From Generating in the Lab to Tutoring Systems in Classrooms

DANIELLE S. MCNAMARA, MATTHEW E. JACOVINA,
ERICA L. SNOW, and LAURA K. ALLEN
Arizona State University

Work in cognitive and educational psychology examines a variety of phenomena related to the learning and retrieval of information. Indeed, Alice Healy, our honoree, and her colleagues have conducted a large body of groundbreaking research on this topic. In this article we discuss how 3 learning principles (the generation effect, deliberate practice and feedback, and antidotes to disengagement) discussed in Healy, Schneider, and Bourne (2012) have influenced the design of 2 intelligent tutoring systems that attempt to incorporate principles of skill and knowledge acquisition. Specifically, this article describes iSTART-2 and the Writing Pal, which provide students with instruction and practice using comprehension and writing strategies. iSTART-2 provides students with training to use effective comprehension strategies while self-explaining complex text. The Writing Pal provides students with instruction and practice to use basic writing strategies when writing persuasive essays. Underlying these systems are the assumptions that students should be provided with initial instruction that breaks down the tasks into component skills and that deliberate practice should include active generation with meaningful feedback, all while remaining engaging. The implementation of these assumptions is complicated by the ill-defined natures of comprehension and writing and supported by the use of various natural language processing techniques. We argue that there is value in attempting to integrate empirically supported learning principles into educational activities, even when there is imperfect alignment between them. Examples from the design of iSTART-2 and Writing Pal guide this argument.

A considerable amount of research in cognitive psychology focuses on the learning process, including how students encode information in memory and subsequently retrieve such information when prompted (Healy & Bourne, 2012; Healy et al., 1993). This research is often conducted in tightly controlled laboratory settings that allow various elements of the learning process to be carefully, and often elegantly, teased apart. Most psychologists would argue that the burgeoning understanding of memory and learning that has resulted from such cognitive studies of memory and learning can and should influence educational practices (see various *General Discussion* sections in this special issue for interesting suggestions about how results might be applied outside the lab). In reality, however, the knowledge gained from these experimental studies permeates educational settings less frequently than might be expected.

There are a number of potential explanations for the infrequent (or at least slow) implementation and adaptation of cognitive principles in classroom settings. One obvious explanation is the imperfect

American Journal of Psychology
Summer 2015, Vol. 128, No. 2 pp. 159–172 •

communication between researchers and educators. For example, teachers rarely read the academic journals in which psychologists publish their findings, and conversely, research conducted by psychologists is not always aligned with high-priority classroom needs. A second explanation is that the conclusions drawn from many scientific studies do not offer practical, concrete, or actionable suggestions that can be implemented in the classroom or other real-world settings (Mayer, 2012).

Although communication between educators and psychologists has encountered problems in the past, researchers remain optimistic about reducing this gap in the future (Mayer, 1992, 2012). For instance, Mayer postulates that a dynamic relationship has been developing, in which educators pinpoint the learning situations and tasks that need to be better understood and psychologists offer methods to study and optimize them. As a recent example, many instructors have used hand-held clickers to quickly administer tests to students, allow students to receive feedback on their performance and discuss their answers with instructors and peers, and refocus instruction based on students' strengths and weaknesses. Research in both controlled laboratory settings (Anderson, Healy, Kole, & Bourne, 2013) and in classrooms (Mayer et al., 2009, Smith et al., 2009) has informed how clickers might be used most effectively and contributed to a richer understanding of the short-term and long-term benefits of using clickers. For example, Smith and colleagues (2009) demonstrated that clickers improved students' deeper conceptual understanding of the material, particularly when their use was combined with peer discussion. Importantly, this work is not being conducted in a vacuum; instead, it responds to real-world questions about the efficacy of a popular educational tool.

In addition to directing research toward pressing educational issues, prominent researchers have published books and review articles to help communicate evidence-based learning principles to a larger audience while highlighting their practical applications. For example, Healy, Schneider, and Bourne (2012) outline a number of empirically based recommendations that can be understood and used by a more general audience to improve the efficacy and efficiency of training. By presenting real-world applications for scientific findings, Healy et al. provide readers with beneficial suggestions and tips, such as including spaced practice sessions and variability during training. Similarly, a recent review article by Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013) provides explicit recommendations for several learning techniques that can be used by both educators and researchers. Importantly, Dunlosky et al. discuss not only the techniques that are well supported by research and easily implemented in classes but also the techniques that either yield mixed findings or have limited findings coming directly from classroom studies. For example, they rate practice testing and distributed (spaced) practice as having high utility, whereas they rate highlighting key sections of texts and rereading as having low utility. The availability of well-balanced review articles such as these can provide suggestions to educators (about what techniques should and should not be adopted) and expose holes in the literature that researchers can work to fill.

In this article, we take an optimistic stance regarding the utility of cognitive learning principles as a guide for instruction and instructional design, despite the acknowledged challenges. We suggest that learning principles can and should be effectively adapted and implemented into educational tasks and systems. Indeed, the adaptation of learning principles is a worthwhile endeavor even when lab findings and desired educational outcomes do not align perfectly. Our goal here is to provide examples of how we have used knowledge and techniques developed by psychologists to guide instructional methods and design. We draw inspiration from Healy et al. (2012) in selecting the principles on which we focus. In their review of empirically based research on training, they discuss several categories of training principles. Here, we focus on three of these principles: the generation effect, deliberate practice and feedback, and antidotes to disengagement. We describe how we have leveraged these cognitive principles in the design of two computer-based intelligent tutoring systems (ITSs), Interactive Strategy Training for Active Reading and Thinking-2 (iSTART-2) and Writing Pal (W-Pal), which aspire to meet the difficult challenge of providing students with adaptive instruction on reading comprehension and writing strategies, respectively. To preface these discussions, we first provide a brief description of each system, including their general educational goals.

