



Does agency matter?: Exploring the impact of controlled behaviors within a game-based environment



Erica L. Snow^{*}, Laura K. Allen, Matthew E. Jacovina, Danielle S. McNamara

Arizona State University, Learning Sciences Institute, Tempe, AZ, USA

ARTICLE INFO

Article history:

Received 19 August 2014

Received in revised form

8 December 2014

Accepted 10 December 2014

Available online 18 December 2014

Keywords:

Agency

Stealth assessment

Entropy

Adaptive environments

Game-based learning

ABSTRACT

When students exhibit control and employ a strategic plan of action over a situation they are said to be demonstrating agency (Bandura, 2001). The current work is comprised of two studies designed to investigate how agency manifests within students' choice patterns and ultimately influences self-explanation quality within the game-based system iSTART-2. In Study 1, 75 college students interacted freely within iSTART-2 for 2 h. Random walk and Entropy analyses were used to quantify the amount of control demonstrated in students' choice patterns, as well as to determine the relation between variations in these patterns and self-explanation performance within iSTART-2. Overall, students who demonstrated more controlled choice patterns generated higher quality self-explanations compared to students who exhibited more disordered choice patterns. This link between performance and controlled choice patterns is hypothesized to be driven, in part, by students' experiences of agency. That is, engaging in controlled patterns should be advantageous only when doing so is a result of students' strategic planning. In Study 2, this hypothesis was tested by assigning 70 students to a choice pattern (i.e., controlled or disordered) that had been yoked to students from Study 1, thus removing students' ability to exert agency over the iSTART-2 system. Results revealed no differences in self-explanation quality between the groups assigned to controlled and disordered choice patterns. Collectively, findings from these studies support the notion that success within game-based systems is related to students' ability to exert agency over their learning paths.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Everyday, people make decisions, set plans, and exert influence over their daily lives. Individuals control a multitude of situations through their choices, decisions, and strategic plans. They exert *agency* (Metcalfe, Eich, & Miele, 2013). Indeed, agency (or a lack thereof) is a pervading aspect of our lives.

According to Bandura (2001), there are four main features of human agency: intentionality, forethought, self-reactiveness, and self-reflectiveness. *First*, agency refers to a *deliberate* action or set of actions, which are purposefully enacted according to a specific plan. *Second*, agency involves forethought. A person exerting high levels of agency will have set goals and plans for how to obtain these goals before carrying out any actions. *Third*, agency is self-reactive, emerging from a person's motivation to succeed and self-regulate. *Finally*, agency is self-reflective, such that the person who is exerting agency is metacognitively aware of the goals, plans, and behavior adjustments necessary to complete a task. Combined, these four components portray agency as a dynamic behavior that involves intentionality, metacognition, and planned sets of behaviors designed to accomplish a goal. Considering agency in this light, it is not surprising that individuals who demonstrate agency over their environment generally lead more successful lives compared to those people who do not (Bandura, 2001; Ford, 1992).

^{*} Corresponding author. Department of Psychology, Arizona State University, Tempe, AZ, USA. Tel.: +1 317 224 3188.
E-mail address: Erica.L.Snow@asu.edu (E.L. Snow).

2. Educational environments and student agency

Within the realm of education, agency emerges as an important factor that influences students' engagement and subsequent learning of academic material (Bandura, 1989; Zimmerman, 2008). Indeed, it is a widely accepted belief in the classroom that affording students control promotes their motivation and subsequent learning outcomes (Flowerday & Schraw, 2000). In support of that assumption, students' emotions while learning have been moderately associated with their perceptions of subjective control (Pekrun, 2006; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010). According to the control-value theory of emotion (Pekrun, 2006; Pekrun et al., 2010), a greater sense of control is related to students' expressed positive attitudes during a learning task. Consequently, engagement is expected to increase when students are afforded autonomy and control during various academic tasks (Calvert, Strong, & Gallagher, 2005; Cordova & Lepper, 1996; Deci & Ryan, 1985; Hidi & Renninger, 2006). Cordova and Lepper (1996), for instance, found that learners who were afforded more opportunities to exert control over a learning task reported stronger motivation and interest and showed better performance on a subsequent math test. Similarly, Calvert et al. (2005) reported that students who were given more control over a computer-based storybook reported greater interest in the task and were more attentive to the material than students who were given more explicit directions by adults. These studies provide growing evidence that agency during learning has the potential to enhance motivation, interest, and attitudes, all of which are associated with positive learning outcomes.

Adaptive learning environments have attempted to leverage these positive effects of agency by incorporating various elements such as customization, games, and “choose your own adventures.” These game-based elements are designed to promote students' feelings of agency and by consequence enhance motivation, performance, and learning outcomes (Cordova & Lepper, 1996; Jackson & McNamara, 2013; Snow, Jackson, Varner, & McNamara, 2013). For example, providing learners with control over when and how long they engage with different lessons (i.e., their learning trajectory) in a system can improve learning outcomes (Tabbers & de Koeijer, 2010). Even giving control over educationally superficial features of a system (e.g., choosing the images that will be depicted by the system) allows learners' experiences to match their personal preferences. This in turn can decrease the effort required to engage in a task and subsequently increase involvement and learning (Corbalan, Kester, & van Merriënboer, 2009; 2011). In this way, students can feel as though they are exerting control over their environment with minimal changes to the learning task itself.

Game-based systems are particularly germane to the issue of a student's sense of agency during learning. By leveraging the mechanics and features found in popular, non-educational games, learning environments infused with games naturally afford students the ability to exert influence on the learning environment (McNamara, Jackson, & Graesser, 2010). Indeed, a number of game features have been adapted from popular games to educational games with the purpose of increasing player engagement and the likelihood of players experiencing agency in a system. For instance, many popular video games allow players to make choices within the environment, follow non-linear paths through the game (e.g., Grand Theft Auto V) or select between several available mini-games (e.g., Nintendo Land). Likewise, educational games often include choices for how to progress through a system as a means to support player agency and increase replayability (Snow, Jacovina, Allen, & McNamara, 2014; Spires, Rowe, Mott, & Lester, 2011). Popular commercial games frequently allow players to customize the visual appearance of game features to their preference (e.g., a player's avatar in World of Warcraft), and this game-based feature (i.e., choice) has been associated with increased immersion and intention to replay a game (Schmierbach, Limperos, & Woolley, 2012; Teng, 2010; Yee, 2006). Similarly, personalization enhances students' motivation and learning outcomes (Cordova & Lepper, 1996).

One of the most effective features that has been incorporated into educational games is user choice. Choices made by individual players have the potential to provide students with a sense of agency, as they are engaging on intellectual or emotional levels and prompt players to persist in their play (Schönau-Fog & Bjørner, 2012). Mechanics and features that promote engagement in these ways are found in educational games such as *Crystal Island* (Lester, Mott, Robison, Rowe, & Shores, 2013) and *Quest Atlantis* (Barab, Pettyjohn, Gresalfi, Volk, & Solomou, 2012), where students are immersed in 3-D game environments. In *Crystal Island*, for example, players control an avatar and explore an island where an illness has recently spread. Players interact with both the environment and other game characters to discover information about this outbreak, and in the process, learn microbiology course content. An important advantage for *Crystal Island* over traditional instruction is that it can promote a strong sense of agency, as students have control over how they obtain knowledge in this environment. A study examining students' performance during the game and on posttest content questions found that students who did well in the game also did well at posttest, and these students were more successful at gathering information during play (Rowe, Shores, Mott, & Lester, 2010). Students who did not do as well in the game, however, also scored lower at posttest and demonstrated less successful information-gathering behaviors. These findings suggest that when students successfully take agency over their learning experience, they achieve higher outcomes in terms of learning the target material.

Despite the generally positive effects of system choices, research suggests that the inclusion of user control may not be universally beneficial for all students (Katz, Assor, Kanat-Maymon, & Bereby-Meyer, 2006). For instance, while some studies have shown added elements of user choice to be associated with positive outcomes (Cordova & Lepper, 1996; Reynolds & Symons, 2001), others have shown inconsistent (Flowerday & Schraw, 2003), neutral (Parker & Lepper, 1992), or negative effects (Flowerday, Schraw, & Stevens, 2004; Iyengar & Lepper, 2000). One reason for these conflicting results may be that users react differently when presented with increased levels of control. Some students may regulate their behaviors and exert control over the environment—inspiring strong feelings of agency—while others may struggle to set goals and make decisions (Zimmerman, 1990). The ability to exert control during a learning task is challenging for many students, as they often struggle to actively monitor their behaviors (Ellis & Zimmerman, 2001). Overall, research suggests that the inclusion of user control (e.g., choice) has the potential to increase learning outcomes among students; however, these effects may vary based on individual differences in users' ability to control their behaviors (McNamara & Shapiro, 2005).

