and interest development in an informal science education program. *International Journal of STEM Education*, *8*(1), 57. https://doi.org/10.1186/s40594-021-00314-y

Halverson, L. R., & Graham, C. R. (2019). Learner engagement in blended learning environments: A conceptual framework. *Online Learning*, *23*(2). https://doi.org/10.24059/olj.v23i2.1481

Hanna, A., & Rounds, J. (2020). How accurate are interest inventories? A quantitative review of career choice hit rates. *Psychological Bulletin*, *146*(9), 765–796. https://doi.org/10.1037/bul0000269

Harackiewicz, J. M., Durik, A. M., & Barron, K. E. (2008). Multiple goals, optimal motivation, and the development of interest. In J. P. Forgas, K. D. Williams, & S. M. Laham (Eds.), *Social Motivation* (1st ed., pp. 21–39). Cambridge University Press. https://doi.org/10.1017/CBO9780511735066.004

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Hardy, J. H., Day, E. A., & Arthur, W. (2019). Exploration-exploitation tradeoffs and information-knowledge gaps in self-regulated learning: Implications for learner-controlled training and development. *Human Resource Management Review*, *29*(2), 196–217. https://doi.org/10.1016/j.hrmr.2018.07.004

Harpstead, E., MacLellan, C. J., Aleven, V., & Myers, B. A. (2015). Replay Analysis in Open-Ended Educational Games. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious Games Analytics* (pp. 381–399). Springer International Publishing. https://doi.org/10.1007/978-3-319-05834-4_17

Hecht, M., Knutson, K., & Crowley, K. (2019). Becoming a naturalist: Interest development across the learning ecology. *Science Education*, *103*(3), 691–713. https://doi.org/10.1002/sce.21503

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review*, *1*(2), 69–82. https://doi.org/10.1016/j.edurev.2006.09.001

Hidi, S. (2016). Revisiting the role of rewards in motivation and learning: Implications of neuroscientific research. *Educational Psychology Review*, *28*(1), 61–93. https://doi.org/10.1007/s10648-015-9307-5

Hidi, S., & Renninger, K. A. (2019). Interest development and its relation to curiosity: Needed neuroscientific research. *Educational Psychology Review*, *31*(4), 833–852. https://doi.org/10.1007/s10648-019-09491-3

Hidi, S., & Renninger, K. A. (2020). On educating, curiosity, and interest development. *Current Opinion in Behavioral Sciences*, *35*, 99–103. https://doi.org/10.1016/j.cobeha.2020.08.002

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hilpert, J. C., Greene, J. A., & Bernacki, M. (2023). Leveraging complexity frameworks to refine theories of engagement: Advancing self-regulated learning in the age of artificial intelligence. *British Journal of Educational Technology*, bjet.13340. https://doi.org/10.1111/bjet.13340

Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in Educational Psychology: Aligning theory and method. *Educational Psychologist*, *53*(3), 185–202. https://doi.org/10.1080/00461520.2018.1469411

Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E. J. (2020). A conceptual introduction to Bayesian model averaging. *Advances in Methods and Practices in Psychological Science*, *3*(2), 200-215.

Hoff, K., Van Egdom, D., Napolitano, C., Hanna, A., & Rounds, J. (2022). Dream jobs and employment realities: How adolescents’ career aspirations compare to labor demands and automation risks. *Journal of Career Assessment*, *30*(1), 134–156. https://doi.org/10.1177/10690727211026183

Høgheim, S., & Reber, R. (2015). Supporting interest of middle school students in mathematics through context personalization and example choice. *Contemporary Educational Psychology*, *42*, 17–25. https://doi.org/10.1016/j.cedpsych.2015.03.006

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Psychological Assessment Resources.

Hong, J., & Perez, T. (2024). Identity and learning: Student and teacher identity development. In P.A. Schutz & K.R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 269-290). Routledge.

Hulleman, C. S., Thoman, D. B., Dicke, A.-L., & Harackiewicz, J. M. (2017). The promotion and development of interest: The importance of perceived values. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The Science of Interest* (pp. 189–208). Springer International Publishing. https://doi.org/10.1007/978-3-319-55509-6_10

Hunsu, N. J., Adesope, O., & Van Wie, B. J. (2017). Engendering situational interest through innovative instruction in an engineering classroom: What really mattered? *Instructional Science*, *45*(6), 789–804. https://doi.org/10.1007/s11251-017-9427-z

Izard, C. E. (1977). *Human Emotions*. Springer US. https://doi.org/10.1007/978-1-4899-2209-0

Kang, J., & Liu, M. (2020). Investigating navigational behavior patterns of students across at-risk categories within an open-ended serious game. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-020-09462-6

Ke, F. (2016). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research and Development*, *64*(2), 219–244. https://doi.org/10.1007/s11423-015-9418-1

Kew, S. N., & Tasir, Z. (2022). Learning analytics in online learning environment: A systematic review on the focuses and the types of student-related analytics data. *Technology, Knowledge and Learning*, *27*(2), 405–427. https://doi.org/10.1007/s10758-021-09541-2

Kier, M. W., & Blanchard, M. R. (2021). Eliciting students' voices through STEM career explorations. *International Journal of Science and Mathematics Education*, *19*(1), 151–169. https://doi.org/10.1007/s10763-019-10042-z

Kim, Y., Dykema, J., Stevenson, J., Black, P., & Moberg, D. P. (2019). Straightlining: Overview of measurement, comparison of indicators, and effects in mail–web mixed-mode surveys. *Social Science Computer Review*, *37*(2), 214–233. https://doi.org/10.1177/0894439317752406

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, *43*, 39-50.

König, C., & van de Schoot, R. (2018). Bayesian statistics in educational research: A look at the current state of affairs. *Educational Review*, *70*(4), 486–509. https://doi.org/10.1080/00131911.2017.1350636

Krapp, A. (1994). Interest and Curiosity. The role of interest in a theory of exploratory action. In H. Keller, K. Schneider, & B. Henderson (Eds.), *Curiosity and Exploration* (pp. 79–100). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77132-3_6

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*(4), 383–409. https://doi.org/10.1016/S0959-4752(01)00011-1

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, *15*(5), 381–395. https://doi.org/10.1016/j.learninstruc.2005.07.007

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Laine, E., Veermans, M., Lahti, A., & Veermans, K. (2017). Generation of student interest in an inquiry-based mobile learning environment. *Frontline Learning Research*, *5*(4), 42–60. https://doi.org/10.14786/flr.v5i4.306

Lajoie, S. P., & Poitras, E. (2024). Technology-Rich Learning Environments: Theories and Methodologies for Understanding Solo and Group Learning. In P.A. Schutz & K.R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 630-653). Routledge.