iSTART-2: A Reading Comprehension Tutor

The iSTART-2 system is designed to improve students' comprehension ability by providing self-explanation and comprehension strategy instruction (Jackson & McNamara, 2013; McNamara, Levinstein, & Boonthum, 2004; Snow, Jacovina, Allen, Dai, & McNamara, 2014). iSTART-2 consists of both a training phase and a practice phase. During the training phase, students watch a series of lesson videos in which a pedagogical agent provides instruction for five types of self-explanation strategies that they can use while reading. Specifically, students learn to paraphrase the content of a text, monitor their comprehension, predict future text content, elaborate using their world knowledge, and bridge new content with information previously encountered in the text. When describing each strategy, the pedagogical agent provides examples of how to use the strategies while reading. After watching each lesson video, students answer checkpoint questions to test their understanding of the presented strategy.

During the practice phase of iSTART-2, students are transitioned into a game-based menu where they are presented with a suite of practice activities. In this menu, students can choose to engage with game-based practice, personalize the system interface, or monitor their progress within the system (Figure 1; these features are later discussed in more detail). Overall, when students receive self-explanation training in iSTART-2 (and its earlier versions, iSTART and iSTART-Motivationally Enhanced [iSTART-ME]), their self-explanation quality and comprehension of difficult science texts improves compared with receiving no self-explanation training or receiving only an introduction to the concept of self-explanations (McNamara et al., 2004; McNamara, O'Reilly, Best, & Ozuru, 2006; McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, 2007); moreover, interacting with the game-based version of the system leads to greater enjoyment than the nongame version (Jackson & McNamara, 2013).

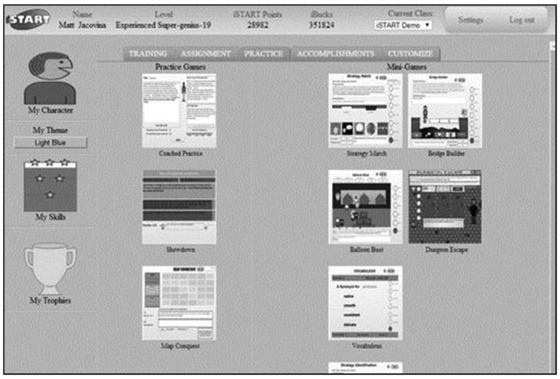

FIGURE 1. Game-based practice menu in iSTART-2. Activities include game-based practice, avatar customization, and achievement screens

Writing Pal: An Automated Writing Strategy Tutor

W-Pal was designed to provide students with writing strategy instruction that covers the entirety of the writing process (Allen, Crossley, Snow, & McNamara, 2014; Roscoe & McNamara, 2013; Roscoe, Allen, Weston, Crossley, & McNamara, 2014; Roscoe, Brandon, Snow, & McNamara, 2013). The system includes eight modules that each correspond to a specific topic within prewriting (Freewriting and Planning), drafting (Introduction Building, Body Building, and Conclusion Building), and revising (Paraphrasing, Cohesion Building, and Revising). Each module contains a series of lesson videos that are delivered by a pedagogical agent. Lessons provide clearly stated strategies and provide examples of how they can be used during the writing process, and checkpoint questions after each lesson help students recognize their level of strategy understanding. The system specifically tailors its strategies to prompt-based, argumentative essays.

Students are able to practice using their writing strategies by playing strategy practice games and writing essays (Figure 2; practice features are described later in more detail). Each module provides one or more educational games; these games provide various goals for students, from generating text using the strategies taught in the lessons to identifying which strategies other essays have successfully implemented. At any time while using the system, students can practice writing entire prompt-based, argumentative essays (alternately, course instructors can assign students to write entire essays). After submitting their essay to W-Pal, students receive formative feedback on their writing that suggests specific strategies they might use to improve their writing. Research from our lab indicates that students improve in their writing proficiency and strategy knowledge over time when using W-Pal (Allen et al., 2014; Crossley, Varner,

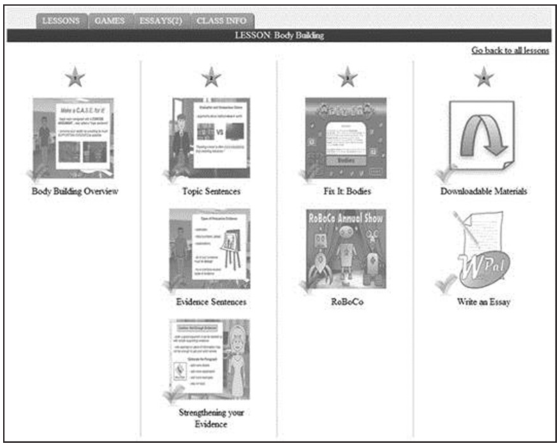

FIGURE 2. Menu for the Body Building module in W-Pal. Activities include lesson videos, practice games, and essay writing practice

Roscoe, & McNamara, 2013; Roscoe, Brandon, et al., 2013).

Designing iSTART-2 and Writing Pal

ITSs designed to provide instruction in ill-defined domains, such as reading comprehension and writing, face particular challenges: End goals are often difficult to define and equally difficult to assess automatically. Additionally, individual readers and writers may find success through the use of different strategies and approaches to reading and writing. But we do not approach this challenge unarmed. As we have suggested, the development of these systems has been guided by well-established learning principles and methods. Roediger and Pyc (2012) make the important point that many traditional learning principles focus on the ability to learn and recall information instead of the ability to acquire complex strategies and skills. However, proficient memory is foundational for sophisticated thinking tasks. We agree, and we add that for complex learning goals (such as comprehending challenging science texts or writing argumentative-based essays), cognitive learning principles can usefully guide instruction for individual components of the overall learning objective.

