Modeling Player Progression in an Educational Game Using Ordered Networks

Xiner Liu[1](✉), Jennifer Scianna[2], Shari J. Metcalf[3], Zhanlan Wei[1], Ryan S. Baker[1], Amanda Barany[1], Luke Swanson[2], and David J. Gagnon[2]

[1] University of Pennsylvania, Philadelphia, PA 19102, USA
xiner@upenn.edu
[2] University of Wisconsin, Madison, WI 53711, USA
[3] Harvard University, Cambridge, MA 02138, USA

Abstract. Understanding the sequence of player decisions in open-ended educational games provides insight into how those decisions influence player persistence or readiness for later challenges. This study uses Ordered Network Analysis to examine how players move between jobs of varying difficulty in the educational game *Wake*. To scaffold players, *Wake* breaks down multi-phase scientific investigations into smaller "jobs". Each job is manually coded based on its difficulty level in *Experimentation*, *Modeling*, or *Argumentation*. We use these difficulty ratings, along with whether the player completed or quit it, as codes to model player progression. We see that players who completed a job with a high quit rate on their first attempt more often followed paths with gradually increasing difficulty prior to accepting that job. In contrast, other players who quit the same job with a high quit rate on their first attempt were more likely to have failed prior jobs requiring basic skills in *Experimentation* or *Modeling* when moving from jobs that did not involve such components. They also tended to remain within jobs without such components across multiple transitions, which may reflect lower preparedness or content knowledge compared to those who completed the later difficult job. Findings also show that players who completed the difficult job on their second attempt spent the time between attempts completing jobs with lower difficulty, which may have helped strengthen foundational skills relevant to the target job or restore confidence. These findings point to opportunities for progression-aware intervention design based on how successful and unsuccessful players move through different types of jobs.

Keywords: Ordered Network Analysis · Educational Games · Player Progression · Quitting Behavior

1 Introduction

Designers of educational systems often make deliberate choices about how to structure and order learning tasks to help students build understanding in a manageable way [6]. Controlling the order in which students encounter new content can support the gradual development of relevant skills and lower the risk of confusion or failure [14].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026

G. Carmona et al. (Eds.): ICQE 2025, CCIS 2677, pp. 369–383, 2026.
https://doi.org/10.1007/978-3-032-12229-2_24

Educational games, on the other hand, are often designed to give students more freedom than traditional instruction [1]. Instead of progressing through a fixed sequence, players may choose which tasks to pursue and in what order (e.g., [16]). This freedom supports autonomy and interest-driven exploration [5], but it also leads to wide variation in the kinds of experiences students bring to later challenges [20]. As a result, players can arrive at the same complex task with very different histories of practice, exposure, and skill development. Prior research has examined player engagement [10], learning gains [24], motivational dynamics [7], and affect states [3] in games and simulations, but analysis on how the structure of player-chosen sequences—especially sequences that vary in cognitive demand—may shape persistence or quitting on difficult tasks affect their readiness for later challenges is less common. In particular, it remains unclear whether the order and difficulty of previously chosen tasks relate to whether a student persists or quits when facing a high-demand scientific problem.

These questions matter in systems where students are expected to manage their own learning trajectories. In open-ended educational games, this risk is especially visible. Struggling or quitting a complex task may reflect not a student's capability, but differences in their preparation stemming from the kinds of prior experiences they bring to that task [22]. Understanding how different sequences of tasks relate to persistence can help identify when and why players are likely to struggle, and what kinds of supports might be introduced in response. In this study, we investigate these questions through an analysis of player behavior in *Wake: Tales from the Aqualab*, a science game that gives students considerable freedom in choosing which research jobs to complete and when to take them on. We focus on one job with an unusually high quit rate and trace the sequences of jobs players take before and between attempts. This selected focus allows us to investigate whether differences in the order and difficulty of prior jobs help explain which players persist and which ones quit. The research questions ask:

– RQ1: What explains why some players quit a challenging job (*Hunting Lions*) on their first attempt, while others complete it?
– RQ2: What explains why some players who quit *Hunting Lions* on their first attempt are able to complete it when they return to it later?

2 Related Work

Researchers have long argued that agency can increase motivation and engagement by meeting learners' need for autonomy and active participation [4]. [13] found that giving students choices to choose their own teams, suggest rule tweaks, or decide on tactics, encouraging self-direction, and minimizing pressure increased children's motivation. Moreover, meaningful choices, such as selecting between two equally effective tools, has been shown to restore a sense of control and motivation after frustrating experiences [21]. Providing players the freedom to choose whether or not to solve programming puzzles in order to remove enemies in a game improves game enjoyment and intrinsic motivation to learn, without reducing learning outcomes [8]. Similarly, the option to choose a non-player character guide has been found to produce better learning outcomes and higher motivation [11].

