UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
Designing for Productive Persistence after Failure

Permalink
https://escholarship.org/uc/item/4wh7z0jn

Author
Tran, Cathy

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Designing for Productive Persistence after Failure

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Cathy Tran

Dissertation Committee:
Distinguished Professor Jacquelynne Eccles, Chair
Assistant Professor AnneMarie Conley, Co-Chair
Associate Professor Susanne Jaeggi
Professor Linda Levine
Professor Jodi Quas

2015

© 2015 Cathy Tran

TABLE OF CONTENTS

		Page
LIST OF FIGURES		iii
LIST OF TABLES		iv
ACKNOWLEDGMENTS		v
ABSTRACT OF THE DISSERTATION		vi
CHAPTER 1:	Introduction	1
CHAPTER 2:	Study 1: Designing for Engagement: Information Seeking at a Science Museum	6
CHAPTER 3:	Study 2: Understanding Goal Change: Theoretical Advances and Leverage for Design	36
CHAPTER 4:	Study 3: The Effects of Miscalibration on Confusion, Learning Behaviors, and Learning Outcomes	64
CHAPTER 5:	Conclusion	98

LIST OF FIGURES

		Page
Figure 2.1	Heat Pump Game During Play and the Display of the Score at the End	33
Figure 2.2	Heat Pump Simulation Game Setup with the Crank and the Heated Handprints	34
Figure 2.3	Heat Pump Simulation Game Setup at the Museum with Two Identical Games Back-to-Back	35
Figure 3.1	Heat Pump Game During Play and the Display of the Score at the End	61
Figure 3.2	Heat Pump Simulation Game Setup with the Crank and the Heated Handprints	62
Figure 3.3	Heat Pump Simulation Game Setup at the Museum with Two Identical Games Back-to-Back	63
Figure 4.1	Screenshot of Icons Used by Students to Rate their Confidence Immediately after Selecting an Answer	94
Figure 4.2	Screenshot of Feedback Interface for Incorrect Answer Shown Immediately after Students' Confidence Rating	95
Figure 4.3	Screenshot of Confusion Rating Scale Shown Immediately after Students Rate their Confusion	96
Figure 4.4	Screenshot of Feedback Symbols Representing Correct and Incorrect Answers and Option Buttons to Request an Explanation or Move onto the Next Problem Immediately	97

LIST OF TABLES

		Page
Table 2.1	Typology of Information-Seeking Behaviors	32
Table 4.1	Six Possible Calibration Categories	85
Table 4.2	Number of Answers for Different Calibration Categories	86
Table 4.3	Mean Confusion Level for Different Calibration Categories	87
Table 4.4	Ordinary Least Squares Regressions Predicting Confusion Rating from Calibration Category	88
Table 4.5	Number of Problems for which an Explanation Box was Requested for Different Calibration Categories	89
Table 4.6	Logistic Regressions Predicting Likelihood to Request an Explanation from Calibration Category	90
Table 4.7	Mean Time Spent on Feedback for Different Categories of Calibration	91
Table 4.8	Ordinary Least Squares Regressions Predicting Time Spent on Feedback from Calibration Category	92
Table 4.9	Regression of Confidence Rating of Incorrect Answers Predicting Correction of Errors on a Posttest	93

ACKNOWLEDGMENTS

This dissertation is a collective effort that emerged from the love, support, and intellectual mind candy that I received over the years.

Many thanks to my dissertation committee and colleagues for sharing their insights as I tackled how to bridge the different perspectives that inform my work and to conduct research that is just as practically valuable as it is theoretically intriguing.

I am incredibly lucky to have had the opportunities to form research-practice partnerships with the MIND Research Institute, the University of Oslo's InterMedia Lab, and the Norwegian Museum of Science and Technology. This dissertation would not have been possible without their collaborations and support.

Thank you to the foundations that funded my research through the National Science Foundation (NSF) Graduate Research Fellowship, the NSF/Norway Research Council Nordic Research Award, and the National Academy of Education/Spencer Dissertation Fellowship. It meant a lot to me to have their vote of confidence.

Last, but not least, lots of hugs and gratitude to my husband Jeff for his support and patience and to our pup Ollie who was the absolute perfect dissertation writing buddy.

ABSTRACT OF THE DISSERTATION

Designing for Productive Persistence after Failure

By

Cathy Tran

Doctor of Philosophy in Education

University of California, Irvine, 2015

Jacquelynne Eccles, Chair; AnneMarie Conley, Co-Chair

This dissertation provides insight on designing learning environments that promote productive persistence. Two projects bridged the fields of educational psychology, cognition, affective science, and game design to examine what influences persistence after failure.

The first project used self-reports, interviews, and observations to explore ways the design of a game in a science museum interacted with and influenced students' motivation during a high school field trip. Analyses illustrated that five design features modified student motivation in shaping behavior, supporting the pursuit of some goals while hindering the pursuit of others. Case studies that traced the information-seeking patterns of four students, who endorsed different motivational profiles, provided an understanding of the reasons for their adoption, maintenance, and change of achievement goals throughout the visit.

The second project provided opportunities for productive confusion by introducing a design intervention that asked students to rate their confidence while solving mathematics problems during a computer task for fourth- and fifth-grade

students. The ratings occurred after students selected answers to multiple-choice questions. Cognitive conflict emerged during instances in which students were not well calibrated; that is, they were highly confident of their answer but incorrect (i.e., high confidence errors) or not confident but correct (i.e., low confidence corrects). Findings underscore that cognitive conflict can elicit confusion and influence students to invest more effort towards better understanding mathematics problems but that their degree of effort can diminish when the feedback does not successfully resolve their confusion.

