International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

Exploring Player Archetypes in a Minecraft-Based Learning Environment

Andres Felipe Zambrano[1], Zhanlan Wei[1], Jaclyn Ocumpaugh[2], Amanda Barany[1], Ryan S. Baker[3], Xiner Liu[1], Yiqiu Zhou[4], Luc Paquette[4], and Jeffrey Ginger[4]

[1]Graduate School of Education, University of Pennsylvania, Philadelphia, United States; [2]College of Education, University of Houston, Houston, United States; [3]Education Futures, Adelaide University, Adelaide, Australia; [4]College of Education, University of Illinois at Urbana-Champaign, Champaign, United States
{azamb13, zhanlanw, xiner} @upenn.edu; {jlocumpaugh, amanda.barany, ryanshaunbaker, luc.paquette} @gmail.com; {yiqiuz3, ginger} @illinois.edu

Keywords:

Student motivation
Player archetypes
Player typologies
Educational games
Minecraft-based learning
Science interest

Received: XXXX
Accepted: XXXX
Published: XXXX
DOI: 10.17083/ijsg.XXXX

Abstract

Games are recognized as valuable tools for learning and interest development. However, the association between behavioral player archetypes and these important outcomes is still underexplored. This study explores the relationship between learning, interest development, and player archetypes (ROAMERS, EXPLORERS, and SCIENTISTS) identified within the WHIMC project, a game-based learning environment where students engage with scientifically accurate hypothetical astronomy scenarios in Minecraft. Grounded in human-computer interaction and player typologies frameworks, we analyze data from 57 participants across four summer camps using Ordered Network Analysis (ONA) and k-means clustering to identify player archetypes emerging from student actions. We then examine how these archetypes relate to learning outcomes and motivational factors. Statistical tests reveal significant differences in in-game actions across archetypes and correlations between player behaviors and learning outcomes. These findings contribute to the design of serious educational games by increasing understanding of how to optimize experiences and enhance science engagement for learners with differing playing styles.

1. Introduction

In well-designed serious games for education, students actively engage with skills, knowledge, identities, values, and practices that contribute to desirable learning outcomes or serve as foundational steps toward educational success [1], [2]. Over the past decade, several meta-analyses and systematic reviews have presented increasingly consistent evidence that games can support student learning [3], [4], [5], [6], [7], [8], often finding the strongest effects on cognitive outcomes (e.g., knowledge acquisition, concept mastery, and test performance), but also finding improvements to affective and motivational outcomes within the game's specific domain, as well as gains in social skills and teamwork [3], [4], [5].

Despite the promising potential of video games for learning, research on who benefits the most—and why—remains limited. Many studies in this area approach the learners in the experimental group as a single population, which is useful for investigating the general impact of games on learning outcomes, but can obscure important differences in motivation, play style, prior knowledge, game literacy, and specific in-game behaviors (see discussion in [9]). These factors can substantially influence students' learning outcomes. Although some studies have examined differences based on traditional demographic categories such as race and gender (e.g., [10], [11], [12], [13]), these factors alone do not capture students' situational interest, prior knowledge, or specific in-game behaviors—all of which may significantly influence learning outcomes (e.g., [14], [15]). Understanding these nuanced differences is critical for designing educational games that adapt to diverse learner and player profiles, optimize all learners' engagement, and ensure that the intended learning benefits are equitably realized.

Research in human-computer interaction (HCI) and video games has introduced various player typologies to explain behavior based on the motivational factors that drive gameplay [16], [17], [18], [19]. These typologies help distinguish whether a player is primarily motivated by social interaction [16], the challenge of completing a game [19], the desire to explore a virtual world or immersive narrative [16], or the enjoyment of learning new knowledge presented in the game [17]. Such motivations shape how students engage with a game and may predispose them to benefit more from certain types of learning opportunities.

Recent studies have applied data mining techniques to identify player archetypes in educational contexts [15], [20], [21], using methods like Epistemic Network Analysis (ENA; [22]) and clustering to both quantify sequences of student actions during gameplay and uncover distinct archetypes in a science inquiry game [15]. These analyses revealed meaningful associations between player archetypes, interest levels, and post-test performance, highlighting the potential of archetype-based approaches to enhance educational game design.

Building on the need to better understand how individual differences in gameplay relate to student interest and learning, in this study, we investigate player archetypes within the WHIMC project—a Minecraft-based serious game where students explore scientifically accurate worlds to investigate hypothetical astronomy questions. Specifically, we address the following research questions: **(RQ1)** What player archetypes emerge within WHIMC? **(RQ2)** How are these archetypes associated with students' interest measures and learning gains? **(RQ3)** How do our findings align with prior research on player archetypes in serious educational games? This work seeks to contribute to the growing literature analyzing relationships between student interest and in-game behaviors. By explicitly linking player archetypes to both interest and learning outcomes, this work aims to advance the growing literature on the interplay between student engagement, in-game behaviors, and learning in educational game contexts.

2. Related Work

2.1 Frameworks of Player Typology

Research has explored how players engage with video games, generating typologies that categorize player motivations [16], [17], [19]. These frameworks aim to provide a structured understanding of why individuals play games and how different motivational factors shape their interactions (see Figure 1).

One of the earliest and most influential models is Bartle's [16] player typology, originally developed for multi-user dungeon games (MUDs). Based on player motivations, Bartle identified four archetypes: ACHIEVERS, who focus on in-game goals, progression, and mastery, often motivated by points and rewards; EXPLORERS, who seek discovery and experimentation, engaging deeply with game mechanics, lore, and hidden elements; SOCIALIZERS, who prioritize interaction and relationship-building, often participating in cooperative play and communication; and KILLERS, who thrive on competition and dominance, favoring player-versus-player activities and skill demonstrations. Expanding on Bartle's framework, Yee [19], [23] proposed ten specific motivational factors that influence player behavior (e.g., competition, teamwork, discovery, and role-playing). Yee used confirmatory factor analysis across multiple MUDs to identify three motivational categories: ACHIEVEMENT, encompassing both ACHIEVERS and KILLERS due to their shared focus on progression and competition; IMMERSION, which includes role-playing and exploration (aligned with Bartle's EXPLORERS); and SOCIALIZING, corresponding to Bartle's SOCIALIZERS.

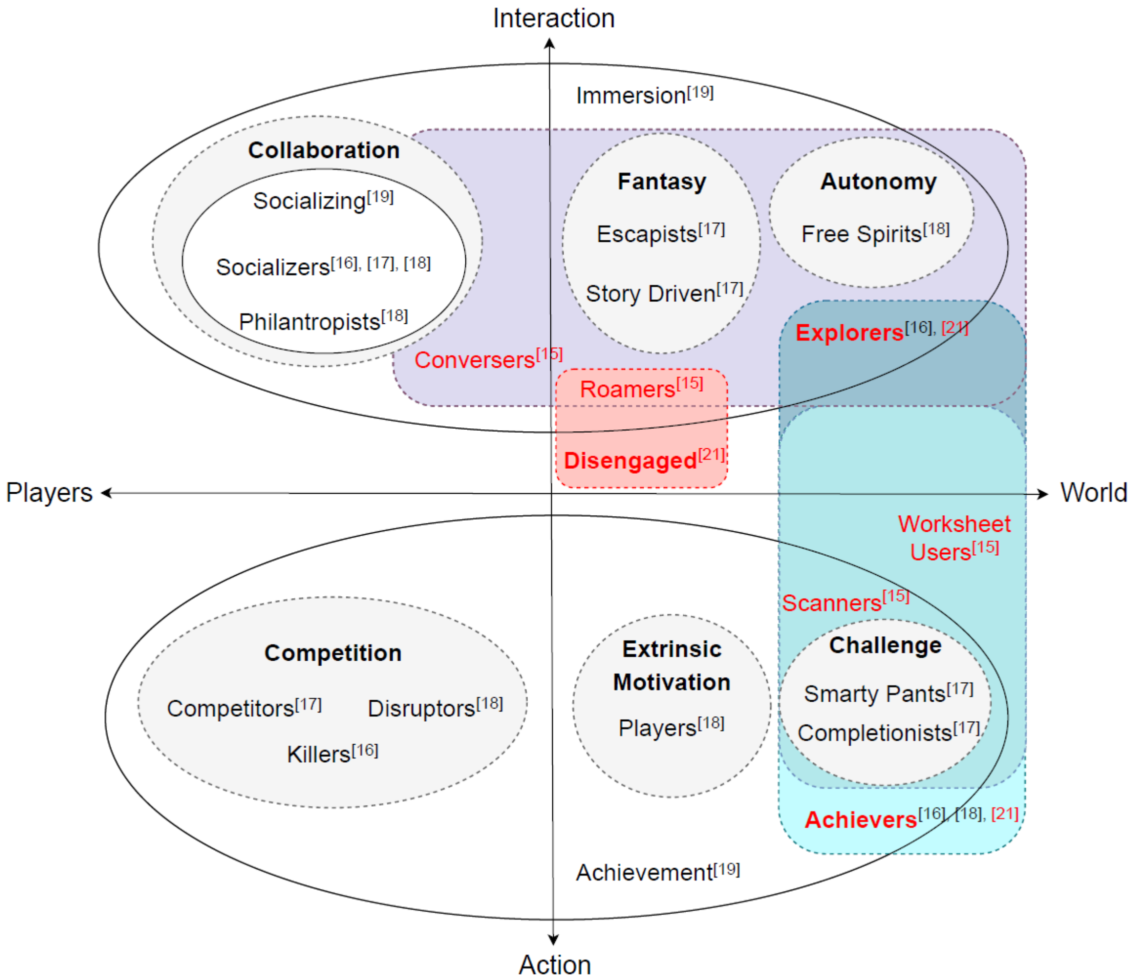

Figure 1. Summary of multiple player typologies mapped to the axes from Bartle's [16] foundational categorization. Black font and red font distinguish between motivational archetypes proposed in HCI frameworks [16], [17], [18], [19] and behavioral archetypes (e.g., differences in in-game actions) identified in serious educational games [15], [21]. Solid ellipses represent Yee's [19], [23] three core motivational archetypes. Gray ellipses with dotted lines are labeled in boldface with Tondello et al.'s [18] six motivational archetypes. Rounded rectangles align conceptually similar behavioral archetypes: red shading indicates students who appear disengaged from the game; purple represents students who explore and show moderate engagement, though not necessarily in science-related activities; blue shows students who repeatedly engage with specific game mechanics and achieve game goals while often performing science-related actions (with the lightest blue corresponding to students primarily focused on in-game achievements); and darkest blue represents the intersection of the purple and light blue groups.

