



Cracking the Code of Learning Gains: Using Ordered Network Analysis to Understand the Influence of Prior Knowledge

Andres Felipe Zambrano¹ , Amanda Barany¹ , Jaclyn Ocumpaugh¹ ,
Nidhi Nasiar¹ , Stephen Hutt² , Alex Goslen³ , Jonathan Rowe³ ,
James Lester³ , Eric Wiebe³ , and Bradford Mott³ 

¹ University of Pennsylvania, Philadelphia, PA, USA
azamb13@upenn.edu

² University of Denver, Denver, CO, USA

³ North Carolina State University, Raleigh, NC, USA

Abstract. Prior research has shown that digital games can enhance STEM education by providing learners with immersive and authentic scientific experiences. However, optimizing the learning outcomes of students engaged in game-based environments requires aligning the game design with diverse student needs. Therefore, an in-depth understanding of player behavior is crucial for identifying students who need additional support or modifications to the game design. This study applies an Ordered Network Analysis (ONA)—a specific kind of Epistemic Network Analysis (ENA)—to examine the game trace log data of student interactions, to gain insights into how learning gains relate to the different ways that students move through an open-ended virtual world for learning microbiology. Our findings reveal that differences between students with high and low learning gains are mediated by their prior knowledge. Specifically, level of prior knowledge is related to behaviors that resemble wheel-spinning, which warrant the development of future interventions. Results also have implications for discovery with modeling approaches and for enhancing in-game support for learners and improving game design.

Keywords: STEM Education · Game-based Learning · Ordered Network Analysis · Epistemic Network Analysis

1 Introduction

Digital games are increasingly prevalent in STEM education [1–3], given their design affordances for engaging learners in authentic scientific processes and STEM experiences that are not otherwise accessible [4–7]. To maximize these benefits, however, game experiences should meet (and where possible, adapt to) diverse student needs and contexts (e.g., prior knowledge). One way to understand such needs is to map student behavior to learning outcomes, a first step in differentiating between a student who is persisting productively and one who might be wheel-spinning [8].

Prior research has explored factors that influence user experiences in virtual environments such as prior knowledge [9], self-regulation [10], interest in the topic of the game [11], and digital literacy [12]. Other studies have examined the in-game experiences of learners through filtered time series analyses [13] and graph theory techniques [14], while comparing students' pathways with an expert problem-solving trajectory [13]. However, new techniques for quantifying learners' trajectories through educational games could help us to better understand how students who are learning more differ in behavior and experiences over time from those who learn less.

This work aims to bridge this gap by employing Epistemic Network Analysis (ENA) [15] to investigate the differences between students' in-game behaviors and their learning gains. We do so in the context of a game developed for microbiology education [16], where rich logfiles of students' actions are explored in conjunction with measures of prior knowledge. We specifically employ an Ordered Network Analysis (ONA), a method derived from ENA that takes into account the temporal order of connections between codes [25]. With this approach, we aim to answer two main research questions: (1) What paths through the game are associated with learning gains? and (2) How does prior knowledge influence these paths?

2 Related Work

Understanding differences in how actions in an educational game are enacted over time by those with higher and lower learning gains is an important component of successful game design. Prior research in this area has applied filtered time series analysis to students' log data [13, 17]. This method uses principal component analysis (PCA) [18] to filter a multivariate time series (defined by a set of constructs coded in the log files data) to a univariate time series [13]. Using PCA as a filtering strategy allows for context-specific analyses about timing, but likely eliminates nuance about which actions are most useful. This is perhaps evidenced by conflicting results in the literature. Researchers have shown that higher rates of student actions in virtual worlds are positively associated with learning in a game-based environment about ecosystems [17], but negatively associated with learning in a game-based environment for microbiology [13], indicating that a more nuanced analysis about students' interactions is required to understand effective learning strategies.

Previous research has used graph theory analysis to examine the locations students visited during gameplay in a science learning environment [14]. Trajectories through virtual locations were compared using metrics of density and similarity among networks. Their results showed that student exploration of the environment might be more beneficial for learning than more efficient pathways focused solely on game completion [14]. However, the employed networks compared transitions completed by at least 50% of the students in each of the two groups (based on high or low learning gains). While this approach highlighted variance in student behaviors related to learning gains, it still used a categorical threshold (transitions done by 50% of students) for defining a binary graph instead of a weighted graph, overlooking the repetitions or weights of those common pathways.

The work reported in this paper builds on this previous research by applying a new method for understanding student trajectories through a virtual world. Specifically, we

use Ordered Network Analysis [25], a method derived from Epistemic Network Analysis (ENA) [15]. ENA measures the relationships between coded elements by quantifying their co-occurrences [15], and then representing them in a network diagram that illustrates the weights of these connections [19]. Each coded element, or unit, is represented as a single point on the Cartesian plane, meaning that the diagram's unit means can be used to determine statistical differences between them [19]. The ability to quantify the connections between codes and determine statistical differences makes ENA a valuable complement to prior methodologies examining learners' game experiences.