3. Assessment of student agency

Students' inconsistent ability to exert control over their environment has posed an assessment problem for researchers, as it is difficult to capture fine-grained behavior variations. Traditionally, students' feelings of agency have been assessed through self-reports (Ellis & Zimmerman, 2001; Zimmerman, 1990). These direct measurements are static in nature and often miss out on behavior patterns that emerge over time. An alternative to self-report metrics is the use of stealth assessments (Shute, 2011; Shute, Ventura, Bauer, & Zapata-

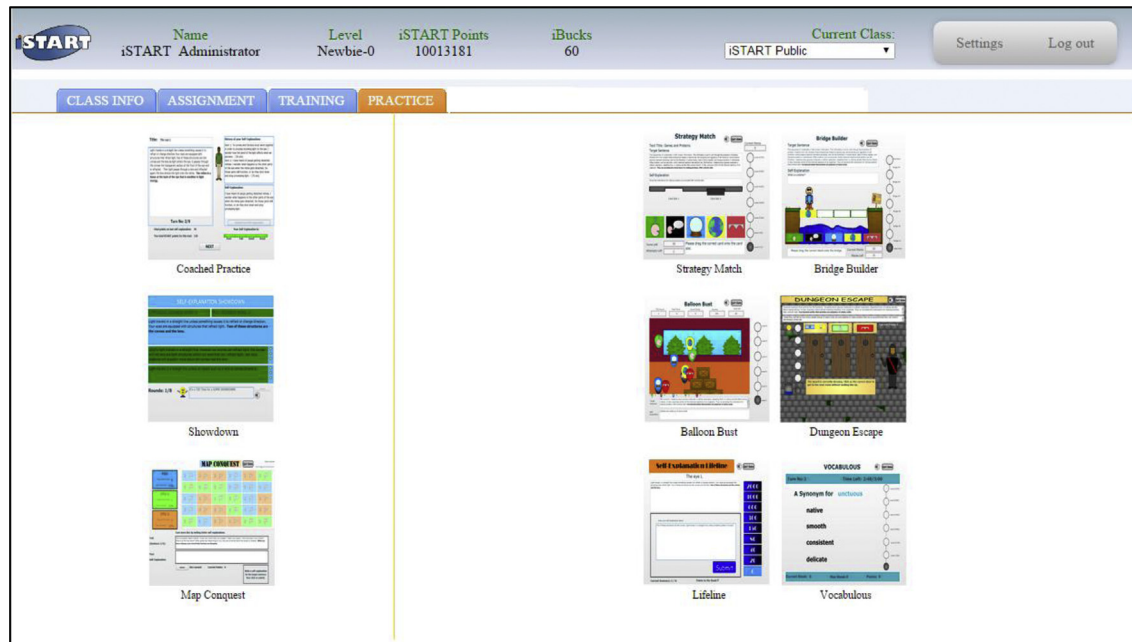


Fig. 1. Screen shot of iSTART-2 Selection Menu.

Rivera, 2009). Stealth assessments are used to covertly measure some attribute or construct without explicitly disrupting the student (Shute, 2011; Shute et al., 2009). These assessments have previously been used to measure a variety of constructs, including students' study habits (Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007) and self-regulation ability (Sabourin, Shores, Mott, & Lester, 2012). Game-based systems offer researchers a novel form of stealth assessment through the use of log data (e.g., keystrokes, mouse clicks, choice patterns).

An important goal for researchers is to devise measurements and methods for analyzing the log data that is generated within these game-based systems. Indeed, vital information may emerge by testing the degree to which variations in students' behavior patterns (as indicated by the system log data) shed light on how students experience agency, and whether those experiences influence learning outcomes. One potential approach to measuring individual differences in controlled behavior is through the use of analyses inspired by dynamical systems theory. Dynamical analysis focuses on the complex behaviors that emerge within a given environment; thus, time is treated as a critical variable in addressing patterns of variation and consistency. Because of the focus on time, dynamical methodologies offer scientists a novel means of classifying variations in students' behavior patterns when they are given agency within an adaptive system. Researchers have previously used dynamical methodologies to investigate variations in behavior patterns within various adaptive systems (Hadwin et al., 2007; Snow, Allen, Russell, & McNamara, 2014; Snow, Jackson, & McNamara, 2014; Snow, Likens, Jackson, & McNamara, 2013; Zhou, 2013). This work has demonstrated the potential for dynamical methodologies to capture nuanced and fine-grained patterns that reveal how students approach learning tasks embedded within adaptive environments.

4. Overview of the current work

The current work is comprised of two studies designed to investigate how agency may manifest within system log data and the ultimate impact that agency has on learning outcomes. In Study 1, we employ three novel dynamical methodologies to investigate how variations in behavioral patterns emerge when students have the potential to exert high levels of control (i.e., they have many choices) within an adaptive environment. Although this level of control should lead to increased levels of perceived agency for students, some may struggle to exert control over such an open environment. In Study 1, we examine *how* students interact with the game-based system iSTART-2, and the subsequent learning outcomes associated with those behavior patterns. In particular, we are interested in the impact that controlled and disordered interaction patterns have on target skill acquisition. We hypothesize that students who experience higher levels of agency will exhibit more controlled patterns of behavior and show higher levels of target skill acquisition compared to students who experience lower levels of agency and act in more disordered patterns. Such findings would support previous work showing that when students exert high levels of agency, they are employing controlled and strategic plans of action (Bandura, 2001) and their subsequent learning performance increases (Zimmerman, 1990). If particular interaction patterns (e.g., controlled patterns) are associated with higher levels of performance, our agency-based account predicts that students will not experience the benefits of those interaction patterns in the absence of choice. Study 2 tests this prediction by removing the opportunity to experience agency by randomly assigning students to an interaction pattern that they must follow. Both Study 1 and Study 2 are presented within the context of the game-based system, iSTART-2.

5. iSTART-2

The iSTART (Interactive Strategy Training for Active Reading and Thinking) program was designed to provide high school students with instruction on the use of self-explanation and comprehension strategies (Jackson & McNamara, 2013; McNamara, Levinstein, & Boonthum, 2004; McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, 2007). Studies have confirmed that students' comprehension and self-

explanation ability is enhanced when they are provided with iSTART strategy instruction (Jackson & McNamara, 2013; McNamara, Boonthum, Levinstein, & Millis, 2007; McNamara, O'Reilly, et al., 2007; O'Reilly, Sinclair, & McNamara, 2004; Taylor, O'Reilly, Rowe, & McNamara, 2006). iSTART-2 (Snow, Jacovina et al., 2014) is the most recent version of iSTART. This system provides students with strategy training within the context of a game-based environment. This environment is designed to enhance students' engagement and persistence during prolonged periods of training. When students are engaged and express interest in a learning task, they are more likely to show positive learning outcomes (Pintrich, 2000). The game-based features included in iSTART-2 are designed to enhance students' motivation and engagement during training and, depending on various circumstances, improve learning outcomes (McNamara et al., 2010).

iSTART-2 consists of two phases: training and practice. Within the training phase, students are introduced to a pedagogical agent who explains and defines the concept of self-explanation. This agent also discusses the iSTART-2 comprehension strategies: comprehension monitoring, predicting, paraphrasing, elaborating, and bridging. Within this phase, students are provided with examples of each comprehension strategy in separate lesson videos. At the end of each lesson video, students are given a short quiz that assesses their understanding of the strategy. During the practice phase of iSTART-2, students are transitioned to an interactive game-based interface. Within this interface, students can read and self-explain new texts, personalize different aspects of the interface and play mini-games (see Fig. 1).

In addition, students can check their personal accomplishments within the system by viewing achievement screens that update students on their current level, number of points earned, and total trophies won. Students increase their level within the system by earning points when interacting with two different types of generative practice where they write their own self-explanations: Showdown, and Map Conquest. These generative practice games were designed to engage students' interest while they practice using strategies by generating self-explanations. For example, in Showdown, students are asked to generate a self-explanation for numerous target sentences while competing against another player. The student's and computer player's generated self-explanations are compared and the highest quality self-explanation wins the round and any subsequent points (see Fig. 2).

As students earn more points within the system, they progress through a series of levels ranging from 0 to 25. Every level progression requires more points to proceed than the previous level, which ensures that students have to exert more effort to progress to higher levels within the system. The points that students earn throughout their interactions with iSTART-2 also serve as a form of currency (iBucks) that can be used to purchase game-based features. Within iSTART-2, each game-based system feature costs 300 iBucks per interaction. There are two ways students can choose to spend their earned iBucks: personalize the system interface and play mini-games. Students can choose to personalize the system interface by editing an avatar or customizing the background theme. When students choose to edit their avatar, they have the option to change the hairstyles and accessories that their avatar displays in the interface. If students choose to edit their background theme, they can modify the color of the interface (24 total color options). These two features were built into the interface to afford students a sense of control over the environment. However, both of these features were designed to be off-task and tangential to the learning goal of the system.