Lane, H. C., Gadbury, M., Ginger, J., Yi, S., Comins, N., Henhapl, J., & Rivera-Rogers, A. (2022). Triggering STEM interest with Minecraft in a Hybrid Summer Camp. *Innovations in Remote Instruction*, 3(4),DOI:https://doi.org/10.1037/tmb0000077.

Lane, H. C., & Yi, S. (2017). Playing with virtual blocks: Minecraft as a learning environment for practice and research. In *Cognitive Development in Digital Contexts* (pp. 145–166). Elsevier. https://doi.org/10.1016/B978-0-12-809481-5.00007-9

Lewalter, D., Gegenfurtner, A., & Renninger, K. A. (2021). Out-of-school programs and interest: Design considerations based on a meta-analysis. *Educational Research Review*, *34*, 100406. https://doi.org/10.1016/j.edurev.2021.100406

Lewalter, D., & Geyer, C. (2009). Motivational aspects of school visits to natural science and technical museums. *Zeitschrift für Erziehungswissenschaft*, *12*, 28-44.

Lim, K. Y., & Lim, L. J. (2020). Redstone Jammin': Conversational analysis from a collaborative music-making activity in Minecraft. *Journal For Virtual Worlds Research*, *13*(2-3), 1-12.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, *70*(4), 647–671. https://doi.org/10.1177/0013164409355699

Linnenbrink-Garcia, L., Patall, E. A., & Messersmith, E. E. (2013). Antecedents and consequences of situational interest. *British Journal of Educational Psychology*, *83*(4), 591–614. https://doi.org/10.1111/j.2044-8279.2012.02080.x

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/10.1016/j.learninstruc.2018.08.002

Loderer, K., Pekrun, R., & Plass, J. L. (2020). Emotional foundations of game-based learning. *Handbook of game-based learning*, 111-151.

Magnussen, R., & Elming, A. L. (2017). Student re-design of deprived neighborhoods in Minecraft: Community-driven urban development. Philadelphia, PA: International Society of the Learning Sciences..

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Marchand, G. C., & Hilpert, J. C. (2024). Contributions of complex systems approaches, perspectives, models, and methods in educational psychology. In P.A. Schutz & K.R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 139-161). Routledge.

Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, *57*(3), 162–177. https://doi.org/10.1080/00461520.2022.2089147

Mayer, R. (2014). *Computer Games for Learning*. Cambridge, MA. The MIT Press

Mayer, R. E. (2019). Computer games in education. *Annual review of psychology*, *70*, 531-549.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, *42*(3), 241–265. https://doi.org/10.2190/EC.42.3.a

Murayama, K. (2022). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychological Review*, *129*(1), 175–198. https://doi.org/10.1037/rev0000349

Nebel, S., Schneider, S., Beege, M., Kolda, F., Mackiewicz, V., & Rey, G. D. (2017). You cannot do this alone! Increasing task interdependence in cooperative educational videogames to encourage collaboration. *Educational Technology Research and Development*, *65*(4), 993–1014. https://doi.org/10.1007/s11423-017-9511-8

Niese, Z. A., Libby, L. K., Eibach, R. P., & Carlisle, C. (2019). I can see myself enjoying that: Using imagery perspective to circumvent bias in self-perceptions of interest. *Journal of*

Experimental Psychology: General, *148*(12), 2258–2276. https://doi.org/10.1037/xge0000612

Novak, E., & Tassell, J. (2015). Using video game play to improve education-majors' mathematical performance: An experimental study. *Computers in Human Behavior*, *53*, 124–130. https://doi.org/10.1016/j.chb.2015.07.001

Nye, C. D., Prasad, J., & Rounds, J. (2021). The effects of vocational interests on motivation, satisfaction, and academic performance: Test of a mediated model. *Journal of Vocational Behavior*, *127*, 103583. https://doi.org/10.1016/j.jvb.2021.103583

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955. https://doi.org/10.1002/asi.20801

O'Keefe, P. A., & Linnenbrink-Garcia, L. (2014). The role of interest in optimizing performance and self-regulation. *Journal of Experimental Social Psychology*, *53*, 70–78. https://doi.org/10.1016/j.jesp.2014.02.004

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal*, *16*(4), 455–474. https://doi.org/10.1080/09585170500384529

Pasco, D., & Roure, C. (2022). Situational interest impacts college students' physical activity in a design-based bike exergame. *Journal of Sport and Health Science*, *11*(2), 172–178. https://doi.org/10.1016/j.jshs.2021.03.003

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In *International handbook of emotions in education* (pp. 120-141). Routledge.

Peterson, B. (2020). Building STEM career interest through curriculum treatments. *Journal of STEM Education. 21*(1), 11.

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology*, *105*(4), 1050–1066. https://doi.org/10.1037/a0032688

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, *50*(1), 85–129. https://doi.org/10.1080/03057267.2014.881626

Pusey, M., & Pusey, G. (2015). Using Minecraft in the science classroom. *International Journal of Innovation in Science and Mathematics Education*, *23*(3).

Quinlan, K. M., & Renninger, K. A. (2022). Rethinking employability: How students build on interest in a subject to plan a career. *Higher Education*, *84*(4), 863–883. https://doi.org/10.1007/s10734-021-00804-6

Renninger, K. A. (2009). Interest and Identity Development in Instruction: An Inductive Model. *Educational Psychologist*, *44*(2), 105–118. https://doi.org/10.1080/00461520902832392

Renninger, K.A., Austin, L., Bachrach, J. E., Chau, A., Emmerson, M. S., King, B. R., Riley, K. R., & Stevens, S. J. (2014). Going beyond the "Whoa! That's Cool!" of inquiry: Achieving science interest and learning with the ICAN intervention. In S. A. Karabenick & T. C. Urdan (Eds.), *Advances in Motivation and Achievement* (Vol. 18, pp. 107–138). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-742320140000018003

Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, *50*(1), 58–69. https://doi.org/10.1080/00461520.2014.999920

Renninger, K. A., Bachrach, J. E., & Hidi, S. E. (2019). Triggering and maintaining interest in early phases of interest development. *Learning, Culture and Social Interaction*, *23*, 100260. https://doi.org/10.1016/j.lcsi.2018.11.007

Renninger, K.A. & Hidi, S. (2016). *The power of interest for motivation and engagement.* Routledge