However, higher levels of player autonomy in games may lead to highly individualized sequences of choices that pose a challenge for analysis. Epistemic Network Analysis

(ENA) and Ordered Network Analysis (ONA) are well-suited to address this challenge because these methods model patterns of connections among coded elements, such as decisions, actions, or discourse, over time. They have been widely used in the quantitative ethnography (QE) community to study learning processes in educational games. One of the earliest examples is [15]'s study of the epistemic urban-planning game *Urban Science*, which showed that players who imitated mentors' modeled thinking internalized an epistemic frame that persisted even after the game. [2] used ENA to analyze student discourse in the collaborative mobile game *School Scene Investigators*, finding that communal language and engaged responses supported the development of scientific practices. [18] used ENA in *Lakeland*, designed to teach students about nutrient cycles, to track how players responded to feedback across sessions. They observed a shift from short-term, reaction-based decisions to more forward-thinking, system-oriented actions.

More recently, researchers have also turned to ONA, which extends ENA by modeling directional transitions between actions or states. [23] used ONA to identify productive behavior patterns and unproductive wheel-spinning in the narrative-based learning game *Crystal Island*. [19] also used ONA to study self-regulated learning in the sustainable development educational game *Futureworld* and found that students who used a planning tool more frequently showed more learning behaviors during gameplay. We extend prior research on directional behavior modeling by applying it to a game with an open-ended structure that allows players to select their own sequences of game tasks (termed jobs). Our study specifically examines how a player's earlier choices impact later completion outcomes on a high-difficulty job. We use this in turn to study how the order of transitions between self-selected jobs relates to persistence in open-choice learning environments.

3 Method

3.1 Context

Wake: Tales from the Aqualab is a science education game designed for middle school students in grades 6–9. In the game, players take on the role of a young marine biologist named Olivia to learn more about four biomes (Kelp Forest, Coral Reef, Bayou, and the Arctic) and their respective ecological systems. Each ecosystem includes multiple research sites where players take on "jobs." A "job" represents a broader scientific assignment or mission within the game, such as studying the impact of an invasive species or monitoring biodiversity in a specific habitat. To complete a job, players perform several smaller "tasks," which may include collecting data, conducting experiments, or analyzing models. They are then expected to construct an argument and report their findings back to the researcher (i.e., the job provider). For example, a job might require players to estimate a species population, simulate environmental changes, or construct food web models, each broken down into discrete tasks.

The game is built around three key scientific practices: *Experimentation* (Exp; the difficulty of observing and measuring relationships of the species collected), *Modeling* (Mod; the difficulty of constructing, using, or refining models to simulate environmental processes), and *Argumentation* (Arg; the difficulty of synthesizing evidence and reporting conclusions back to the researcher to complete the job). Players use tools like a digital tablet to document species and ecosystems, experimentation tanks to test variables and species interactions, and modeling labs to visualize ecological relationships. Upgrades to the ship and equipment provide access to additional tools and research areas, which allow players to tackle more complex jobs (requiring more advanced scientific skills) over time. The design introduces scientific tools and methods incrementally in a way that provides opportunities for students to learn new skills without being overwhelmed. Before accepting any job, players can view a short description, the required upgrades, and a difficulty rating (from 0 to 4) to help guide their decision. To scaffold players, *Wake* also offers a guiding character, V1ct0r, whom players can summon at any point during gameplay for help. V1ct0r provides contextually relevant guidance based on the current task, available tools, and recent discoveries through dialogue with the player.

Wake is implicitly structured into 54 jobs. Students have the freedom to select from available jobs in any order, though some jobs remain locked until specific prerequisites are met, such as completing earlier jobs or acquiring ship and tool upgrades from the shop. Once a job has been started, players can leave it and switch to another job before completing the original one. Before players are able to choose jobs independently, they are required to complete a fixed sequence of three tutorial jobs that introduce the basic game mechanics.

3.2 Data Collection, Preparation, and Automated Code Application

Wake uses the Open Game Data infrastructure to track and log player activities. The opengamedata-unity package connects the game to a cloud-based server, which records player actions throughout each session. Each player receives a unique login that allows tracking across multiple sessions and days. This setup produces a dataset of individual player interactions collectively referred to as telemetry data.

The game organizes telemetry data into three categories: *Player Actions* (e.g., visiting research sites, entering data into scientific tools, selecting evidence during argumentation), *System Events* (e.g., feedback from the game, introduction of guiding characters, narrative elements), and *Progression Events* (e.g., completing tasks, discovering new species for a particular job). Each event includes metadata that specifies the timing, sequence, player identification, and game state at the time of the event. These events are transmitted to the Open Game Data server in real time, creating a chronological log of gameplay, articulated in the game's specific language.

The study includes players who started the game in January 2024, as indicated by their completion of the three required tutorial jobs. We retrieved log data for each of these players from January through May 2024. We restricted the sample to those who started in January to reduce confounding factors related to differences in when and how the game is introduced across schools. The five-month window was selected to allow sufficient time for players to progress through the game. A shorter period might exclude players who are still actively progressing, while a longer period could introduce variation in the data due to players pausing and resuming the game after long gaps, which makes it harder to interpret patterns of engagement consistently across participants. For our analysis, we segmented the chronologically organized log data into distinct jobs, with each job representing a player working on a single job until they either completed it or switched to another. We then aggregated the data into a single row per job, which includes the player's username (i.e., the user login assigned to them), the job they accepted, whether they completed the job (determined by the presence of a *complete_job* event in the logs), and the difficulty ratings for the three dimensions associated with the job: *Experimentation*, *Modeling*, and *Argumentation*. We limited the analysis to players who completed all three tutorial jobs, which led to a sample of 482 users.