Students can also choose to spend their earned iBucks on a suite of four mini-games. These mini-games were added to iSTART-2 as a means of identification practice for the self-explanation strategies that the students learned. Each mini-game allows students to engage in play while at the same time practicing reading comprehension strategies. Mini-games are designed to be on-task and act as an extension of the learning goal of iSTART-2. Although these games vary in their game mechanics, the strategy identification task is very similar in each. One example of an iSTART-2 mini-game is Balloon Bust. In Balloon Bust, students are presented with a text and a self-explanation. Students decide which previously learned strategy was used to generate the self-explanation and click on the corresponding balloon to pop it (see

Self Explanation SHOWDOWN!

Player 1 (You)	Sentences Won: 1	Player 2	Sentences Won: 0
WINS!	iSTART Points: 50		

Diagnosis of Genetic Disorders

Text: Once it is clearly understood where a gene is located and the gene's DNA sequence is known, a diagnosis of a genetic disorder may be made before birth. **The DNA of people with and without the disorder is analyzed for common patterns that may be associated with the disorder.** Text

Opponent: this says that the dna of people with disorders is mutated. this causes the genes to go haywire. Opponent

You: I know about DNA. DNA is the building block for all the proteins in our bodies. It is our genetic map. You

50

Player 1 wins this point.

Flag this SE

Fig. 2. Screen shot of Showdown.

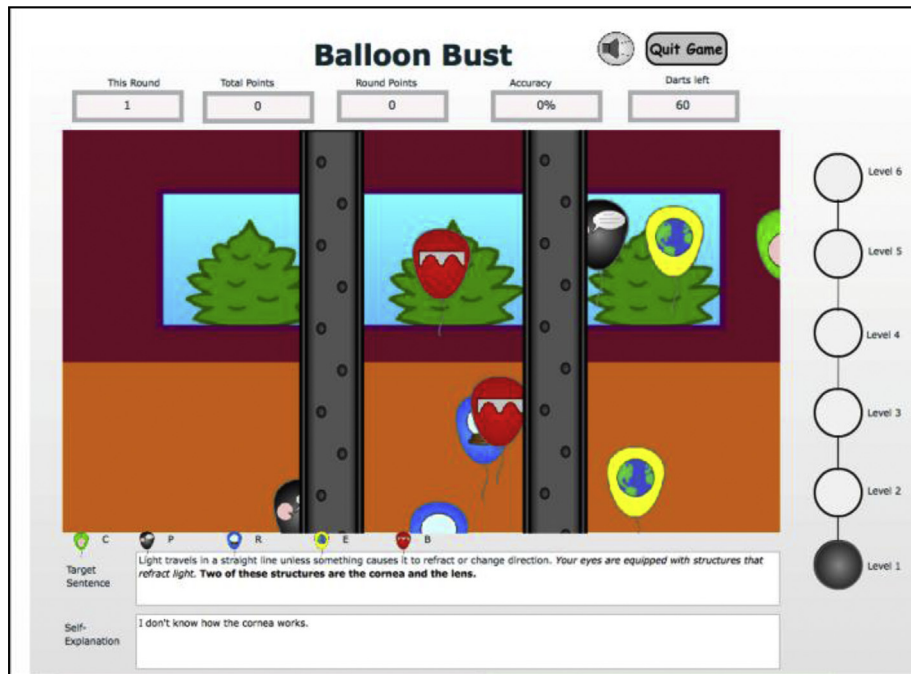


Fig. 3. Screen shot of Balloon Bust.

Fig. 3). In each game, students are given a score based on their performance and a corresponding trophy, if applicable. Students can accumulate trophies throughout their time in the system and view their performance on the achievement screen tab (i.e., points, levels, and trophies) on the main interface menu.

Each of these game-based features (i.e., personalizable features and mini-games) is available in the iSTART-2 interface. The system allows features to become unlocked as students progress to higher levels within the system. For instance, the Mohawk hairstyle for avatars can be locked until students reach level 11. The system was designed this way to ensure that some features can act as incentives to enhance students' performance and effort within the system (for more information about the iSTART-2 design, please see Jackson & McNamara, 2013). However, for the current study, all features were unlocked and accessible to the students from the outset.

6. Study 1

Game-based learning environments frequently include elements designed to increase feelings of agency and to provide students with control over their learning trajectory. Increasing students' agency generally has a positive effect on learning; however, the benefits accrued by students depend on their ability to successfully take control over their environment. Dynamical analyses provide the means to assess students' choice patterns within learning environments and to capture fluid changes in fine-grained behavior patterns. These fine-grained measures potentially afford a deeper understanding of how students vary in their ability to exert agency over their environments and how those behavior patterns relate to learning outcomes. Study 1 uses three dynamical techniques (random walks, Euclidean distances, and Entropy scores) to visualize and quantify students' choice patterns within the iSTART-2 interface. Using these methodologies, we investigate how variations in choice patterns emerge and ultimately impact students' learning outcomes (i.e., self-explanation quality) within the context of iSTART-2.

7. Method

7.1. Participants

This study included 75 college students from a large university campus in Southwest United States. These students were, on average, 18.8 years of age (range: 18–24 years), with a mean reported grade level of college freshman. Of the 75 students, 57% were male, 56% were Caucasian, 23% were Asian, 5% were African-American, 11% were Hispanic, and 5% reported other nationalities.

7.2. Procedure

This study included one 3-h session consisting of a pretest, strategy training (via iSTART-2), extended game-based practice within iSTART-2, and a posttest. At pretest, students were asked to answer a battery of questions that assessed their prior motivation and attitudes. During training, all students watched the iSTART training videos, which instructed them on the application of self-explanation strategies. After students watched the training videos, they were transferred into the game-based practice menu embedded within iSTART-2. During their time within the game-based practice menu, students were free to interact with the system interface anyway they chose. Students spent

approximately 2 h within the game-based interface. After they finished game-based practice, all students were transitioned to the posttest, where they completed attitude questionnaires similar to those in the pretest.

7.3. Measures

7.3.1. Strategy performance

During game-based practice, students' generated self-explanation quality was measured using an algorithm that combines both Latent Semantic Analysis (LSA; Landauer, McNamara, Dennis, & Kintsch, 2007) and word-based measures (Jackson, Guess, & McNamara, 2009; McNamara, Boonthum, et al., 2007; McNamara, O'Reilly, et al., 2007). Within iSTART-2, all self-explanations are scored on a scale from 0 to 3. Students are assigned a training self-explanation score by averaging the scores of all their generated self-explanations.

7.3.2. System interaction choices

In this study, students were free to interact within the iSTART-2 system. The game-based features fall into one of four types of game-based feature categories. Each type of game-based feature category represents a different functionality within the system interface.

1. *Generative practice games* include game-based practice environments (i.e., Map Conquest, and Showdown) where students generate their own self-explanations. Within each of these practice games students receive feedback regarding their self-explanations. This provides each student the opportunity to apply and revise comprehension strategies to challenging texts.
2. *Identification mini-games* include four game-based practice environments designed to reinforce learning strategies by asking the students to identify the self-explanation used to generate example self-explanations. These mini-games are designed to provide an alternative form of strategy practice (i.e., strategy recognition) to students.
3. *Personalizable features* (i.e., avatar and background customization) provide students with the opportunity to customize the system interface. These elements were designed to afford students a feeling of personal investment during long-term practice.
4. *Achievement screens* provide students with the opportunity to view their earned trophies, self-explanation scores, levels, and points within the system. These menus were embedded within the system to provide students with the opportunity to monitor their progress during strategy practice.

7.3.3. Game performance

In both the generative practice and identification mini-games, students earn trophies based on their performance. Students can earn bronze, silver, and gold trophies throughout their time in iSTART-2 and can view these accomplishments on the achievement screen.

7.3.4. Posttest attitudes

At posttest, students were asked to answer three questions that assessed their confusion, feelings of control, and boredom while engaged within the system (see Table 1). These single-item measures have been used in previous studies to assess individuals' motivation and enjoyment within adaptive systems (e.g., Jackson & McNamara, 2013; Snow, Jackson, et al., 2013).

The use of single-item versus multi-item measurements has been subject to much debate within the psychological literature (De Boer et al., 2004). Indeed, proponents of multi-item measures (i.e., 3 or more items designed to measure one construct) argue that the use of multidimensional scales improves the validity and reliability of the intended measurement (McIver & Carmines, 1981; Nunnally & Bernstein, 1994). However, the use of single-item measures has been shown to have practical advantages. For instance, Robins, Hendin, and Trzesniewski (2001) argued that the use of single-item measures can help reduce the fatigue, frustration, and boredom that is typically associated with high redundancy in multi-item scales. This practicality has led to the wide spread use of single-item measures as a way to assess constructs such as, intelligence (Paulhus, Lysy, & Yik, 1998), agency (Metcalfe et al., 2013), and anxiety (Davey, Barratt, Butow, & Deeks, 2007).