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Renninger, K. A., & Hidi, S. E. (2020). To level the playing field, develop interest. *Policy Insights from the Behavioral and Brain Sciences*, *7*(1), 10–18. https://doi.org/10.1177/2372732219864705

Renninger, K. A., & Hidi, S. E. (2021). Interest development, self-related information processing, and practice. *Theory Into Practice*, 1–12. https://doi.org/10.1080/00405841.2021.1932159

Renninger, K. A., & Hidi, S. E. (2022). Interest: A unique affective and cognitive motivational variable that develops. In *Advances in Motivation Science* (Vol. 9, pp. 179–239). Elsevier. https://doi.org/10.1016/bs.adms.2021.12.004

Renninger, K. A., & Riley, K. R. (2012). Interest, Cognition, and the Case of L- and Science. In S. Kreitler (Ed.), *Cognition and Motivation* (1st ed., pp. 352–382). Cambridge University Press. https://doi.org/10.1017/CBO9781139021463.021

Renninger, K. A., & Su, S. (2012). Interest and its development. In R. M. Ryan (Ed.), *The Oxford Handbook of Human Motivation* (1st ed., pp. 167–188). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0011

Rodríguez-Aflecht, G., Jaakkola, T., Pongsakdi, N., Hannula-Sormunen, M., Brezovszky, B., & Lehtinen, E. (2018). The development of situational interest during a digital mathematics game. *Journal of Computer Assisted Learning*, *34*(3), 259–268. https://doi.org/10.1111/jcal.12239

Rotgans, J. I. (2015). Validation study of a general subject-matter interest measure: The Individual Interest Questionnaire (IIQ). *Health Professions Education*, *1*(1), 67–75. https://doi.org/10.1016/j.hpe.2015.11.009

Rotgans, J. I., & Schmidt, H. G. (2017). The relation between individual interest and knowledge acquisition. *British Educational Research Journal*, *43*(2), 350–371. https://doi.org/10.1002/berj.3268

Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Rovers, S. F. E., Stalmeijer, R. E., Van Merriënboer, J. J. G., Savelberg, H. H. C. M., & De Bruin, A. B. H. (2018). How and Why Do Students Use Learning Strategies? A Mixed Methods Study on Learning Strategies and Desirable Difficulties With Effective Strategy Users. *Frontiers in Psychology*, *9*, 2501. https://doi.org/10.3389/fpsyg.2018.02501

Rowe, J., R, S. L., W, M. B., & C, L. J. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education*, *1–2*, 115–133. https://doi.org/10.3233/JAI-2011-019

Ruff, C. (2016). Finding relevance, competence, and enjoyment: The development of domain identification and interest in first-year science majors. *International Journal of Teaching and Learning in Higher Education*, *28*(3), 335-346.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*.

Ryan, R. M., Deci, E. L., Vansteenkiste, M., & Soenens, B. (2021). Building a science of motivated persons: Self-determination theory's empirical approach to human experience and the regulation of behavior. *Motivation Science*, *7*(2), 97–110. https://doi.org/10.1037/mot0000194

Sardá-Espinosa, A. (n.d.). *Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package*.

Saw, G. K., & Agger, C. A. (2021). STEM Pathways of rural and small-town students: Opportunities to learn, aspirations, preparation, and college enrollment. *Educational Researcher*, *50*(9), 595–606. https://doi.org/10.3102/0013189X211027528

Schank, R. C. (1979). Interestingness: Controlling inferences. *Artificial Intelligence*, *12*(3), 273–297. https://doi.org/10.1016/0004-3702(79)90009-2

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3–4), 299–323. https://doi.org/10.1080/00461520.1991.9653136

Schiefele, U. (2009). Situational and individual interest. In *Handbook of motivation at school* (pp. 211-236). Routledge.

Shin, D. D., & Kim, S. (2019). Homo curious: Curious or interested? *Educational Psychology Review*, *31*(4), 853–874. https://doi.org/10.1007/s10648-019-09497-x

Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion*, *5*(1), 89–102. https://doi.org/10.1037/1528-3542.5.1.89

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science*, *17*(1), 57–60. https://doi.org/10.1111/j.1467-8721.2008.00548.x

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

Snow, E. L., Allen, L. K., & McNamara, D. S. (2015). The Dynamical Analysis of Log Data Within Educational Games. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious Games Analytics* (pp. 81–100). Springer International Publishing. https://doi.org/10.1007/978-3-319-05834-4_4

Sobolewska, E. (2019). Dynamic Time Warping (DTW) as a means to cluster time series. RPubs by RStudio. https://rpubs.com/esobolewska/dtw-time-series

Starkweather, J. (2011). Sharpening Occam's Razor: Using Bayesian Model Averaging in R to separate the wheat from the chaff. *Research and Statistical Support*.

Steinkuehler, C., & Squire, K. (2024). Gaming in educational contexts. In P.A. Schutz & K.R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 674-695). Routledge.

Stoll, G., Einarsdóttir, S., Song, Q. C., Ondish, P., Sun, J. J., & Rounds, J. (2020). The roles of personality traits and vocational interests in explaining what people want out of life. *Journal of Research in Personality*, *86*, 103939. https://doi.org/10.1016/j.jrp.2020.103939

Su, R. (2020). The three faces of interests: An integrative review of interest research in vocational, organizational, and educational psychology. *Journal of Vocational Behavior*, *116*, 103240. https://doi.org/10.1016/j.jvb.2018.10.016

Tan, A. L., Jamaludin, A., & Hung, D. (2021). In pursuit of learning in an informal space: A case study in the Singapore context. *International Journal of Technology and Design Education*, *31*(2), 281–303. https://doi.org/10.1007/s10798-019-09553-1

Tanaka, A., & Murayama, K. (2014). Within-person analyses of situational interest and boredom: Interactions between task-specific perceptions and achievement goals. *Journal of Educational Psychology*, *106*(4), 1122–1134. https://doi.org/10.1037/a0036659

Thoman, D. B., Sansone, C., Robinson, J. A., & Helm, J. L. (2020). Implicit theories of interest regulation. *Motivation Science*, *6*(4), 321–334. https://doi.org/10.1037/mot0000160

Thoman, D. B., Smith, J. L., & Silvia, P. J. (2011). The resource replenishment function of interest. *Social Psychological and Personality Science*, *2*(6), 592–599. https://doi.org/10.1177/1948550611402521

Tissenbaum, M. (2020). I see what you did there! Divergent collaboration and learner transitions from unproductive to productive states in open-ended inquiry. *Computers & Education, 145*, 103739. https://doi.org/10.1016/j.compedu.2019.103739

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, *64*(1), 37-54.