In the following sections, we broadly describe three learning principles (generation effect, deliberate practice and feedback, and antidotes to disengagement) and how they have influenced the design of iSTART-2 and W-Pal. These three principles are among a larger set discussed by Healy et al. (2012) in their review of empirically based research on training. We focus on a set of principles that help to guide learners' cognitive resources and effort. Alone, these principles might not be sufficient for the formidable challenge of improving students' reading and writing skills. But when applied in learning environments, such as iSTART-2 and W-Pal, each principle augments the success of the systems' educational content.

The Generation Effect

The generation effect is the finding that people have a stronger memory for content that they actively generate compared with content that they read or copy in a more passive fashion (Slamecka & Graf, 1978; for a review, see Bertsch, Pesta, Wiscott, & McDaniel, 2007). For example, if participants are asked either to read a pair of related target words (e.g., "rapid–fast") or to generate one of the target words (e.g., "rapid–f___"), they typically perform better on a subsequent memory tasks if they had generated the second word. Although this finding has theoretical implications for the nature of memory, it may not, on the surface, seem applicable to many educational goals (with several exceptions, such as when students are tasked with memorizing new words or facts).

Nonetheless, studies that have expanded on this work have progressively elucidated the educational relevance and boundaries of the generation effect. Nearly two decades ago, Healy and colleagues (Crutcher & Healy, 1989; McNamara & Healy, 1995a, 1995b, 2000) argued for a procedural account as a theoretical explanation for the generation effect. This account proposes that generation advantages are attributable to the greater likelihood of engaging in cognitive operations during the generation process that link target information with information stored in memory. If mental processes engaged during encoding are then reinstated during retrieval, performance should be high (also see deWinstanley, Bjork, & Bjork, 1996). For example, when participants are generating the word "fast" in response to "rapid," they might be engaging in associative cognitive processes. If these same participants again engage these processes during retrieval, there is a greater chance that they will successfully recall or recognize the target words. However, if participants passively read the target words, there is a lower chance they will meaningfully link the target words through cognitive processes, thus yielding low retention.

Tests of the procedural account of the generation effect have often used mathematical stimuli because the cognitive processes involved in completing a math problem are easy to pinpoint and predict compared with the processes involved in generating verbal content (but see Bjork & Storm, 2011, and McNamara & Healy, 1995a, for work demonstrating benefits of adopting processing strategies that involve generation during reading). When successfully generating answers to multiplication problems, students are almost certainly engaging in the mental procedure of multiplication. An important advantage arises through studying the specific mental processes engaged while generating; researchers can assess the retention of the generated content and the retention and acquisition of the skills afforded by those mental processes. For example, McNamara and Healy (1995a) found that compared with students who practiced reading multiplication problems and answers, participants who practiced generating answers to multiplication problems subsequently showed better performance on difficult (but not easy) multiplication problems. They concluded that participants who generated answers to multiplication problems had acquired the skills to solve difficult multiplication problems by repeating those mental procedures. There was no difference in performance for easy multiplication problems because most participants entered the study with those skills already well developed.

Across many domains, an important educational goal is for students to learn and apply strategies that will eventually become regularly used cognitive skills. Based on findings emerging from research on the generation effect, educators can support this goal by crafting generative activities that encourage use of the desired cognitive strategies and processes. Simply reading about strategies is probably not enough to encourage the mental activities that will eventually lead to strategy use becoming automatic. This idea has profoundly influenced the design of both the lesson content in iSTART-2 and W-Pal and the practice activities offered to students. Because both systems attempt to teach strategies, a passive learning experience would be unsuccessful; generative practice that involves strategic mental procedures is necessary for achieving system goals.

THE GENERATION EFFECT AND ISTART-2 AND W-PAL.

The procedural account of the generation effect was in part an inspiration for developing the self-explanation strategies taught in iSTART-2. A precursor to these strategies was ample work demonstrating that self-explaining while reading improves comprehension for many readers (Chi, de Leeuw, Chiu, & LaVancher, 1994). Not all readers benefit equally, however, because individuals differ in their ability to successfully and spontaneously self-explain. In response to this problem, McNamara (2004) developed Self-Explanation Reading Training (SERT) to provide students with specific strategies to guide their self-explanations and, in turn, improve their use of comprehension strategies.

McNamara (2004) reasoned that the benefits of self-explanation stemmed primarily from the integration of new information from the text with prior knowledge. It is this construction of a coherent, well-integrated mental representation of a text that is key to comprehension (see McNamara & Magliano, 2009). For example, in response to a science text about evaporation, a reader explained that "evaporation occurs when all of the water boils out of the pot and the bottom of the pot begins to burn." By integrating personal experience with boiling water and associating this experience with the new information about evaporation in the text, the reader is engaging in integrative processes that are more likely to lend to a coherent understanding of evaporation. However, some readers do not use comprehension strategies that afford generating the inferences necessary to make these connections. Readers without sufficient, easily accessible prior knowledge are likely to generate explanations that are incomplete, superficial, or simply paraphrases of the text.

SERT was designed to provide students with instruction and practice to use comprehension strategies such as generating bridging inferences and elaborations using domain knowledge when available but using world knowledge, common sense, and logic when directly relevant knowledge was not readily available. These strategies encourage students to be active readers and thinkers rather than passive recipients of information. That is, the lesson content prompts students to generate responses while they read, consistent with suggestions stemming from research on the generation effect. The success of this one-on-one training depended on students' generating their own self-explanations using the prescribed strategies. In addition to learning from the texts presented in the study, students practiced engaging the mental procedures necessary to use the self-explanation strategies. In this way, they were also acquiring the skills necessary to regularly and spontaneously self-explain successfully while reading challenging content.