In the current study, all single-item measures were presented as forced choice survey scales ranging from 1 to 6 (1 = strongly disagree; 6 = strongly agree; see Table 1). Even-numbered scales remove the opportunity for students to adopt a middle, *neutral* stance (e.g., no opinion, I do not know, not applicable). That is, selecting 1 to 3 indicates some level of disagreement with the statement and selecting 4 to 6 indicates some level of agreement. The use of a neutral option within Likert-scales presents a variety of empirical issues (Moors, 2008). First, when middle or neutral options are explicitly offered, participants can be more likely to choose them (Bishop, 1987). Furthermore, it has been argued that interpretations of middle responses vary aside from true neutrality (e.g., the item is not applicable, or the respondent is unwilling to answer; Stone, 2004). This ambiguity has proven problematic for researchers attempting to establish construct validity (Klopfer & Madden, 1980; Stone, 2004). One goal of the current work is begin to establish the validity of the relation between self-report measures of agency and actual observed behaviors within an adaptive learning environment. Thus, in this work, we employ a forced-choice design in an attempt to eliminate the ambiguity and validity concerns associated with a neutral option.

Table 1
Posttest attitude measures.

Dependent measure	Response statement	Response scale
Confusion	"I was confused about what I should be doing"	1–6
Control	"I felt like I had no control over the system"	1–6
Boredom	"I felt bored in the system"	1–6
1 (strongly disagree) to 6 (strongly agree)		

7.4. Data processing

During students' time in the game-based practice menu, all choices (e.g., to play a mini-game or update their avatars) were recorded. These data logs were then used to categorize each interaction as one of the four game-based feature category types (i.e., generative practice games, identification mini-games, personalizable features, and achievement screens).

8. Quantitative method

Variations in students' behavior patterns were assessed using three dynamical methodologies: random walks, Euclidean distances and Entropy analyses. These methodologies afford the opportunity to quantify variations in students' choice patterns and examine how these different trajectories impacted students' learning outcomes (i.e., self-explanation quality) within the context of iSTART-2. A description and explanation of random walks, Euclidean distances, and Entropy analyses are described below.

8.1. Random walks

Random walks generated a representation of each student's unique interaction trajectory within iSTART-2. This mathematical tool provides a spatial representation of patterns within categorical data as they manifest across time (Benhamou & Bovet, 1989; Lobry, 1996). Each student's trajectory within the system was represented by first examining the sequential order of interactions with various game-based features. Each game-based feature category was assigned an orthogonal vector along an X, Y scatter plot (see Table 2). The assignment of these vector locations is random and not associated with any qualitative value. Random walks have previously been used to trace students' interaction patterns within the game-based system, iSTART-ME and within the writing tutor, Writing Pal (Allen, Snow, & McNamara, 2014; Snow, Allen, Jackson, & McNamara, 2014; Snow, Jacovina, et al., 2014; Snow, Likens, et al., 2013). Each student's random walk began at the origin (0,0). Then, using the system log data to examine the sequential order of students' choices, the particle moves in a manner consistent with the vector assignment of the specified choice. The culmination of the movement of the particle results in a continual trajectory or "walk" that visually represents each student's time within the iSTART-2 system.

To illustrate what a random walk might look like for a student who made four choices within the iSTART-2 system, see Fig. 4. For all students, the starting point of their walk is (0,0); this is where the horizontal and vertical axes intersect. In this example, the first interaction the student chose was a generative practice game; so, the particle moves one unit to the left along the X-axis (see # 1 in Fig. 4). The second interaction choice that the student made was to play a mini-game; thus, the particle moves one unit up along the Y-axis (see # 2 in Fig. 4). The student's third interaction choice was with another practice game, which moves the particle one unit left along the X-axis (see # 3 in Fig. 4). The student's fourth interaction choice was with an achievement screen, which moves the particle one unit down along the Y-axis (see # 4 in Fig. 4). Using these simple rules, a random walk was calculated for every student ($n = 75$) who interacted with iSTART-2.

Figs. 5 and 6 are students' actual walks generated to represent their time spent within iSTART-2. The student represented in Fig. 5 reveals a walk with an upward trajectory; we can then infer that this student interacted predominantly with the identification mini-games. Conversely, the student represented in Fig. 6 demonstrated a rightward walk trajectory. Thus, we can infer that this student spent the majority of the time interacting with the personalizable features. Interestingly, both walks reveal that students interacted with a multitude of features, as evidenced by the fluctuations in the figures. Overall the use of random walks provides researchers with a means to visualize trajectories and patterns of choices.

8.2. Euclidian distance

As illustrated in Fig. 5, random walks capture fluctuations in students' choice patterns that may inform researchers on the degree to which a student is controlled or deliberate when approaching a learning task. To quantify these fluctuations, distance time series were calculated for each student by using a measure of Euclidean distance. For each student, Euclidian distance was measured from the origin (coordinates 0,0) to each step within his or her walk (see Equation (1)). In the Euclidian distance equation, y and x represent the particle's place on the y -axis and x -axis, respectively, and the i represents the i th step within each walk:

$$Distance = \sqrt{(y_i - y_0)^2 + (x_i - x_0)^2} \quad (1)$$

A Euclidean distance was calculated for each step in a student's walk, which provides a measure of how far the particle moved from the origin. When these distance steps are combined, they produce a distance time series. Distance time series reveal patterns of movement and potentially reveal systematic patterns in learning trajectories through coordinated "steps."

Table 2
Vector assignments.

Game-based feature	Vector assignment
Generative practice games	-1 on X-axis (move left)
Identification mini-games	+1 on Y-axis (move up)
Personalizable features	+1 on X-axis (move right)
Achievement screens	-1 on Y-axis (move down)

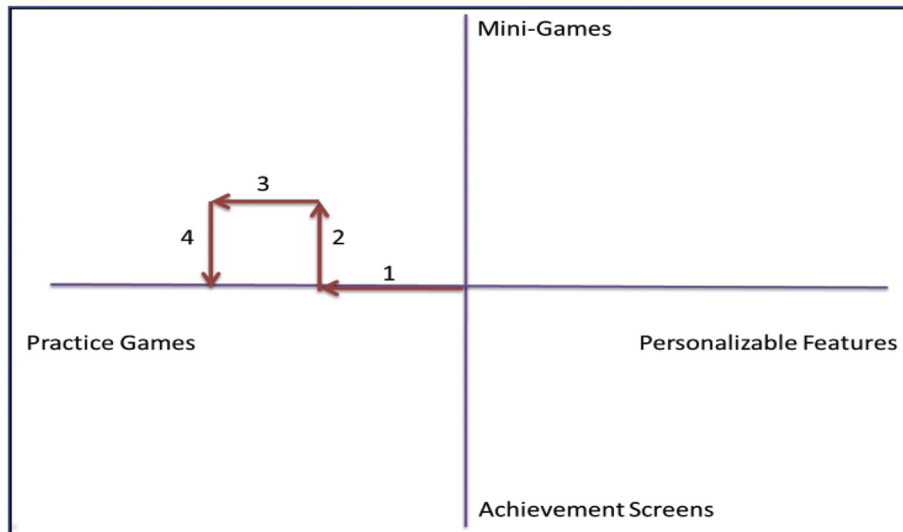


Fig. 4. Example random walk.

8.3. Entropy

Random walks provide a visualization of students' learning trajectories and Euclidian distances help quantify the movements depicted with these walks. In turn, Entropy analyses were conducted to quantify the degree to which these fluctuations are controlled or disordered. Entropy analysis is a statistical measure that has been used previously to measure random, controlled, and ordered processes (Fasolo, Hertwig, Huber, & Ludwig, 2009; Grossman, 1953; Shannon, 1951; Snow, Allen, Russell, et al., 2014; Snow, Jacovina et al., 2014). In the current study, Shannon Entropy (Shannon, 1951) is used to gain a deeper understanding of how students' choice patterns reflect controlled and ordered processes. To calculate Entropy, we used the distance time series produced from each student's random walk (see Equation (2)). Within the Entropy equation, $P(x_i)$ represents the probability of a given interaction. For instance, the Entropy for student X is the inverse of the sum of products calculated by multiplying the probability of each distance by the natural log of the probability of that distance. This formula captures the amount of control and order presented within the distance time series generated within each student's random walk.

$$H(x) = - \sum_{i=0}^N P(x_i)(\log_e P(x_i)) \tag{2}$$

In this study, when a student's choice pattern produces a low Entropy score, it suggests that they demonstrated a highly organized pattern. Conversely, when a student's choice pattern produces a high Entropy score, it suggests that the student demonstrated a

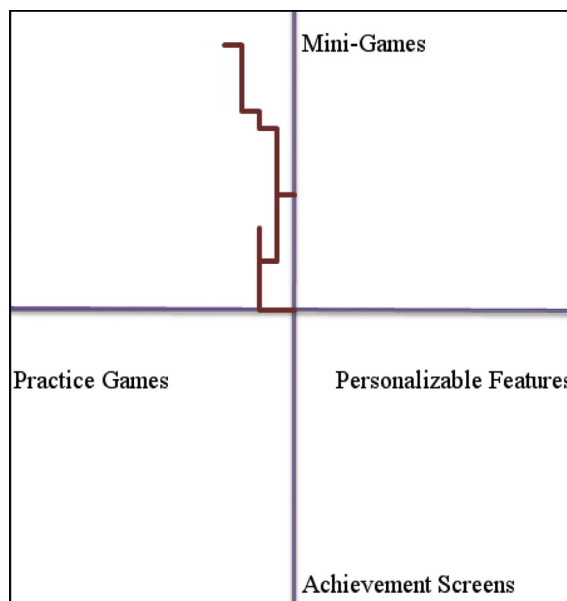


Fig. 5. Actual random walk from study 1 with an upward trajectory.