In the context of game-based learning and virtual worlds, ENA was first used in previous research on language data to investigate identity exploration [20], scientific practices [21], and student goals [22], but has also been applied to trace data of student interactions with virtual environments. For example, Karumbaiah et al. [23] applied ENA to clickstream data to uncover the trajectories that lead to quitting behaviors in a learning game called Physics Playground. Similarly, Scianna and Knowles [24] used ENA on player log files to identify how they responded differently to game events during second playthroughs of a video game, showing improvements in student planning skills. These initial examples demonstrate the valuable potential of traditional quantitative ethnographic techniques such as ENA (which typically models patterns in complex discourse data) with detailed log files of player actions.

Ordered Network Analysis (ONA) [25] uses the same principles as ENA but takes directionality of connections between codes over time into account. We do so in order to understand how different trajectories through a virtual game relate to differences in learning gains. Specifically, we are interested in finding ways to identify students who are productively exploring the game in ways that improve their learning, as opposed to those who might be wheel-spinning, a phenomenon that occurs when students become stuck because they lack the prerequisite knowledge necessary to advance [26, 27].

3 Methods

The goal of this work is to identify potentially productive problem-solving actions in a game-based learning environment called CRYSTAL ISLAND [16]. We employ Ordered Network Analysis (ONA) [25], which has been used to offer a more nuanced understanding of complex cognition and behavior across contexts (e.g., [25, 28]), and holds particular promise in problem-solving contexts, particularly those such as CRYSTAL ISLAND in which the orders of certain actions may contribute to in-game success, learning, or stagnation. As with ENA, the ONA algorithm uses a *moving window* to identify connections between lines in students logfile data (e.g., in-game actions or locations) within the recent temporal window. However, ONA accounts for the order in which connections might occur in the data by constructing an asymmetric adjacency matrix for each unit. That is, it calculates both the strength of associations between students first completing Action A (e.g., Reading) and then completing Action B (e.g., Moving) as well as Action B followed by Action A. In this way, the conventional visualization of an ENA model is expanded to include bi-directional edges between each pair of connected nodes. These edges denote the strength of these ordered associations, offering a more nuanced view of students' problem-solving behaviors within the game environment.

3.1 Dataset

We analyzed data from 92 students from an urban school in the southeastern United States (i.e., from [29]) who used an inquiry-based game to learn microbiology. In CRYSTAL ISLAND [16], players adopt the role of a researcher tasked with diagnosing the cause of a disease outbreak affecting an island-based research team. To successfully complete the game, players must explore several locations, interact with non-player characters (NPCs), collect information from in-game reading materials, and test their hypotheses using laboratory equipment. Players are provided with a concept matrix to help them organize the information obtained from the readings, as well as a worksheet to organize their hypotheses and results. The game interface, including the “golden pathway” (i.e., the most efficient path to game completion) proposed by [13], is presented in Fig. 1.

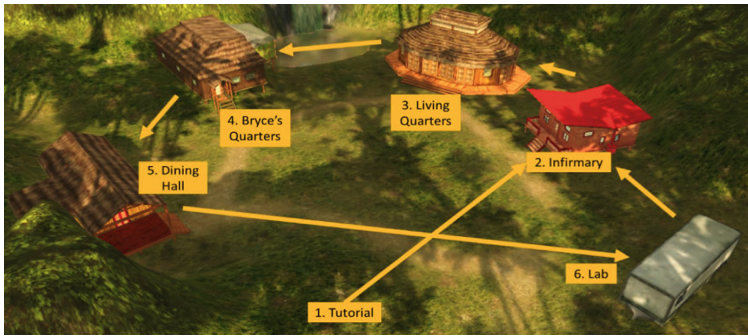


Fig. 1. Overview of CRYSTAL ISLAND with the expert “golden pathway” for game completion as operationalized by [13].

3.2 Learning Measures

To evaluate students’ knowledge of microbiology, identical pre- and post-tests were administered. Normalized learning gains were calculated based on the equation proposed in [30]. This approach compares differences between post- and pre-test scores with potential improvements students could achieve based on their initial scores. Data from 26 students with incomplete test responses were excluded, leaving a total of 66 students for the current analyses.

3.3 Unit Variables and Other Divisions

The *unit variables* (i.e., the variables that organize the dataset into meaningful groups for comparison and visualization) for ONA in this study were based on student learning gains and prior knowledge. Specifically, we segmented and unitized the students into two groups based on learning gains because the difference models of ONA do not handle further division. We analyzed the data from students with high and low prior knowledge separately. Both learning gains and prior knowledge groupings were generated using

median values ($Mdne = 8$ for pretest or prior knowledge and $Mdn = 0.15$ for learning gains).

We present comparative ordered network difference models for high and low learning gain groups for each prior knowledge group to identify potential variations in behaviors that may lead to differences in learning. We also conducted a correlation analysis (spearman) to assess the impact of these behaviors on learning. Table 1 shows the distribution of students based on learning gains and prior knowledge.

Table 1. Distribution of students based on learning gains and prior knowledge.