Following the success of SERT, the self-explanation strategies were adapted to an ITS, which currently exists as iSTART-2. A key advantage of iSTART-2 over SERT is its ability to provide students with practice generating self-explanations and automated feedback about the quality of their self-explanations. Individualized, immediate feedback is crucial because students must be consistently encouraged to engage in the appropriate mental processes that can ultimately lead to the acquisition of reading strategies and skills. For example, feedback in the Coached Practice activity suggests that students use bridging and elaboration strategies when their self-explanations are identified by the system as being underdeveloped (see the section on deliberate practice and feedback for further discussion of practice activities).

In sum, iSTART-2 first provides students with overt strategies to use while generating self-explanations. It then requires students to generate responses in practice activities and provides encouragement to generate strong links between text information and prior knowledge. Results from our lab have demonstrated the effectiveness of the training lessons and practice activities in increasing students' ability to generate self-explanations while reading compared with students who only receive an introduction to the concept of self-explanation without explicit strategy practice (McNamara et al., 2006).

W-Pal also teaches students strategies and provides generative practice that encourages students to directly apply those strategies. Similar to iSTART-2, the purpose of this practice extends beyond generating high-quality responses; it also provides students an opportunity to compose texts using the strategies taught in the W-Pal lesson videos. For example, in the lessons that cover body paragraphs, the pedagogical agent teaches the C.A.S.E. strategy, which encourages students to begin body paragraphs with a concise argument and then to support that argument with sufficient evidence. In the practice game *Roboco,* students are assigned essay prompts and a thesis statement, and then they are asked to write a topic sentence and supporting sentences for the given thesis. After submitting their responses, students receive suggestions for improving their work that correspond to the C.A.S.E. strategy and the lesson videos in the Body Building module.

By focusing specifically on writing body paragraphs, students are more likely to engage in the cognitive processes involved with using the C.A.S.E. strategy while generating their responses (e.g., retrieving examples from memory in support of their topic sentence). On the other hand, were students to immediately attempt to compose an entire essay after viewing the lessons on the C.A.S.E. strategy, they might fail to incorporate the newly encountered strategies into their writing. Thus, our assumption in designing W-Pal is that the system will be more effective by providing generative practice and encouraging the use of a small number of newly introduced strategies rather than solely providing didactic instruction and expecting students to incorporate the strategies while writing complete essays. Results from our lab support this assumption, suggesting that students using the W-Pal system that included this type of generative practice experience greater strategy acquisition than students who engaged only in essay-writing practice (Roscoe, Brandon, et al., 2013; Roscoe, Snow, Brandon, & McNamara, 2013).

DELIBERATE PRACTICE AND FEEDBACK.

The successful development of complex cognitive skills, such as reading comprehension and writing, depends on a learner's ability to efficiently allocate cognitive resources and direct effort to specific components of the knowledge or skill to be acquired. In the previous section on the generation effect, we discussed how particular practice activities could support a generative advantage for students. In this section, we expand on our discussion of the benefits of practice as it is used to enhance skill acquisition. When people engage in practice, they are able to become more automatized and strategic about their actions in learning tasks. Importantly, however, all methods of practice are not equal. The most effective form of practice is *deliberate,* in that it is effortful, highly focused, and highly motivated (Healy et al., 2012). When learners engage in deliberate practice, they are expected to focus on their own weaknesses, target improvement in these areas, and monitor their own progress through self-evaluations (Ericsson, 1996; Ericsson, Krampe, & Tesch-Römer, 1993). Deliberate practice is more effective than other forms of practice for the promotion of skill acquisition and expertise (Ericsson et al., 1993). Without engaging in this form of practice over prolonged periods of time, even the most talented learners will fail to achieve their highest levels of performance.

A key component of deliberate practice is feedback. The uptake of information during practice tasks can be strongly facilitated through the provision of individualized, targeted, and actionable feedback from experts, particularly at the earliest stages of skill acquisition (Ericsson et al., 1993; Ericsson, 2008). The receipt of feedback allows learners to understand the criteria by which they are being assessed and to identify the factors that contribute to the quality of their performance. Importantly, the type of feedback received by learners is crucial to its effectiveness, and this often varies according to the specific context of the learning task (Schmidt & Bjork, 1992). For instance, trial-by-trial feedback can increase the rate of learning in certain situations. However, if a learner is already accurately able to self-monitor, feedback can be distracting and harm performance; in this case, more periodic, summarization feedback may be useful (Healy et al., 2012; Schmidt, Young, Swinnen, & Shapiro, 1989). In general, the primary purpose of feedback is to provide learners with information about *what* needs to be improved in their work and *how* to improve it. The ultimate goal of this form of feedback, then, is for learners to become self-regulated and ultimately have the ability to accurately monitor their own progress during learning (Duvivier et al., 2011).

When incorporated into the classroom, deliberate practice and formative feedback can lead to strong learning gains by students. Teachers can reinforce the information taught in their classrooms by assigning students to practice tasks and paying careful attention to the feedback they provide to students. The development of the ITSs in our lab has been motivated largely by these principles of deliberate practice and feedback. Specifically, both iSTART-2 and W-Pal place a strong emphasis on providing students with multiple forms of practice, and in all these forms of practice, students receive individualized feedback that targets specific areas for improvement.

DELIBERATE PRACTICE AND FEEDBACK IN ISTART-2 AND W-PAL. In the iSTART-2 system, we have specifically integrated components that provide students with opportunities to engage in deliberate practice and receive individualized feedback. The system includes two forms of practice. The first practice module is housed in the initial training portion of iSTART-2, immediately after the introduction and demonstration modules. In this module, students practice using the self-explanation strategies while reading two complex science texts. For each self-explanation a student submits to the iSTART-2 system, individualized feedback is provided on its overall quality.