Although Bartle's and Yee's frameworks offer valuable insights into player behaviors and motivations, these frameworks have been criticized for oversimplifying motivations and assuming broad generalizability, despite being developed primarily from data on MUD players. To address these limitations, researchers have proposed alternative typologies that extend to other game genres and capture a broader range of motivations. For instance, Kahn et al. [17] conducted a confirmatory factor analysis of players of multiplayer online battle arena (MOBA) and massively multiplayer online (MMO) games, generating a six-category typology that introduced four new player archetypes that can still be mapped into Bartle's and Yee's frameworks. ESCAPISTS, who play to relax and disconnect, and STORY DRIVEN players, who are deeply engaged in the game's narrative, resemble Bartle's EXPLORERS through their high immersion, but extend this category by differentiating between these two different drivers for exploration and engagement. COMPLETIONISTS, motivated by mastering game mechanics, and SMARTY PANTS players, who see games as intellectual challenges that promote learning and cognitive growth (e.g., puzzle game players seeking to sharpen analytical skills), align with Bartle's ACHIEVERS but are more clearly driven by intrinsic motivation and a desire for high-level challenge. Similarly, Tondello et al. [18] expanded previous models with the HEXAD archetypes, adding categories such as PHILANTHROPISTS, who enjoy helping others and sharing knowledge, and DISRUPTORS, who take pleasure in subverting game mechanics and exploring alternative ways to play, often disturbing other players when realizing these goals. These categories parallel Bartle's ACHIEVERS or KILLERS but specify distinct motivators such as altruism and a desire to test a system's boundaries, thereby refining the broader original categories. This typology was validated through survey data and has been widely used in gamification and serious game research.

Players' underlying motivations can offer valuable insights into how they engage with serious games. In educational games, in particular, some students may align more naturally with certain learning objectives depending on their motivations for playing. For example, SOCIALIZERS and PHILANTHROPISTS may benefit most from collaboration and interaction, strengthening teamwork skills. EXPLORERS, driven by curiosity about the virtual world or narrative, may develop greater motivation or interest in the academic content embedded in the storyline. Meanwhile, ACHIEVERS and SMARTY PANTS, focused on progression and intellectual challenge, may maximize cognitive outcomes by striving for mastery. Understanding how player archetypes relate to learning outcomes can inform the design of more effective educational games and interventions—allowing experiences to be tailored to students' diverse traits and leveraging intrinsic motivation to improve learning.

2.2 Logs Analysis in Serious Educational Games

Another approach to clustering students involves analyzing the specific actions they take within the game. This shifts the concept of archetypes from a motivational to a behavioral perspective (the definition adopted in this study) based on the idea that similar in-game behaviors may achieve similar learning outcomes, regardless of what motivated those actions. For instance, in a game designed to foster interest through interactions with non-playable characters (NPCs), EXPLORERS, SOCIALIZERS, and ACHIEVERS may all benefit from these interactions, even if their initial motivations differ. To give another example, if disciplinary content is delivered through optional in-game notes that are not required for progression or social interaction, SOCIALIZERS and ACHIEVERS may overlook them, missing valuable learning opportunities.

Focusing on specific in-game actions opens two valuable avenues for research. First, researchers can examine associations between specific actions and learning gains. Multiple studies have explored these relationships using techniques such as correlation mining, sequence analysis, and clustering analysis (e.g., [24], [25], [26], [27]). For example, Novolsetseva et al. [27] identified multiple behavioral clusters based on students' in-game actions. They found that students who appeared to develop a strategy frequently inspected and analyzed in-game supportive materials, and engaged in a higher number of purposeful actions—behaviors resembling those of systematic EXPLORERS. These students tended to achieve better outcomes than those who lacked a clear strategy or explored less.

Kang et al. [24] also used clustering to investigate how performance relates to the similarity between an student's actions and those of their group members. They showed that students whose behaviors were either too similar (like Bartle's and Tondello et al.'s SOCIALIZERS) or too dissimilar from their peers (like Tondello et al.'s DISRUPTORS or Bartle's KILLERS) performed more poorly, whereas those who had moderate similarity to their peers (like Tondello et al.'s PHILANTHROPISTS) achieved greater gains. One possible explanation is that moderate similarity in behaviors allows students to collaborate productively—sharing enough common approaches to coordinate effectively—while still maintaining diverse strategies that promote exploration and complementary skills. In contrast, very high similarity may lead to redundancy or over-reliance, whereas very low similarity can result in misalignment or unproductive competition. Although these studies did not explicitly reference player typologies, their findings indicate that identifying behavioral archetypes and the in-game behaviors that define them may help explain variations in learning outcomes.

The second avenue of research that emerges from analyzing specific in-game actions involves understanding why these behaviors develop. In addition to the motivational typologies described above, other factors that may shape gameplay behavior include student interest in the game's topic or mechanics [25]; prior knowledge of the subject matter [28], [29]; skills with the game's mechanics [15], [30]; self-efficacy [31], [32]; and self-regulation [33]. For instance, Nasiar et al. [28] found that students with lower prior knowledge of the subject explored less and deviated more from the expected sequence of in-game actions than their more knowledgeable peers. If differences in in-game actions are associated with multiple causes (e.g., both motivational typologies and prior knowledge), then analyzing such actions directly could be beneficial.

2.3 Archetypes in Serious Educational Games

Parallels can be drawn between player typology frameworks—which emphasize motivational factors—and students' in-game actions in serious educational games, both of which may influence learning outcomes. However, few studies have attempted to directly distill player archetypes from behavioral patterns or study how archetypes relate to educational outcomes. One early attempt was made by Slater et al. [21], who used clustering analysis to identify distinct player archetypes based on in-game actions in a single-player, level-based game. Their analysis, which conceptualized archetypes the same way we do in this paper, revealed three clear player archetypes: ACHIEVERS, EXPLORERS, and DISENGAGED, each associated with different levels of in-game achievement (the latter showing low levels of both immersion and achievement; see Figure 1). Similarly, Swanson et al. [20] employed clustering to identify archetypes in a single player, resource management game, uncovering groups such as CAPITALISTS, PLANNERS, and INACTIVES, which closely align with Slater et al.'s ACHIEVERS, EXPLORERS, and DISENGAGED categories. However, neither study included external measures of learning or interest, limiting their ability to assess how these archetypes relate to broader educational outcomes.

More recently, in the context of an open world, single player scientific inquiry serious game, Zambrano et al. [15] proposed a combined approach using Clustering Analysis and Ordered Network Analysis (ONA; [34])—a technique derived from Epistemic Network Analysis (ENA; [22]) that quantifies ordered transitions between consecutive actions or codes. In the context of an open-world, role-playing STEM game, Zambrano et al. applied these methods to log data capturing students' in-game interactions. Their analysis identified four distinct player archetypes: ROAMERS, who spent extended time outdoors, engaging only in movement without performing additional actions (an indicator of disengagement); CONVERSERS, who held long conversations with NPCs but did not engage in science-related tasks; SCANNERS, who concentrated on solving the game's central mystery by actively testing hypotheses; and WORKSHEET USERS, who systematically documented their findings throughout gameplay.

These groups differed not only in their gameplay behaviors but also in levels of situational interest, self-efficacy, perceived enjoyment, and post-test scores.

Although the archetypes identified by Zambrano et al. [15] differ from those observed by Slater et al. [21] and Swanson et al. [20]—primarily due to fundamental differences in the games studied—meaningful parallels can still be drawn across the typologies. All three studies identified two broad categories of students: engaged and disengaged. Within the engaged group, each study further distinguished students focused on achievement from those focused on exploration, with the latter consistently representing the largest subgroup. This suggests that while specific archetypes may vary depending on game mechanics, these higher-level behavioral patterns remain consistent across contexts and game genres and are likely shaped by students' initial interest in the game's domain.

By analyzing the primary strategies students use, examining how motivational and cognitive factors—such as initial interest and prior knowledge—relate to different player archetypes, and evaluating how various forms of game interaction influence learning outcomes, game designers can better understand the diverse characteristics of learners. This understanding can inform the design of educational games and the development of targeted interventions tailored to distinct learner profiles. Building on this motivation, the present study aims to identify and study player archetypes within a different educational game, examine their associations with interest development, learning gains, and other motivational measures, and assess the generalizability of the behavioral typologies identified in these three prior studies.

3. Methods

3.1 Educational Context

The data analyzed in this paper come from four 5-day summer camps held in different locations across the United States as part of the WHIMC project [35]. WHIMC uses Minecraft's Java Edition to create simulations that engage learners in exploring hypothetical astronomy scenarios through *What-if* questions, such as "*What if Earth had no moon?*" or "*What if Earth orbited a colder sun?*" During the first three days of the camp, learners were guided by pedagogical agents (NPCs) and human facilitators as they evaluated the habitability of various hypothetical worlds and real exoplanets modeled in Minecraft. Using scientific tools, students measured key habitability factors such as temperature, air pressure, radiation, gravity, and atmospheric composition, and made evidence-based judgments about each world's potential to support life. Their conclusions were grounded in scientific data, direct observation, and prior knowledge of astronomy and environmental science. In the final two days, students explored a Mars map built using real Martian terrain data, where they were tasked with designing and constructing a shelter capable of sustaining human life on Mars.

During the 5-day camp, students completed several motivational and learning assessments (see Table 1). On Day 1, they took an interest development scale [36] and an astronomy knowledge assessment [37]. On Day 3, they completed an astronomy and Minecraft interest survey [35], [37]. On the final day, students completed a situational interest scale [38], a self-efficacy scale [39], and repeated the Day 1 instruments. Notably, for the first of the four summer camps, the situational interest scale and the astronomy/Minecraft interest surveys were not administered; these measures were introduced beginning with the second camp.