The difficulty ratings for each job were developed using a predefined set of criteria (see Table 1) and were assigned through human coding based on how each job is played in the game. Although the original scale ranged from 1 to 5, ratings of 4 and 5 are both shown as 4 in the game, due to the limited number of level-5 jobs. We applied the same adjustment in our analysis to match the difficulty levels as presented to players. Coders independently applied the criteria, and any disagreements were resolved through social moderation. This process involves human judgment and interpretation, but it differs from most qualitative analysis (cf. [17]), where coding typically occurs after data collection, is closely tied to the research question, and often evolves through iterative analysis. The difficulty ratings used here were created prior to data collection and independent of the specific analytic goals of this study. As such, these difficulty ratings can be seen as design-time aspects of the game's structure rather than as being initially intended as interpretive codes.

Even though these codes might initially have been designed for analysis, these ratings can still be used as meaningful codes for quantitative ethnography analysis because they reflect systematic human judgments about core features of the tasks players' encounter. The ratings are grounded in the mechanics and goals defined by the game designers and capture variation in the cognitive and procedural complexity across jobs. This variation shapes how players experience and navigate the game. These codes do not necessarily represent actual difficulty and do not infer meaning from behavior after the fact; rather, we use these codes to model how different players respond to structured elements that are present in the design of the game levels.

For the *Argumentation* dimension, we simplified the five original levels into three broader categories. This decision was based on the observation that transitions between the original Levels 2 and 3 did not consistently reflect increased difficulty. Instead, they only indicated a longer or more elaborate argument process without a meaningful increase in the complexity of thinking or reasoning. As such, we found it insufficient to separate these levels within the analysis.

3.3 Selecting a Job for Analysis

Among all jobs, *Hunting Lions* showed the highest quit rate. It was accepted 151 times by 87 players, but the completion rate is only 21%-3% lower than the next most frequently quit job and 20% lower than the third. Even when including players who quit and later returned to complete it, the completion rate only reached 36%, compared to an average completion rate of 70% across the next ten most frequently quit jobs. This observation matches the finding in [12].

This pattern may be explained by the nature of the *Hunting Lions* job. The job requires players to investigate food web relationships involving lionfish at the Coral Reef site, conduct experiments, and then construct a predictive model of the future of that ecosystem if human hunting is permitted. The job carries the maximum difficulty rating (5) across all three dimensions and involves 11 detailed tasks and the collection of

Table 1. Coding Scheme for Job Difficulty Ratings Across Three Dimensions

Dim.	Definition and Example
Experimentation	**Level 0**: No experimentation component is present in the job.
	Level 1 (**Observe**): Identify species behaviors and direct interactions by observing static or dynamic visualizations (e.g., urchins eat kelp).
	Level 2 (**Relate-scaf.**): With guidance, identify relationships between species and abiotic factors by comparing outputs across conditions (e.g., the coral needs more light to photosynthesize)
	Level 3 (**Relate-unscaf.**): Same process as Level 2, but without guidance.
	Level 4 (**Quantify**): Measure interactions among species and between species and abiotic factors (e.g., eating rate, reproduction rate, environmental impact, etc., of a microorganism).
Modeling	**Level 0**: No modeling component is present in the job.
	Level 1 (**Recreate**): Use a provided concept map and predefined starting conditions to recreate known facts (e.g., the food web of a turtle species).
	Level 2 (**Simulate-scaf.**): With guidance, build a simulation model showing interactions between species that cannot be confirmed by existing data or are not yet known (e.g., model interactions between species in a specific site).
	Level 3 (**Simulate-unscaf.**): Same process as Level 2, but without guidance.
	Level 4 (**Predict**): Use the model to explore how changes in starting conditions affect or predict future outcomes (e.g., whether the rig disrupts the reef ecosystem and should be removed).
Argumentation	**Level 1** (**Respond**): Respond to a single question using a single type of evidence.
	Level 2 (**Given** Claim): Use multiple types of evidence or respond to multiple related questions in order to support a given claim.
	Level 3 (**Choose** Claim): Compare and evaluate multiple competing claims using multiple types of evidence.

32 necessary facts. A more detailed description of *Hunting Lions* and other jobs in the game can be found in the Wake Teacher Guide (sites.google.com/wisc.edu/waketeacherguide). Due to its particularly high quit rate, *Hunting Lions* was selected as the target job for subsequent analysis.

3.4 Player Grouping Based on Target Job Completion

Among the 87 players who accepted *Hunting Lions*, 17 completed it on their first attempt, 19 completed it on their second attempt, and 2 required more than two attempts to finish the job. To answer the first research question (i.e., why some players are able to complete *Hunting Lions* the first time they accept it), we grouped the 87 players into two categories: Complete (17 players who completed the job on their first attempt) and Quit (70 players who did not).