Feedback on the quality of students' self-explanations is driven by a natural language processing algorithm. This algorithm uses both latent semantic analysis (Landauer, McNamara, Dennis, & Kintsch, 2007) and word-based measures to provide each self-explanation a score that ranges from 0 to 3. A score of 0 is assigned to a self-explanation that is too short to accurately assess or consists of information that is irrelevant to the given text. A score of 1 is given to a self-explanation that relates directly to the target sentence but does not elaborate using any additional information. A self-explanation that is given a score of 2 incorporates information from within the text beyond the target sentence, and a score of 3 suggests that information has been incorporated into the self-explanation at a global level. These self-explanations may provide information related to the overall purpose of the text, or they may contain elaborations that incorporate more general world knowledge outside of the target text. The accuracy of this iSTART algorithm has been shown to be comparable to human raters (Jackson, Guess, & McNamara, 2010; McNamara, Boonthum, Levinstein, & Millis, 2007).

Students are provided with feedback on their explanations, based on a number of factors including the quality of the explanation (as assessed by the automated algorithm), their prior performance, and the characteristics of the text. When students generate satisfactory explanations, they are sometimes provided positive feedback such as "good job" or "great work!" only. Other times, students may be asked what strategies were used in their explanations. Under some circumstances, students are asked to modify unsatisfactory self-explanations. For example, in response to a sentence in a text about thunderstorms ("Each new surge of warm, moist air rises higher than the last, continually adding to the height of the cloud"), a student generated the explanation "I really don't understand what this sentence is saying," to which the agent responded, "Please try to add information that is related to the sentence. Explain what the sentence means and how it relates to what you already know." Essentially, the pedagogical agents in iSTART provide feedback and encouragement during practice by assessing the degree to which the explanation goes beyond the text using various types of inferences. To the same sentence, a different student responded, "New surges of moist air rise higher than the surge before, gradually adding to the height of the cloud." This explanation was coded as a paraphrase that was similar to the original text; therefore, the pedagogical agent responded, "Hmm, this sounds familiar. Try to add in more information that helps explain the text." After several more attempts with feedback, the student finally generated, "To gradually add height to a cloud new surges of warm air must rise higher than the last surge, without a continual supply of moist air the cloud will be small and will evaporate in 1 to 15 minutes." This was coded as a text-based response (that followed multiple attempts), and so the agent responded, "O.K. If you add a little more next time, it will be even better." Hence, the student was not further pushed to edit that particular explanation (which has negative consequences) but was provided with feedback on how to perform better in the next round.

Because deliberate practice should take place over a prolonged time period (Ericsson et al., 1993; Healy et al., 2012), the second practice module in iSTART-2 was specifically developed to provide an extended practice environment where students could engage in practice over weeks or even months. Similar to the first practice module, this extended practice portion of iSTART-2 allows students to generate self-explanations (see the previous section on the Generation Effect for more details) and receive individualized feedback. Additionally, students can play "identification" mini-games (see the following section on *Antidotes to Disengagement*) to practice identifying certain strategies in previously generated self-explanations. Finally, this module of iSTART-2 contains a number of features that allow students to monitor their own progress, such as trophies that represent students' level of performance and a graphic representation of students' scores over time (note that students also receive feedback during practice activities). Overall, these components of deliberate practice appear to have been effectively integrated in the iSTART system, as previous research indicates that the extended practice module improves students' self-explanation performance over time (Jackson, Boonthum, & McNamara, 2010; Jackson & McNamara, 2013), whereas students who do not engage in practice improve less (Jackson, Boonthum-Denecke, & McNamara, 2012).

Similar to iSTART-2, the W-Pal system contains a number of features that specifically encourage students to engage in deliberate practice and receive effective feedback. W-Pal contains two primary forms of practice: strategy specific and whole essay. Strategy-specific practice occurs in the context of mini-games, where students are asked to generate text or identify certain writing techniques or strategies (Allen et al., 2014; see section on *Antidotes to Disengagement* for more information on the mini-games in W-Pal). This strategy-focused practice component allows students to practice specific phases of the writing process without becoming overwhelmed by the demands of composing an entire essay (Roscoe et al., 2014). W-Pal also contains an automated writing evaluation component, which allows students to compose entire essays and then receive both summative and formative feedback on their progress (Crossley et al., 2013; for an overview of automated writing evaluation systems, see Allen, Jacovina, & McNamara, in press).

Natural language processing algorithms drive the essay feedback in W-Pal. The algorithm in W-Pal evaluates the quality of essays based on the calculation of lower- and higher-level linguistic features of text using both Coh-Metrix (McNamara & Graesser, 2012; McNamara, Graesser, McCarthy, & Cai, 2014) and the Writing Analysis Tool (McNamara, Crossley,

& Roscoe, 2013). Hierarchical classification is then used to model the quality of essays (see McNamara, Crossley, Roscoe, Allen, & Dai, 2015, for more information).

Importantly, these algorithms allow the W-Pal system to provide both summative and formative feedback on their submitted essays. In terms of summative feedback, students receive a holistic rating that ranges from *Poor* to *Great* on a 6-point scale. The formative feedback in W-Pal is scaffolded feedback and emphasizes the use of writing strategies. Specifically, this feedback is meant to reinforce the strategies taught in the lessons and uses reflective questions to remind students of important goals (Roscoe et al., 2011). Overall, the nature of the feedback in W-Pal was designed to teach students about higher-level aspects of high-quality writing that can, ideally, be transferred to new contexts. One of the key aspects of deliberate practice is to allow students to develop in such a way that they are eventually capable of accurately monitoring their own progress (Duvivier et al., 2011). In W-Pal, students are prompted to evaluate their own writing after each essay that is submitted for evaluation. Although novice student writers tend to perform poorly on self-assessments of writing quality (Varner, Roscoe, & McNamara, 2013), recent research suggests that W-Pal can promote students' understanding of criteria and allow them to assess their own work more accurately after training than before training (Allen, Crossley, Snow, Jacovina, Perret, & McNamara, in press).