A total of 61 students from both urban and rural settings across four different states participated in the summer camps. The sample included 40 male students, 17 female, 1 non-binary, and 3 who preferred not to disclose their gender. Participants represented diverse ethnic backgrounds: 23 identified as White, 14 as African American, 7 as Hispanic/Latinx, 2 as Native American, 9 as Other, and 6 preferred not to disclose their ethnicity. Participation in the study was entirely voluntary, with written consent obtained from all students and their parents. Four students who either did not complete any surveys or showed no recorded activity in the virtual world were excluded from the analysis.

Table 1. Interest and knowledge assessments

Instrument	Day	Camps	N
Boeder et al.'s Interest Development [36]	1	4	56
Gadbury et al.'s Astronomy Knowledge Assessment [37]	1	4	52
Gadbury et al.'s Astronomy Interest [35], [37]	3	3	36
Gadbury et al.'s Minecraft Interest [35], [37]	3	3	36
Boeder et al.'s Interest Development [36]	4	3	39
Linnenbrink-Garcia et al.'s Situational interest [38]	4	3	39
Britner & Pajares' Self-efficacy [39]	4	4	52
Gadbury et al.'s Knowledge Assessment [37]	4	4	48

3.2 Ordered Network Analysis

The purpose of this research is to identify student archetypes–operationally defined in line with [15, 20, 21] as patterns of behavior that characterize a substantial subset of students' gameplay styles, derived from in-game action data within WHIMC. We then examine the relationships between their in-game actions, knowledge, learning outcomes, and motivational measures. To this end, we employ Ordered Network Analysis (ONA; [34]). ONA has been previously used to analyze log data from various game-based learning environments and to identify differences in gameplay patterns between high- and low-learning students [15], [40], [41].

Epistemic Network Analysis (ENA), the foundation upon which ONA is built [22], constructs relationship models from unit variables, grouping variables, conversation variables, and stanzas. Initially developed to analyze discourse features that frequently co-occur, ENA uses these variables to identify patterns of co-occurrence among time-grouped constructs. The unit variable defines the primary level of analysis, (e.g., a single student), while the grouping variables organize units into broader categories for comparison (e.g., experimental conditions, cohorts, or teams). ENA selects cases to consider together by applying a moving window, referred to as a stanza, to a coded ordered dataset, counting co-occurrences between pairs of constructs for each unit of analysis (in this case, individual students) within the stanzas. These stanzas are grouped into conversation variables that capture the broader context (i.e., the overall task or timeframe which is segmented into stanzas by the moving window). ENA then constructs a weighted network for each unit, depicting each construct as a node in a visualization and showing connections between each of those as weighted edges to offer insight into broader behavioral or cognitive patterns across the dataset. ONA extends this foundational method (ENA) to also consider the order in which constructs appear in the data (i.e. A → B is treated differently than B → A). Additionally, ONA calculates connections involving self-transitions (e.g., A → A and B → B). Thus, ONA visualizations illustrate the direction and strength of connections (represented by bi-directional edges between nodes) and the frequency of construct repetition (indicated by node sizes).

In this study, both the unit and conversation variables were defined at the level of each individual student's gameplay session, ensuring that data from one student was not linked to another's. We also included the day as an additional conversation variable, since actions from different days are unlikely to be connected in the same way as consecutive actions within a single session. Although previous ENA studies analyzing non-linguistic gameplay behaviors have used game levels as stanzas [40], we chose not to use that finer-grained approach, as recent actions in one game world can still be meaningfully related to actions taken in a subsequent world. We tested several moving window lengths—which control how many previous lines of coded data are considered for co-occurrence. After finding no substantial differences in lengths ranging from 2 to 10, we used the standard value of 4, commonly adopted in ENA and ONA studies [42].

3.3 Coding Logs

This study focuses on the first three days of the summer camp, during which students explored various *what if* worlds to identify physical variables, observe environmental changes, and assess the potential for human habitability. These days were selected because the worlds involved similar expected behaviors rooted in scientific exploration, allowing a single codebook to be consistently applied across all three. The structured nature of these activities also ensured greater consistency across camps held in different locations. In contrast, the final two days centered on habitat-building tasks, which involved a different set of student actions that could not be effectively captured using the same codebook. These activities were also more open-ended, making them harder to analyze through interaction logs alone and more difficult to compare across camps.

The interaction logs captured each student's X, Y, and Z coordinates every three seconds, along with the in-game commands they used. These commands included actions such as measuring physical variables with scientific tools, recording observations visible to other players, and teleporting to the location of another player or object. The constructs identified from the log data—defined through collaborative discussions among the authors and an initial exploration of the logs including correlating behaviors to astronomy interest and examining in-game action distributions—corresponded to the specific actions students performed during this period. Students were free to explore the virtual worlds at their own pace—moving quickly or slowly, alone or with others—while observing areas of interest, collecting measurements, and sharing findings through written messages within the game. These behaviors formed the basis of the codebook presented in Table 2.

Table 2. Codebook

Code	Definitions
Non-stopping	The student has stopped less than 3 times during the last minute. A stop corresponds to moving less than three blocks during a period of three seconds.
Slow Exploration	The student has moved less than 30 blocks during the last minute (and more than 0).
Social Movement	The student was less than 20 blocks from another player during all of the last minute.
Individual Movement	The student was more than 35 blocks away from any other player during all of the last minute.
Teleport	The student has teleported to another location
Point of Interest	The student is inside of a point of interest for 10 seconds or more. Every 10 seconds within the Point of Interest trigger this code again.
Talk to NPCs	The student is at a distance of 4 blocks or less to an NPC for 10 seconds or more. Every 10 seconds close to the NPC trigger this code again.
Science Tool	The student uses a scientific tool to measure a physical variable.
Scientific Description	In-game observations in which students describe the virtual world without making an additional analysis or questioning the implications of that observed phenomenon. For example, "There is a tree and a cow," or "Temperature is -10 F."
Scientific Inquiry	In-game observations in which students ask a science-related question. Ex: "Can humans survive without the moon."
Scientific Reasoning	In-game observations in which students use logical reasoning and attempt to comprehend scientific concepts based on their observations (e.g., "theory: the biodome roof opens to regulate sunlight.")
Non-scientific Observations	In-game observations that include social, cultural, or emotional references that are not related to science or astronomy.

For constructs that could be directly identified from the logs without requiring interpretive judgment (e.g., *Non-Stopping*, *Slow Exploration*, *Individual Movement*, *Teleport*, *Talk to NPCs*, and others), no manual coding or inter-rater reliability evaluation was necessary. Each time a student's logged actions met the definition of any construct, a new coded line was generated. For example, when a student teleported to a new location, a new line coded as *Teleport* was added. If the student's subsequent behavior matched the criteria for another construct—such as *Slow Exploration*, defined as moving fewer than 50 blocks during a minute—a new line labeled with that construct was added. In this way, when the ONA algorithm is applied, a connection from *Teleport* to *Slow Exploration* is recorded. We adopted this event-based coding approach instead of using fixed time intervals, as some constructs correspond to brief actions lasting only a few seconds, while others span several minutes. A fixed granularity would obscure the temporal order of actions occurring within the same segment. Each coded line was assigned to a single construct. When actions co-occurred (e.g., *Using a Tool* while *Visiting a Point of Interest*), two separate coded lines were created: one for the initially triggered construct and another for the co-occurring behavior.

The thresholds for the *Non-Stopping* and *Slow Exploration* codes were determined by analyzing the distributions of students' movement speeds and pause durations, taking into account the sampling time of the location data (every 3 seconds). We operationalized the *Non-Stopping* definition in this way (vs. alternative measures of rapid movement) because WHIMC requires students to pause briefly to take measurements or make observations. For the *Individual Movement* code, the threshold was based on the maximum in-game field of view (35 blocks); in contrast, a reduced threshold of 20 blocks was used for the *Social Movement* code to identify instances when students were in close proximity to others, rather than simply observing from a distance. *Points of Interest* refer to specific in-game locations students are expected to visit; a 10-second threshold was applied because 95% of visits to these areas lasted at least that long. Conversations with NPCs occur through text boxes that automatically appear when a student is within 4 blocks of an NPC; this distance was therefore used as the threshold for the *Talk to NPC* code. The 10-second duration threshold for this code mirrors the rationale used for *Points of Interest*.

The four codes used to analyze students' in-game observations—*Scientific Description*, *Scientific Inquiry*, *Scientific Reasoning*, and *Non-scientific Observations*—were adapted from prior research on student interactions in a similar environment [43]. Although [43] proposes a more comprehensive coding scheme, we select a reduced set of constructs in order to streamline both the Ordered Network Analysis and Clustering Analysis (see Sections 3.3 and 3.4). Inter-rater reliability between two human coders was established for these four codes (Kappa ≥ 0.75) using a set of 100 student observations. After reaching agreement, the coders divided the remaining dataset (1033 observations in total) for manual annotation.

3.4 Cluster Analysis

We employed k-means clustering to identify player archetypes, following the approach proposed by Zambrano et al. [15]. The clustering features consisted of the strengths of directed transitions between different activities (e.g., *Science Tool → Scientific Description*) and the frequency of repeated actions (e.g., *Science Tool → Science Tool*), aggregated at the student level. These transition strengths, or connection weights, were extracted from the ONA model developed using WebENA [44]. The variance explained by the first two principal components used in the ONA visualization was 26.1% and 19.4%, respectively—typical for ENA and ONA plots (e.g., [45]). We conducted cluster analysis using all connection weights rather than relying solely on the first two components, thereby avoiding the information loss associated with dimensionality reduction.

We used silhouette analysis [46] with the Sci-Kit Learn library in Python [47] to determine the optimal number of clusters. Silhouette values, which range from -1 to 1, measure the similarity of an object to its own cluster (cohesion) relative to other clusters (separation). We calculated silhouette values for cluster counts ranging from 2 to 20 and selected N=3 as it produced the highest average silhouette score. After assigning each student to a cluster, these clusters were used as grouping variables to create the ONA models using the WebENA tool [44]. This grouping variable aggregates the strength of each identified transition for all students within a group, enabling comparisons of average patterns across groups.