	Low Learning	High Learning	Subtotal
Low Prior Knowledge	16	14	30
High Prior Knowledge	18	18	36
Subtotal	34	32	66

3.4 Codes

Two sets of codes were applied to the log files of students used for the analyses in this study: location codes ($N = 7$) and actions codes ($N = 7$). Action codes were selected based on student behaviors that have proven relevant in previous work [13]. Location codes correspond to the 7 sites on the virtual island in the game. Both are automatically

Table 2. Definition of codes (Actions).

Code	Definition
Movement	The action of moving from one location to another is logged if the student performed an action in the new location <i>or</i> spent at least 10 s there
NPC Interaction	Talking with an NPC. Each message is counted as one action
Reading	Observing a book or research article (found in several game locations) for at least 5 s. Each minute of reading is counted as one action
Concept Matrix	Each submission of a concept matrix (where students summarize what they just read) counts as one action
Object	Picking up an object that would be used for testing a hypothesis in the future. Each object counts as one action
Hypothesis Testing	Scanning an object using laboratory equipment to determine if it contains the virus or bacteria the student hypothesized is affecting the island. Each scan counts as one action
Diagnosis Worksheet	Providing an entry in the game’s worksheet, which helps students to systematize the hypothesis testing process. Each entry counts as one action

recorded in Crystal Islands log files. Given the small number of classifications and their direct and unambiguous nature, it is unnecessary to report IRR for these codes.

Table 2 provides an overview of the action codes which were extracted from each line of the interaction logs. Although individual lines of data in the interaction logs do not account for the duration of each action, many actions have a relatively fixed duration that is consistent across students. For instance, hypothesis testing in the lab typically takes around 5 s. For other actions that lack a fixed duration, we employ a repetition strategy for considering the duration. For example, when students interact with NPCs, we consider each message they sent as a separate action, rather than grouping the entire interaction as a single data line. Thus, students who exchange more messages with the same NPC will have more lines representing this action. Similarly, in the case of reading and movement, if the student spends more than a minute on a particular page or moving between two locations, we duplicate the line to indicate that the student was engaged in that action for a longer duration.

Table 3 shows the location codes that were extracted from the students' interaction logs. In this study, location changes were coded either when the player performed an action other than "Movement" in the new location or spent at least 10 s there. This threshold differentiates between locations where the student was actively engaged and those where the student was simply moving between locations. Similarly, the code "Outside" was reserved for students who spent more than 10 consecutive seconds outside, suggesting deliberation or uncertainty about where to go next.

Table 3. Definition of codes (Locations).

Code	Definition
Tutorial	Students begin the game at this location (a beach), where they receive initial instructions, interact with an NPC, learn how to pick up objects, and are introduced to the concept matrix. They are instructed to go to the infirmary at the end of the tutorial
Infirmary	Students are instructed to go to this location after the tutorial to receive the only additional instructions in the game. These come from an NPC (the nurse), but they can also interact with NPC patients. They return later to provide the suggested treatment for the disease
Laboratory	Students bring objects to this location to test them for contamination with a virtual scanner (i.e., with the virus or bacteria causing the island inhabitants to get sick). They can also read research articles at this location. Because they can only carry 3 objects at the same time, students often require multiple visits to this location to solve the game
Living Quarters	In this location , students can acquire information by interacting with NPCs and reading disciplinary content

(continued)

Table 3. (continued)

Code	Definition
Bryce’s Quarters	Students who enter this location can interact with NPCs and read pages and articles associated with microbiology. This location has more disciplinary content than any other individual location in the game
Dining Hall	Location where students can find most of the objects to scan, including the one necessary for solving the mystery
Outside	This location composes the largest region of the game. All students must move through this area to get from one location to another, but it contains no additional information or objects to test. Therefore, it is only analyzed if the student spends at least 10 s without entering another location

In the current analysis, action codes and location codes are presented in separate difference models. While actions and locations are meaningfully connected, we took this approach for two reasons. First, in some cases, actions are bounded by location, which would artificially inflate the appearance of certain action/location connections. Second, ordered and epistemic networks become practically and visually overwhelming with higher numbers of codes present, often necessitating techniques to make existing models more parsimonious (e.g., [31]). In this case, presenting two separate models of different code sets avoids both issues and provides two corresponding perspectives on student experience.

3.5 Segmentation and Stanza Selection

ENA models complex systems by creating a series of adjacency matrices that connect the codes applied to one line to codes in prior lines (in this case, a moving window). *Conversation variables* serve as boundaries for these calculations by grouping data into stanzas across which associations are not calculated (e.g., all data associated with student 1 in one stanza and student 2 in separate stanza). For this study, we set student play sessions as the conversation variable. Extant research using non-linguistic data has used conversation variables such as game level [23] or complete playthrough [24]. Because the duration of CRYSTAL ISLAND play was short ($Avg = 65.1\ min$, $SD = 13.5$) and no natural break points were observed, we refrained from segmentation into more fine-grained stanzas.