To address the second research question (i.e., why some players who quit *Hunting Lions* on their first attempt are able to complete it when they return), we looked more closely at players who accepted the job at least twice. We identified two groups: (1) players who quit *Hunting Lions* the first time and completed it on their second attempt (Quit–Complete), and (2) players who quit both times (Quit–Quit). We included only those who played at least two other jobs between their first and second *Hunting Lions* attempts. This filter allowed for at least one job-to-job transition for meaningful ONA modeling (details in Sect. 3.5). It also helps exclude cases where players who may have been adequately prepared from the first attempt but left the job for unrelated reasons, such as random distraction. This filtering resulted in 14 players in the Quit–Quit group and 8 players in the Quit–Complete group.

3.5 Ordered Network Analysis

To model how players progress through the game, we constructed ordered networks based on sequences of accepted jobs. Ordered Network Analysis (ONA) models the directional transitions between coded events—in this case, difficulty codes tied to each accepted job. To capture transitions between two consecutive jobs, we set the moving stanza window size to 2. This configuration captures how players progress from one job to the next and makes it possible to identify which difficulty codes tend to follow others, and in what direction, across individual play sequences. ONA is particularly suitable for this study because it preserves the order of consecutive job transitions and supports direct comparison between groups based on the structure of those transitions. This allows us to examine whether certain patterns of progression are more common among players with different outcomes on the selected high-difficulty job.

Our choice of a window size of 2 is inspired by the approach in [9], where the authors used Epistemic Network Analysis (ENA) to model learner affect transitions in ASSISTments, prior to the development of ONA. To represent directionality in transitions, they modified the coding scheme by manually appending sender and receiver suffixes to each code. ONA, on the other hand, supports directionality as a built-in feature and was designed to model ordered transitions explicitly. It also restricts connections to codes that appear within the same line of data, which is particularly important in this study, as difficulty codes within one job do not follow a temporal sequence and should not be treated as transitions. Since our analysis focuses on how players move from one job to the next, ONA captures transitions across distinct jobs in the order players accepted them and produces a more accurate representation of sequential gameplay.

Ordered networks were generated using the ENA Web Tool. We created two sets of difference models to compare player behavior in relation to the job *Hunting Lions*. First, we compared players who completed *Hunting Lions* on their first attempt (Complete) to those who quit the job on their first attempt (Quit). This model includes only the jobs each player completed prior to their first *Hunting Lions* attempt (See discussion in 5.1). Second, we compared players who initially quit *Hunting Lions* but completed it on a second attempt (Quit-Complete) to those who quit it both times (Quit-Quit). This comparison is based on only the jobs they completed between their first and second attempts at *Hunting Lions* (See discussion in 5.2).

Data were segmented by player ID so that job sequences were tracked only within individual players. Separate ONA models were created for each of the three difficulty dimensions by including only the relevant codes from that dimension in each network. To improve interpretability, transitions with line weights less than 0.02 ($lw < 0.02$) were excluded from the visualizations.

4 Results

In this paper, we model player progression in the educational game *Wake* using job difficulty ratings paired with completion or quitting outcomes as the codes in Ordered Network Analysis to track transitions between jobs. In this section, we present the results of the ordered network models and the differences in observed transition patterns for the player groups defined in Sect. 3.4.

4.1 Difference Models for Quit and Complete Groups

Figure 1 presents the difference models comparing players who completed *Hunting Lions* on their first attempt (Complete) to those who did not (Quit), based on the jobs accepted prior to their first attempt. Across all three difficulty dimensions, Mann-Whitney tests reveal significant differences between the two groups along the X-axis (*Experimentation*: $U = 1016$, $p < 0.01$, $r = -0.71$; *Modeling*: $U = 985$, $p < 0.01$, $r = 0.66$; *Argumentation*: $U = 324$, $p < 0.01$, $r = 0.46$). These results suggest that the two groups followed significantly different progression patterns before their first attempt at *Hunting Lions*, which may have contributed to whether they were able to complete the job.

Descriptive statistics show that players in the Quit group accepted a mean of 28 jobs before their first *Hunting Lions* attempt, four fewer than the Complete group, but this difference was not significant ($U = 685$, $p = 0.33$). The two groups also spend a similar amount of time on average per job (Quit group: 19 min; Complete group: 17 min; $U = 576$, $p = 0.84$), after excluding jobs where students spent more than three hours, as these instances more likely reflect idle time rather than active gameplay. However, the proportion of prior jobs completed was significantly higher in the Complete group (92.7%) than in the Quit group (86.9%) ($t = 2.2$, $p = 0.03$).

The difference network for *Experimentation* indicates that players who completed *Hunting Lions* (Complete group) show stronger and more frequent transitions from completing *Level 0* to completing *Level 1* (line weight difference $lw_{Diff} = 0.06$), from completing *Level 2* to completing *Level 3* ($lw_{Diff} = 0.06$), and from completing Level 3 to completing *Level 4* ($lw_{Diff} = 0.02$), although the last transition is not shown in the difference plot due to its lower weight. These transitions indicate that, compared to the Quit group, Complete players more frequently completed sequences of jobs where each new job represented an adjacent increase in *Experimentation* difficulty. We also observe stronger transitions from completing *Level 0* to *Level 2* ($lw_{Diff} = 0.06$) and from *Level 1* to *Level 3* ($lw_{Diff} = 0.05$). While these involve increases of two difficulty levels rather than one, they still reflect relatively gradual shifts and can be considered incremental progressions along the *Experimentation* dimension—especially the latter transition, as *Level 3* requires identifying the same types of relationships across conditions as *Level*

2, with the primary difference being the removal of scaffolding. Descriptive statistics further support these findings. Among Complete players, 25% had at least 20% of their transitions between completed jobs involve a one- or two-level increase in *Experimentation* difficulty, compared to 14.8% among Quit players; at the 75th percentile, the rates were 33% and 25%, respectively. These results suggest that gradual progression through increasing levels of difficulty was more characteristic of players in the Complete group.