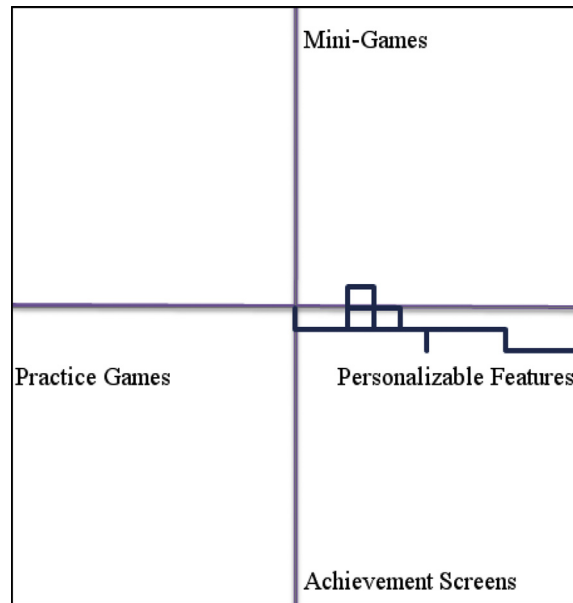


Fig. 6. Actual random walk from study 1 with a rightward trajectory.

disorganized choice pattern. The differences between controlled and disordered behaviors can be visualized using gait patterns. For example, in Fig. 7, the footsteps illustrate a systematic and controlled gait. These steps are evenly spaced and ordered. Thus, each step in this pattern comprises a controlled and systematic series analogous to low Entropy. Conversely, Fig. 8 reveals a random and disordered gait pattern where the steps are unevenly distributed and unpredictably jump around, illustrative of disorder and thus high Entropy.

8.4. Statistical analyses

To examine the influence of students' behavior patterns within iSTART-2 on their daily performance and posttest attitudes, we conducted Pearson correlation and regression analyses using Entropy scores, average self-explanation quality during training, game performance, and posttest survey responses. In addition, a regression analysis was conducted to examine the degree to which students' Entropy scores accounted for variance in their daily self-explanation quality. Finally, Pearson correlation analyses were conducted to examine the relation between students' Entropy scores and their in-game performance (i.e., trophies won) and also the relation between students' Entropy scores and posttest attitudes.

9. Results

9.1. Entropy

The current study examined the impact of variations in students' behavior patterns on the in-system performance within iSTART-2. An Entropy analysis was calculated to quantify the fluctuations that manifest within each student's random walk. In the current study, Entropy scores varied considerably, suggesting that students' behavior patterns ranged from controlled to disordered (range = 1.32–2.32, $M = 1.83$, $SD = .24$; skew = $-.22$; kurtosis = -1).

9.2. Interaction choices

The relation between Entropy scores and students' frequency of interaction choices was calculated by using Pearson correlations. Results from this analysis revealed no significant relations between students' Entropy scores and the frequency of interactions with generative practice games ($r = -.04$, $p = .73$), identification mini-games ($r = .17$, $p = .13$), personalizable features ($r = -.05$, $p = .66$), or achievement screen views ($r = -.11$, $p = .37$). This suggests that controlled patterns of interactions were not related to any specific feature within iSTART-2.



Fig. 7. Example of controlled pattern (i.e., low Entropy).



Fig. 8. Example of disordered pattern (i.e., high Entropy).

9.3. In-system performance

9.3.1. Self-explanation quality

To examine the effects of controlled interaction patterns on self-explanation quality, a regression analysis was conducted. In this analysis, we examined how Entropy was related to students' average self-explanation quality score from their time within the iSTART-2 system. This analysis revealed a significant relation between students' Entropy scores and their average self-explanation quality scores, $F(1, 74) = 4.33$, $p = .041$, $R^2 = .06$. These results indicate that students who engaged in more controlled interaction patterns generated higher quality self-explanations relative to students who engaged in more disordered interaction patterns.

9.3.2. In-game performance

Within iSTART-2, students also earned trophies based on their performance within the practice games. To examine the relation between students' patterns of interactions and game-performance, Pearson correlations were conducted. Results from these analyses revealed a significant negative relation between Entropy scores and generative practice trophies ($r = -.26$, $p = .02$). This means that students with a higher Entropy score (indicative of disordered choice pattern) tended to earn fewer trophies within the system. Again, students who interacted in a more ordered and controlled way showed more success within the system.

9.4. Posttest attitudes

To assess how Entropy related to students' self-reported feelings of confusion, boredom and control, Pearson correlation analyses were conducted. Results from these analyses revealed no significant relation between students' Entropy and their self-reported confusion ($r = .03$, $p = .78$) or their self-reported boredom ($r = .07$, $p = .55$). However, Entropy was significantly positively related to students' feeling of lack of control ($r = .26$, $p = .02$). Thus, students' who engaged in more controlled interaction patterns also reported higher feelings of control. Importantly, the lack of control was not related to other factors such as confusion or boredom.

10. Discussion

Enhanced feelings of agency have been found to increase students' engagement and ultimately improve learning outcomes (Flowerday & Schraw, 2000). Learning environments attempt to leverage these findings by adding elements of control for students. However, students vary in their ability to effectively control and regulate their behaviors when presented with opportunities to exert control (Zimmerman, 1990). Thus, some students may experience high levels of agency and thrive in an environment where they are presented with many choices, whereas others may struggle to effectively exert control over their behaviors and develop a plan of action (McNamara & Shapiro, 2005). Study 1 captured these fluctuations in students' ability to control their behaviors through the use of dynamical methodologies.

Entropy was explored as a means to provide a stealth assessment of students' patterns of interactions within iSTART-2. The use of stealth assessments is important for researchers because they can obviate the need for intrusive and explicit questions (e.g., "How in control are you right now?") which may disrupt students' flow within game-based systems and their subsequent ability to accurately report feelings of agency. In Study 1, we interpreted students with a low Entropy score as interacting with the system with purpose and control. Conversely, when students' choice patterns produced a high Entropy score, they demonstrated a lack of purpose or control. Each student's Entropy score reveals a trend across time, suggestive of the degree to which each student exerted agency within the iSTART-2 system. Our interpretation of the Entropy scores is further supported by students' posttest self-reported feelings of control. Specifically, students who had a more disordered interaction pattern also reported feeling less in control. This relationship between Entropy and feelings of control begins to provide concurrent validity that Entropy may be one way to covertly assess students' feeling of agency over their environment.

In addition to students' attitudes, Entropy scores were also related to students' performance within the iSTART-2 system. In particular, when students demonstrated a more controlled interaction pattern (i.e., low Entropy) within the system, they generated higher quality self-explanations and performed better in the practice games during training compared to students who exhibited a disordered interaction pattern (i.e., high Entropy). It is important to note that Entropy was not significantly related to any specific game-based feature. That is, students who tended to behave in more controlled manners did not engage with similar activities within the system interface. This suggests that there are multiple ways to succeed within iSTART-2 and that the impact of students' interactions on learning has less to do with *what* they choose, but *how* they choose to do it. Open interface game-based systems are designed to afford students opportunities to create their own learning trajectory by exploring different activities and features. The results of Study 1 suggest that to be successful within these interfaces, students need to take agency over their learning paths and make meaningful choices. These results are supported by previous work that reveals a positive relation between students' ability to regulate and control their learning behaviors during a task and learning outcomes (Calvert et al., 2005).