CHAPTER 1

Introduction

We could not learn without failure. In fact, trial and error may be the most fundamental learning mechanism in nature. It is how babies explore and get to know the world. What tastes good? What makes us laugh? What hurts? We try. We stumble. We revise. We try again. If we only did things we already knew how to do, we would not learn anything. This kind of productive failure is prevalent in children's play and inherent in their games, particularly video games. Failure is the norm in most video games—success is the exception. How many replays does it take to complete a level of Nintendo's *Super Mario Bros* successfully? Players experiment with one strategy after another, making adjustments, keeping what works, and discarding what proves to be unsuccessful. It is a wonderful iterative learning process.

Sadly, those same players who show such resilience in games often view failure in school as confirmation of their inability. This dissertation tackles that challenge of how to design learning environments that embrace failure and the benefits of learning from failure. Deep knowledge is more likely to be developed through failures than successes since we are more likely to search for causes for failure than causes for success (for a review, see Weiner, 1985). For instance, research with tutors shows that students learn more from explanations given by tutors after, rather than before, they have experienced a lack of requisite knowledge to solve a given task (VanLehn, Siler, Murray, Yamuachi, & Baggett, 2003). Research on productive failure shows that students are more prepared to learn after they generate several solutions to a complex, open-ended problem even when those solutions are rarely accurate (Kapur, 2008). Finally, work in the area of desirable difficulties provides evidence that instructional manipulations that introduce difficulties can be strategically designed to slow down the rate of

knowledge acquisition but enhance recall and transfer (Bjork, 2013). This body of work illustrates that learning profits from instances of failure, but with the caveat that to reap those benefits, one must have adequate resources to resolve his or her lack of understanding and must be motivated to persist in the face of failure.

This dissertation provides insight on how to design learning environments in a manner that elicits productive persistence which occurs when, after failure, students choose the strategies and behavioral moves that drive towards successful completion of the learning goal (Chase, 2011). Within this realm, the impact of video games on motivating players to persist in overcoming failure has generated substantial interest in the potential for using games and game principles to motivate more "serious" undertakings. Literature reviews and meta-analyses about the effects of educational games have concluded that their impact on learning and motivation is mixed, however, with many studies revealing that educational games are often no more effective than other instructional methods (Clark, Yates, Early, & Moulton, 2010; Ke, 2009; Sitzmann, 2011; Vogel et al., 2006; Wouters, Nimwegen, Oostendorp, & van der Spek, 2013). In explaining the reason for this, Wouters et al. (2013) suggest that a greater integration between the fields of game design and instructional design is needed. Steps toward this integration require bridging the work in educational psychology, cognition, affective science, and game design to understand what influences persistence after failure and how knowledge about those influential factors can inform the development of educational environments.

Designing effective instructional game and non-game environments that promote productive persistence goes beyond imposing game elements such as points and badges onto existing educational tasks or integrating educational elements such as multiple-choice questions into existing games. This research was inspired by the need for identifying the underlying

elements, structures, and contexts that are likely to promote productive persistence in games in order to aid in the investigation of how to effectively apply those elements to the challenges of more formal learning environments. Studies 1 and 2 of this dissertation emerged from my research stay at the University of Oslo's InterMedia Lab and were conducted in collaboration with the Norwegian Museum of Science and Technology. My colleagues and I explored ways in which the design of a game in a science museum exhibit interacts with and influences students' goals during a high school field trip. A micro-level analysis that combined self-reports, interviews, and videos allowed us to assess the process of goal pursuit and learning. The first study illustrated ways in which particular design features of the game supported and hindered the pursuit of different goals by affording particular types of information-seeking behaviors. The match and mismatch between those design affordances and visitors' goals influenced the players' level of persistence and satisfaction with the exhibit. The second study detailed existing methodological limitations of investigating task goal change and highlighted alternative methods to better understand that process. Using case studies of four students, we presented video and interview data that illustrated the goal change processes of four students, including shifts in goal pursuit intensity and goal type. In a separate project that was a collaboration with the MIND Research Institute (Irvine, CA), Study 3 introduced a design intervention in math software that was intended to provide opportunities for productive confusion. We assessed ways in which that intervention influenced students' level of attention to feedback, information-seeking behaviors, and likelihood of correcting previous errors. The complementary perspectives from the three studies in this dissertation provide practical insights and theoretical contributions regarding what makes failure appealing, how failure relates to learning, and ways to design for productive responses to failure.

References

Bjork, R.A. (2013). Desirable difficulties perspective on learning. In H. Pashler (Ed.), *Encyclopedia of the mind.* Thousand Oaks, CA: Sage Reference.

Chase, C.C. (2011). *Motivating persistence in the face of failure: The impact of an ego-protective buffer on learning choices and outcomes in a computer-based educational game* (Doctoral dissertation). Stanford University, CA.

Clark, R. E., Yates, K., Early, S., & Moulton, K. (2010). An analysis of the failure of electronic media and discovery-based learning: Evidence for the performance benefits of guided training methods. In K. H. Silber & R. Foshay (Eds.), *Handbook of training and improving workplace performance: Instructional design and training delivery* (pp. 263–297). Somerset, NJ: Wiley.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379–424.