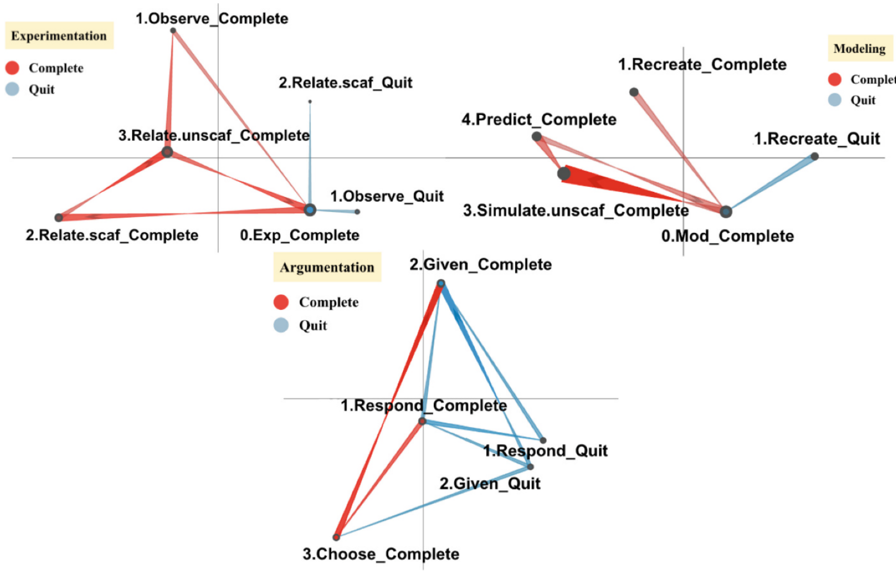

Fig. 1. Difference models for the Quit group (blue) and the Complete group (red) across the three difficulty dimensions: *Experimentation* (top left), *Modeling* (top right), and *Argumentation* (bottom). (Color figure online)

In the *Experimentation* plot, we see that players who quit *Hunting Lions* on their first attempt (Quit group) show only two connections with higher weights than those seen in the Complete group. That is, transitions from completing a job with no experimentation component (*Level 0*) to quitting a job at *Level 1* ($lw_{\text{Diff}} = 0.02$), which suggests that moving from jobs with no experimentation to those requiring basic observation might have presented a point of difficulty for the Quit players. Similarly, transitions from quitting a *Level 2* job to completing a *Level 0* job are also more prominent ($lw_{\text{Diff}} = 0.02$), which indicates that, after not finishing a task with moderate experimentation demands (e.g., identifying relationships between species and environmental factors), Quit players more often selected subsequent jobs with no experimentation requirements. These patterns suggest that providing additional scaffolds as players move from jobs without experimentation to those involving basic observation might help reduce the chances of quitting later on a harder job. More guidance to help players persist with jobs that involve interpreting species-environment relationships, rather than returning to jobs with no experimentation component after an unsuccessful attempt, might also support their progression toward completing harder jobs later on. Players in the Quit group also show a stronger self-connection at *Level 0* (Diff = 0.08), which indicates that they tend to complete multiple jobs with no experimentation component in sequence without moving on to jobs at higher experimentation levels. This pattern may suggest that players are at risk of struggling later; recognizing it as it emerges could create an opportunity to prompt players to begin practicing collecting and interpreting evidence before they encounter more difficult tasks. For example, in the Quit group, 2 players completed six jobs with no *Experimentation* component in a row; 4 players completed seven such jobs, and one player even completed eight in a row.

In the difference network for the *Modeling* dimension, the only transition that appears more frequently in the Quit group is from completing a *Level 0* job to quitting a *Level 1* job ($lw_{Diff} = 0.03$). The fact that this transition stands out for the Quit group suggests that they were more likely to quit when moving from completing a job with no modeling component to a job that required basic modeling (i.e., recreating known relationships). The Quit group also shows a stronger self-connection at *Level 0* (Diff = 0.03), which, similar to what was observed for *Experimentation*, indicates that they frequently completed several jobs that did not include any modeling components.

In contrast, the Complete group shows stronger transitions from completing *Level 0* to completing *Level 1* ($lw_{Diff} = 0.03$), as well as from completing *Level 3* to completing *Level 4* ($lw_{Diff} = 0.04$), which reflects more frequent movement between adjacent levels of modeling difficulty. Direct transitions from completing *Level 0* to completing *Level 3* ($lw_{Diff} = 0.08$) also appear in the Complete group. However, upon inspection of the data, these transitions typically occur later in gameplay, after players in the Complete group have already accepted a number of other jobs and just before they attempt *Hunting Lions* (this is possible because *Level 0* jobs can reappear later in the game or in biomes the player has not yet explored). Such a connection does not suggest that they were able to take on complex modeling jobs without intermediate steps. Rather, it points to the possibility that earlier experiences, such as completing jobs at consecutive modeling levels, may have contributed to their ability to complete higher-level modeling jobs later in the game, even when not preceded immediately by lower-level ones.