To define the *moving window* length (which defines how far back codes in one line are associated with codes applied to prior lines) we drew on previous studies applying ENA to log files [23, 24]. In these cases, wider moving windows were needed to capture the contexts of fine-grained log data [32]. In Karumbaiah and colleagues’ [23] work on a level-based game with a limited action set, authors set the moving window to encompass approximately 20 s of gameplay that represented how far back a player might routinely connect a single in-game action to prior actions.

In CRYSTAL ISLAND, the duration of different actions varies widely; reading an article, for example, may take more time than picking up an object. To account for this,

we set the moving window width based on the average number of actions performed by students in two consecutive locations. This approach ensures that the associations in the networks are based on meaningful stanzas of gameplay, while also accounting for variations in the length of actions. It also aligns with the design of the game, in which sets of problem-solving actions are situated in specific in-game locations and students must choose to move between them in certain orders to access actions that can complete in-game goals [13, 17].

Students performed an average of 188.3 actions ($SD = 45.9$) and 22.5 visits to locations ($SD = 8.1$) throughout the gameplay. On average, students conducted 8.3 actions ($SD = 2.5$) during each visit to a given location before moving on to the next. In line with prior work on this topic [14], we connected the actions performed in one location to those in the following location, defining a moving window size of two times this average number of actions during each visit (17). We opted for this choice instead of the infinite stanza, as the initial actions were similar among all players, and recent actions are more relevant to current actions than the earlier ones. Moreover, we apply a threshold at 0.021 in the visualization of connections to ensure interpretability of the models.

4 Results

As discussed above, we divide the students in these analyses into high and low prior knowledge groups. We then consider, separately, how trajectories through different locations, and trajectories through different actions are associated with learning.

4.1 Low Prior Knowledge Students

This work explores differences between students who demonstrated high and low learning gains in both the high and low prior knowledge groups. For the low prior knowledge group, Mann-Whitney U tests showed significant differences ($alpha = 0.05$) between those with high and low learning gains for both: (1) location patterns ($U = 192.00, p < 0.001, r = -0.71$) and (2) action patterns ($U = 22.00, p < 0.001, r = 0.80$). Goodness of fit, measured using both Pearson and Spearman correlation indexes, was greater than 0.82 for both axes in each model. The ordered networks shown in this study only display connections with a weight higher than 0.02. We employ this threshold to prevent the figures from being overloaded with additional connections among codes that do not correspond to the most substantial differences between the compared groups.

Figure 2a and b displays the difference models between students with high (blue) and low (red) learning gains for the low prior knowledge group. We have tailored all the ONA model visualizations in this study to enhance the readability of transitions between nodes using the visualization software *Diagrams.net*, and strategically omitted line weights for visual clarity (see below). While our ONA visualizations deviate from the formats common to prior studies, the node positioning and arrow line weights are analogous to the ordered networks generated in R for this data. In these visualizations, the size of each node signifies the frequency of each action or location, and which group (high or low learning gains) undertakes that action or visits that location more often.

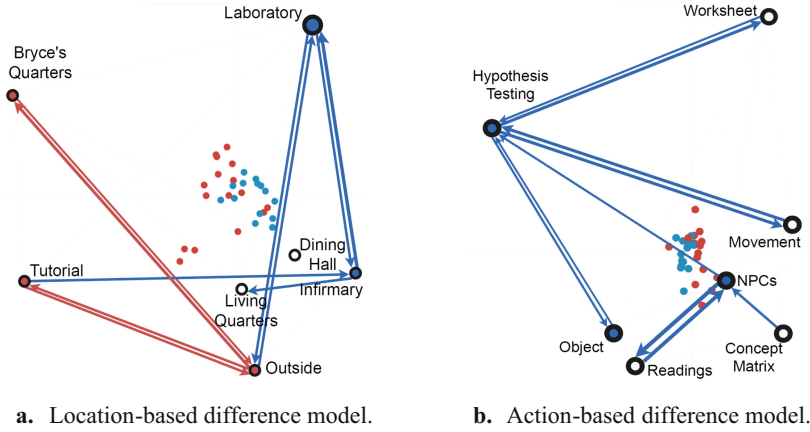


Fig. 2. Location-based and action-based difference models between students with high learning gains (blue) and low learning gains (red) for *low prior knowledge* students. Line weights below $|0.2|$ are omitted for visual clarity.

Students with higher learning gains in the low prior knowledge group were observed to engage in more transitions between testing hypotheses and filling out the worksheet with the corresponding results of the tests (*normalized line weight*, $lw = 0.048$), indicating a greater emphasis on understanding and framing the test results within the context of the game's goal. Additionally, students with higher learning gains were observed to transition more frequently between movement and hypothesis testing ($lw = 0.050$). This pattern of movements might be related to the ultimate goal of discovering the solution for the game, rather than seemingly undirected movement among locations.

The patterns of locations visited by students in the high learning gain condition also show similar results. Students with higher learning consecutively visited the laboratory (the place for hypothesis testing) and infirmary (the place where students receive game instructions and interact with patients suffering from the disease) more than students in the low learning group ($lw = 0.040$). This result aligns with a positive correlation between the number of hypotheses tested and learning gains for the low prior knowledge group ($\rho = 0.404$, $p = 0.027$). Similarly, students with higher learning gains transition directly from the tutorial to the infirmary, a place where they receive more productive information.