An alternative explanation for the results found in Study 1 is that controlled learning trajectories will enhance learning outcomes regardless of students' agency in determining those trajectories. This explanation does not rely on students experiencing a sense of agency. According to this hypothesis, as long as students interact with the system in a controlled pattern, they should benefit regardless of their perceived level of agency. If, indeed, agency is of secondary importance, assigning students with a predetermined interaction pattern should

yield similar results. In Study 2, we tested this alternative hypothesis by removing students' agency within the system and assigning them to previously generated interaction patterns. This alternative hypothesis would be supported if students who engage in controlled patterns still outperform students who engage in disordered patterns, despite the patterns being assigned rather than chosen. On the other hand, an agency-based hypothesis would be supported if performance differences do not emerge between students who engage in controlled and disordered interaction patterns.

11. Study 2

Study 1 examined how students' self-selected interaction patterns influenced system performance and perceptions of control. Study 2 builds upon this work by examining whether the positive influence of controlled behavior patterns (i.e., based on Entropy scores) on target skill acquisition (i.e., self-explanation quality) persists when the ability to choose that pattern is removed. A main component of agency is students' ability to take control over their learning path and set their own strategic plan (Zimmerman, 2008). This suggests that if agency contributed to the learning benefits observed in Study 1, those benefits should be attenuated when students have little control over their learning. Study 2 is designed to test that assumption by taking away students' ability to experience agency and instead randomly assigning them a learning trajectory within iSTART-2. This manipulation will begin to tease apart the role of agency and specific interaction patterns on students' learning in the system.

12. Method

12.1. Participants

The current study included 70 college students from a large university campus in the Southwest United States. These students were, on average, 19.8 years of age (range: 18–24 years), with a mean reported grade level of college freshman. Of the 70 students, 64% were male, 57% were Caucasian, 23% were Asian, 4% were African-American, 7% were Hispanic, and 9% reported other nationalities.

12.2. Procedure

The procedures for Study 2 were identical to the Study 1 procedures for pretest, strategy training, and posttest. The sole difference occurred when students were transitioned into the game-based practice portion of the experiment. In this section, instead of having free choice over their interaction trajectory, students in Study 2 were randomly assigned an interaction trajectory that was previously identified as controlled or disordered. Students received a list of actions that they were asked to complete sequentially. At the beginning of the practice section of the experiment, an experimenter gave a brief explanation of each of the four types of actions that appeared on the list.

12.3. Design

Study 2 included two conditions designed to examine the effect of interaction patterns (controlled or disordered) on target skill acquisition. A median split was conducted on the Entropy scores for participants in Study 1 to create two groups of interaction patterns (i.e., controlled interaction condition and disordered interaction condition). Interaction trajectories generated by students who participated in Study 1 were yoked to students in Study 2. Students were randomly assigned one of these previously generated interaction patterns and were instructed to replicate the list of behaviors while they engaged within the iSTART-2 environment.

12.4. Measures

All of the measures for Study 2 were identical to the measures within Study 1.

12.5. Data processing

All students in Study 2 were assigned an interaction trajectory within the iSTART-2 system. Log data was used to validate that students followed their assigned interaction path. All students who participated in Study 2 demonstrated at least a 90% adherence rate. This reveals that most students followed their assigned interaction path with a high level of accuracy.

12.6. Statistical analyses

To examine differences between controlled and disordered students' in-system performance and posttest attitudes, we conducted ANOVA and Bayesian factor analyses using average self-explanation quality during training, game performance (i.e., trophies won), and posttest survey responses including the between-subjects factor of assigned interaction pattern (i.e., controlled or disordered). In addition, separate one-way between subjects ANOVAs were conducted on students' self-explanation quality during training and total trophies won to assess the extent to which the pre-assigned controlled or disordered patterns influenced performance within iSTART-2. Bayes factor analyses were conducted for each of these one-way ANOVAs to assess the probability of a null hypothesis being accurate over an alternative hypothesis. Additionally, three separate one-way ANOVAs were conducted on students' posttest survey responses to examine students' attitudes regarding the system as a function of whether their assigned interaction pattern was controlled or disordered. Finally, to compare self-explanation quality between students in Study 1 and Study 2, a two-way ANOVA was conducted using average self-explanation quality and Entropy scores.

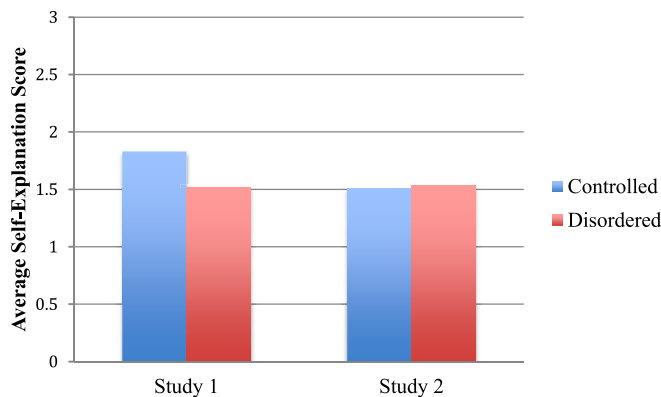


Fig. 9. Self-explanation scores across studies 1 and 2.

13. Results

Results from Study 1 revealed that students who engaged in a more controlled pattern of interaction performed better within iSTART-2. To investigate whether these results were related to students' ability to take agency over their learning path, we manipulated students' interaction paths in Study 2. Specifically, all students were assigned a previously generated interaction path from the students in Study 1. A median split was calculated on these paths to distinguish between controlled ($M = 1.55$, $SD = .29$) and disordered interaction patterns ($M = 2.10$, $SD = .10$). Using this median split, we examined differences in system performance between students assigned controlled and disordered interaction patterns.

13.1. In-system performance

13.1.1. Self-explanation quality

Results from Study 1 revealed that when students displayed more controlled interaction patterns, they also generated higher quality self-explanations. In Study 2, we calculated a one-way ANOVA to examine if this trend was similar when students were assigned an interaction pattern. Results from this analysis revealed no significant difference between the self-explanation scores for students in the controlled and disordered interaction pattern groups, $F(1,69) = .11$, $p = .74$.

This finding supports the null hypothesis. However, traditional statistics are not designed to support the null; thus these results may be overestimated. To test the probability of the extent to which the data supports this null hypothesis over an alternative hypothesis, a two-sample Bayesian t -test analysis was calculated. Bayesian factor analyses are designed to specifically test how probable the support of a null hypothesis is given the t -value and sample size of the test group. This Bayesian factor analysis was conducted using the web-based platform developed by Rouder, Speckman, Sun, Morey, and Iverson (2009). For this analysis, the sample sizes of the both conditions (controlled interaction pattern group, $n = 35$, and disordered interaction pattern group, $n = 35$), along with the t -value of the above analysis ($t = .33$), were used to produce a JZS (Jeffreys – Zellner – Siow) Bayes factor (Rouder et al., 2009) of 5.26. This JZS Bayes factor suggests that there is substantial evidence to support the null hypothesis (i.e., five times more likely) over the alternative.

13.1.2. Game-performance

A similar one-way ANOVA analysis was conducted to examine whether there was a difference in total trophies won between the two interaction pattern groups. Findings from these analyses revealed that there were no significant differences in total trophies won between the controlled and disordered interaction pattern groups, $F(1,69) = 2.25$, $p = .12$. This result suggests that when students were assigned to a controlled interaction pattern, they did not perform more successfully than students assigned to a disordered interaction pattern.

To examine the extent to which the data supports this null hypothesis over the alternative hypothesis, two-sample Bayesian t -test analysis was calculated. Again, sample sizes from both conditions were used (controlled interaction group, $n = 35$, and disordered interaction group, $n = 35$) along with the t -value from the above analysis ($t = .35$). This analysis revealed a JZS Bayes factor of 5.21. This JZS Bayes factor suggests that there is substantial evidence to support this null hypothesis (i.e., five times more likely) compared to the alternative hypothesis.

13.2. Posttest attitudes

Three one-way ANOVAs were calculated to examine differences between the controlled and disordered groups of students' self-reported confusion, boredom, and lack of control. Results from this analysis show a significant difference in students' self-reported confusion at posttest, $F(1,69) = 9.63$, $p = .003$. Students who were assigned to controlled interaction patterns reported significantly higher amounts of confusion about what they were doing ($M = 3.37$, $SD = 1.14$) compared to students who were assigned disordered interaction patterns ($M = 2.49$, $SD = 1.25$). However, there was no significant difference in self-reported feeling of lack of control between the two interaction pattern groups, $F(1,69) = 2.72$, $p = .11$, or self-reported boredom, $F(1,69) = .73$, $p = .39$.