The difference plot for *Argumentation* shows that players in the Quit group were more likely to return to and complete jobs at the same argumentation difficulty level after previously quitting jobs at that level, compared to those in the Complete group. For example, we observe heavier transitions from quitting a *Level 2* job to completing a *Level 2* job ($lw_{Diff} = 0.04$), and from quitting a *Level 1* job to completing a *Level 1* job ($lw_{Diff} = 0.01$). Even if they did not succeed on a specific difficulty level, players in the Quit group are relatively likely to attempt the same difficulty level again and successfully complete it on a subsequent try. At the same time, the plot reveals a pattern where Quit players completed a higher-level job but then quit a job at a lower level. For instance, transitions from completing a *Level 2* job to quitting a *Level 1* job ($lw_{Diff} = 0.02$), and from completing *Level 3* to quitting *Level 2* ($lw_{Diff} = 0.02$), appear more prominently in the Quit group. This indicates that completing more challenging jobs did not necessarily support success on jobs requiring less argumentation—for example, the ability to use multiple types of evidence to construct one's own claim did not always translate to using multiple types of evidence to respond to a given prompt, and using multiple types of evidence to respond to a given prompt did not necessarily support the ability to use a single type of evidence when required for a similar prompt. In other words, different argumentation jobs, even at lower levels, posed distinct challenges that were not addressed through prior completion of more difficult ones.

4.2 Difference Models for Quit-Complete and Quit-Quit Groups

Of the 22 players who accepted the *Hunting Lions* job at least twice, 8 completed it on their second attempt. We constructed a difference-ordered network showing transitions between jobs accepted between the two attempts for the Quit-Complete and Quit-Quit groups (see Fig. 2) to examine differences in their progression.

Across all three difficulty dimensions, Mann–Whitney tests reveal significant differences between the two groups along the X-axis (*Experimentation*: $U = 98.5$, $p < 0.01$, $r = -0.76$; *Modeling*: $U = 96.5$, $p = 0.01$, $r = -0.72$; *Argumentation*: $U = 26$, $p = .04$, $r = 0.54$). On average, players in both groups accepted 8 jobs between their first and second *Hunting Lions* attempts. Quit-Complete players spent less time per job (9 min) than Quit-Quit players (18 min), though this difference was not significant ($U = 45$, $p = 0.48$). The Quit-Complete group also completed a higher percentage of these jobs (95%) compared to the Quit-Quit group (81%), but this difference was also not significant ($t = 1.9$, $p = 0.07$).

When examining the three difference plots together, one unexpected pattern stands out: the nodes representing the highest difficulty levels in each dimension are more strongly connected for players in the Quit-Quit group than for those in the Quit-Complete group. Specifically, transitions from completing a *Level 3 Modeling* job to a *Level 4 Modeling* job ($lw_{Diff} = 0.07$), and from completing a *Level 2* to a *Level 3 Argumentation* job ($lw_{Diff} = 0.07$), appear more frequently in the Quit-Quit group. This insight prompted us to examine the distribution of difficulty levels among the jobs accepted by the two groups. The data shows that 27% of the jobs accepted by Quit-Quit players had a *Level 3 Argumentation* rating, compared to only 17% for the Quit-Complete group. Similarly, 13% of the jobs accepted by the Quit-Quit group were rated *Level 4* in *Modeling*, compared to just 4% in the Quit-Complete group. This indicates that players who eventually completed *Hunting Lions* were not necessarily taking on more of the most difficult jobs.

Instead of progressing through more difficult jobs, the Quit-Complete group shows more transitions between jobs at similar or lower difficulty levels. For example, from *Level 2* to *Level 1* in both *Argumentation* ($lw_{Diff} = 0.08$). Revisiting less complex jobs may have helped reinforce earlier skills or improve fluency with the core demands of each dimension. In other words, Quit-Complete players' ability to complete *Hunting Lions* on the second attempt may have come from consolidating prior knowledge or rebuilding confidence.

Another pattern we observe is the stronger self-connection of nodes representing quitting at *Level 1*—for example, quitting *Level 1 Modeling*, which involves replicating known relationships ($Diff = 0.04$), and quitting *Level 1 Argumentation*, which involves selecting one piece of evidence to answer a given question ($Diff = 0.04$)—in the Quit-Quit group. Quitting at Level 1 reflects difficulty with jobs that involve only the most fundamental elements of each dimension. This suggests possible continued uncertainty about the basic ecological content (e.g., species interactions), confusion about task expectations, or difficulty identifying relevant information and deciding how to apply it. In