Towhidnejad, M., Kestler, C., Jafer, S., & Nicholas, V. (2014). Introducing computational thinking through stealth teaching. *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 1–7. https://doi.org/10.1109/FIE.2014.7044407

Van de Schoot, R., & Miocević, M. (2020). *Small sample size solutions: A guide for applied researchers and practitioners*. Taylor & Francis.

Vytasek, J. M., Patzak, A., & Winne, P. H. (2020). Analytics for student engagement. In M. Virvou, E. Alepis, G. A. Tsihrintzis, & L. C. Jain (Eds.), *Machine Learning Paradigms* (Vol. 158, pp. 23–48). Springer International Publishing. https://doi.org/10.1007/978-3-030-13743-4_3

Wang, J., Hazari, Z., Cass, C., & Lock, R. (2018). Episodic memories and the longitudinal impact of high school physics on female students' physics identity. *International Journal of Science Education*, *40*(13), 1543–1566. https://doi.org/10.1080/09500693.2018.1486522

Wang, Z., & Adesope, O. (2016). Exploring the effects of seductive details with the 4-phase model of interest. *Learning and Motivation*, *55*, 65–77. https://doi.org/10.1016/j.lmot.2016.06.003

Wiedbusch, M., Dever, D., Li, S., Amon, M. J., Lajoie, S., & Azevedo, R. (2023). Measuring multidimensional facets of SRL engagement with multimodal data. In V. Kovanovic, R. Azevedo, D. C. Gibson, & D. Lfenthaler (Eds.), *Unobtrusive Observations of Learning in Digital Environments* (pp. 141–173). Springer International Publishing. https://doi.org/10.1007/978-3-031-30992-2_10

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

Woodward, M. (2024, Feb. 15). *Minecraft user statistics: How many people play Minecraft in 2024?*. Search Logistics. https://www.searchlogistics.com/learn/statistics/minecraft-user-statistics/

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Wouters, P., & van Oostendorp, H. (Eds.). (2017). *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*. Springer International Publishing. https://doi.org/10.1007/978-3-319-39298-1

Yaden, J. (28 January, 2021). *The bestselling games of all time.* Digital Trends. https://www.digitaltrends.com/gaming/bestselling-games-of-all-time/

Yi, S. (2021). The impacts of a science-based videogame intervention on interest in STEM for adolescent learners. 30684736. [Doctoral dissertation, University of Illinois Urbana-Champaign].

Yi, S., Gadbury, M., & Lane, H. C. (2021, June). Identifying and coding STEM interest triggers in a summer camp. In *15th International Conference of the Learning Sciences (ICLS) 2021*.

Zepeda, C. D., Martin, R. S., & Butler, A. C. (2020). Motivational strategies to engage learners in desirable difficulties. *Journal of Applied Research in Memory and Cognition*, *9*(4), 468–474. https://doi.org/10.1016/j.jarmac.2020.08.007

Zhang, M., Du, X., Hung, J.-L., Li, H., Liu, M., & Tang, H. (2022). Analyzing and interpreting students' self-regulated learning patterns combining time-series feature extraction, segmentation, and clustering. *Journal of Educational Computing Research*, *60*(5), 1130–1165. https://doi.org/10.1177/07356331211065097

Zheng, R. Z., & Gardner, M. K. (Eds.). (2017). *Handbook of Research on Serious Games for Educational Applications:* IGI Global. https://doi.org/10.4018/978-1-5225-0513-6

APPENDIX A: STEM INTEREST, ASTRONOMY INTEREST, & MINECRAFT INTEREST SURVEY ITEMS

STEM Interest Items

Please tell us how interested you are in these general activities:

Viewing maps of cities when visiting them

Rearranging spaces to make them more efficient and interesting

Understanding why buildings were constructed the way they were

Learning how to make slime in science class

Experimenting with different proportions in a cooking recipe

Learning about how different plastics, resins, and powders change during 3D printing

Finding the boiling point of different liquids

Learning what happens when a plumbing system breaks

Observing traffic patterns and finding solutions for congestion

Figuring out the best place to build a bridge

Naming as many countries on an unmarked globe as I can

Learning how different ecosystems work

Studying why earthquakes happen

Dating the age of rocks based on their composition

Understanding what makes a volcano erupt

Learning why the continents are shaped the way they are

Using numbers to confirm ideas and solve problems

Programming a robot through trial and error

Figuring out how to protect eggs from a high fall

Determining the force required to karate chop a wooden board

Astronomy Interest Items

Please indicate how much you agree or disagree with the following statements. -

I think Astronomy is interesting.

I like thinking about "what if" questions, such as "what if the earth had two moons".

I want to know more about exoplanets.

I search for information about space in my free time.

I would consider working in a space-related job someday.

Minecraft Interest Items

Please tell us how interested you are in these Minecraft activities. -

Finding or capturing certain kinds of animals and plants

Using a combination of decorative and structural blocks

Planning, renovating, growing, or protecting villages with structures

Brewing many kinds of potions

Crafting foods and making/using dyes

Learning which foods fill me up the most and last the longest

Learning what colors each biome has and why

Building underwater

Making creeper/fire-proof houses

Creating systems to transport resources

Conserving or reducing use of local resources

Making a tree farm

Exploring to find specific biomes (mesa, flower forest, basalt deltas, etc.)

Repairing or reusing tools to avoid making new ones

Mapping out my world and/or using a compass

Creating timer-based redstone contraptions like traps or doorbells

Using redstone and pistons

Using slime blocks to create automations (walker, flying machine, etc.)

Seeing how far I can get a TNT cannon to shoot

Gliding through an Elytra course, racing boats, or running parkour tracks

APPENDIX B: SELF-EXPLANATION PROMPTS & POSSIBLE ANSWERS FOR SCORING

Lunar Crater Questions	Considering measurements you took, explain at least 2 reasons you must wear a spacesuit on the moon.	Why would a psychologist be needed at a moon base?	How would you define habitability?
Possible Answers	They take measurements for pressure, temperature, oxygen and airflow; explanations would relate to the lack of air or the extreme cold	The explanation given in the game is to help people take care of their mental health while isolated for long periods of time	Many ways to interpret this, we expect learners to talk about science variable related conditions like atmosphere or radiation, but could also include food or water
No Moon Questions	Why is the wind speed faster without a moon?	What is one consequence you observed of having higher wind speeds?	Could humans survive on No Moon? Why?
Possible Answers	Wind flows are caused primarily by heating from oceans and from the Earth's rotation rate. Faster rotation would lead to different ocean heating patterns and faster winds	Wind turbine, short stubby plants that are more distributed, areas outside of valley stripped of growth	They see evidence that yes they could but this is critical thinking about what people need. If we went there from here, we would have great difficulty surviving because our biological clocks are synched only to 24 hours.
Colder Sun Questions	Why is part of the world always hot and the other always cold?	What might be a good way to generate energy on Colder Sun?	Could humans survive on Colder Sun? Why?
Possible Answers	It is a tidally-locked planet, meaning that one side always faces the star. That side always receives heat from the star, while the opposite (night) side never receives heat from the star.	Solar panels on hot side, or more creative answers like geothermal	Maybe in the habitable strip because it could have appropriate climate conditions and life
Tilted Earth Questions	Why might a polar bear be in the jungle on Tilted Earth?	Why is half the world in complete darkness for 3 months?	Could humans survive on Tilted Earth? Why?