Students in the high learning group tended to visit the living quarters after visiting the infirmary more often than the low learning students ($lw = 0.028$). Overall, learning gains are positively correlated with the connection between interactions with NPCs and reading material ($\rho = 0.340$, $p = 0.066$), which can be completed in the infirmary and living quarters. These types of actions tend to be performed more often by students who understand the game's dynamics. The total number of actions, which are also associated with a better understanding of the game's logic, are correlated with learning gains for the low prior knowledge group ($\rho = 0.328$, $p = 0.077$). This aligns with the results shown by the actions-based difference model (Fig. 2b) for this group, where students who learned more tend to have more connections between actions than their peers in

the low learning gain condition, explaining why the line omission procedure resulted in a difference model that only showcases higher learners' higher rates of action. These results suggest that students may require the instructions provided in the infirmary to effectively learn from the disciplinary content taught in the living quarters.

In contrast, students in the low learning gain group tended to move between the tutorial and outside locations more often compared to their high learning gain peers ($lw = 0.036$). Low learning is also associated with Bryce's Quarters. This seems surprising, given that this is the location where most of the educational reading materials can be found. Although, as observed in Fig. 2b, the students with lower learning gains are not engaging more in the actions that can be conducted in Bryce's quarters than their peers who are learning more.

4.2 High Prior Knowledge Students

Statistical differences were also found among the high prior knowledge group, where a third and fourth Mann-Whitney tests revealed that the behavior of students with high learning gains ($Mdn = -0.10$, $N = 18$) was significantly different from the behavior of the students in the low learning gain group in terms of the visited locations ($Mdn = 0.10$, $N = 18$, $U = 288.00$, $p < 0.001$, $r = -0.78$), and their actions ($U = 231.50$, $p = 0.003$, $r = -0.43$) at the alpha = 0.05 level. The goodness of fit, calculated with both Pearson and Spearman correlation indexes, was greater than 0.86 for both axes.

Figure 3a and b shows the difference models between students with high (blue) and low (red) learning gains for the low prior knowledge condition. The behavior of students with high prior knowledge substantially differs from those with low prior knowledge. Results from Fig. 3a suggest that the transitions related to the infirmary are conducted

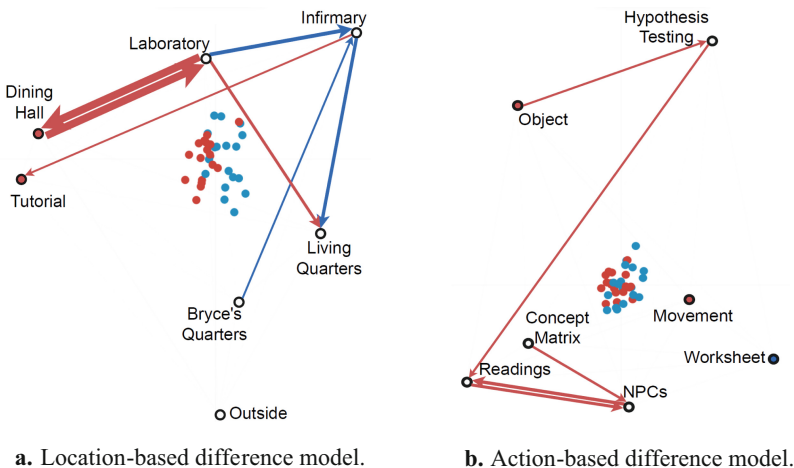


Fig. 3. Location-based and action-based difference models between students with high learning gains (blue) and low learning gains (red) for *high prior knowledge* students. Line weights below 0.21 are omitted for visual clarity.

more by the high learning group in the high prior knowledge condition. However, transitions having this place as starting point do not suppose a substantial difference between the two learning groups, as observed in the low prior knowledge case. Although some students with high prior knowledge and low learning gains tend to return from the infirmary to the tutorial ($lw = 0.022$), this connection is weaker than the observed back and forth between the tutorial and outside, in the low prior knowledge students.

For the high prior knowledge group, visits to the living quarters from the laboratory instead of the infirmary were also more common for the low learning gain group ($lw = 0.029$). Notably, while the laboratory was mainly visited by students in the high learning gain group for the low prior knowledge condition, the strongest connection observed in this study appeared between the laboratory (where testing occurs) and the dining hall (where most of the testable objects are located) for the low learning gain group among the high prior knowledge learners ($lw = 0.082$). Some students in this group may be repeatedly moving between the laboratory and the dining hall to obtain more objects, test more hypotheses, and solve the challenge quickly without engaging in reading or deeper understanding. This result aligns with a negative significant correlation between learning gains and the number of objects picked up by students in the high prior knowledge group ($\rho = -0.375$, $p = 0.024$).