13.3. Self-explanation comparison for study 1 and 2

To examine how the influence of interaction patterns on self-explanation quality was influenced by students' agency, a 2×2 ANOVA was conducted including the between-subjects factors of study (Study 1 vs. Study 2) and interaction pattern (controlled vs. disordered). This analysis revealed a marginally significant main effect of study on self-explanation quality, $F(3,136) = 3.45, p = .065$, with students in Study 1 generating higher quality self-explanations ($M = 1.67, SD = .54$) than students in Study 2 ($M = 1.52, SD = .45$). There effect of interaction pattern (controlled vs. disordered) was not significant, $F(3,136) = 2.83, p = .95$. However, there was a significant interaction between study and interaction pattern, $F(3,136) = 4.46, p = .036$. Students generated higher quality self-explanations when they self-selected to use a controlled interaction pattern ($M = 1.83, SD = .57$) compared to those who self-selected a disordered interaction pattern ($M = 1.52, SD = .46$) or were assigned to either controlled ($M = 1.51, SD = .46$) or disordered interaction patterns ($M = 1.53, SD = .44$; see Fig. 9). This finding suggests that the degree of control in students' interaction patterns had a greater impact on performance, with the potential to enhance performance, when students were free to make their own choices within the system.

14. Discussion

Students who exhibit high levels of agency are said to be engaging in a form of strategic planning (Zimmerman & Schunk, 2001). These students approach a learning task with a goal and subsequently set a plan of action to accomplish this goal. Study 2 was designed to test the impact of agency on learning outcomes in a game-based system. Within iSTART-2, students are afforded high levels of agency and the results from Study 1 revealed that students approached the system in various ways. Some students acted in a controlled manner (i.e., low Entropy) whereas others acted in a more disordered fashion (i.e., high Entropy). Results showed that when students took agency over their learning path and interacted in a controlled manner within the system, they generated higher quality self-explanations and performed better within the games. Study 2 was designed to tease apart the impact of students' agency and the level of control in their interaction patterns. In particular, we aimed to determine whether the relation of controlled interaction patterns to learning outcomes held when students' agency was removed.

The findings presented here indicate that when students' agency was limited (i.e., they were assigned an interaction pattern), the performance differences between students who engaged in controlled and disorder patterns disappeared. This indicates that students' ability to exert agency and choose their own learning path was related to target performance. Interestingly, in Study 2, students assigned to the controlled condition reported higher levels of confusion than those assigned to the disordered condition. Thus, when agency is inhibited, a controlled pattern may seem redundant and the purpose of such a pattern may be obscured from the student.

This assumption is further supported by the between-experiment analysis on self-explanation quality. This analysis revealed a significant interaction between study (Study 1 vs. Study 2) and interaction pattern (controlled vs. disordered). Specifically, students who engaged in a controlled interaction pattern based on their own choices generated higher quality self-explanations than all other students. This suggests that a controlled pattern only mattered when the student set their own plan and self-selected that interaction path within the system. Thus, assigning students a statistically controlled path without any context may seem just as disorganized to the student as a statically disorganized path or a self-selected random path. Combined, results from Studies 1 and 2 support our hypothesis that agency is an important component of students' success within adaptive environments.

15. Conclusion

Log data from game-based systems have the strong potential to provide researchers with an opportunity to covertly assess students' ability to exert agency over a learning task (Sabourin et al., 2012; Snow, Allen, Russell, et al., 2014; Snow, Likens, et al., 2013). These systems often allow students to exhibit various levels of control, which influences the interaction patterns that manifest while they interact within the system. This study explored the use of three dynamical methodologies as potential forms of stealth assessment for how students exerted agency within the game-based environment, iSTART-2. Additionally, we examined the impact of these individual interactions on students' attitudes and performance. These analyses approach agency in a novel way by examining nuances in students' log data to capture tendencies in their choice selections and behaviors across time. These methodologies may prove useful for the improvement of student models that rely on understanding the relation between students' abilities and performance. Indeed, the tracking and modeling of behavioral trends and patterns over time is critical to our understanding of the various ways in which students exert agency over their environment.

Students vary in their ability to exert agency during learning tasks (Zimmerman, 1990). Results from the current study build upon this work by revealing that students who chose to engage in controlled interaction patterns generated higher quality self-explanations than students who did not. These findings suggest that agency is a key component of success within a game-based system. Thus, researchers and system designers might strive to find the "sweet spot" of how many choices to include within game-based environments, such that it maximizes the number of students who will experience high levels of agency and minimize the number of students who are overwhelmed and thus respond with disordered choices.

This study builds upon previous work examining how tracing and classifying students' interactions within adaptive systems can provide information about their ability to regulate and control behaviors (Hadwin et al., 2007; Snow, Jacovina, et al., 2014). The analyses presented here are intended to provide further evidence that dynamical methodologies are valuable tools that can shed light upon various behavioral trends that may manifest within students' log data. Of course this study is not the end of the story. Future confirmatory studies are needed to further demonstrate the relation between traditional measures of control and regulation (i.e., self-reports) and these statistical techniques. Such work will potentially establish the respective utility of dynamical and traditional measures of learning behaviors (and the subsequent intent behind those behaviors). In Study 1, we found a significant relation between disordered interaction patterns and students' feelings of lack of control. These results begin to establish the link between agency and controlled interaction patterns. However, one potential weakness that merits further investigation is the use of single-item self-report scales to assess constructs such as control, confusion, and boredom. While previous work has shown that single-item measures can be reliable and potentially increase validity (e.g., Stone, 2004), the affective constructs presented here may indeed be multidimensional and therefore single-item responses may not provide a holistic view of

these constructs. Examining relations between multidimensional self-reports and online choice patterns may further illuminate our understanding of these constructs. Such an approach may afford a more in-depth understanding of how feelings of agency and subsequent behavior patterns emerge overtime.

The ultimate purpose of using dynamical measures is to capture students' behavioral trends they emerge in real time and relate those trends to learning outcomes. Thus, the true measure of the applicability of these measures lies within researchers' ability to implement them in real-time as a means to inform student models. For instance, one critical research question for system developers is how to develop optimal learning trajectories for each individual student. Through the use of visualization and dynamical techniques, systems may be able to recognize non-optimal patterns and steer students toward more effective behaviors. For instance, if students are engaging in disordered behavior patterns, these techniques may be useful in augmenting adaptive environments through the recognition of non-optimal patterns and subsequently prompting students toward a more controlled behavior trajectory.

In conclusion, students' ability to exert agency over a learning task and act in a decisive manner has been shown to be a critical skill for academic success (Bandura, 1989; Hadwin et al., 2007). The current study builds upon this work by revealing that agency is a key component of success within game-based systems. The analyses presented here are among the first attempts to covertly examine how variations in students' interaction patterns (both self-selected and assigned) influence target skill performance. Game-based systems frequently incorporate multiple instructional choices and learning trajectories that afford students the potential for high levels of agency. We expect that the findings presented here will help shed light upon the impact that interface design can have on students' performance. Overall, these findings support the notion that in order to be successful within game-based systems students need to take agency over their learning paths and make meaningful choices.

Acknowledgments

This research was supported in part by the Institute for Educational Sciences (IES R305G020018-02, R305G040046, R305A080589, R305A130124) and National Science Foundation (NSF REC0241144, IIS-0735682). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES or NSF. We are particularly grateful to the many members of the SoLET Lab who have contributed to the completion of this study and this project.