Possible Answers	It's not jungle all year round, much of the year it's cold and still snowy, the scenario implies rapid yearly climate shifts and plant growth	Prompts kids to talk about the way the earth rotates (spins on its axis) always pointing in the same direction while orbiting the Sun. Tilted that far over, the Sun would be below the horizon for long periods.	In our scenario potentially, prompts kids to think about what humans need to survive. Probably would lead to mass migrations to stay warm.
Mynoa Questions	Why are the sides of Mynoa so different?	How is Mynoa different from the other What-If Earths?	Could humans survive on Mynoa? Why?
Possible Answers	They can pick any number of answers related to the influence of Tyran, like eclipse or tidal bulge or how life could develop differently	Many ways, they're probably point out Tryan	In our scenario potentially, prompts kids to think about what humans need to survive
Two Moons Questions	Why are tides higher with two moons?	Besides the number of moons, what is different from No Moon?	Why is one moon's surface covered in volcanoes?
Possible Answers	Each moon contributes to the tides. The closer one, when half the distance as our present moon, will by itself create tides 8 times higher than tides today.	Massive tidal shifts, more light at night	Changing gravitational pulls of Earth and other moon would cause that moon to change shape as it orbits. The resulting friction from changing shape would heat the planet, and volcanoes would form to release the internal liquid rock
Exoplanets Questions	What is an exoplanet?	Why is Kepler red?	Why is the water toxic on Trappist?
Possible Answers	An 🪐 Exoplanet is a planet that orbits a star other than the Sun	It has a different atmospheric composition that filters out a different range of light	It's like the geothermal composition of water found in yellowstone, volcanic and potential bacterial impact
Exoplanets Questions cont.	A brown dwarf is most like what type of planet?	What is happening to the ice on Gliese? Why?	Why is one side of Cancri covered in lava and the other side covered with a rocky surface?

Possible Answers	Gas giant!	High pressure keeps it ice solid even at high temperatures, it is "on fire" as an exothermic reaction, sublimating rapidly	The temperature differences between the two sides, which is a consequence of this exoplanet being tidally locked

APPENDIX C: FULL INTERVIEW PROTOCOL

Interview Protocol 2022

Make sure to ask if it is ok to record their responses. Their responses are confidential and will only be heard by people on the research team. Be sure to let each student know that the interview is voluntary and if there is any question they do not feel comfortable answering they can ask to "pass".

School and Technology

1. How do you like school? Which class is your favorite and why?
2. How do you feel about learning with technology?
3. Have you played any educational video games? Do you feel you can learn from playing video games? Why?

STEM Identity

4. What does science mean to you?
5. Do you see yourself as a science person? Why or why not?

Long-term Interest

6. Has WHIMC camp helped answer any questions you have about space?
7. Is there anything new that interests you about space?
8. Have you watched any YouTube videos about space or Googled any additional information regarding space?
9. Have you visited a planetarium? Would you consider visiting the local planetarium in your free time?
10. Have you talked about any camp activities with friends? If so, what did you talk about?

Minecraft Play

11. How often do you play Minecraft? About how long do you play each time? Where and how do you play it? Why? Do you play alone, or with friends?
12. What's your favorite thing you have ever built in Minecraft? Why?

Astronomy Knowledge

11. Can you tell me about the Moon? How does it affect us here on Earth?

    a. TRANSFER: What do you think Earth would be like if the Moon were twice as large as it is now?

    b. What would Earth be like if we had two moons?

c. What are the "must have's" you'll need to survive on the moon?

12. Can you tell me about what the Earth's axis of rotation is (in degrees)? Why is it important?

a. TRANSFER: What do you think Earth would be like if it didn't rotate?

13. Can you describe what a "habitable zone" is? A "habitat" is a natural home of an animal, plant, or another living organism like you and me.

    a. TRANSFER: What do you think would happen if the Earth, as it is now, were closer to the sun, like Venus? Follow-up: Why would we not be able to survive?

14. What is an exoplanet? Why might the color of the [exoplanet] landscape be different from what we see on Earth?

Camp Feedback

15. How did you feel about interacting with the agent?

16. What aspects did you find helpful/fun? What was not helpful/fun?

17. What would you like to see the agent be able to do?

18. If you could choose to go more in-depth about any of the hypothetical worlds we explored, which one would you choose and why?

19. What is a space-related "what-if question" you have?

20. Would you be interested in playing on our public server after the camp?

21. (for those applicable) Would you want to join our Minecraft camp again if we offer it again?

APPENDIX D: EXAMPLE WHIMC SUMMER CAMP SCHEDULE 2022

Day 1: Introductions, Rocket Launch & Moon base

12:00 p.m. - Preparation

Computer setup

1:00 – 1:10 p.m. - Research: Surveys – Minecraft & STEM surveys (Have survey up on screen before kids sit down).

1:10 – 1:30 p.m. Presentation: Welcome and introductions

What-if question (examples given first)

What do you want to do when you grow up?

What is WHIMC, why do we do research? You are helping us!

Is Earth special? (show ISS video; explain in terms of astronomy, but also climate change)

What is STEM (or STEAM?) What does NASA do? Who works there?

Discussion: Camp ground rules

Same computer each week

Be respectful of others

1:30 – 1:50 p.m. Play: Rocket launch

1:50 – 2:00 p.m. - Research: Astronomy Knowledge Baseline Quiz

2:00 – 2:10 p.m. Discussion: What was the coolest part of Rocket Launch?