The ordered network analysis for the high prior knowledge students also reveals a notable connection starting from hypothesis testing to readings for the low learning gain group ($lw = 0.022$), indicating that students in this group tend to read the disciplinary content after testing the hypothesis. This is the opposite of the desired order of these two actions and was also observed in the negative significant correlation between learning gains and the number of transitions from testing a hypothesis to reading ($\rho = -0.332$, $p = 0.048$).

Furthermore, as shown in Fig. 3b, students with lower learning gains also show a stronger connection between readings and interactions with NPCs (the two alternatives they have in the game to receive disciplinary content; $lw = 0.042$) than their high-learning peers. These students might be returning more to these two activities after picking an object and rejecting a hypothesis in the lab to receive more insights about other possible solutions. Although they are engaging more in these two activities, this does not imply that they are reflecting on this information, and therefore, they would not be learning as much as their peers.

5 Discussion

Understanding the relationship between prior knowledge and students' in-game behavioral patterns is an important step for developing new ways to support struggling learners, especially in an open-ended learning environment. While it is perhaps not surprising that students with lower prior knowledge likely need different learning experiences than those who come in with higher domain knowledge, understanding how to quantify those experiences has historically been challenging in such open-ended environments. This study uses an ordered network analysis (ONA) to better understand the trajectories of students in a middle-school science game and to gain insight into which behaviors are more likely to lead to learning (e.g., productive persistence, as opposed to wheel-spinning) even if they do not reflect the most efficient route through the game.

Results from this ONA show that overall, students with low prior knowledge require more location visits and actions in order to learn, while those with high prior knowledge are the opposite. It seems likely that this might extend to other learning contexts as well (e.g., [33]). That is, we should expect students with low levels of prior knowledge to have different questions and to require a higher level of repetition than students who already know quite a bit about the domain.

Beyond these differences, we also see that the same trajectories can be associated with the opposite effects on learning among high and low prior knowledge learners. Such results have important implications for supporting learners, whether through larger changes to the game design or through detector-driven prompts that are delivered when a student appears to be struggling.

Game design changes might explore ways to mitigate the ways in which the layout of the game influences student behavior. For example, ONA suggests that students with low learning gains may sometimes ignore the instructions to go directly to the infirmary (which is farther away from the tutorial than some of the other buildings in the game). While major game changes (e.g., moving the location of the infirmary closer to the tutorial) might be antithetical to the self-regulation the game is designed to foster, it might be useful to prompt students who have ignored these instructions a second time. Likewise, students who are ignoring reading materials in favor of what they think is a more efficient strategy (e.g., the loop of picking up new objects and testing them) might be prompted to make better use of the available resources in much the same way that a resident at a hospital might have a supervising physician suggest they read relevant articles. Finally, students might also be asked to externalize their hypotheses and their rationale for selecting the objects to scan before testing to see if they are relying on prior knowledge or disengaging from the educational material.

Other game design changes might be more subtle. For example, we currently see a strong pattern among high prior knowledge students with low learning gains. Many of these students seem to be looping between the dining hall (where most of the testable objects are) and the laboratory (where tests can be performed). It is possible that a more even distribution of testable objects and reading material throughout the game, might increase the chances that a student comes across (and makes use of) these resource articles. However, we note that this same looping behavior (between the dining hall and the laboratory) leads to higher learning gains for students with low prior knowledge. Therefore, any game redesigns should be carefully prioritized and tested among students with both high and low prior knowledge, which could be operationalized by integrating the pre-test into the game's adaptability algorithms.

Another possible solution is to implement wheel-spinning detectors (e.g., [8, 34–36]), and features based on the results of this ONA might help us to better predict this construct. For example, features related to multiple trips to the tutorial or excessive amounts of time wandering outside might be important indicators that a student does not have the skills to move forward even if they have not fully disengaged from the system. Likewise, we might develop features related to students who are testing without reading or reading but not leveraging that information to go to the next most productive location. Such detectors have been implemented in a range of learning software, but it can be more challenging to detect unproductive behaviors in less-linear systems. The results

here, however, suggest a path forward for such research using features discovered with ONA, akin to the discovery with models approach used in previous research [37].

6 Conclusions and Future Work

We have shown that ONA, a particular derivation of ENA, of student log files illustrates important patterns in open-ended learning environments, but this work is not without limitations. A potential limitation of this approach is that an ONA collapses a large number of student trajectories across time. That is, while ONA offers the advantage of directionality (compared to more traditional ENA methods) it does not show the specific moment when these connections occur during the student's game trajectory. Future work could work to address this limitation, either through data segmentation that could provide more nuanced insights into the trajectory of student behavior over time [38], or by supplementing ONA results with other techniques that examine time-series data. Future work might also leverage this approach to examine player differences based on other learner metrics such as content-related interest or in-game affect.

We have also shown that learning trajectories differ considerably among high and low prior knowledge learners, but future research should consider other student-level characteristics, since disciplinary content is not the unique factor that influences students' understanding of game dynamics. A few characteristics that may be especially important include prior game literacy [39] as well as reading skills and visual attention [40]. Students who are unaccustomed to exploration-based games or who struggle to comprehend written instructions (and written content in general) may encounter greater difficulty in engaging with and persisting in the game. Both lines of inquiry have the potential to yield valuable insights for improving game design, and supporting students' learning.