References

- Allen, L. K., Snow, E. L., & McNamara, D. S. (2014). The long and winding road: Investigating the differential writing patterns of high and low skilled writers. In J. Stamper, S. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 304–307). London, UK.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, 44(9), 1175–1184.
- Bandura, A. (2001). Social cognitive theory: an agentic perspective. *Annual Review of Psychology*, 52(1), 1–26.
- Barab, S., Pettyjohn, P., Gresalfi, M., Volk, C., & Solomou, M. (2012). Game-based curriculum and transformational play: designing to meaningfully positioning person, content, and context. *Computers & Education*, 58(1), 518–533.
- Benhamou, S., & Bovet, P. (1989). How animals use their environment: a new look at kinesin. *Animal Behaviour*, 38, 375–383.
- Bishop, G. F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly*, 51(2), 220–232.
- Calvert, S., Strong, B., & Gallagher, L. (2005). Control as an engagement feature of young children's attention to and learning of computer content. *The American Behavioral Scientist*, 48, 578–589.
- Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2009). Combining shared control with variability over surface features: effects on transfer test performance and task involvement. *Computers in Human Behavior*, 25, 290–298.
- Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2011). Learner-controlled selection of tasks with different surface and structural features: effects on transfer and efficiency. *Computers in Human Behavior*, 27, 76–81.
- Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88, 715–730.
- Davey, H. M., Barratt, A. L., Butow, P. N., & Deeks, J. J. (2007). A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *Journal of Clinical Epidemiology*, 60, 356–360.
- De Boer, A. G. E. M., Van Lanschot, J. J. B., Stalmeier, P. F. M., Van Sandick, J. W., Hulscher, J. B. F., De Haes, J. C. J. M., et al. (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research*, 13(2), 311–320.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Ellis, D., & Zimmerman, B. J. (2001). Enhancing self-monitoring during self-regulated learning of speech. In H. J. Hartman (Ed.), *Metacognition in learning and instruction. Theory, research, and practice* (pp. 205–228). Dordrecht: Kluwer Academic Press.
- Fasolo, B., Hertwig, R., Huber, M., & Ludwig, M. (2009). Size, entropy, and density: what is the difference that makes the difference between small and large real-world assortments? *Psychology & Marketing*, 26, 254–279.
- Flowerday, T., & Schraw, G. (2000). Teachers' beliefs about instructional choice: a phenomenological study. *Journal of Educational Psychology*, 92, 634–645.
- Flowerday, T., & Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *Journal of Educational Research*, 96, 207–215.
- Flowerday, T., Schraw, G., & Stevens, J. (2004). The role of choice and interest in reader engagement. *Journal of Experimental Education*, 72, 93–114.
- Ford, M. E. (1992). *Motivating humans: Goals, emotions, and personal agency beliefs*. Newbury Park, CA: Sage Publications.
- Grossman, E. R. F. W. (1953). Entropy and choice time: the effect of frequency unbalance on choice-response. *Quarterly Journal of Experimental Psychology*, 5(2), 41–51.
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2, 107–124.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41, 111–127.
- Iyengar, S., & Lepper, R. (2000). When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79, 995–1006.
- Jackson, G. T., Guess, R. H., & McNamara, D. S. (2009). Assessing cognitively complex strategy use in an untrained domain. *Topics in Cognitive Science*, 2, 127–137.
- Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, 105, 1036–1049.
- Katz, I., Assor, A., Kanat-Maymon, Y., & Bereby-Meyer, Y. (2006). Interest as a motivational resource: feedback and gender matter, but interest makes the difference. *Social Psychology of Education*, 9, 27–42.
- Klopfer, F. J., & Madden, T. M. (1980). The middlemost choice on attitude items: ambivalence, neutrality, or uncertainty? *Personality and Social Psychology Bulletin*, 6(1), 97–101.
- Landauer, T. K., McNamara, D. S., Dennis, S. E., & Kintsch, W. E. (2007). *Handbook of latent semantic analysis*. Lawrence Erlbaum Associates Publishers.
- Lester, J. C., Mott, B. W., Robison, J. L., Rowe, J. P., & Shores, L. R. (2013). Supporting self-regulated science learning in narrative-centered learning environments. In *International handbook of metacognition and learning technologies* (pp. 471–483). New York: Springer.
- Lobry, J. R. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Molecular Biological Evolution*, 13, 660–665.
- McIver, J. P., & Carmines, E. G. (1981). *Unidimensional scaling*. Thousand Oaks, CA: Sage.
- McNamara, D. S., Boonthum, C., Levinstein, I. B., & Millis, K. (2007). Evaluating self-explanations in iSTART: comparing word-based and LSA algorithms. In T. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 227–241). Mahwah, NJ: Erlbaum.
- McNamara, D. S., Jackson, G. T., & Graesser, A. C. (2010). Intelligent tutoring and games (ITaG). In Y. K. Baek (Ed.), *Gaming for classroom-based learning: Digital role-playing as a motivator of study* (pp. 44–65). Hershey, PA: IGI Global.

- McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, & Computers*, 36, 222–233.
- McNamara, D. S., O'Reilly, T., Rowe, M., Boonthum, C., & Levinstein, I. B. (2007). iSTART: a web-based tutor that teaches self-explanation and metacognitive reading strategies. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 397–420). Mahwah, NJ: Erlbaum.
- McNamara, D. S., & Shapiro, A. M. (2005). Multimedia and hypermedia solutions for promoting metacognitive engagement, coherence, and learning. *Journal of Educational Computing Research*, 33, 1–29.
- Metcalfe, J., Eich, T. S., & Miele, D. B. (2013). Metacognition of agency: proximal action and distal outcome. *Experimental Brain Research*, 229, 485–496.
- Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *Quality & Quantity*, 42(6), 779–794.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*, 3. New York: McGraw-Hill.
- O'Reilly, T. P., Sinclair, G. P., & McNamara, D. S. (2004). Reading strategy training: automated versus live. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Cognitive Science Society* (pp. 1059–1064). Mahwah, NJ: Erlbaum.
- Parker, L. E., & Lepper, M. R. (1992). Effects of fantasy contexts on children's learning and motivation: making learning more fun. *Journal of Personality and Social Psychology*, 62, 625–633.
- Paulhus, D. L., Lysy, D. C., & Yik, M. S. (1998). Self-report measures of intelligence: are they useful as proxy IQ tests? *Journal of Personality*, 66, 525–554.
- Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychological Review*, 18, 315–341.
- Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102, 531–549.
- Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology*, 25, 92–104.
- Reynolds, P. L., & Symons, S. (2001). Motivational variables and children's text search. *Journal of Educational Psychology*, 93, 14–22.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and social psychology bulletin*, 27, 151–161.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237.
- Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2010). Individual differences in gameplay and learning: a narrative-centered learning perspective. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (pp. 171–178). ACM.
- Sabourin, J., Shores, L. R., Mott, B. W., & Lester, J. C. (2012). Predicting student self-regulation strategies in game-based learning environments. In *Intelligent tutoring systems* (pp. 141–150). Berlin Heidelberg: Springer.
- Schmierbach, M., Limperos, A. M., & Woolley, J. K. (2012). Feeling the need for (personalized) speed: how natural controls and customization contribute to enjoyment of a racing game through enhanced immersion. *Cyberpsychology, Behavior, and Social Networking*, 15(7), 364–369.
- Schönauf-Fog, H., & Bjørner, T. (2012). "Sure, I would like to continue": a method for mapping the experience of engagement in video games. *Bulletin of Science, Technology & Society*, 32, 405–412.
- Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, 30, 50–64.
- Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games and Instruction*, 55, 503–524.
- Shute, V. J., Ventura, M., Bauer, M., & Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning. In *Serious games: Mechanisms and effects*, 2 pp. 295–321. Philadelphia, PA: Routledge/LEA.
- Snow, E. L., Allen, L. K., Jackson, G. T., & McNamara, D. S. (2014). Tracking choices: computational analysis of learning trajectories. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 316–319). Heidelberg, Berlin, Germany: Springer.
- Snow, E. L., Allen, L. K., Russell, D. G., & McNamara, D. S. (2014). Who's in control?: categorizing nuanced patterns of behaviors within a game-based intelligent tutoring system. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 185–192). Heidelberg, Berlin, Germany: Springer.
- Snow, E. L., Jackson, G. T., & McNamara, D. S. (2014). Emergent behaviors in computer-based learning environments: computational signals of catching up. *Computers in Human Behavior*, 41, 62–70.
- Snow, E. L., Jackson, G. T., Varner, L. K., & McNamara, D. S. (2013). Investigating the effects of off-task personalization on system performance and attitudes within a game-based environment. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 272–275). Heidelberg, Berlin, Germany: Springer.
- Snow, E. L., Jacovina, M. E., Allen, L. K., Dai, J., & McNamara, D. S. (2014). Entropy: a stealth assessment of agency in learning environments. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 241–244). Heidelberg, Berlin, Germany: Springer.
- Snow, E. L., Likens, A., Jackson, G. T., & McNamara, D. S. (2013). Students' walk through tutoring: using a random walk analysis to profile students. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 276–279). Heidelberg, Berlin, Germany: Springer.
- Spires, H. A., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). Problem solving and game-based learning: effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, 44(4), 453–472.
- Stone, M. H. (2004). Substantive scale construction. In E. V. Smith, Jr., & R. M. Smith (Eds.), *Introduction to Rasch measurement* (pp. 201–225). Maple Grove, MN: JAM Press.
- Tabbers, H. K., & de Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38, 441–453.
- Taylor, R. S., O'Reilly, T., Rowe, M., & McNamara, D. S. (2006). Improving understanding of science texts: iSTART strategy training vs. web design control task. In R. Sun, & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 2234–2239). Mahwah, NJ: Erlbaum.
- Teng, C. I. (2010). Customization, immersion satisfaction, and online gamer loyalty. *Computers in Human Behavior*, 26, 1547–1554.
- Yee, N. (2006). Motivations for play in online games. *Cyberpsychology & Behavior*, 9(6), 772–775.
- Zhou, M. (2013). Using traces to investigate self-regulatory activities: a study of self-regulation and achievement goal profiles in the context of web search for academic tasks. *Journal of Cognitive Education and Psychology*, 12, 287–305.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: an overview. *Educational Psychologist*, 25, 3–17.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183.
- Zimmerman, B. J., & Schunk, D. H. (2001). Reflections on theories of self-regulated learning and academic achievement. In *Self-regulated learning and academic achievement: Theoretical perspectives*, 2 pp. 289–307).