2:10 – 2:20 Presentation: Tutorial: Observe Like a Scientist

2:20 – 2:40 BREAK

2:40 – 3:00 Activity: Take measurements outside with barometers, UV cards

3:00 – 3:20 p.m. Play: Lunar crater, Explore Hub if time permits

3:20 – 3:30 p.m. Assessment: LC Self-explanations

3:30 – 3:45 p.m. Play: Baseline Earth

3:45 – 4:00 p.m Activity: Design skins

Day 2: No Moon, Colder Sun, & Tilted Earth

1:05 – 1:15 p.m. Presentation: No Moon

1:15 – 1:35 p.m. Play: Earth with No Moon

1:35 – 1:40 p.m. Assessment: Self-explanations No Moon

1:40 – 1:50 p.m. Discussion: What did you notice about Earth with no moon?

Presentation: No Moon wrap up

1:50 – 2:00 p.m. Presentation: Colder Sun

2:00 – 2:20 p.m. Play: Colder Sun

2:20 – 2:25 p.m. Assessment: Self-explanations CS

2:25 – 2:45 p.m. BREAK

2:50 – 3:00 p.m. Discussion: What did you notice about Earth with a Colder Sun

Presentation: Colder Sun wrap up

3:00 – 3:10 p.m. Presentation: Tilted Earth

3:10 – 3:30 p.m. Play: Tilted Earth

3:30 – 3:35 p.m. Assessment: Tilted Earth SEs

3:35 – 3:45 p.m. Discussion : What did you notice about Tilted Earth?

Presentation: Tilted Earth wrap up

3:45 – 4:00 p.m. Play: Free play on build world (exoplanet contest world)

Day 3: Mynoa, Two Moons, Exoplanets

1:00 – 1:05 p.m. Set up & handout posters

1:05 – 1:15 p.m. Presentation: What if Earth was a moon (Mynoa)

1:15 – 1:35 p.m. Play: Mynoa

1:35 – 1:40 p.m. Assessment: Self-explanations Mynoa

1:40 – 1:50 p.m. Discussion: What did you notice about Mynoa?

Presentation: Mynoa wrap up

1:50 – 2:00 p.m. Presentation: Two Moons

2:00 – 2:20 p.m. Play: Two Moons

2:20 – 2:25 p.m. Assessment: Self-explanations Two Moons

2:25 – 2:45 p.m. BREAK

2:50 – 3:00 p.m. Discussion: What did you notice about Two Moons

Presentation: Two Moons wrap up

3:00 – 3:10 p.m. Presentation: Exoplanets, methods for finding exoplanets

3:10 – 3:20 p.m. Presentation: AI Agent Intro

3:20 – 3:40 p.m. Play: Trappist, Gliese, Kepler, Cancri, Brown Dwarf

3:40 – 3:45 p.m. Assessment: Exoplanet SEs

3:45 – 4:00 p.m. Design: Draw an exoplanet and features of that world. Let students present exoplanets, if time permits

Day 4: Habitation Challenge

1:00 – 1:10 p.m. Settle down & handout posters

1:10 – 1:30 Presentation: What is AI and how is it developed?
You guys are helping us to develop a teaching/learning agent!
AI Explanation with CPG Gray
Talk to Jay on Zoom; use linktr.ee/whimc for kids to go to the data input page

1:30 – 1:50 Presentation: Habitation challenge

1:50 – 2:10 p.m. Design: Draw your Habitat!

2:10 – 2:30 p.m. BREAK

2:30 – 4:00 p.m. Play: Build Habitats on Mars!

Research: Semi-structured Interviews

Day 5:

Habitation Continued

1:05 – 1:15 p.m. Research: STEM and Minecraft Survey

1:15 – 1:30 p.m. Discussion: Review survival concerns, presentation checklist

1:30 – 2:30 p.m. Design: Continue habitats on Mars

2:30 – 2:50 p.m. BREAK

2:50 – 3:00 p.m. Research: Knowledge Assessment

3:00 – 3:30 p.m. Design: Continue habitats on Mars

3:30 – 4:00 Showcase: Living on Mars!

Research: Semi-structured Interviews

APPENDIX E: ASTRONOMY KNOWLEDGE ASSESSMENT SCORING RUBRIC

Scores:

0 - "I don't know" or response is off-topic
1 - answers prompt but answer is vague or generally incorrect
2 - answers prompt with one or more key ideas indicated by black bullets. May not fully answer the question. Partial credit.
3 - answers prompt with examples and provides reasoning why it is the case or what would happen - inference (one or more uses of white bullets).

Q1: Can you tell me about the Moon? (Focus on this question) → How does it affect us here on Earth?

- Tides
  - o Result of gravitational pull of moon
  - o High tide on side facing moon and opposite side
  - o Two high tides and low tides at all times as Earth spins
  - o Pulls moon forward, giving moon energy, causes moon to spiral away
- Earth loses energy it gives to moon by slowing down
  - o Days are getting longer
  - o Reduced wind speed
  - o This allows humans and nature to grow upwards
- Light at night for nocturnal animals

Q2 TRANSFER: What do you think Earth would be like if the Moon were twice as large as it is now?

- Gravitational force of moon on Earth stronger.
- Tides higher
  - o Threatening to coastal cities
- Rotation slowed even more
  - o Days and nights longer
- Brighter nights
  - o Less stars visible

Q3: What would Earth be like if we had two moons?

1. Days would be longer
2. Tidal variance would be greatly increased
3. Night would be a lot brighter (you can read a book outdoors at night!)
4. Eventually the moons would collide, creating a ring around the Earth.

o With two moons there would be a greater gravitational force on the Earth pulling the tides and slowing rotation.

Q4: What are the "must have's" you'll need to survive on the moon? (Easy: Max 2 points)

- Shelter from radiation
- Dirt, seeds, water for farming
- Source for electricity

- Oxygen supply
- Suit and helmet
- Air pressure
- Regulated temperature
- Communication
- Transportation
- Medical supplies
- Other people
- Entertainment

Q5: Can you tell me what the Earth's axis of rotation is? (Focus on this question) → Why is it important?

- Line from North Pole to South Pole that Earth rotates around
- 23.5 degree tilt
- Results in seasons
  - o Seasons are not proximity but amount of light hitting Earth's surface and for how long
  - o Summer sees more concentrated sunlight for longer times
  - o More concentrated light provides more heat
- Results in day and night
  - o If tilted more parts of earth would not see sun for up to six months at a time

Q6: TRANSFER: What Earth you think Earth would be like if it didn't rotate?

- Every place on earth would have months of only darkness and months of only light
  - o Habitability would be difficult on dark side.

Q7: Can you describe what a habitable zone is?