The coded in-game actions (and combinations of actions) and interest and knowledge measures were compared across the resulting clusters using a Kruskal-Wallis test. We also compared normalized learning gains across clusters. Normalized learning gains were calculated as the ratio of the observed improvement (or decrease) to the maximum possible improvement (or decrease) for each student. To control for the false discovery rate, we applied the Benjamini-Hochberg correction to the significance level of each individual test [48].

We also analyzed correlations between interest measures and in-game actions (codes) derived from the groupings to identify associations not directly observable from the clusters, which do not fully capture the continuous distribution of variables. Specifically, we used Spearman's rank correlation to analyze the correlations between the interest measures and the transitions that showed significant differences in strength across the three clusters, applying the Benjamini-Hochberg correction to these results as well. To address the loss of statistical power caused by multiple hypothesis testing—particularly important with a sample size of 57 students—we limited the correlation analysis to the transitions that represented the primary differences among the three archetypes. For the same reason, we focused on initial interest and normalized learning gains, as these instruments were completed across all four camps and reflected the key differences between the clusters. Finally, we conducted a Monte Carlo analysis [49], with 10,000 runs, to establish a 95% confidence interval for the number of statistical tests that could be significant by chance, given the total number of tests conducted.

4. Results

4.1 Clustering and Ordered Network Analysis

The silhouette analysis identified three main player archetypes or typologies. Inspired by Bartle's archetypes [16], we refer to these clusters as EXPLORERS, ROAMERS, and SCIENTISTS. Table 3 presents the average code frequencies per student for each cluster. Statistical differences were identified for seven codes after applying the Benjamini-Hochberg correction. A Monte Carlo analysis (95% confidence interval) suggests that only 0 to 2 significant results would be expected by chance. Figure 2 illustrates the individual models for these three archetypes. Table 4 lists the weights of all the connections for which a Kruskal-Wallis test (with Benjamini-Hochberg adjustment of alphas) revealed statistically significant differences across the three groups.

Table 3. Average code frequencies (and SD) per student for each cluster. Significant differences across the three groups, determined using a Kruskal-Wallis test with a Benjamini-Hochberg correction of alphas, are highlighted in bold.

Code	Roamers	Explorers	Scientists	p-val
Science Tool	**7.8 (6.2)**	**20.3 (14.1)**	**40.4 (20.0)**	**<0.001**
Scientific Description	**9.7 (8.1)**	**8.1 (5.2)**	**22.1 (15.2)**	**0.017**
Non-Stopping	**22.1 (12.0)**	**26.5 (10.5)**	**15.4 (11.2)**	**0.018**
Point of Interest	**6.3 (4.4)**	**38.2 (19.3)**	**14.5 (11.7)**	**<0.001**
Social Movement	**13.5 (7.8)**	**24.5 (10.5)**	**14.3 (9.3)**	**0.002**
Talk to NPC	**4.4 (4.4)**	**17.4 (10.2)**	**6.9 (4.6)**	**<0.001**
Slow Exploration	**4.1 (4.0)**	**9.2 (5.7)**	**5.0 (4.6)**	**0.003**
Teleport	3.2 (4.3)	10.8 (13.5)	11.5 (12.9)	0.218
Ind. Movement	11.5 (8.1)	14.4 (9.7)	10.1 (8.1)	0.334
Non-Scientific Observation	2.4 (3.6)	2.2 (2.7)	1.7 (2.6)	0.457
Scientific Reasoning	2.5 (2.3)	2.6 (3.3)	3.8 (3.7)	0.595
Scientific Inquiry	1.3 (1.6)	1.8 (4.6)	1.0 (1.0)	0.658

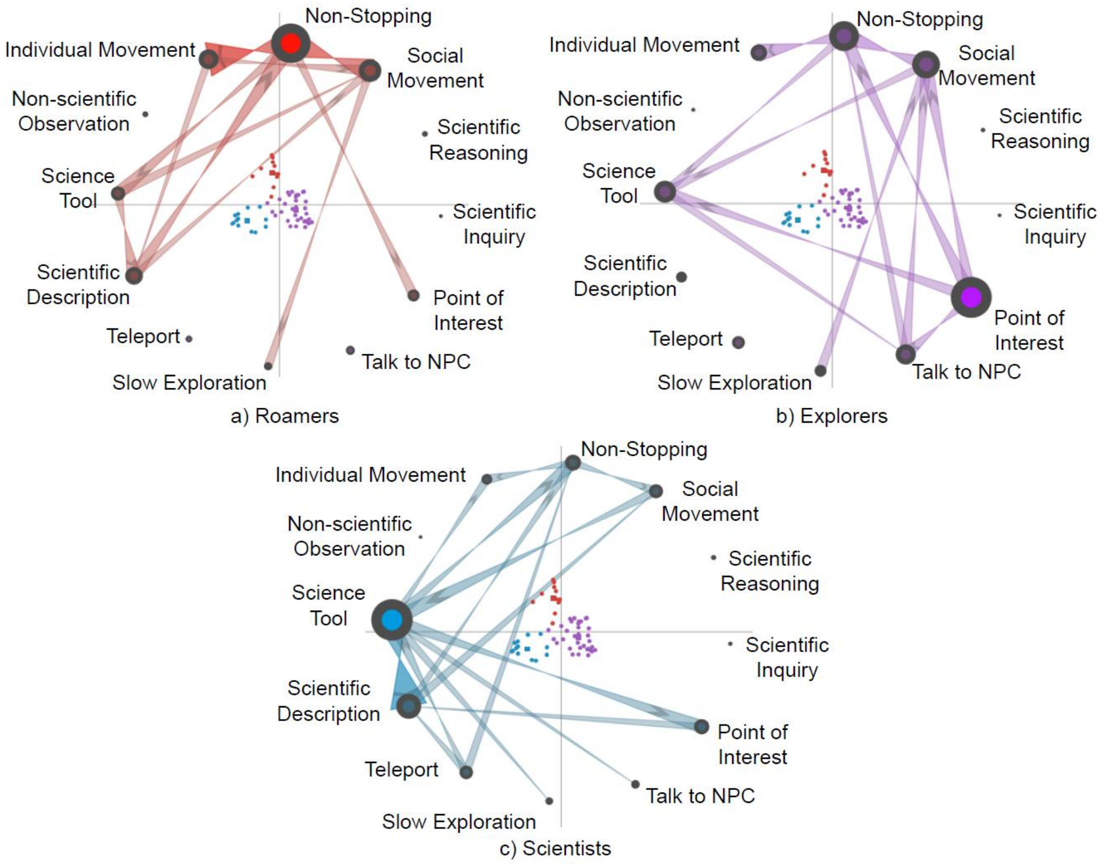

Figure 2. Individual ONA models of the identified player archetypes: a) Roamers, b) Explorers, and c) Scientists.

Table 4. Connection weights (CW) for the three player archetypes. Data is sorted by transition weights, all of which showed significant differences across the three groups based on a Kruskal-Wallis test (with Benjamini-Hochberg correction of alpha). Grayscale highlights larger transition weights.

Archetypes	Transition	Explorers	Roamers	Scientists	p-val
Explorers > Roamers > Scientists	Point of Interest → Non-Stopping	**0.113**	0.071	0.047	<0.001
	Non-Stopping → Point of Interest	**0.106**	0.077	0.042	0.001
	Point of Interest → Social Movement	**0.106**	0.044	0.036	<0.001
	Social Movement → Talk to NPC	**0.083**	0.046	0.027	0.006
	Talk to NPC → Social Movement	**0.083**	0.040	0.033	0.007
	Non-Stopping → Talk to NPC	**0.060**	0.049	0.023	0.004
	Talk to NPC → Slow Exploration	**0.045**	0.022	0.009	0.001
	Talk to NPC → Non-Stopping	**0.045**	0.019	0.014	<0.001
Explorers > Scientists > Roamers	Point of Interest → Point of Interest	**0.584**	0.109	0.159	<0.001
	Talk to NPC → Talk to NPC	**0.107**	0.038	0.040	<0.001
	Point of Interest → Talk to NPC	**0.100**	0.027	0.035	<0.001
	Social Movement → Point of Interest	**0.097**	0.003	0.033	<0.001
	Talk to NPC → Point of Interest	**0.096**	0.014	0.021	<0.001
	Slow Exploration → Talk to NPC	**0.045**	0.009	0.013	0.002
	Point of Interest → Slow Exploration	**0.022**	0.004	0.008	0.003
Roamers > Explorers > Scientists	Non-Stopping → Non-Stopping	0.188	**0.476**	0.071	<0.001
	Social Movement → Non-Stopping	0.155	**0.239**	0.061	<0.001
	Non-Stopping → Individual Movement	0.139	**0.285**	0.073	0.001
	Non-Stopping → Social Movement	0.130	**0.202**	0.072	<0.001
	Individual Movement → Non-Stopping	0.118	**0.231**	0.063	0.004
	Individual Movement → Social Movement	0.035	**0.087**	0.024	0.001
	Scientific Description → Non-Stopping	0.040	**0.144**	0.065	<0.001
Scientists > Explorers > Roamers	Science Tool → Science Tool	0.149	0.053	**0.537**	<0.001
	Point of Interest → Science Tool	0.090	0.015	**0.097**	0.004
Scientists > Roamers > Explorers	Science Tool → Scientific Description	0.043	0.116	**0.292**	<0.001
	Scientific Description → Science Tool	0.033	0.082	**0.238**	<0.001
	Scientific Description → Scientific Description	0.031	0.115	**0.176**	0.001

The EXPLORERS are characterized by visiting more *Points of Interest* (Avg = 38.2, $p < 0.001$) and interacting with more *NPCs* (Avg = 17.4, $p < 0.001$) than any other group. Additionally, they spend more time in these areas, as reflected in the higher connection weights for the self-transitions of both codes (CW = 0.584 for *Points of Interest* → *Point of Interes*t and CW = 0.107 for *Talk to NPC* → *Talk to NPC*). This pattern is also evident in their higher connection weights across most transitions involving either *Points of Interest* or *Talk to NPC*. However, despite actively visiting the expected locations and engaging with NPCs more than the other archetypes, EXPLORERS are less likely to use *Science Tool*s or make science-related observations—both desired behaviors in the game—compared to the SCIENTISTS. Notably, the EXPLORERS produce the fewest *Scientific Descriptions* after using a *Science Tool* (CW = 0.043, $p < 0.001$) and make fewer *Scientific Descriptions* overall (Avg = 8.1, $p = 0.017$).