ANTIDOTES TO DISENGAGEMENT.

As described in the earlier sections, W-Pal and iSTART-2 both require students to engage in repetitive actions across multiple training sessions. As a consequence, students may begin to feel bored and subsequently disengage from the learning tasks (Bell & McNamara, 2007). Such disengagement is problematic, because learning new skills requires prolonged practice over an extended time period (Anderson, Conrad, & Corbett, 1989; Newell & Rosenbloom, 1981). Indeed, students' ability to acquire new skills often involves multiple steps that range from lower-level learning (i.e., learning individual principles or strategies) to higher-level learning (i.e., applying and combining newly learned principles to a task; VanLehn, 1996).

The need to maintain students' engagement during learning led to research searching for cognitive antidotes to disengagement (Healy et al., 2012). The antidote to disengagement principle calls for researchers to add a cognitive component or element to learning tasks as a means to fight disengagement. These cognitive elements should be designed to counter disengagement by prompting students to interact in a new way within the learning task. Thus, these components can break up the monotony of a task by requiring a new skill or offering students a new way to interact with the learning task.

One such cognitive component that can be used as an antidote to disengagement is educational games (and game-based features). By their nature, games are often designed to increase players' depth of cognitive engagement, which has previously been shown to improve long-term retention (Hannafin & Hooper, 1993). Such results make game elements a particularly enticing antidote to cognitive disengagement. A common way many educational games promote cognitive engagement is by offering students high levels of agency and personalization (Jackson & McNamara, 2013; Snow, Likens, Jackson, & McNamara, 2013). This added level of control is designed to promote engagement across multiple training sessions by affording the students a sense of personal investment in their learning progress (Jackson & McNamara, 2013; Snow et al., 2014). Therefore, well-designed educational games afford researchers an opportunity to frame content in an environment that encourages sustained engagement.

One core principle that guides the development of our ITSs is that educational games have the potential to act as an antidote to disengagement and thus to promote learning in a long-term setting. This design principle has become a cornerstone in our pedagogical philosophy in both iSTART-2 and W-Pal. Indeed, both of our ITSs embed educational games and game-based features as a way to combat students' disengagement. The inclusion of these features has been shown to be especially successful at decreasing disengagement across multiple training sessions (Jackson & McNamara, 2013).

ANTIDOTES TO DISENGAGEMENT IN ISTART-2 AND W-PAL.

One of the original iterations of the iSTART program was a non–game-based system designed to teach self-explanation strategies to high school students. This system had very similar pedagogy to the current iSTART-2 system; however, in the original iSTART system there were no game-based practice activities. Instead, students were simply asked to generate self-explanations, after which they received formative feedback from a pedagogical agent about the quality of their self-explanations. This process was repeated many times, and subsequently many students reported high levels of boredom and disengagement (Bell & McNamara, 2007). To create an antidote for this disengagement, iSTART-ME and, subsequently, iSTART-2 were developed. In iSTART-2, educational games (and game-based features) are embedded in the practice menu (Figure 1). This practice menu is designed to promote long-term self-explanation strategy practice without causing students to disengage from the system.

In iSTART-2's practice menu, students can interact with two types of game-based practice (i.e., generative and identification) and also personalize the system interface (edit the background color or customize an avatar). These features are designed to promote cognitive engagement, thus preventing disengagement and potential negative consequences on learning gains. These game-based features have been shown to sustain students' engagement over long periods of time. For instance, Jackson and McNamara (2013) examined how game-based features in the first game-based version of iSTART (iSTART-ME) influenced students' self-reported engagement and motivation compared with the non–game-based version of iSTART. Their results indicated that students who engaged with the game-based version reported higher levels of sustained engagement and motivation than those who interacted with the non–game-based version of iSTART. Furthermore, students in the game condition who interacted more with the games and the game-based features (e.g., editing their avatars) reported higher levels of engagement and motivation and also showed higher performance in terms of generating better self-explanations (Snow, Jackson, Varner, & McNamara, 2013a, 2013b). Thus, the use of games in iSTART seems to serve as an antidote to disengagement both for long-term practice (Jackson & McNamara, 2013) and short-term practice (Snow et al., 2014).

W-Pal also uses games as an antidote to disengagement (Allen et al., 2014; Roscoe, Brandon, et al., 2013; Roscoe, Snow, et al., 2013). However, unlike iSTART-2, where students can choose which games to play and when, W-Pal is modular and thus uses instructional scaffolding when presenting games to students. A typical module in W-Pal contains three or four lesson videos, one or two games, and a prompt-based essay to write (Figure 2). Each game in W-Pal is designed to provide students with an opportunity to practice applying the strategies they have just learned from the lesson videos. These games are designed to be cognitively engaging by including various game-based features such as a narrative, competition, points, levels, personalization, and characters. Overall, students have rated the game-based features embedded in W-Pal quite favorably (Allen et al., 2014; Roscoe, Brandon, et al., 2013; Roscoe, Snow, et al., 2013). For instance, Roscoe and colleagues (Roscoe, Brandon, et al., 2013; Roscoe, Snow, et al., 2013) found that students who interacted with the W-Pal system reported high levels of enjoyment and engagement while also perceiving game-based practice as a helpful tool for mastering the use of writing strategies. These results suggest that the game-based features in W-Pal not only engage students' interest during strategy practice but also provide perceived pedagogical assistance.