Ke, F. (2009). A qualitative meta-analysis of computer games as learning tools. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 1–32). Hershey, PA: Information Science Reference.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64,* 489–528.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*, 209-249.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34,* 229–243.

Weiner, B. (1985). Spontaneous causal thinking. *Psychological Bulletin, 97*(1): 74–84.

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van der Spek, E.D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2): 249-265.

CHAPTER 2

Study 1: Designing for Engagement: Information Seeking at a Science Museum

As informal science institutes increasingly partner with schools, a question that has become prevalent is how to measure the types of learning happening in those out-of-school environments (Bevan et al., 2010). Traditional pre- and post- measures of content knowledge do not align with museum educators' goals of sparking curiosity and increasing students' future desire to learn. The amount of time spent at an exhibit, a common marker of engagement, also has limited assessment value because decades of research on student motivation suggest that deep and continuous learning does not depend only on *amount* of engagement but also on *reasons* for engagement (Schunk, Pintrich & Meece, 2007). Understanding ways in which a museum exhibit influences students' goals and goal pursuit behaviors would provide insight on students' reasons for engagement, allowing for a more complete assessment of the impact of the exhibit. Goals are an important consideration in the design of science museums because goals influence what students attend to and how they respond to feedback from both the physical and digital features of the exhibit and from other visitors. The design features of the exhibit interact with learners' goals to influence their behaviors, and reciprocally, the outcomes of those behaviors dynamically shape their goals during the visit. For students who enter without clearly defined goals, the way the exhibit is designed can influence them to adopt particular goals. These goals include, for example, mastering content knowledge, avoiding looking incompetent, and outperforming others. Understanding how the design of an exhibit can influence students' goals matters because goals have behavioral, cognitive, and emotional consequences.

Though students' reasons for engagement are cognitive, those reasons behaviorally manifest in the choices that students make about which information to attend to and the types of

information that they seek. These information-seeking behaviors provide insight about students' learning processes and their progress or lack thereof towards content mastery. To provide the premise for these connections, we first introduce achievement goal theory (Dweck & Leggett, 1988) and explain why its emphasis on the reasons individuals are motivated, as well as the standard against which they judge their success, is appropriate for studying interactive games at a science museum. Next, we explain how students' goals and learning processes can be understood through their behaviors by providing an overview of the research on the relation between achievement goals and help-seeking behaviors. The aim of this paper is to build on and extend the literature in help seeking and learning to better understand the ways the design of an exhibit can influence and provide resources for different types of information-seeking behaviors and goals. This is especially of relevance to museums because the free-choice environment, as Allen (2004) puts it, is "the single greatest constraint underlying exhibit design." Whereas in school, teachers can use targeted strategies to regulate different students' progress, in a museum, this regulation mostly occurs through the affordances that the design of the exhibit provides for the students, who enter with different background knowledge and goals.

Achievement goals = Reason for motivation + Standard for judging success

We draw on achievement goal theory (Dweck & Leggett, 1988) of motivation to offer researchers, educators, and designers theoretically-grounded constructs that have implications for both studying and designing digital games in museum settings. Achievement goals reflect the reasons individuals are motivated as well as the standard against which they judge their success. Researchers have distinguished two achievement goal orientations toward learning: mastery and performance. A mastery goal focuses on *developing* skills and knowledge whereas a performance goal focuses on *demonstrating* competence by outperforming others (Anderman &

Wolters, 2006). Though students typically endorse multiple goals, one of those goals is often more dominant than the others (Conley, 2012). The adoption of mastery goals has been linked to positive effects, including deeper cognitive strategies and greater interest in the task (Dweck & Leggett, 1988). Performance goals have varying consequences with an important distinction being whether students adopt a performance-approach focus on appearing highly competent or a performance-avoid focus on not appearing dumb. For instance, performance-approach students may want the highest score in the class whereas performance-avoid students will choose to not study for an exam to justify potential low grades being a result lack of effort rather than ability. Research has generated mixed results on the associations between achievement and the performance-approach goal orientation, indicating that such orientations are linked to unrealistic high goals in some cases but also to high achievement in others (Harackiewicz et al., 1997; Kaplan & Maehr, 2007). The endorsement of the achievement goal orientation of high performance-approach in conjunction with high mastery goals has been linked to both high interest and high achievement in a task (Conley, 2012). Performance-avoid goals are often linked to negative consequences such as procrastination, putting in little effort, and cheating (Elliot & Church, 1997; Harackiewicz, Barron, & Elliot, 1998; Urdan, 2004).

The context of our study, a heat pump game at a science museum, is particularly appropriate for studying motivation through the lens of achievement goal theory because it provides affordances for both mastery and performance goals. For instance, students can pursue mastery goals of understanding the scientific mechanisms that underlie the functioning of the heat pump through the digital simulations of the inner workings of the heat pump (Figure 2.1, left). As players rotate a physical metal crank right or left, the simulation shows the resulting changes in the pressure and flow of heat in the depiction of the interior of the heat pump

displayed on a digital screen. For students oriented towards performance goals, the game scores provide affordances for pursuing those goals. The scores reflect how successfully the players were able to operate the heat pump to keep a house at an optimal temperature throughout the year. Since we have two games occurring simultaneously and in close proximity to each other, students can compare scores. These game features allow us to gain insight about how students use the exhibit in different ways to align available resources to their specific goals as well as how the consequences of their behaviors influence their subsequent goal pursuits during the game.