- Conditions are suitable for life "livable"
  - o Any mention of the right distance from star, temperature, oxygen levels, pressure, etc…

Q8: What do you think would happen if the Earth was closer to the sun, like Venus? Why would we not be able to survive?

- We are at a distance where temperatures are not extreme
  - o If closer, temperatures would be unbearable for life
- We are at a distance where liquid does not evaporate
  - o If closer, oceans would evaporate because of heat

Q9: What is an exoplanet? (Easy: 0 for idk; 1 for wrong; 2 for correct)

- An exoplanet is a planet orbiting a star different from our own.

Q10: Why might the color of an exoplanet landscape be different from what we see on Earth? (Hard: 3 points for one or more bullets).

- o The color of the start the planet orbits may be different than ours (i.e. red dwarf).
- o The atmosphere of the planet may spread light particles in a differently (wavelengths) than our atmosphere.

APPENDIX F: CLUSTER VALIDITY INDICES

Sil = Silhouette (maximize), SF = Score Function (maximize), CH = Calinski-Harabasz Index (maximize), DB = Davies-Bouldin Index (minimize), DBstar = modified Davies – Bouldin Index (minimize), D = Dunn Index (maximize), COP = COP Index (minimize)

Exploration Clusters CVI

Sil	0.12428280	0.12614420	0.09578725	0.03697897	0.07606464
SF	0.00061280	0.00000424	0.00000017	0.00000001	0.00000000
CH	62.13934000	23.49275000	26.22972000	20.50850000	15.76888000
DB	1.56394800	2.20311000	2.06425600	1.60168100	2.28375900
DBstar	1.56394800	2.28452500	2.30741500	1.80570000	2.52900600
D	0.11330510	0.11031460	0.09843364	0.09160607	0.11635250
COP	0.51606390	0.41685390	0.38387920	0.41205510	0.34089640

Observation Clusters CVI

Sil	0.40269456	0.31270460	0.29112510	0.16099990	0.26603670
SF	0.00202981	0.00058460	0.00002507	0.00000677	0.00000046
CH	24.42692336	43.15898000	29.51455000	16.57096000	22.48151000
DB	1.18514271	1.14052800	1.19591900	2.53931600	1.32263800
DBstar	1.18514271	1.28033700	1.59507500	4.96910700	1.89602700
D	0.06310513	0.05026945	0.06343594	0.01533677	0.06871507
COP	0.27133504	0.19521640	0.17934440	0.19859160	0.14813550

Tools Clusters CVI

Sil	0.44595162	0.29630590	0.12036030	0.13870320	0.10965070
SF	0.00695430	0.00020339	0.00005326	0.00002810	0.00000002
CH	57.20877651	35.81963000	24.75375000	20.35093000	14.81024000
DB	1.20404506	1.36625800	1.68012900	1.95561900	2.33634400
DBstar	1.20404506	2.17098400	3.91638300	6.05981900	6.60999400
D	0.02308281	0.00826239	0.01246446	0.01246446	0.00964445
COP	0.31647824	0.19113290	0.18583020	0.17861770	0.10752070

Nodes Clusters CVI

Sil	0.29849979	0.14339890	0.12889500	0.12825290	0.08754844
SF	0.00146638	0.00001340	0.00000092	0.00000001	0.00000000
CH	68.02162004	24.10970000	26.00942000	18.99705000	17.72658000
DB	1.31032725	2.13696800	1.71207600	1.96523800	1.58287500
DBstar	1.31032725	2.25087200	1.85863400	2.25145800	1.69281400
D	0.07576973	0.08296868	0.09386474	0.13443420	0.10700810
COP	0.41022008	0.29989770	0.28776700	0.25771450	0.22755920

APPENDIX G: BAYESIAN STATISTICS FOR BMA

STEM Interest Models

```
             p!=0    EV         SD        model 1     model 2     model 3     model 4     model 5
  Intercept  100.0  -7.821e-02  1.578478       .      -9.727e-01  -5.935e-01  -4.294e-16   1.282e-16
  se_sum       7.3   1.206e-02  0.060823       .           .           .           .       1.398e-01
  obs_sum      3.6  -1.719e-03  0.016414       .           .           .           .           .
  nodes_sum    9.0   1.521e-02  0.067421       .           .           .           .           .
  exp_sum      2.9  -9.379e-06  0.006738       .           .           .           .           .
  kqs_int      8.5   3.202e-02  0.143639       .           .           .       3.010e-01       .
  tools_sum    3.3   1.253e-03  0.015149       .           .           .           .           .
  engBH        8.1  -3.225e-01  1.523778       .           .           .           .           .
  engBL       13.3   6.823e-01  2.218809       .       3.696e+00       .           .           .
  engEXP       2.9   3.224e-03  0.763505       .           .           .           .           .
  engINT       2.9  -7.632e-03  0.694996       .           .           .           .           .
  campma       3.4  -7.562e-02  0.848413       .           .           .           .           .
  campmp       3.8  -1.108e-01  0.965249       .           .           .           .           .
  campsdp      4.0  -1.056e-01  0.835409       .           .           .           .           .
  campsk       3.7  -9.800e-02  0.922190       .           .           .           .           .
  campub       2.9  -8.256e-03  0.727202       .           .           .           .           .
  campug       8.2   4.620e-01  1.994636       .           .       5.012e+00       .           .

 nVar                                          -1          1           1           1           1
 r2                                          0.000       0.019       0.019       0.012       0.012
 BIC                                      0.000e+00   2.867e+00   2.883e+00   3.414e+00   3.447e+00
 post prob                                   0.255       0.061       0.060       0.046       0.045
```

Minecraft Interest Models

```
  22  models were selected
 Best  5  models (cumulative posterior probability =  0.4763 ):
           p!=0    EV         SD        model 1     model 2     model 3     model 4     model 5
Intercept  100.0  -2.702e-02  1.687913  -5.682e-15   7.021e-01  -6.113e-15  -6.346e-15  -5.332e-15
se_sum      81.9   3.747e-01  0.238790   4.694e-01   4.827e-01       .       4.018e-01   4.089e-01
obs_sum      5.2   3.507e-03  0.028124       .           .           .           .           .
nodes_sum    8.4   1.353e-02  0.058238       .           .           .           .       1.245e-01
exp_sum      2.9   2.656e-05  0.007799       .           .           .           .           .
kqs_int     11.1   5.979e-02  0.215974       .           .           .       4.219e-01       .
tools_sum    3.0   6.340e-04  0.015402       .           .           .           .           .
engBH        6.2   2.379e-01  1.393749       .           .           .           .           .
engBL        3.4  -6.650e-02  0.751341       .           .           .           .           .
engEXP       3.3  -7.635e-02  1.010634       .           .           .           .           .
engINT       3.1   4.615e-02  0.871735       .           .           .           .           .
campma       3.0   2.452e-02  0.828405       .           .           .           .           .
campmp       7.5  -4.106e-01  1.918803       .      -5.336e+00       .           .           .
campsdp      4.2   1.326e-01  0.996747       .           .           .           .           .
campsk       2.9  -6.464e-03  0.818100       .           .           .           .           .
campub       3.0  -4.109e-02  0.922283       .           .           .           .           .
campug       3.0   3.768e-02  0.870951       .           .           .           .           .

nVar                                       1           2           0           2           2
r2                                       0.092       0.108       0.000       0.107       0.105
BIC                                     -2.997e+00  -5.604e-02   0.000e+00   9.298e-02   2.638e-01
post prob                                0.256       0.059       0.057       0.055       0.050
```