Furthermore, although the EXPLORERS visit *Points of Interest* and interact with *NPCs*, they are also the group that most frequently engages in *Non-Stopping* (Avg = 26.5, $p = 0.018$). This behavior is particularly notable after reaching *Points of Interest* (CW = 0.113, $p < 0.001$) or *Talking to NPCs* (CW = 0.045, $p < 0.001$). This suggests that, although they visit these locations, they might not take the time to stop and perform actions such as making observations or using tools. However, this does not imply that these students always rush through the environment without analysis. On the contrary, EXPLORERS also engage the most in *Slow Exploration* (Avg = 9.2, $p = 0.003$) and tend to explore in a social manner (*Social Movement*), staying close to other players (Avg = 24.5, $p = 0.002$). This indicates that, while they slow down to observe or follow others—suggested by the predominance of *Social Movement*—they are less likely to pause to perform other actions beyond exploration, such as posting observations.

Like EXPLORERS, ROAMERS are also characterized by frequent movement without stopping at specific locations to complete other tasks (*Non-Stopping* Avg = 22.1, $p = 0.018$). However, unlike the EXPLORERS, ROAMERS visit significantly fewer *Points of Interest* (Avg = 6.3, $p < 0.001$) and interact with fewer *NPCs* (Avg = 4.4, $p < 0.001$). This group performs few actions beyond movement and exhibits the highest connection weight for *Non-Stopping* self-transitions (CW = 0.476, $p < 0.001$) as well as for all transitions involving *Non-Stopping*. While ROAMERS make slightly more *Scientific Descriptions* than EXPLORERS (Avg = 9.7), they use *Science Tools* the least (Avg = 7.8, $p < 0.001$). Additionally, they engage the least in recurrent conversations with *NPCs* (CW = 0.038, $p < 0.001$) and rarely use multiple *Science Tools* repeatedly (CW = 0.053, $p < 0.001$) or after arriving at a *Point of Interest* (CW = 0.015, $p=0.004$). These findings suggest that ROAMERS may require more support to develop behaviors aligned with the game's desired learning outcomes than the other groups.

The SCIENTIST group aligns most closely with the desired/designed pattern of interactions with the game as defined by course instructors and WHIMC designers. This group is characterized by significantly higher use of *Science Tools* (Avg = 40.4, $p < 0.001$) and more frequent *Scientific Descriptions* (Avg = 22.1, $p = 0.017$). Notably, they are also the group that most often makes *Scientific Descriptions* immediately after using a *Science Tool* (CW = 0.292, $p < 0.001$). Although SCIENTISTS visit *Points of Interest* less often than EXPLORERS (Avg = 14.5, $p < 0.001$), they are the group most likely to use *Science Tools* upon arriving at these locations (and staying there for at least 10 seconds; CW = 0.097, $p = 0.004$). Additionally, SCIENTISTS engage in *Non-Stopping* the least (Avg = 15.4, $p = 0.018$) and exhibit the lowest connection weights for transitions involving *Non-Stopping*. This suggests that SCIENTISTS tend to pause more frequently to perform other actions instead of continuously moving through the game. No significant differences were observed for the remaining codes (e.g., *Teleport*, *Individual Movement*, and *Non-scientific Observations*) across the three groups.

4.2 Associations with Knowledge and Motivation Measures

4.2.1 Measures per Player Archetype

Table 5 presents the average knowledge and motivation measures for each group. We observed differences between the initial level of interest across the three groups below the significance threshold of 0.05 (p = 0.027), with ROAMERS showing the lowest initial interest. However, none of these possible high-level group differences remain significant after applying the Benjamini-Hochberg correction, and this set of analyses does not have more significant results than could be expected by chance according to a Monte Carlo analysis (95% confidence interval of 0 to 2 significant tests).

Table 5. Average knowledge and motivation measures per archetype. Standard deviations are shown in parentheses.

Measure	Explorers	Roamers	Scientists	p-val
Number of Students	33	11	13	NA
Boeder et al.'s Interest Development (*initial*)	3.45 (1.41)	2.25 (1.69)	3.98 (1.07)	0.027
Boeder et al.'s Interest Development (*final*)	3.50 (1.79)	2.67 (1.62)	4.06 (1.57)	0.202
Gadbury et al.'s Astronomy Interest	2.41 (1.16)	1.85 (0.60)	3.06 (1.24)	0.136
Gadbury et al.'s Minecraft Interest	2.25 (1.67)	2.28 (0.61)	2.62 (0.83)	0.555
Linnenbrink-Garcia et al.'s Situational Interest	2.24 (0.91)	2.20 (0.41)	2.50 (0.62)	0.556
Britner & Pajares' Self-Efficacy	4.42 (1.28)	4.41 (1.21)	4.36 (1.22)	0.953
Pre-test	11.65 (2.20)	11.33 (1.50)	11.67 (3.06)	0.929
Post-test	12.55 (2.93)	10.50 (3.16)	13.82 (2.86)	0.090
Normalized Learning Gains	0.12 (0.22)	-0.02 (0.22)	0.24 (0.26)	0.069

4.2.2 Correlation Mining

To complement the analysis of interest and knowledge measures across player archetypes, we calculated Spearman correlation coefficients between the transitions listed in Table 4 (along with the individual codes showing significant group differences in Table 3) and two variables: initial interest (measured on Day 1) and normalized learning gains. In total, we conducted 68 correlations (27 transitions and 7 individual codes, each tested against 2 outcome measures). Of these, 12 yielded p-values below the 0.05 significance threshold (see Table 6), but none of them remain significant after the Benjamini-Hochberg correction. To assess whether this number of significant results could be due to chance, we conducted a Monte Carlo analysis, which estimated the 95% confidence interval for false positives under 68 independent tests to be between 0 and 7. This suggests that while individual correlations should be interpreted with caution, the overall pattern of results is unlikely to have occurred by chance.

Table 6. Spearman correlation coefficients between actions vs. initial interest development and normalized learning gains. Sig. p-values (0.05) are highlighted in bold. Only actions with a statistically significant correlation with Interest or Learning gains are shown in this table. The first column indicates the player archetype who most frequently performed each action.

Most freq. Archetype	Action	Boeder et al.' Interest Development (initial)	Learn Gains
Explorers	Point of Interest	**0.325 (0.014)**	0.163 (0.272)
	Talk to NPC	**0.314 (0.019)**	0.066 (0.657)
	Social Movement	**0.302 (0.024)**	0.018 (0.905)
	Talk to NPC → Non-Stopping	**0.283 (0.034)**	0.163 (0.273)
	Non-Stopping	**0.280 (0.037)**	0.138 (0.356)
	Talk to NPC → Slow Exploration	0.039 (0.775)	**-0.316 (0.030)**
Scientists	Scientific Description	**0.301 (0.024)**	0.120 (0.422)
	Science Tool	0.269 (0.053)	**0.346 (0.017)**
	Science Tool → Science Tool	0.132 (0.332)	**0.329 (0.024)**
	Point of Interest → Science Tool	0.120 (0.379)	**0.430 (0.003)**
	Science Tool → Science Description	0.021 (0.879)	**0.284 (0.045)**
Roamers	Non-Stopping → Non-Stopping	**-0.285 (0.034)**	-0.050 (0.737)

Six of the actions that showed statistical differences across the three player archetypes also exhibited apparent positive correlations with initial interest: *Point of Interest* (ρ = 0.325, p = 0.014), *Talk to NPC* (ρ = 0.314, p = 0.019), *Social Movement* (ρ = 0.302, p = 0.024), *Scientific Description* (ρ = 0.301, p = 0.024), *Talk to NPC → Non-Stopping* (ρ = 0.283, p = 0.034), and *Non-Stopping* (ρ = 0.280, p = 0.037). Again, we cannot be certain about individual correlations, but the overall pattern is highly unlikely to be due to chance. Interestingly, five of these six actions were predominantly performed by EXPLORERS. These actions are particularly associated with student movement within the virtual worlds and suggest that students who move more frequently began the camp with higher levels of interest.

Certain actions—such as visiting *Points of Interest* or *Talking to NPCs*—are intuitively linked to student interest. However, it is notable that *Non-Stopping* appeared in two of the positively correlated behaviors (*Non-Stopping* itself and *Talk to NPC → Non-Stopping*), challenging our initial hypothesis that *Non-Stopping* is inherently undesirable. Importantly, repeated *Non-Stopping* (i.e., *Non-Stopping → Non-Stopping*), which was more common among ROAMERS than EXPLORERS, was negatively correlated with interest. This distinction suggests that although EXPLORERS also engage in *Non-Stopping*, they typically do so after a meaningful event—such as interacting with an *NPC* or *Visiting a Point of Interest* (see Table 4)—and are more likely to follow it with *Slow Exploration* (see Tables 3 and 4), a behavior potentially associated with higher interest. In contrast, while ROAMERS engage in fewer total *Non-Stopping* instances, they tend to sustain it for longer stretches. This sustained *Non-Stopping* is negatively associated with initial interest, highlighting a key difference in exploration patterns across archetypes and their relationship to student motivation.