Goals influence information-seeking behaviors

We extend the existing research literature on achievement goals and help seeking in the classroom to the context of an educational game at a science museum. Research on achievement goal orientations suggests that both individual goals and student perceptions of the classroom contribute to different help-seeking behaviors including choice in source of help (e.g., teachers vs. peers) and choice in type of help (e.g., seeking to quickly obtain the correct answer vs. seeking to understand the reasons for the correct answer) (Schenke, Chang, Conley & Karabenick, 2014). For instance, classrooms led by teachers who encourage questions and are available for assistance typically have students who engage in adaptive help seeking with the goal of understanding how to do a problem (Karabenick & Sharma, 1994). However, classrooms led by teachers who are focused on absolute performance and competition result in students seeking more maladaptive forms of help such as seeking assistance with the goal of quickly getting the answer (Newman & Schwager, 1993). These differences in help-seeking actions in part explain why students learn and progress at different rates when encountering the same obstacles.

To build on the classroom help-seeking research to better understand how the design of a game exhibit influences students' information-seeking behaviors at a science museum, we first broadened the definition of help seeking, which typically includes only the social elements, to also include the non-social components of the process. This is necessary because the physical and digital aspects of an educational game exhibit are significant parts of the learning experience. Our definition of help seeking therefore encompasses any "process in which humans purposefully engage in order to change their state of knowledge" (Marchionini, 1995, p. 5), which involves goal-directed interactions with any information system including humans, computers, and physical manipulative. Because help seeking is often used in the education research literature to refer to social help seeking, we adopt the term *information seeking* to reflect the notion that both social and non-social help seeking play crucial roles in interactions with digital and physical installations at science museums.

Games provide a naturalistic way for assessing the effects of the design of a task on information-seeking behaviors in a short timespan because games are deliberately designed to set players up for failure and allow them to experience it, experiment with it, and learn from it (Juul, 2013). Players engage in rapid and complex interactions as a result of the responsive, adaptive, and interactive components of a given game. This is conducive to studying the directed, effortful process of learners setting goals for their learning and then attempting to monitor, regulate, and control their behavior, motivation, and cognition. In the heat pump game, the frequent patterns of information-seeking behaviors that emerged during gameplay allowed us to study how the design of the game affects information seeking and goal regulation. Students incorporate feedback from their successes and failures into their evolving game strategies and goals.

We used self-reports, interviews and videos to trace learners' goals and behaviors as they interacted with a science game that is part of an exhibit that our team developed for a museum. Specifically, our research questions are as follow:

1) What types of information-seeking behaviors do students employ while playing an educational game at a science museum?
2) In what ways do students' goals interact with the design features to influence their information-seeking behaviors?

Method

Context of learning environment

Data presented were collected during a science museum visit during which Norwegian high school students explored an interactive digital game about heat pumps as part of a class field trip during the school day. The game is part of a larger exhibit about future energy sources at the Norwegian Science and Technology Museum. The exhibit resembles a carnival with game booths that each focus on different energy sources such as sun, wind, and ocean waves.

The heat pump game booth designed by our university lab and the museum aimed to familiarize students about energy transfer and the relation among pressure, condensation, evaporation, and temperature. Players learn about those physics properties through the general function of a heat pump, which moves heat from inside to outside and vice versa through processes of condensation and evaporation. When the game begins, the start screen describes the function of a heat pump before transitioning to a screen of the magnified heat pump with details of its inner workings at the center of the screen. A house and its internal temperature meter are on the left side of the simulation and a meter of the outdoor temperature is on the right side (Figure 2.1, left). Players are challenged to keep the house temperature consistently warm

throughout the year by operating the heat pump through physically rotating a metal crank underneath the screen in the appropriate direction using the appropriate speed. Next to the crank are two physical handprints that students can touch; the handprints change temperature to align with the movement of heat on the screen in the simulation of the heat pump (Figure 2.2). As students manipulate the heat pump compressor using the physical crank to heat up or cool down the house, the heat pump's inner workings dynamically move in real time in the middle of the screen (Figure 2.1, left). Those visualizations are linked to player's operation of the metal crank; the two chambers in the simulation react to clockwise and counter-clockwise turnings of the crank, illustrating conditions of boiling and condensation. The result screen at the end of the game shows the percentage of time the house stayed within the desirable zone of warmth as well as the amount of energy saved by using the heat pump, when compared to alternative forms of temperature regulation (Figure 2.1, right). Each game cycle is approximately one minute, and each student in this study visited the heat pump exhibit for about 15 minutes, as part of a group of five students.

Participant selection

The four key participants in our study were chosen from 32 first-year students (16 of each gender, ages 15 and 16) in one science classroom at a moderately selective high school in Oslo, Norway. We administered self-report measures of achievement goal orientations for science learning prior to the museum visit to identify students with different motivational profiles. We chose four participants who endorsed one of each of the following motivation profiles: 1) predominantly mastery oriented; 2) predominantly performance-approach oriented; 3) predominantly performance-avoid and mastery oriented; and 4) similarly performance-approach, performance-avoid and mastery oriented. Our analyses focused on these students' interactions

with classmates in the groups they were assigned to during a field trip to a science museum. The researchers were blind to the profiles of the selected students while they analyzed the data. Selecting our participants based on motivational profiles allowed us to assess how the design features of the exhibit related to information-seeking behaviors across students who endorsed different motivational goals.