In summary, we envision the visualization and quantification capabilities of ONA as a significant contribution to closing the interpretation loop when examining the interactions between students and open-ended, game-based learning environments. Combining this analysis with a range of other data on student characteristics and in-game behaviors can yield a broad range of insights for game designers and educators to enhance game-based learning environments. In particular, it would be good to use these insights as part of a discovery with models approach [37] to better capture potential features for modeling differences between students who are wheel-spinning versus those that are persisting productively.

Acknowledgments. This work was funded by a grant from the National Science Foundation (NSF Cyberlearning #2016943). We would also like to thank Kristy Boyer for access to this data and Yuanru Tan for developing the ONA package used in this analysis. 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.

References

1. Clark, D.B., Sengupta, P., Brady, C.E., Martinez-Garza, M.M., Killingsworth, S.S.: Disciplinary integration of digital games for science learning. *Int. J. STEM Educ.* **2**(1), 2 (2015)
2. Gao, F., Li, L., Sun, Y.: A systematic review of mobile game-based learning in STEM education. *Educ. Tech. Res. Dev.* **68**, 1791–1827 (2020)
3. Saricam, U., Yildirim, M.: The effects of digital game-based STEM activities on students' interests in STEM fields and scientific creativity: minecraft case. *Int. J. Technol. Educ. Sci.* **5**(2), 166–192 (2021)
4. Shaffer, D.W., Gee, J.P.: *How Computer Games help Children Learn*. Palgrave Macmillan, New York (2006)
5. Annetta, L.A., Minogue, J., Holmes, S.Y., Cheng, M.T.: Investigating the impact of video games on high school students' engagement and learning about genetics. *Comput. Educ.* **53**(1), 74–85 (2009)
6. Anderson, J., Barnett, M.: Using video games to support pre-service elementary teachers learning of basic physics principles. *J. Sci. Educ. Technol.* **20**, 347–362 (2011)
7. Kanematsu, H., Kobayashi, T., Barry, D.M., Fukumura, Y., Dharmawansa, A., Ogawa, N.: Virtual STEM class for nuclear safety education in metaverse. *Procedia Comput. Sci.* **35**, 1255–1261 (2014)
8. Owen, V.E., et al.: Detecting wheel spinning and productive persistence in educational games. In: *Proceedings of the 12th International Conference on Educational Data Mining*, pp. 378–383 (2019)
9. Aleven, V., Stahl, E., Schworm, S., Fischer, F., Wallace, R.: Help seeking and help design in interactive learning environments. *Rev. Educ. Res.* **73**(3), 277–320 (2003)
10. Niefeld, J.L., Shores, L.R., Hoffmann, K.F.: Self-regulation and gender within a game-based learning environment. *J. Educ. Psychol.* **106**(4), 961 (2014)
11. Zhang, J., et al.: Investigating student interest and engagement in game-based learning environments. In: *Artificial Intelligence in Education: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part I*, pp. 711–716. Springer International Publishing (2022)
12. Steinkuehler, C.: Video games and digital literacies. *J. Adolesc. Health.* **54**(1), 61–63 (2010)
13. Sawyer, R., Rowe, J., Azevedo, R., Lester, J.: Filtered Time Series Analyses of Student Problem-Solving Behaviors in Game-Based Learning. *International Educational Data Mining Society* (2018)
14. Nasiar, N. et al.: It's good to explore: investigating silver pathways and the role of frustration during game-based learning. In: Wang, N., Rebolledo-Mendez, G., Dimitrova, V., Matsuda, N., Santos, O.C. (eds.) *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky*. AIED 2023. *Communications in Computer and Information Science*, vol. 1831. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-36336-8_77
15. Shaffer, D.W., et al.: Epistemic network analysis: a prototype for 21st-century assessment of learning. *Int. J. Learn. Media* **1**(2), 33–53 (2009)
16. Rowe, J.P., Shores, L.R., Mott, B.W., Lester, J.C.: Integrating learning, problem solving, and engagement in narrative-centered learning environments. *Int. J. Artif. Intell. Educ.* **21**(1–2), 115–133 (2011)
17. Reilly, J.M., Dede, C.: Differences in student trajectories via filtered time series analysis in an immersive virtual world. In: *Proceedings of the 9th International Conference on Learning Analytics and Knowledge*, pp. 130–134 (2019)