Overall, four patterns showed strong positive correlations with learning gains: *Science Tool* (ρ = 0.346, p = 0.017), *Science Tool → Science Tool* (ρ = 0.329, p = 0.024), *Point of Interest → Science Tool* (ρ = 0.430, p = 0.003), and *Science Tool → Scientific Description* (ρ = 0.284, p = 0.045). Notably, these actions were predominantly performed by the SCIENTISTS and reflected desired behaviors in the game. These correlations are likely driven by the overall positive relationship between using science tools and learning. However, these specific patterns—(1) using a tool after visiting a point of interest (a stronger correlation than the direct association between tool usage and learning), (2) employing multiple science tools, and (3) making scientific observations after using a science tool—suggest that learning is not solely linked to the use of the tools themselves. Instead, it is closely tied to using the tools under the right conditions to achieve the core objective of the learning experience: learning science by evaluating planetary habitability through scientific exploration and experimentation.

In contrast, the action *Talk to NPC → Slow Exploration*, which was primarily performed by the EXPLORERS rather than the SCIENTISTS, was negatively correlated with learning gains. These findings suggest that while some of the EXPLORERS' actions reflect high interest, this interest may not necessarily translate into behaviors that enhance learning gains. On the other hand, the SCIENTISTS, who also began with high interest, exhibited behaviors more aligned with the game's intended design, contributing to higher learning gains. ROAMERS overall tended to not perform any of these actions associated with higher learning or interest. Again, any individual correlation among this set remains uncertain, so replication will be important.

5. Discussion & Conclusion

5.1 Typologies across Games

This study develops player archetypes using a different approach than most of the previous literature, clustering based on behaviors rather than on self-reported motivation (e.g., [16], [17], [18], [19]). We identified three distinct player archetypes (EXPLORERS, ROAMERS, and

SCIENTISTS), which show substantial overlap with the 4 archetypes identified by Zambrano et al. [15], who analyzed behavior in a different STEM-focused open-world game (see Figure 3). Our ROAMERS, like Zambrano et al.'s, are characterized by disengaged behaviors, including movement through the virtual environment without performing discipline-related actions, and by lower interest in STEM. Similarly, our EXPLORERS match Zambrano et al.'s CONVERSERS—who visit multiple locations and frequently interact with NPCs, but do not deeply engage in science-related behaviors or achieve high learning gains. Notably, EXPLORERS and CONVERSERS were the most common archetype in each studies, representing over 40% of participants. Finally, our SCIENTISTS, who align with Zambrano et al.'s WORKSHEET USERS and SCANNERS, consistently engaged in the game's core scientific behaviors (in our study, frequent hypothesis testing and systematic data collection and in [15], frequent use of scientific tools and observations). In both studies, these actions were associated with the highest post-test scores and learning gains.

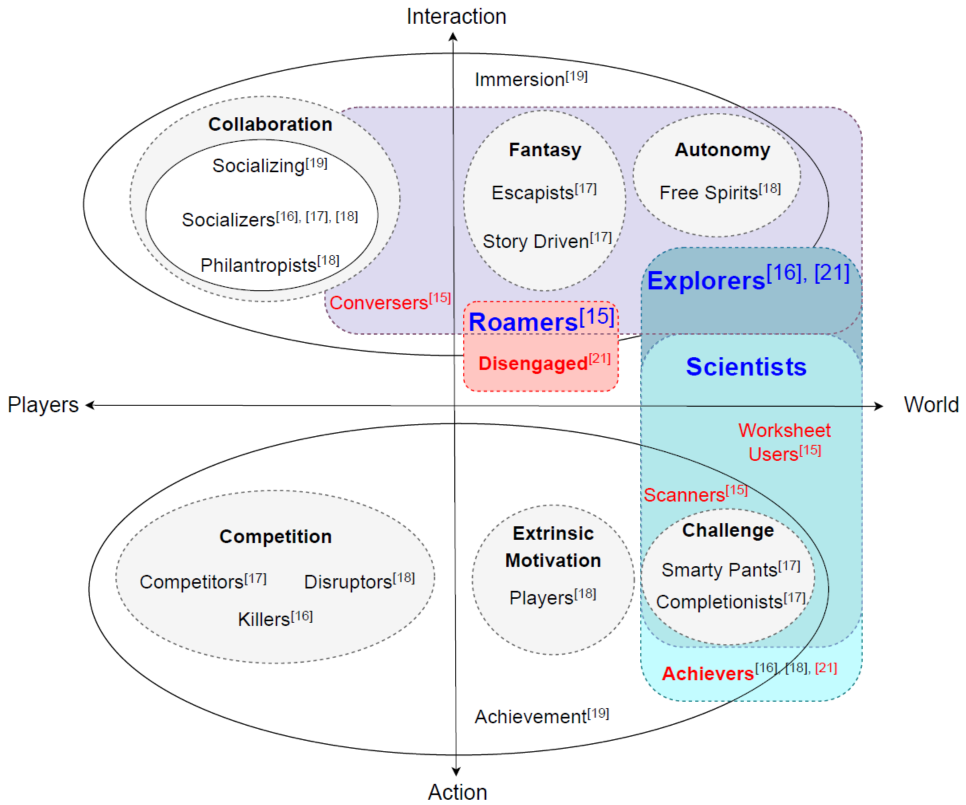

Figure 3. Placement of the three identified archetypes (ROAMERS, EXPLORERS, and SCIENTISTS) within the broader framework proposed across multiple typologies. This figure replicates Figure 1, with the addition of the three archetypes we identified positioned according to their observed conceptual similarities.

These parallels between behavioral typologies also hold across different game genres. For instance, the DISENGAGED and EXPLORER archetypes identified by Slater et al. [21] in a level-based puzzle game closely resemble ROAMERS and EXPLORERS in our study. Likewise, Slater et al.'s ACHIEVERS and our SCIENTISTS are characterized by performing fewer but focusing more deliberate actions that are aligned with the game's core learning objectives. Although Slater et al. did not examine these archetypes using external measures, the similarities with our typology suggest that these archetypes may be capturing generalizable patterns of engagement, at least among the kinds of populations that have been using the games in these studies.

5.2 Implications for Game Design

The consistency of player typologies across multiple serious games—and their alignment with multiple archetypes proposed in HCI frameworks of player typology (e.g., [16], [17], [19])—

can help game designers anticipate how these archetypes may emerge. This insight enables designers to better predict the specific actions students are likely to perform and to proactively address undesired behaviors. Among our identified archetypes, ROAMERS appear to be the group most in need of intervention. They showed no positive learning gains and exhibited overall disengagement, performing fewer of the key actions designed for learning than the other two archetypes. These behaviors were also associated with lower motivational measures, suggesting that this group may have less interest in STEM and science overall—potentially influencing their level of engagement with the game.

Since continuous movement without stopping was the primary behavior associated with low interest, a behavior also observed for this most disengaged group across different games (e.g., [15]) it could serve as a trigger for facilitators or interviewers to explore students' engagement through qualitative methods—for example, through in-the-moment interviews [50]. This approach could help uncover why some students continue moving without pause, even when they are not reaching new points of interest. Once such students are identified, capturing the specific moments when they stop to contemplate something within the game may also be valuable. These instances could reveal elements that momentarily sparked their interest, offering developers insights into features worth amplifying in future versions of the game. Incorporating more attention-catching features—such as more animated or interactive objects, surprising visual effects [51], references to familiar pop culture themes [52] or intriguing hypothetical scenarios (e.g., seeing two moons, [35])—could also be incorporated (if not already present) to test whether their presence encourages students to slow down or increase their engagement and situational interest.

Beyond these design elements, developers might also consider introducing new narrative elements to stimulate situational interest. Richer storytelling—such as supplementary quests or short narrative arcs delivered through NPCs that present a mystery to solve or evoke empathy in the player (e.g. [53], [54])—has been shown to positively influence student engagement [53], [55]. Additionally, recognizing that some students continue roaming without visiting key points of interest, designers could implement a badge or reward system tied to specific locations (e.g., [56]). Such mechanics could encourage exploration of meaningful areas rather than aimless movement across the map, thereby aligning students' actions more to the intended tasks of the game. However, effectively implementing these strategies will require a deeper understanding of what could genuinely engage ROAMERS—whether it is the visual design, narrative content, topic, level of challenge, or specific mechanics—and which factors are most likely to trigger their curiosity.

The other two types of players, EXPLORERS and SCIENTISTS, appeared more engaged and interested in the game. However, EXPLORERS may still benefit from additional support. Although they seem to be actively engaged, as reflected in their higher number of visits to points of interest compared to any other group in the context of this game, they do not consistently engage in key learning-related behaviors, such as using scientific tools or making systematic observations. Given the established link between these science-oriented actions and learning outcomes, also observed in [15], the mechanics that already engage EXPLORERS could be adapted to encourage them to slow down and participate more deliberately in the actions associated with learning. An instructional design could, for instance, provide students with a set of missions or goals that maintains a degree of freedom while smoothly guiding them toward meaningful engagement with the game's learning mechanics, an approach associated with both improved learning outcomes and a more positive affective experience during gameplay [57].

A concrete example of this type of intervention within our game would be to encourage scientific reflection and tool use by introducing checkpoints at points of interest that require students to make an observation or use a tool before leaving each location. Another approach could be to integrate brief quizzes or micro-tasks related to these locations, similar to those implemented in other educational games (e.g., [54]). These tasks could be triggered at moments when students interact with mechanics that provide domain-specific content, prompting them to reflect on what they are reading or observing and to engage with the intended game mechanics to gather the information needed to answer the questions.

Furthermore, the EXPLORER group also tend to exhibit more social behaviors, including frequent interactions with NPCs and social movement patterns, observed within the CONVERSER archetype in the Zambrano et al's study [15]. Since both behaviors are positively correlated with interest, developers can leverage these tendencies not only to enhance learning gains but also to reinforce students' existing interest in science. One approach is to incorporate additional NPC interactions that highlight the specific actions scientists would perform at various points of interest visited by these students. Facilitators can also encourage deeper engagement by prompting students to interact with NPCs and discuss the scientific work represented in each world or biome. By explicitly linking scientific practices, NPC interactions, and students' in-game observations, educators can strengthen both STEM learning and situational interest—building on students' natural curiosity and social engagement.