Analytical methods

Our analytical process combined data from the self-reports about motivation, video analyses during the museum visit, and student interviews after the visit. The self-reports allowed us to identify students who endorsed different levels and types of goals in order to assess how the design features of the exhibit related to interaction patterns across students who endorsed different motivational profiles. The videos of the students' visit at the heat pump game booth allowed us to identify the different types of information-seeking behaviors that emerged. Finally, there were instances in which there were insufficient details to determine from the observed behaviors if there was an intent to seek information, and if so, what information was being sought. The one-on-one student interviews after the visit allowed us to show those ambiguous video clips to students so that, if they were able to recall their thinking, they could explain what prompted their behaviors during those episodes. This helped us determine whether to classify those episodes as information-seeking behaviors.

Achievement goals measure. The nine-item self-report measure was a Norwegian translation of the Achievement Goals Questionnaire (Elliot & McGregor, 2001), adapted to focus on achievement goals during science class. Items assessing mastery, performance-approach, and performance-avoid goals were scored on a 7-point Likert scale ranging from 1 (*not at all true for me*) to 7 (*very true for me*). Mastery goals focused on learning and understanding (e.g., "My goal

in science class is to learn as much as I can"); performance-approach goals focused on demonstrating ability and outperforming others (e.g., "My goal in science class is to look smarter than other students"); and performance-avoid goals focused on not looking dumb (e.g., "My goal in science class is to avoid looking like I can't understand the material"). There was acceptable internal consistency for the performance-approach ($\alpha = 0.81$), performance-avoid ($\alpha = 0.76$), and mastery scales ($\alpha = 0.83$). Previous factor analyses with larger samples support that the Norwegian translation of the achievement goals measure resulted in one-dimensional constructs (Bjørnebekk & Diseth, 2010).

Interaction analysis of video recordings. Interaction analysis (Jordan & Henderson, 1995) is a method that emphasizes the patterns of the interaction of human beings with one another and with objects of their environments, noting which resources and conversations get taken up and how. This includes talk, nonverbal interaction, and the use of physical (e.g., heat pump crank) and digital (e.g., heat pump screen animation) artifacts. The aim is to uncover activity patterns that emerge as the analysis proceeds to identify routine practices and problems within the game and the resources for their solution. Using principles of interaction analysis, we looked at two groups of five students interacting with back-to-back heat pump games (Figure 2.3). Each group spent about 15 minutes at the heat pump game and all interactions with our four target students were included in the analyses.

Interaction analysis aligns with our goals to identify how design features and the social environment interact to support or hinder the pursuit of different types of goals. The bottom-up procedure is appropriate given that our context of a game at a science museum includes resources that are very different from the resources in contexts commonly reviewed in the help seeking and information seeking literature such as classroom environments and Internet search

engines. As such, in categorizing information-seeking behaviors, we remained open to how the categories might emerge within our data. The patterns of behaviors emerged as we identified key activity episodes that promoted or inhibited goal pursuit and analyzed the interactions prior to, during, and after those episodes. We determined how particular game design elements facilitated information-seeking behaviors and how patterns of information-seeking behaviors relate to student goals.

Our analysis of video recordings of two groups of students interacting with the heat pump exhibit followed the precepts of interaction analysis for understanding everyday work practices and human-machine interactions (Jordan & Henderson, 1995; Suchman, 1987). First, a project staff prepared transcripts of utterances for the recorded footage. Two researchers then viewed the videos and read the transcripts repeatedly to formulate tentative assertions. In subsequent viewings and transcript readings, we refined the transcriptions of information-seeking episodes to include other interaction factors including gestures, pauses, and overlaps of utterances. In cycles of interpretation of particular episodes and testing of categorical constructions of information seeking, we discarded, modified, or adopted emerging understandings. During this process, the researchers who conducted the analyses were blind to the motivational profiles of the students who were selected by another researcher on the team so that this knowledge would not bias their analyses.

Post-visit video recall and interviews. Interviews were conducted with students one-on-one with about three weeks after the museum field trip, allowing sufficient time for the analyses of results from the video data to guide the interview structure. The interviews lasted about half an hour and were video recorded. The interview guide was semi-structured and contained three sections. First, the interviewer asked students about their goals for the museum visit. Second, the

interviewer showed two short video clips, less than one minute each, of the interviewee's behaviors at the heat pump game. After each video episode, the interviewer asked the students to explain why they behaved as they did if they are able to recall those reasons. To our knowledge, video recall has not been used with high school students and in our context. However, video-recall procedures have been frequently used in family psychology research and were developed to elicit participants' subjective understanding of their interactions, behaviors, or experiences in conjunction with traditional observer coding systems (Welsh & Dickson, 2005). Previous research has supported the validity of video-recall procedures by simultaneously recording a variety of physiological measures during the original conversation and again during the video-recall procedure. The physiological data during the recall session were significantly related to participants' physiological data during the original interaction, suggesting that participants were "reliving" their experiences (Levenson & Gottman, 1985).

The video clips used during the interview were flagged as ones that lacked sufficient details, based on our interaction analyses, to include or rule out as information-seeking behaviors. For example, in one simulated video recall episode, a student moved from his game to watch another student play the same game for about five seconds before returning to his game. There were no actions prior to or after that episode that provided details for identifying if he was seeking information and if so what information. There was no talk, as he simply went to the other side to watch when his game cycle ended and returned to his side to start a new game. When shown this video clip during the interview, the student indicated that he wanted to know the purpose of the physical handprints and if the other group had strategies for incorporating them into their gameplay. Because the other group also did not use the handprints, the student decided that they were not important and continued to not use them after returning to the game

on his side. The supplementary data from the interview allowed us to document the clip as an information-seeking behavior.