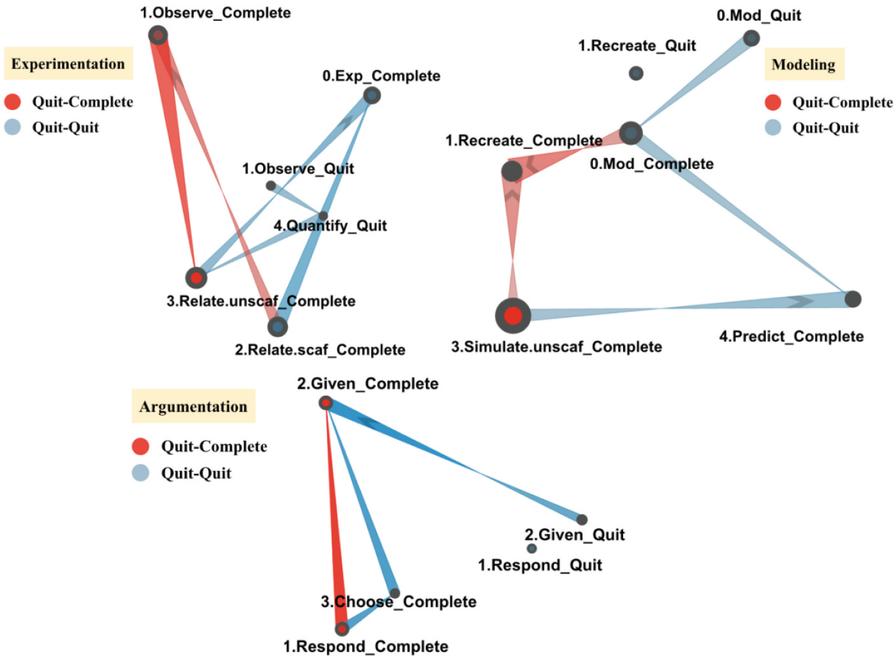

Fig. 2. Difference models for the Quit-Quit group (blue) and the Quit-Complete group (red) across the three difficulty dimensions: *Experimentation* (top left), *Modeling* (top right), and *Argumentation* (bottom). (Color figure online)

some cases, this may also reflect their disengagement from the job. Without the necessary foundational knowledge, skills, or motivation to complete these early-stage jobs, players may not gain the clarity or confidence needed to return to and complete more complex tasks like *Hunting Lions*. This may both hold back their immediate progress in the game and also make it harder for the player to successfully complete later tasks that build on earlier understandings.

5 Discussion and Conclusions

This study used Ordered Network Analysis (ONA) to model how players navigated jobs of varying difficulty in the educational game *Wake*. Each accepted job is represented in the network as a node defined by its difficulty level in one of three dimensions (*Experimentation*, *Modeling*, or *Argumentation*), paired with whether the player completed or quit the job. ONA models the transition between these codes based on the order in which jobs were accepted, with a window size of two to capture immediate sequential shifts in player behavior. We compared the resulting trajectory networks across different player groups to identify patterns in prior job selection that may explain why some players successfully completed the job *Hunting Lions* (the job with the highest quit rate) while others quit on the first attempt, and why some who initially quit were later able to return and complete it on a second attempt.

Our results show that players who completed *Hunting Lions* on their first attempt more often transitioned between jobs with gradually increasing levels of difficulty (especially in *Experimentation* and *Modeling*). While this does not necessarily indicate that making such transitions directly causes successful completion of the target job, the presence of these patterns suggests that gradual exposure to increasing challenge may support the development of relevant skills needed for completing more complex jobs. These findings point to potential design strategies for supporting students in the games. Systems might encourage players to select jobs that build on previous ones with small increases in difficulty or offer additional scaffolds when players shift between disconnected difficulty levels. Doing so may help players develop an internal model of how scientific practices build in complexity and what it looks like to improve across them.

In contrast, players who quit *Hunting Lions* on their first attempt often failed when they moved directly from jobs without *Experimentation* or *Modeling* components (*Level 0*) to jobs that required basic skills in those areas (*Level 1*). They also repeatedly selected *Level 0* jobs in both dimensions, which suggests that they more frequently stayed in simpler tasks without building relevant scientific practices. This pattern indicates minimal exposure to core scientific practices and suggests that these players may not have received adequate preparation for higher-level jobs. The game could respond by identifying players who complete two or more *Level 0* jobs in succession and recommending a *Level 1* job that introduces a foundational skill in a single dimension. When a player quits a *Level 1* job, the game could direct them to a short, structured task that addresses the same core skill with additional guidance. These targeted supports would help redirect players toward gradual skill development rather than extended repetition of low-level tasks.

Interestingly, Players who quit *Hunting Lions* on their first attempt but completed it on their second did not select more difficult jobs between attempts. Instead, they more often moved between jobs at similar or lower difficulty levels across all three dimensions. These transitions suggest that players in the Quit-Complete group may have used the time in between attempts to reinforce or consolidate their foundational skills. When a player quits a high-difficulty job, the system could therefore identify the primary dimension of difficulty and immediately recommend a set of jobs at lower levels within that dimension. The goal is to create multiple opportunities to apply similar forms of reasoning within a more manageable context, which may help players internalize key strategies and recognize recurring patterns, increasing their readiness when they re-encounter high-difficulty jobs later in the game.