The third main group, SCIENTISTS, parallels the WORKSHEET USERS and SCANNERS in Zambrano et al.'s study [15], tending to have the best motivation and highest scores on knowledge measures. Because this group generally follows the intended actions of the game and achieves the highest learning outcomes, their behaviors can serve as a model for actions that should be encouraged across other archetypes. In the context of our game, scientific tool usage—the behavior most strongly correlated with learning gains—could be promoted by incorporating additional opportunities to use these tools or by prompting students to engage with them as they explore the virtual worlds. For example, developers might introduce NPCs, pop-up messages, or quests that request specific measurements or encourage a more systematic approach to data collection. Such interventions align with those proposed for the other two archetypes and may help guide all students toward deeper scientific engagement.

Lastly, while designing educational games that enhance interest and provide targeted scaffolding to guide students toward behaviors that improve learning is essential, the designers of interventions cannot assume that students are uniform across contexts or remain unchanged during their interaction with the platform. Any technology introduced into a classroom inevitably interacts with the classroom's culture, making it necessary for human educators to continually review and adapt interventions based on students' needs [58], [59]. Many of the most effective and widely used intelligent tutoring systems have integrated dashboards that help facilitators identify when students may need additional support [58]. For example, in Reasoning Mind, teachers received real-time information when a student is struggling with a specific concept, and teacher professional development emphasized using that information to immediately engage in proactive remediation, which became a common classroom practice [60]. Similarly, in ASSISTments, teachers review reports of the previous night's homework before class and adjust their planned lessons based on the questions students found most difficult [61]. These dashboard-based interventions have been shown to improve teachers' situational awareness [62], regardless of their age, gender, years of experience, or technological self-efficacy [63]. In the context of educational games, similar dashboards could allow teachers and facilitators to detect when students deviate from expected behaviors or learning pathways. Teachers and facilitators might then intervene when a student leaves a point of interest without completing the intended actions, prompting students to reflect on the game content. Likewise, if the game detects excessive roaming or signs of struggle, the system could alert teachers or facilitators to approach the student, identify potential issues, and provide timely support.

5.3 Limitations & Future Work

Although correlations between students' initial STEM interest and certain in-game behaviors offer potential explanations for the observed archetypes, a deeper understanding of why students adopt these patterns—and how to better support them—likely requires a qualitative approach that incorporates students' perspectives, such as in the moment interviews when key behaviors occur [50]. These interviews can shed light on students' strategies, actions, and motivations in response to specific in-game events (e.g., continuous movement without stopping, which was associated with low interest). By incorporating this qualitative lens, researchers and designers can gain a more holistic understanding of student engagement and develop targeted interventions to better support diverse learners.

Another limitation of this study's methodology is its assumption that students maintain a consistent archetype across different games and throughout the entire gameplay experience. In reality, students may adopt different behavioral archetypes depending on their interest in a particular genre or topic. Investigating these shifts more longitudinally, across games and game domains, would be a valuable direction for future research. Additionally, even within the same game, a player's archetype may shift in response to major changes in game mechanics. In our study, such a shift occurred after the third day of camp, when students moved from conducting scientific observations and evaluating planetary habitability to designing a biome for survival. Although this transition motivated our decision to focus on the initial three-day period, future research should explore how students' archetypes evolve within the course of longer games, what triggers these changes, and how they relate to different learning outcomes.

Despite these limitations, the findings of this study contribute to the literature on educational game design by examining the relationship between behavior-based archetypes and their associated outcomes. By better understanding how students engage with games, developers and educators may be better able to identify and support struggling or disengaged learners, build on the actions of already engaged students to improve their outcomes, and integrate in-game or in-person strategies to optimize the learning experience for learners across different gameplay approaches.

Acknowledgments

This work was supported by the National Science Foundation (Grant No. DRL2301173). Andres Felipe Zambrano thanks the Ministerio de Ciencia, Tecnología e Innovación and the Fulbright-Colombia commission for supporting his doctoral studies through the Fulbright-MinCiencias 2022 scholarship.

Conflicts of interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] J. P. Gee, "Learning by Design: Good Video Games as Learning Machines," *E-Learning and Digital Media*, vol. 2, no. 1, pp. 5–16, Mar. 2005, doi: 10.2304/elea.2005.2.1.5.

[2] D. W. Shaffer, K. R. Squire, R. Halverson, and J. P. Gee, "Video Games and the Future of Learning," *Phi Delta Kappan*, vol. 87, no. 2, pp. 105–111, Oct. 2005, doi: 10.1177/003172170508700205.

[3] D. B. Clark, E. E. Tanner-Smith, and S. S. Killingsworth, "Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis," *Review of Educational Research*, vol. 86, no. 1, pp. 79–122, Mar. 2016, doi: 10.3102/0034654315582065.

[4] C. Cole, R. H. Parada, and E. Mackenzie, "A scoping review of video games and learning in secondary classrooms," *Journal of Research on Technology in Education*, vol. 56, no. 5, pp. 544–577, Sept. 2024, doi: 10.1080/15391523.2023.2186546.

[5] R. L. Lamb, L. Annetta, J. Firestone, and E. Etopio, “A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations,” *Computers in Human Behavior*, vol. 80, pp. 158–167, Mar. 2018, doi: 10.1016/j.chb.2017.10.040.

[6] N. Barz, M. Benick, L. Dörrenbächer-Ulrich, and F. Perels, “The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis,” *Review of Educational Research*, vol. 94, no. 2, pp. 193–227, Apr. 2024, doi: 10.3102/00346543231167795.

[7] W. Mao, Y. Cui, M. M. Chiu, and H. Lei, “Effects of Game-Based Learning on Students’ Critical Thinking: A Meta-Analysis,” *Journal of Educational Computing Research*, vol. 59, no. 8, pp. 1682–1708, Jan. 2022, doi: 10.1177/07356331211007098.

[8] Y. Wang *et al.*, “A systematic review on affective computing: Emotion models, databases, and recent advances,” *Information Fusion*, vol. 83, pp. 19–52, 2022, doi: 10.1016/j.inffus.2022.03.009.

[9] A. All, E. P. Nunez Castellar, and J. Van Looy, “Measuring Effectiveness in Digital Game-Based Learning: A Methodological Review.,” *International Journal of Serious Games*, vol. 1, no. 2, June 2014, doi: 10.17083/ijsg.v1i2.18.

[10] H. Lukosch, S. Kurapati, D. Groen, and A. Verbraeck, “Gender and Cultural Differences in Game-Based Learning Experiences,” *Electronic Journal of e-Learning*, vol. 15, no. 4, pp. pp310-319, 2017.

[11] H. A. Nguyen, X. Hou, J. E. Richey, and B. M. McLaren, “The Impact of Gender in Learning With Games: A Consistent Effect in a Math Learning Game,” *International Journal of Game-Based Learning (IJGBL)*, vol. 12, no. 1, pp. 1–29, 2022, doi: 10.4018/IJGBL.309128.

[12] K. J. M. Kiili, A. Perttula, A. Lindstedt, S. Arnab, and M. Suominen, “Flow Experience as a Quality Measure in Evaluating Physically Activating Collaborative Serious Games,” *International Journal of Serious Games*, vol. 1, no. 3, Sept. 2014, doi: 10.17083/ijsg.v1i3.23.

[13] S. T. Yong, I. Harrison, and P. Gates, “Using Digital Games to Learn Mathematics – What students think?,” *International Journal of Serious Games*, vol. 3, no. 2, June 2016, doi: 10.17083/ijsg.v3i2.113.

[14] J. Hamari, D. J. Shernoff, E. Rowe, B. Coller, J. Asbell-Clarke, and T. Edwards, “Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning,” *Computers in Human Behavior*, vol. 54, pp. 170–179, 2016, doi: 10.1016/j.chb.2015.07.045.

[15] A. F. Zambrano *et al.*, “Unlocking Gameplay Insights with Epistemic (Ordered) Network Analysis: Understanding the Potential of Video Games to Foster Authentic Scientific Practices in STEM Education,” *Journal of Science Education and Technology*, 2025, doi: 10.1007/s10956-025-10213-4.

[16] R. Bartle, “Hearts, clubs, diamonds, spades: Players who suit MUDs,” *Journal of MUD research*, vol. 1, no. 1, p. 19, 1996.

[17] A. S. Kahn *et al.*, “The Trojan Player Typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations,” *Computers in Human Behavior*, vol. 49, pp. 354–361, Aug. 2015, doi: 10.1016/j.chb.2015.03.018.

[18] G. F. Tondello, R. R. Wehbe, L. Diamond, M. Busch, A. Marczewski, and L. E. Nacke, “The Gamification User Types Hexad Scale,” in *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, in CHI PLAY ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 229–243. doi: 10.1145/2967934.2968082.

[19] N. Yee, “Motivations for play in online games,” *CyberPsychology & behavior*, vol. 9, no. 6, pp. 772–775, 2006, doi: 10.1089/cpb.2006.9.772.

[20] L. Swanson *et al.*, “Leveraging cluster analysis to understand educational game player styles and support design,” in *GLS 13.0 Conference Proceedings*, 2022.

[21] S. Slater, A. Bowers, S. Kai, and V. Shute, “A typology of players in the game physics playground,” in *Proceedings of DiGRA 2017 Conference*, 2017.

[22] D. W. Shaffer, W. Collier, and A. R. Ruis, "A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data," *Journal of Learning Analytics*, vol. 3, no. 3, pp. 9–45, 2016, doi: 10.18608/jla.2016.33.3.

[23] N. Yee, "The Demographics, Motivations, and Derived Experiences of Users of Massively Multi-User Online Graphical Environments," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 3, pp. 309–329, June 2006, doi: 10.1162/pres.15.3.309.

[24] J. Kang, D. An, L. Yan, and M. Liu, "Collaborative Problem-Solving Process in a Science Serious Game: Exploring Group Action Similarity Trajectory," *International Educational Data Mining Society*, 2019.

[25] Y. Zhou and L. Paquette, "Investigating Student Interest in a Minecraft Game-Based Learning Environment: A Changepoint Detection Analysis," in *Proceedings of the 17th International Conference on Educational Data Mining*, B. Paaßen and C. D. Epp, Eds., Atlanta, Georgia, USA: International Educational Data Mining Society, July 2024, pp. 396–404. doi: 10.5281/zenodo.12729844.