Findings

Typology of information-seeking behaviors

Our documentation and categorization of the types of information-seeking behaviors that students employ resulted in a typology for information-seeking behaviors that was appropriate for our context. This typology emerged after the authors sorted the behaviors, proposing a variety of categories for consideration. The typology that we chose represented all information-seeking behaviors that were identified and is general enough such that it can be adapted to other museum and game environments. The two dimensions of our typology of information-seeking behaviors are: 1) social to non-social; and 2) more directed to less directed. Examples of the different information-seeking behaviors that we found within the different combinations of the dimensions are in Table 2.1.

Within the *social to non-social dimension*, we categorized the behaviors into three buckets: a) two-way interactions, b) one-way interactions, and c) no interactions with another individual. We define two-way interactions as social information-seeking behaviors intended to elicit a response such as a question directed at a specific person or group (e.g., "do you know why these colors changed?"). One-way interactions are in the middle of the dimension as somewhat social and include a general announcement (e.g., "this is hard!") or watching others interact but not engaging in that interaction. We define no interactions as non-social information-seeking behaviors that do not involve another person (e.g., reading text on the wall next to the game). For the *more directed to less directed dimension*, we also categorized the behaviors into three groups: a) creating opportunities to seek information; b) taking advantage of available

opportunities to seek information; and c) exploratory behaviors. We define more directed information-seeking behaviors as being purposeful and less directed information-seeking behaviors as those exhibited without little to no intention to act on the new information. For example, creating opportunities to seek information is very directed, taking advantage of available opportunities to seek specific information is somewhat directed, and exploratory behaviors are less directed. Table 2.1 includes more detailed examples of these different categories.

Student goals interact with game design features to influence information-seeking behaviors

Highlighting five key design features, we illustrate how students used resources in the environment, including the game and each other, to seek information to pursue their goals. In this section, we highlight the interactions of all five students in one group and of three students in a second group. Though both groups contained five members, we focused only on the students that interacted with the four target students that were previously selected based on their motivation profiles. Those four students (*as indicated by asterisks) are also the students who were interviewed after the visit.

Group A:

1) *Frank** reported **similar goal orientations** across mastery (6.67 out of 7), performance-avoid (5 out of 7) and performance-approach (4.7 out of 7) goal orientations.
2) *Neil** reported **predominantly a mastery goal orientation** (6.7 out of 7) and **low performance goal orientations for both approach** (2 out of 7) **and avoid** (1.5 out of 7).

3) *Amy* reported a predominantly mastery goal orientation (5 out of 7) and low performance-approach (2.7 out of 7) and performance-avoid (1 out of 7) goal orientations.
4) *Mary* reported a predominantly mastery goal orientation (5 out of 7) and a low performance-avoid (3 out of 7) orientation, and a very low performance-approach (0.3 out of 7) goal orientation.
5) *Ashley* reported a predominantly mastery goal orientation (7 out of 7), a relatively low performance-approach orientation (3.3 out of 7), and a very low performance-avoid orientation (1 out of 7).

Group B:

1) *Carly** reported a **predominantly performance-approach goal orientation** (6 out of 7) and **moderate performance-avoid** (4.5 out of 7) **and mastery** (4.7 out of 7) goal orientations.
2) *Linda** reported **high performance-avoid** (5.5 out of 7) and **high mastery** (6 out of 7) goal orientations and a **low performance-approach** (3.3 out of 7) goal orientation.
3) *George* reported a high mastery goal orientation (6.3 out of 7) and low performance-approach (1 out of 7) and performance-avoid (1 out of 7) goal orientations.

Below we detail ways in which specific design features are related to students' information-seeking behaviors and the extent to which these behaviors are related to learners' previously reported achievement goals. Our findings illustrate that the interactive network of social others, physical objects, and digital displays partially determines, constrains, and supports the types of needs and inquires that arise from the learner during their interactions with the exhibit.

Design feature #1: Close proximity of games. The proximity of two identical back-to-back games at the heat pump exhibit (Figure 2.3) and the design of the players facing each other allows for players on opposite sides to hear each other's conversations and move back and forth. Therefore, individuals can pursue their goals by weaving in and out of conversations that align and misalign with their goals. This is illustrated in an episode in which Amy and Ashley announced that they did not understand what was going on. Neil, who endorsed a predominantly mastery goal orientation, pointed to the heat pump simulation in the middle of the screen to ask them what that visual conveyed. In response to his question, Ashley shrugged and Amy continued to crank while laughing and not responding to his question. When shown this episode during the video-recall portion of the interview, Neil explained that he still did not know how the heat pump worked at this point and was therefore engaging in *more directed* and *social* information seeking by asking the other students to decipher the simulation with him. Although Amy and Ashley did not engage in that conversation with him, another student from the other side popped her head around her game screen to answer. Mary, who endorsed a high mastery goal orientation similarly to Neil, chimed in with her three predictions about how the heat pump works by making references to the visual simulation as well as to the physical handprints. Neil then turned away from Amy and Ashley and engaged in a conversation with Mary. Although Neil's information-seeking behavior was unsuccessful when directed at Amy and Ashley, he received the help from Mary as a result of the close proximity of the two games.