In our lab we embed educational games in our ITSs as an antidote to disengagement. Healy et al. (2012) argue that antidotes to disengagement should be cognitively engaging and provide students with new ways to interact with learning tasks. This principle has guided the creation and refinement of both iSTART-2 and W-Pal. However, as with many design principles, successful implementation presents challenges. Foremost has been finding the balance between engagement and learning. For instance, although we have found that games have positive effects on motivation and engagement, other researchers have argued that games can act as seductive distractors that can pull students' attention away from the learning task, ultimately decreasing performance (Harp & Mayer, 1997). Thus, an ongoing goal of our lab is to investigate the "sweet spot" between disengagement and learning in game-based environments. For instance, we are conducting experiments to investigate *when* students should be exposed to games and game-based features in our systems. For example, games may be deployed immediately or delayed until after an initial non–game-based practice has been completed. This ongoing study is designed to examine when games should be available during practice and when they potentially distract from the learning process. We are conducting a series of studies to investigate this question and others on how particular game features relate to both motivation and learning. Our ultimate goal is to provide further information about the impact of game-based practice on learning across multiple time intervals.

Conclusions

In both iSTART-2 and W-Pal, our overarching goal is to provide students with automated strategy training to improve their comprehension and writing abilities. In this article we have outlined how the design of these systems has been guided by learning principles from cognitive and educational psychology. Each system encourages *generative*, *deliberate* practice and provides students with *feedback* on their performance. Extended practice is supported by the inclusion of educational games intended to act as *antidotes to disengagement*. These three learning principles have all been supported by empirical studies, lending evidence to their potential efficacy. The adaptation of each principle is not necessarily straightforward; for example, using strategies to generate body paragraphs in a prompt-based essay is quite different from the early word generation studies that explored the generation effect. We present the examples from our own lab to argue the value in attempting to apply learning principles to meet complex educational goals that are relevant to both students' and educators' needs. By leveraging the wealth of findings from work on memory and learning, researchers and educators can design successful interventions and techniques. Testing these designs with additional empirical work can then enrich the understanding of learning principles and provide concrete, actionable suggestions that can be implemented in educational settings outside the lab.

This objective exemplifies the work of Alice Healy, whom we honor in this special issue. Alice has devoted her career to conducting research to better understand learning and memory and applying those principles to settings outside the laboratory, to real-world settings such as military training and the classroom. Alice has embraced the philosophy that basic laboratory-proven principles are subject to the test of real-world applications, and in turn, those tests provide a feedback loop, back into our theoretical understandings of cognition. We similarly embrace the principle that a strong test of theory is the extent to which it holds outside the lab and can be successfully applied in the real world, as messy as it sometimes is.

NOTES

This work was supported by the Institute of Education Sciences (IES), U.S. Department of Education, through grants R305A120707 and R305A130124 to Arizona State University. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily represent views of the IES. We are grateful to the many colleagues and students who have contributed to the work described in this article.

Address correspondence about this article to Danielle McNamara, Learning Sciences Institute, Psychology Department, Arizona State University, P.O. Box 872111, Tempe, AZ 85287-2111 (e-mail: Danielle.McNamara@asu.edu).

REFERENCES

Allen, L. K., Crossley, S. A., Snow, E. L., & McNamara, D. S. (2014). Game-based writing strategy tutoring for second language learners: Game enjoyment as a key to engagement. *Language Learning & Technology, 18,* 124–150.

Allen, L. K., Jacovina, M. E., & McNamara, D. S. (in press). Computer-based writing instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed.).

Allen, L. K., Crossley, S. A., Snow, E. L., Jacovina, M. E., Perret, C., & McNamara, D. S. (in press). Am I wrong or am I right? Gains in monitoring accuracy in an intelligent tutoring system. *Proceedings of the 17th International Conference on Artificial Intelligence in Education (AIED).*

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science, 1,* 467–505.

Anderson, L. S., Healy, A. F., Kole, J. A., & Bourne, L. E. Jr. (2013). The clicker technique: Cultivating efficient teaching and effective learning. *Applied Cognitive Psychology, 234,* 222–234.

Bell, C., & McNamara, D. S. (2007). Integrating iSTART into a high school curriculum. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 809–814). Austin, TX: Cognitive Science Society.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition, 35,* 201–210.

Bjork, E. L., & Storm, B. C. (2011). Retrieval experience as a modifier of future encoding: Another test effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1113–1124.

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

Crossley, S. A., Varner (Allen), L. K., Roscoe, R. D., & McNamara, D. S. (2013). Using automated cohesion indices as a measure of writing growth in intelligent tutoring systems and automated essay writing systems. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED)* (pp. 269–278). Heidelberg, Berlin: Springer.

Crutcher, R. J., & Healy, A. F. (1989). Cognitive operations and the generation effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 669–675.

deWinstanley, P. A., Bjork, E. L., & Bjork, R. A. (1996). Generation effects and the lack thereof: The role of transfer-appropriate processing. *Memory, 4,* 31–48.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14,* 4–58.

Duvivier, R. J., van Dalen, J., Muijtjens, A. M., Moulaert, V. R. M. P., van der Vleuten, C. P. M., & Scherpbier, A. J. J. A. (2011). The role of deliberate practice in the acquisition of clinical skills. *BMC Medical Education, 11,* 101.

Ericsson, K. A. (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Erlbaum.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15,* 988–994.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Hannafin, M. J., & Hooper, S. R. (1993). Learning principles. In M. Fleming & W. H. Levie (Eds.), *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed., pp. 191–231). Hillsdale, NJ: Educational Technology Publications.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89,* 92–102.