[26] M. Liu, Y. Cai, S. Han, and P. Shao, "Understanding Student Navigation Patterns in Game-Based Learning," *Journal of Learning Analytics*, vol. 9, no. 3, pp. 50–74, 2022, doi: 10.18608/jla.2022.7637.

[27] D. Novoseltseva, C. P. Lelardeux, and N. Jessel, "Examining Students' Behavior in a Digital Simulation Game for Nurse Training," *International Journal of Serious Games*, vol. 9, no. 4, pp. 3–24, Nov. 2022, doi: 10.17083/ijsg.v9i4.543.

[28] N. Nasiar *et al.*, "It's Good to Explore: Investigating Silver Pathways and the Role of Frustration During Game-Based Learning," in *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky*, N. Wang, G. Rebolledo-Mendez, V. Dimitrova, N. Matsuda, and O. C. Santos, Eds., Cham: Springer Nature Switzerland, 2023, pp. 497–503. doi: https://doi.org/10.1007/978-3-031-36336-8_77.

[29] X. Yang, S. Rahimi, V. Shute, R. Kuba, G. Smith, and C. Alonso-Fernández, "The relationship among prior knowledge, accessing learning supports, learning outcomes, and game performance in educational games," *Educational Technology Research and Development*, vol. 69, no. 2, pp. 1055–1075, Apr. 2021, doi: 10.1007/s11423-021-09974-7.

[30] M. Yilmaz Soylu and R. H. Bruning, "Exploring Self-regulation of More or Less Expert College-Age Video Game Players: A Sequential Explanatory Design," *Frontiers in Psychology*, vol. 7, 2016, doi: 10.3389/fpsyg.2016.01441.

[31] K. Christoph and H. Tilo, "Effectance, self-efficacy, and the motivation to play video games," in *Playing video games*, Routledge, 2012, pp. 153–168.

[32] D. Pavlas, K. Heyne, W. Bedwell, E. Lazzara, and E. Salas, "Game-based Learning: The Impact of Flow State and Videogame Self-efficacy," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 54, no. 28, pp. 2398–2402, Sept. 2010, doi: 10.1177/154193121005402808.

[33] A. Gabbiadini and T. Greitemeyer, "Uncovering the association between strategy video games and self-regulation: A correlational study," *Personality and Individual Differences*, vol. 104, pp. 129–136, Jan. 2017, doi: 10.1016/j.paid.2016.07.041.

[34] Y. Tan, A. R. Ruis, C. Marquart, Z. Cai, M. A. Knowles, and D. W. Shaffer, "Ordered network analysis," in *International Conference on Quantitative Ethnography*, Springer, 2022, pp. 101–116, doi: 10.1007/978-3-031-31726-2_8.

[35] H. C. Lane *et al.*, "Triggering STEM interest with Minecraft in a hybrid summer camp.," *Technology, Mind, and Behavior*, vol. 3, no. 4, 2022, doi: 10.1037/tmb0000077.

[36] J. D. Boeder, E. L. Postlewaite, K. A. Renninger, and S. E. Hidi, "Construction and validation of the Interest Development Scale.," *Motivation Science*, vol. 7, no. 1, pp. 68–82, 2021, doi: 10.1037/mot0000204.

[37] M. Gadbury, "Addressing Misconceptions in Science through Minecraft," in *Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024, pp. 1370-1373*, International Society of the Learning Sciences, 2024. doi: 10.22318/icls2024.357916.

[38] L. Linnenbrink-Garcia *et al.*, “Measuring situational interest in academic domains,” *Educational and psychological measurement*, vol. 70, no. 4, pp. 647–671, 2010, doi: 10.1177/0013164409355699.

[39] S. L. Britner and F. Pajares, “Sources of science self-efficacy beliefs of middle school students,” *Journal of Research in Science Teaching*, vol. 43, no. 5, pp. 485–499, 2006, doi: 10.1002/tea.20131.

[40] S. Karumbaiah, R. S. Baker, A. Barany, and V. Shute, “Using Epistemic Networks with Automated Codes to Understand Why Players Quit Levels in a Learning Game,” in *Advances in Quantitative Ethnography*, vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds., in Communications in Computer and Information Science, vol. 1112. , Cham: Springer International Publishing, 2019, pp. 106–116. doi: 10.1007/978-3-030-33232-7_9.

[41] A. F. Zambrano *et al.*, “Cracking the code of learning gains: Using Ordered Network Analysis to Understand the Influence of Prior Knowledge,” in *International Conference on Quantitative Ethnography*, Springer, 2023, pp. 18–33, doi: 10.1007/978-3-031-47014-1_2.

[42] A. L. Siebert-Evenstone, G. Arastoopour Irgens, W. Collier, Z. Swiecki, A. R. Ruis, and D. Williamson Shaffer, “In Search of Conversational Grain Size: Modeling Semantic Structure using Moving Stanza Windows,” *Journal of Learning Analytics*, vol. 4, no. 3, pp. 123–139, Dec. 2017, doi: 10.18608/jla.2017.43.7.

[43] X. Liu *et al.*, “Investigating Learner Interest and Observation Patterns in a Minecraft Virtual Astronomy Environment,” in *Advances in Quantitative Ethnography*, Y. J. Kim and Z. Swiecki, Eds., Cham: Springer Nature Switzerland, 2024, pp. 19–34. doi: 10.1007/978-3-031-76332-8_2.

[44] C. L. Marquart, C. Hinojosa, Z. Swiecki, B. Eagan, and D. W. Shaffer, “Epistemic network analysis (Version 1.7.0) [Software].” 2021. [Online]. Available: http://app.epistemicnetwork.org

[45] J. Yu, D. Hu, and C.-H. Wang, “Development of ENA 3D: A Tool for Epistemic Network Analysis in Three-Dimensional Space,” in *Advances in Quantitative Ethnography*, Y. J. Kim and Z. Swiecki, Eds., Cham: Springer Nature Switzerland, 2024, pp. 152–165. doi: 10.1007/978-3-031-76335-9_11.

[46] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.

[47] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[48] Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995, doi: 10.1111/j.2517-6161.1995.tb02031.x.

[49] N. Metropolis and S. Ulam, “The Monte Carlo Method,” *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, Sept. 1949, doi: 10.1080/01621459.1949.10483310.

[50] R. S. Baker *et al.*, “Detector-driven classroom interviewing: focusing qualitative researcher time by selecting cases in situ,” *Educational technology research and development*, vol. 72, no. 5, pp. 2841–2863, Oct. 2024, doi: 10.1007/s11423-023-10324-y.

[51] N. Whitton, *Learning with digital games: A practical guide to engaging students in higher education*. Routledge, 2009.

[52] X. Zhang, Q. Huang, Z. Shi, and K. Zhang, “How cultural elements shape game evaluations: The role of cultural authenticity and perceived effort,” *Computers in Human Behavior*, vol. 162, p. 108452, Jan. 2025, doi: 10.1016/j.chb.2024.108452.

[53] A. Dörringer, M. Klopp, and R. Rossmann, “Digital Storytelling in Serious Games: Empirical Research on the Impact of Narrative in 3D Learning Worlds,” in *Proceedings of the 6th European Conference on Software Engineering Education*, in ECSEE ’25. New York, NY, USA: Association for Computing Machinery, 2025, pp. 96–105. doi: 10.1145/3723010.3723027.

[54] J. P. Rowe, L. R. Shores, B. W. Mott, and J. C. Lester, “Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments,” *International Journal of Artificial Intelligence in Education*, vol. 21, no. 1–2, pp. 115–133, 2011, doi: 10.3233/JAI-2011-019.
[55] S. W. McQuiggan, J. P. Rowe, S. Lee, and J. C. Lester, “Story-Based Learning: The Impact of Narrative on Learning Experiences and Outcomes,” in *Intelligent Tutoring Systems*, B. P. Woolf, E. Aïmeur, R. Nkambou, and S. Lajoie, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 530–539. doi: 10.1007/978-3-540-69132-7_56.
[56] D. Kao and D. F. Harrell, “The Effects of Badges and Avatar Identification on Play and Making in Educational Games,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, in CHI ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–19. doi: 10.1145/3173574.3174174.
[57] M. Taub, R. Sawyer, A. Smith, J. Rowe, R. Azevedo, and J. Lester, “The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment,” *Computers & Education*, vol. 147, p. 103781, Apr. 2020, doi: 10.1016/j.compedu.2019.103781.
[58] R. S. Baker, “Stupid Tutoring Systems, Intelligent Humans,” *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 600–614, June 2016, doi: 10.1007/s40593-016-0105-0.
[59] J. W. Schofield, *Computers and classroom culture*. Cambridge University Press, 1995.
[60] W. L. Miller, R. S. Baker, M. J. Labrum, K. Petsche, Y.-H. Liu, and A. Z. Wagner, “Automated detection of proactive remediation by teachers in reasoning mind classrooms,” in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, in LAK ’15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 290–294. doi: 10.1145/2723576.2723607.
[61] M. Feng and N. T. Heffernan, “Informing teachers live about student learning: reporting in the assistment system,” *Technology, Instruction, Cognition and Learning*, vol. 3, pp. 1–14.
[62] R. Kasepalu, P. Chejara, L. P. Prieto, and T. Ley, “Studying teacher withitness in the wild: comparing a mirroring and an alerting & guiding dashboard for collaborative learning,” *International Journal of Computer-Supported Collaborative Learning*, vol. 18, no. 4, pp. 575–606, Dec. 2023, doi: 10.1007/s11412-023-09414-z.
[63] A. van Leeuwen, C. A. N. Knoop-van Campen, I. Molenaar, and N. Rummel, “How Teacher Characteristics Relate to How Teachers Use Dashboards: Results From Two Case Studies in K-12,” *Learning Analytics*, vol. 8, no. 2, pp. 6–21, Sept. 2021, doi: 10.18608/jla.2021.7325.