Design feature #2: Short, low-stakes game cycles. The heat pump exhibit's feature of a low-stakes, one-minute game provided opportunities for students to learn through experimentation. For instance, at the outset of his gameplay, Frank cranked the heat pump clockwise very quickly while watching the screen feedback and results. Then he cranked the heat

pump counterclockwise very quickly. In the post-visit interview, a researcher showed Frank a clip of him during these episodes in which his actions resulted in low game scores. Frank, who endorsed a high mastery goal orientation, explained his action by saying, "I really didn't know what was going on with the heat pump. I had some hypotheses that I came up with so it was a bit of guessing. My goal was trying to understand the heat pump and what it did." These goal-driven behaviors that aligned with the mastery goal orientation that Frank endorsed reflect a *more directed* and *non-social* information-seeking pattern that was supported by the design of the game.

In a later incident, however, this same design feature of a short game cycle disrupted Frank's information seeking. While his classmates played the back-to-back heat pump games, Frank stood to the side in the middle and did not participate in the discussions that were predominantly focused on scores and not on science understanding. Classmates laughed and shouted directions at each other to crank left or right, faster or slower, to get the highest possible score. In the post-visit interview, Frank said that he felt the purpose of the museum visit was to learn and not to play. He chose to stay out of the discussions focused on high scores, but when he overheard the word "evaporation" being spoken, he joined the conversation and said, "I heard something about evaporation and…" However, he was interrupted by the results screen appearing and Amy exclaiming her score and pumping her fist in the air as Ashley asked her if she was going to try again to get a score of 100.

Design feature #3: Start of new game cycle. Related to the short cycle, the start of a new game provides an opportunity for a natural switch in conversational topics, a design feature that can shift the goal of the group. For example, in Group A's interactions with the game, Amy and Ashley strategized to get a high score while Frank stood to the side of the game. The

simulation of the scientific mechanism of the heat pump at the center of the screen was not a point of focus for Amy and Ashley, as indicated by them looking at the side temperature meters and shouting out the numbers. When the screen transitioned from the results screen to a new game screen, however, their conversation about scoring high points subsided and they were quiet. As the new game started, Frank walked from the side to the front of the game screen to join them and shifted the conversation towards one about scientific understanding rather than high scores:

> Frank: How many degrees is it?
>
> Amy: 19 there *((points to temperature meter on the left where the house is))* and it is 26 there *((points to the temperature meter on the right))*, no 23, 24, now it is almost as hot *((keeps her finger pointing at the temperature meter on the right as it changes))*
>
> Frank: What is it that happens, really?
>
> Amy: I have no idea.
>
> Frank: What is that installation there? *((points to the simulation of the interior of the heat pump dynamically moving))*
>
> Amy: It is cold there, isn't it?
>
> Mary: *((walks from her side to the other side with Frank and Amy))* Here is heat pump. When this here gets colder, this needs to be warmer. *((puts one finger on the top and one finger on the bottom of the right blue chamber of the heat pump simulation))*

The conversation between Frank, Amy, and Mary proceeded to be about understanding how the heat pump scientifically works for the remainder of the game cycle, taking on a mastery goal orientation. During the post-visit interview, when shown the clip of this interaction, Frank said he did not understand how the heat pump worked and therefore asked others to decipher the

simulation with him. Frank engaged in *more directed* and *social* information seeking in this episode.

Design feature #4: Ease of seeing other players' screens. The screen for the heat pump game was large and slanted because it was designed to function as both an individual and a group activity during which several visitors surrounding the heat pump game could see the screen. At one point, Neil left his game screen and stood by the one on the other side to watch his classmates play. When shown this video clip and asked to explain what was going on during the post-visit interview, Neil said that he wanted to know if they had other strategies of playing the game and if they used the handprints in a helpful manner. This vicarious learning behavior of gaining knowledge from the gameplay of others is an example of a *one-way social interaction* and *more directed* information seeking. Through this action, Neil said that he determined that the other team was not using the handprints either and decided that the prints did not have an important role. In Group B, Linda also pursued vicarious learning and said in the interview that she declined when asked whether she wanted to play the game because she did not want to look stupid. This behavior aligns with her endorsement of a high performance-avoid orientation in her self-report. Rather than participating by taking an active role in playing the game, she opted to learn through others' gameplay and their mistakes.

Design feature #5: Instructional aims. The written instructions for the heat pump game oriented students to a goal of keeping the appropriate temperature in the house: *Use the heat pump to keep the house temperature at about eighteen degrees throughout the year.* Carly, Linda, and George were very fixated on the numbers and sought information on the easiest way to keep the house within the optimal temperature to score well in the game. Although Carly endorsed a moderate mastery goal orientation in conjunction with her high performance-

approach orientation, in her interview, she said that since the goal of the game was to get a high score, she did not mind not understanding the science to get a high score given the context. As indicated by their behaviors and talk, those three students were able to recognize a pattern between the meter, calendar, and crank direction that allowed them to get high scores without paying attention to the scientific simulation of the inner workings of the heat pump.

Towards the end of the heat pump exhibit visit, students were asked to make a video using an iPod that displayed another set of instructions: *Try the heat pump game. Describe what you see and feel. Take a short video and some pictures where you try to explain why the house gets warmer and colder despite temperature changes there.* These instructions were intended by the team to guide students towards adopting a mastery goal of understanding how the heat pump works and making connections between the different elements of the game such as the crank, heat pump simulation screen, and handprints. After Carly read the iPod directions, these interactions followed:

George: I do not know why. It's just like that.

George: That's it?

Carly: Wait.

George: The heat pump is made to do it.