Astronomy Interest Models

```
 27  models were selected
 Best  5  models (cumulative posterior probability =  0.4721 ):
           p!=0    EV         SD        model 1     model 2     model 3     model 4     model 5
Intercept  100.0  -2.123e-02  0.467863   5.547e-16   7.071e-16   6.294e-16  -2.988e-01   4.823e-16
se_sum      27.9   2.459e-02  0.048557       .       7.859e-02        .           .           .
obs_sum      3.9  -9.525e-05  0.005557       .           .            .           .           .
nodes_sum    4.7  -1.171e-03  0.009554       .           .            .           .           .
exp_sum      8.7  -1.447e-03  0.006053       .           .      -1.629e-02        .           .
kqs_int     93.9   2.915e-01  0.130269   3.165e-01   2.602e-01   3.502e-01   3.468e-01   3.563e-01
tools_sum    6.3  -1.623e-03  0.008954       .           .            .           .      -2.543e-02
engBH        3.9  -4.706e-03  0.216338       .           .            .           .           .
engBL        7.2   8.363e-02  0.407716       .           .            .      1.135e+00        .
engEXP       2.9   1.601e-02  0.261445       .           .            .           .           .
engINT       4.2   2.837e-02  0.308249       .           .            .           .           .
campma       2.9   1.485e-02  0.237901       .           .            .           .           .
campmp       2.7  -7.764e-03  0.222018       .           .            .           .           .
campsdp      2.9  -1.178e-02  0.192242       .           .            .           .           .
campsk       4.1   2.347e-02  0.294658       .           .            .           .           .
campub       5.1  -5.861e-02  0.414369       .           .            .           .           .
campug       2.7   6.116e-03  0.241943       .           .            .           .           .

nVar                                       1           2           2           2           2
r2                                       0.115       0.143       0.133       0.129       0.127
BIC                                     -4.927e+00  -3.038e+00  -2.217e+00  -1.853e+00  -1.619e+00
post prob                                0.230       0.089       0.059       0.049       0.044
```

Knowledge Question Models

```
  p!=0    EV         SD         model 1    model 2    model 3    model 4    model 5
Intercept  100.0   1.0265515  0.738494    0.92962    1.80189    0.44471    0.93965    1.06435
obs_sum     55.9   0.0387108  0.040703    0.06882       .       0.09060       .          .
nodes_sum    9.4   0.0049628  0.019987       .          .          .          .          .
exp_sum      3.5   0.0003124  0.002979       .          .          .          .          .
tools_sum   47.6   0.0345217  0.043099       .       0.09235       .       0.05362    0.05809
engBH        3.2  -0.0178277  0.261876       .          .          .          .          .
engBL        3.4  -0.0224536  0.226832       .          .          .          .          .
engEXP       2.3   0.0115083  0.225517       .          .          .          .          .
engINT      14.0   0.2357972  0.765884       .          .          .          .          .
campma       2.4  -0.0106470  0.207102       .          .          .          .          .
campmp       3.8   0.0253026  0.296153       .          .          .          .          .
campsdp     28.2  -0.7520615  1.395850       .      -2.82692       .          .          .
campsk       4.5  -0.0358179  0.378780       .          .          .          .          .
campub      78.6  -2.9763178  2.048916   -3.39497   -4.64496       .      -3.33820   -4.20645
campug      97.9  -4.4270154  1.573247   -4.45519   -5.23130   -3.75531   -4.59664   -4.78137
se_sum      25.0   0.0236571  0.049535       .          .          .       0.09786       .

nVar                                         3          4          2          4          3
r2                                        0.337      0.368      0.282      0.357      0.320
BIC                                     -18.29849  -17.56049  -16.51731  -16.26311  -16.26229
post prob                                 0.113      0.078      0.046      0.041      0.041
```

Self-Explanation Models

```
 23  models were selected
 Best  5  models (cumulative posterior probability =  0.5571 ):
           p!=0    EV         SD        model 1    model 2    model 3    model 4    model 5
Intercept  100.0   1.3766196  1.380249    1.22852    1.38932    0.78206    2.05502    1.67238
obs_sum     24.3   0.0206830  0.046052      .        0.08124      .          .          .
nodes_sum    7.0   0.0045635  0.028246      .          .          .          .          .
exp_sum      6.7   0.0017159  0.009542      .          .          .          .        0.02689
tools_sum    4.5  -0.0005029  0.014560      .          .          .          .          .
engBH      100.0   9.7174697  2.304914   10.27538    8.43578   10.62595    9.71627    9.56335
engBL        3.9  -0.0532416  0.481454      .          .          .          .          .
engEXP      15.5   0.5884219  1.737010      .          .        3.58816      .          .
engINT       3.6  -0.0516074  0.559988      .          .          .          .          .
campma       7.8  -0.1912208  0.998876      .          .          .          .          .
campmp       5.0  -0.0697761  0.673514      .          .          .          .          .
campsdp    100.0  -9.3134930  2.293555   -9.31228   -8.80001   -9.09266   -9.77701  -10.03719
campsk      96.7  -7.5756171  2.900322   -7.80869   -6.96754   -8.11492   -8.57928   -8.09969
campub     100.0  -9.0575677  2.805605   -9.22560   -7.98887   -9.57651  -10.05210   -9.54059
campug      11.6  -0.3689176  1.390683      .          .          .       -3.05210      .

nVar                                        4          5          5          5          5
r2                                        0.447      0.464      0.461      0.457      0.455
BIC                                     -27.72400  -25.72027  -25.29761  -24.75123  -24.45092
post prob                                 0.267      0.098      0.079      0.060      0.052
```