Methodologically, using ONA with a window size of 2 enabled us to straightforwardly identify how players move across difficulty levels within specific dimensions and to detect differences in progression strategies between groups. However, the interpretive power of the resultant models remains incomplete. The plots reflect patterns in observable job transitions but cannot account for unlogged influences on player choices, such as confusion about instructions, external interruptions, or loss of motivation. As such, our analysis provides a structured summary of behavioral tendencies, but it does not explain why players make particular decisions. Future work should introduce additional instruments, such as short surveys or post-session check-ins, to ask players about their reasons for leaving specific jobs. These self-reports can be combined with behavioral markers from log data, such as time spent idle, repeated failed attempts, or absence of tool use, to help infer whether quitting followed a period of difficulty or occurred without signs of struggle. With this additional information, we can better distinguish between meaningful and incidental quitting and inform the design of more targeted, evidence-based support strategies.

Acknowledgments. We would like to thank NSF #DRL-1907437 for support of this project.

References

1. Boeker, M., Andel, P., Vach, W., Frankenschmidt, A.: Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. PLoS One. **8**(12), e82328 (2013)
2. Bressler, D.M., Bodzin, A.M., Eagan, B., Tabatabai, S.: Using epistemic network analysis to examine discourse and scientific practice during a collaborative game. J. Sci. Educ. Technol. **28**, 553–566 (2019)
3. Chen, X., Zou, D., Kohnke, L., Xie, H., Cheng, G.: Affective states in digital game-based learning: thematic evolution and social network analysis. PLoS One. **16**(7), e0255184 (2021)
4. Cordova, D.I., Lepper, M.R.: Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization, and choice. J. Educ. Psychol. **88**(4), 715 (1996)
5. Deen, M.: GAME, Games autonomy motivation & education: how autonomy-supportive game design may improve motivation to learn (2015)
6. Gagne, R.M.: Educational technology and the learning process. Educ. Res. **3**(1), 3–8 (1974)
7. Habgood, M.J., Ainsworth, S.E.: Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. J. Learn. Sci. **20**(2), 169–206 (2011)
8. Joshi, A., Mousas, C., Tokey, S.S., Kao, D.: Exploring player agency in educational video games. Entertain. Comput. **52**, 100754 (2025)
9. Karumbaiah, S., Baker, R. S.: Studying affect dynamics using epistemic networks. In: International Conference on Quantitative Ethnography, pp. 362–374 (2021)
10. Ke, F., Xie, K., Xie, Y.: Game-based learning engagement: a theory- and data-driven exploration. Br. J. Educ. Technol. **47**(6), 1183–1201 (2016)
11. Kinzer, C.K., et al.: The impact of choice and feedback on learning, motivation, and performance in an educational video game. Games Learn. Soc. Conf. **2**, 175–181 (2012)
12. Liu, X., Slater, S., Swanson, L., Metcalf, S.J., Gagnon, D. J., Baker, R.S.: Identifying when and why students choose to quit jobs in a science exploration game. In: Joint International Conference on Serious Games, pp. 56–69 (2024)
13. Mandigo, J., Holt, N., Anderson, A., Sheppard, J.: Children's motivational experiences following autonomy-supportive games lessons. Eur. Phys. Educ. Rev. **14**(3), 407–425 (2008)
14. Merrill, M.D.: First Principles of Instruction. Wiley (2012)
15. Nash, P., Shaffer, D.W.: Mentor modeling: the internalization of modeled professional thinking in an epistemic game. J. Comput. Assist. Learn. **27**(2), 173–189 (2011)
16. Nguyen, H., Harpstead, E., Wang, Y., McLaren, B.M.: Student agency and game-based learning: a study comparing low and high agency. In: 19th Conf. of Artificial Intelligence in Education, vol. 19, pp. 338–351 (2018)
17. Saldana, J.: Fundamentals of Qualitative Research. Oxford University Press (2011)
18. Scianna, J., Gagnon, D., Knowles, M.A.: Counting the game: visualizing changes in play by incorporating game events. In: Intl. Conf. on Quantitative Ethnography, pp. 218–231 (2021)
19. Shin, E., Carpenter, D., Rowe, J.: Analysis of self-regulated learning behaviors in game-based learning using ordered network analysis. J. Educ. Technol. **40**(2), 307–330 (2024)

20. Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., Lester, J.: The agency effect: the impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. Comput. Educ. **147**, 103781 (2020)
21. Tyack, A., Wyeth, P.: The small decisions are what makes it interesting: autonomy, control, and restoration in player experience. In: ACM on Human-Computer Interaction, pp. 1–26 (2021)
22. Yang, X., Rahimi, S., Shute, V., Kuba, R., Smith, G., Alonso-Fernández, C.: The relationship among prior knowledge, accessing learning supports, learning outcomes, and game performance in educational games. Educ. Technol. Res. Dev. **69**, 1055–1075 (2021)
23. Zambrano, A.F., et al.: Cracking the code of learning gains: using ordered network analysis to understand the influence of prior knowledge. In: Intl. Conf. on Quantitative Ethnography, pp. 18–33 (2023)
24. Zhao, D., Muntean, C.H., Chis, A.E., Rozinaj, G., Muntean, G.M.: Game-based learning: enhancing student experience, knowledge gain, and usability in higher education programming courses. IEEE Trans. Educ. **65**(4), 502–513 (2022)