Healy, A. F., & Bourne, L. E. Jr. (Eds.). (2012). *Training cognition: Optimizing efficiency, durability, and generalizability*. New York, NY: Psychology Press.

Healy, A. F., Clawson, D. M., McNamara, D. S., Marmie, W. R., Schneider, V. I., Rickard, T. C., . . . Bourne, L. E. Jr. (1993). The long-term retention of knowledge and skills. In D. Medin (Ed.), *The psychology of learning and motivation* (pp. 135–164). New York, NY: Academic Press.

Healy, A. F., Schneider, V. I., & Bourne, L. E. Jr. (2012). Empirically valid principles of training. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Training cognition: Optimizing efficiency, durability, and generalizability* (pp. 13–39). New York, NY: Psychology Press.

Jackson, G. T., Boonthum, C., & McNamara, D. S. (2010). The efficacy of iSTART extended practice: Low ability students catch up. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 349–351). Berlin, Germany: Springer.

Jackson, G. T., Boonthum-Denecke, C. B., & McNamara, D. S. (2012). A comparison of gains between educational games and a traditional ITS. In P. M. McCarthy & G. M. Youngblood (Eds.), *Proceedings of the 25th International Florida Artificial Intelligence Research Society (FLAIRS) Conference* (pp. 444–449). Menlo Park, CA: AAAI Press.

Jackson, G. T., Guess, R. H., & McNamara, D. S. (2010). 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.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.

Mayer, R. E. (1992). Cognition and instruction: Their historic meeting within educational psychology. *Journal of Educational Psychology, 84,* 405–412.

Mayer, R. E. (2012). Advances in applying the science of learning to education: An historical perspective. *Journal of Applied Research in Memory and Cognition, 1,* 249–250.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., . . . Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34,* 51–57.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38,* 1–30.

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., Crossley, S. A., & Roscoe, R. D. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods, 45,* 499–515.

McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., & Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing, 23,* 35–59.
McNamara, D. S., & Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P. M. McCarthy & C. Boonthum (Eds.), *Applied natural language processing and content analysis: Identification, investigation, and resolution* (pp. 188–205). Hershey, PA: IGI Global.
McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, England: Cambridge University Press.
McNamara, D. S., & Healy, A. F. (1995a). A generation advantage for multiplication skill and nonword vocabulary acquisition. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills* (pp. 132–169). Thousand Oaks, CA: Sage.
McNamara, D. S., & Healy, A. F. (1995b). A procedural explanation of the generation effect: The use of an operand retrieval strategy for multiplication and addition problems. *Journal of Memory and Language, 34,* 399–416.
McNamara, D. S., & Healy, A. F. (2000). A procedural explanation of the generation effect for simple and difficult multiplication problems and answers. *Journal of Memory and Language, 43,* 652–679.
McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavior Research Methods, Instruments, & Computers, 36,* 222–233.
McNamara, D. S. & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation*. New York, NY: Elsevier Science.
McNamara, D. S., O'Reilly, T., Best, R., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research, 34,* 147–171.
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. 397–421). Mahwah, NJ: Erlbaum.
Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.
Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition, 1,* 242–248.
Roscoe, R. D., Allen, L. K., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2014). The Writing Pal intelligent tutoring system: Usability testing and development. *Computers and Composition, 34,* 39–59.
Roscoe, R. D., Brandon, R. D., Snow, E. L., & McNamara, D. S. (2013). Game-based writing strategy practice with the Writing Pal. In K. Pytash & R. Ferdig (Eds.), *Exploring technology for writing and writing instruction* (pp. 1–20). Hershey, PA: IGI Global.
Roscoe, R. D., & McNamara, D. S. (2013). Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology, 105,* 1010–1025.
Roscoe, R. D., Snow, E. L., Brandon, R. D., & McNamara, D. S. (2013). Educational game enjoyment, perceptions, and features in an intelligent writing tutor. In C. Boonthum-Denecke & G. M. Youngblood (Eds.), *Proceedings of the 26th Annual Florida Artificial Intelligence Research Society (FLAIRS) Conference* (pp. 515–520). Menlo Park, CA: AAAI Press.
Roscoe, R. D., Varner, L. K., Cai, Z., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2011). Internal usability testing of automated essay feedback in an intelligent writing tutor. In R. C. Murray & P. M. McCarthy (Eds.), *Proceedings of the 24th international Florida Artificial Intelligence Research Society (FLAIRS) Conference* (pp. 543–548). Menlo Park, CA: AAAI Press.
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207–217.
Schmidt, R. A., Young, D. E., Swinnen, S., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 352–359.
Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.
Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Sui, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science, 323,* 122–124.
Snow, E. L., Jackson, G. T., Varner, L. K., & McNamara, D. S. (2013a). The impact of system interactions on motivation and performance. In *Proceedings of the 15th International Conference on Human–Computer Interaction (HCII)* (pp. 103–107). Berlin, Germany: Springer.
Snow, E. L., Jackson, G. T., Varner, L. K., & McNamara, D. S. (2013b). 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). Retrieved from http://www.educationaldatamining

.org/conferences/index.php/EDM/2013/schedConf/presentations
Snow, E. L., Jacovina, M. E., Allen, L. K., Dai, J., & McNamara, D. S. (2014). Entropy: A stealth measure 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). Retrieved from http://www.educationaldatamining.org/conferences/index.php/EDM/2014/schedConf/presentations
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). Retrieved from http://www.educationaldatamining.org/conferences/index.php/EDM/2013/schedConf/presentations
VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47,* 513–539.
Varner (Allen), L. K., Roscoe, R. D., & McNamara, D. S. (2013). Evaluative misalignment of 10th-grade student and teacher criteria for essay quality: An automated textual analysis. *Journal of Writing Research, 5,* 35–59.