18. Abdi, H., Williams, L.J.: Principal component analysis. *WIRs: Comput. Stat.* **2**(4), 433–459 (2010)
19. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: analyzing the structure of connections in cognitive, social, and interaction data. *J. Learn. Analytics* **3**(3), 9–45 (2016)
20. Barany, A., Foster, A.: Examining identity exploration in a video game participatory culture. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) *ICQE 2019. CCIS*, vol. 1112, pp. 3–13. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33232-7_1
21. 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)
22. Bressler, D.M., Annetta, L.A., Dunekack, A., Lamb, R.L., Vallett, D.B.: How STEM game design participants discuss their project goals and their success differently. In: *Advances in Quantitative Ethnography: 3rd International Conf., ICQE 2021, Virtual Event, November 6–11, 2021, Proceedings 3*, pp. 176–190. Springer International Publishing (2022). https://doi.org/10.1007/978-3-030-93859-8_12
23. Karumbaiah, S., Baker, R.S., Barany, A., Shute, V.: Using epistemic networks with automated codes to understand why players quit levels in a learning game. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) *Advances in Quantitative Ethnography. ICQE 2019. Communications in Computer and Information Science*, vol. 1112. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33232-7_9
24. Scianna, J., Gagnon, D., Knowles, B.: Counting the game: visualizing changes in play by incorporating game events. In: Ruis, A.R., Lee, S.B. (eds.) *Advances in Quantitative Ethnography. ICQE 2021. Communications in Computer and Information Science*, vol. 1312. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_15
25. Tan, Y., Ruis, A.R., Marquart, C., Cai, Z., Knowles, M., Shaffer, D.W.: Ordered network analysis. In: Damşa, C., Barany, A. (eds.) *Advances in Quantitative Ethnography: Fourth International Conference, ICQE 2022. Springer International Publishing* (2022). https://doi.org/10.1007/978-3-031-31726-2_8
26. Wan, H., Beck, J.B.: Considering the Influence of Prerequisite Performance on Wheel Spinning. *International Educational Data Mining Society* (2015)
27. Palaoag, T.D., Rodrigo, M.M.T., Andres, J.M.L., Andres, J.M.A.L., Beck, J.E.: Wheel-spinning in a game-based learning environment for physics. In: Micarelli, A., Stamper, J., Panourgia, K. (eds.) *ITS 2016. LNCS*, vol. 9684, pp. 234–239. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39583-8_23
28. Fan, Y., et al.: Dissecting learning tactics in MOOC using ordered network analysis. *J. Comput. Assist. Learn.* **39**(1), 154–166 (2023)
29. Min, W., et al.: Multimodal goal recognition in open-world digital games. *Proc. AAAI Conf. Artif. Intell. Interact. Dig. Entertainment* **13**(1), 80–86 (2021)
30. Vail, A.K., Grafsgaard, J.F., Boyer, K.E., Wiebe, E.N., Lester, J.C.: Predicting Learning from student affective response to tutor questions. In: Micarelli, A., Stamper, J., Panourgia, K. (eds.) *ITS 2016. LNCS*, vol. 9684, pp. 154–164. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39583-8_15
31. Wang, Y., Swiecki, Z., Ruis, A.R., Shaffer, D.W.: Simplification of epistemic networks using parsimonious removal with interpretive alignment. In: Ruis, A.R., Lee, S.B. (eds.) *ICQE 2021. CCIS*, vol. 1312, pp. 137–151. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_10
32. Arastoopour Irgens, G., Shaffer, D.W., Swiecki, Z., Ruis, A.R., Chesler, N.C.: Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *Int. J. Eng. Educ.* **32**, 1492–1501 (2015)

33. Beck, J.E., Gong, Y.: Wheel-spinning: Students who fail to master a skill. In: *Artificial Intelligence in Education: 16th International Conference AIED 2013, Memphis, TN, USA, July 9–13, 2013*, pp. 431–440. Springer, Berlin Heidelberg (2013)
34. Yang, J.C., Quadir, B.: Effects of prior knowledge on learning performance and anxiety in an English learning online role-playing game. *J. Educ. Technol. Soc.* **21**(3), 174–185 (2018)
35. Kai, S., Almeda, M.V., Baker, R.S., Heffernan, C., Heffernan, N.: Decision tree modeling of wheel-spinning and productive persistence in skill builders. *J. Educ. Data Min.* **10**(1), 36–71 (2018)
36. Botelho, A.F., Varatharaj, A., Patikorn, T., Doherty, D., Adjei, S.A., Beck, J.E.: Developing early detectors of student attrition and wheel spinning using deep learning. *IEEE Trans. Learn. Technol.* **12**(2), 158–170 (2019)
37. Baker, R.S., Yacef, K.: The state of educational data mining in 2009: a review and future visions. *J. Educ. Data Min.* **1**(1), 3–17 (2009)
38. Brohinsky, J., Marquart, C., Wang, J., Ruis, A.R., Shaffer, D.W.: Trajectories in epistemic network analysis. In: Ruis, A.R., Lee, S.B. (eds.) *ICQE 2021. CCIS*, vol. 1312, pp. 106–121. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_8
39. Squire, K.D.: Video-game literacy: A literacy of expertise. In: *Handbook of Research on New Literacies*, pp. 635–669. Routledge (2014)
40. Kress, S., Neudorf, J., Borowsky, B., Borowsky, R.: What’s in a game: video game visual-spatial demand location exhibits a double dissociation with reading speed. *Acta Physiol. (Oxf)* **232**, 103822 (2023)