The act of needing to record a video prompted them to seek a deeper understanding of how the heat pump functions and to reflect on the various components of the game, resulting in *more directed, social* information seeking:

George: Wait a minute. I just have to think a bit about what to say first.

Carly: Say that when we cranked this *((pointing at the physical rank)).*

George: Yes.

Carly: And tried to keep the inside temperature stable.

George: Yes.

Carly: We noticed according to season it was colder and then it was warmer.

George: And we had to adjust ourselves to the environment in order to get a stable temperature inside.

Carly: Yes.

Carly: Because, yes.

George: Yes.

Carly: Because since it gets so hot in the summer this *((pointing to inside the thermometer))* will be warmer but you try to make it colder. We need it to be cooler.

The students continue the discussion and read the text on the walls to develop their iPod video. These interactions show that even though students may endorse a particular goal at one point in the visit, the instructional text in environment can reshape these goals.

Discussion

Understanding the ways in which the design of an exhibit at a science museum can impact learners' experience requires recognizing the types of interactions that the design affords. Learners' behaviors are influenced by both the environment and their personal motivation—some come in with goals while others do not but later adopt goals during the visit. Some change those goals during their visit. Documenting visitors' goals, the ways they pursue those goals, and the consequences of those pursuits provides insight as to why a visitor may stay engaged or leave an exhibit.

Our paper focused on students who endorsed different types of goals for science learning to cast a wide net to capture a variety of information-seeking behaviors. Much of the previous

research on help seeking and information seeking have been conducted in the context of classrooms and Internet searches, and categorizations from that work were not applicable our context. We derived a typology for the information-seeking behaviors that emerged at the science exhibit and categorized the observed behaviors along the two behavioral dimensions of: 1) social to non-social; and 2) more directed to less directed. Our hope is that this typology can inform future research on visitor engagement with games and at science museums as well as guide future research about adaptive and non-adaptive information-seeking behaviors and how to design for adaptive paths.

We uncovered five design features that were associated with information-seeking behaviors: 1) short, low-stakes game cycles; 2) start of a new game cycle; 3) close proximity of games; and 4) ease of seeing other players' screens; and 5) instructional aims. Several notable patterns emerged in our detailed analysis of the ways in which these design features interact with student goals to influence students' information-seeking behaviors.

First, a design feature is not inherently "good" or "bad" and can have both consequences. For example, the short, low-stakes game cycles were very conductive to experimenting and mastery-oriented learning. However, conceptual discussions about the display of the scientific simulation of the heat pump were hampered by the short cycles that transition away from the simulation display back to the start screen. Second, students with different goals are able to adapt design features to their needs. For example, the ease of seeing other players' screens was beneficial to Neil, who endorsed a high mastery goal orientation, because he could watch others to learn from new techniques they used that he had not considered in his own gameplay. For Linda, who endorsed a very different goal orientation of being high performance-avoid, that same design feature allowed her to watch others and learn from their mistakes and the game

feedback when she opted to not play the game herself for fear of getting a low score. Finally, the design features of instructional aims influence shift in goals, such as in the case of the two different instructional texts that resulted in a shift from a conversation between Carly and George about high scores to a conversation about the scientific process of the heat pump.

Understanding the ways these design features affect students' information-seeking behaviors allows for data-driven conversations about how one can redesign exhibits to be more effective. For example, to address the previously mentioned short game cycles cutting off productive conversations, one could create two modes: one being the described timed game and the other being an untimed simulation for experimenting with the components of the heat pump and pausing, replaying, and discussing the animation in depth. In another case, one disadvantage of the ease of seeing other people's screens, based on our analyses, may be that educators may want students such as Linda to participate and learn more actively, rather than only vicariously, due to the students' fear of failure. The untimed simulation mode, therefore, may be a productive stepping stone for such students to become more comfortable with being involved because no scores are displayed. Finally, we found that students were able to get a high score by gaming the system. For example, Carly and George memorized which direction to turn the crank during the different months, illustrating that future versions of the exhibit need to better integrate the science learning with the game mechanics such that understanding the science would be a stronger contingency for receiving high scores.

Our study also provided insight about the mismatch between designers' intentions behind various design features and the actual experiences of students. A helpful way of understanding this is through the notions of inscription and translation. Broadly, inscriptions are the intentions behind the design of the game and translations are how the users actually perceive and use the

game features. Latour (2005) uses the term *inscription* to describe how researchers, designers, and curators inscribe certain features in the construction of the material environment in order to facilitate human action. For example, our team intended to include multiple areas of focus (e.g., simulation, temperature meter, handprints) for learners to choose the resources that they decided most aligned with their goals. To assess how those intentions unfolded, we analyzed students' *translations* (Latour, 1987) of the design features to understand the processes through which they selected and used the material resources in ways that were relevant to their needs and goals. In our case, the competing features that were inscribed into the design affected the group dynamics as students needed to wait for opportunities to shift the focus of the conversation, such as during a new game cycle. In some cases, students waited for opportunities to move out of one conversation that was a mismatch with their goals and into another that was more related to their goal. Being cognizant about the relation between inscriptions and translations would help designers and educators understand their misconceptions about how certain features of a learning environment are related to information-seeking patterns and goal regulation, allowing for information to guide revisions of a current exhibit as well as future development.

We hope that the insights that emerged from our analyses illustrate for researchers and designers how motivation theories provide a helpful lens for understanding the nuances of student engagement, and in particular, the relations between goals, design, and information-seeking behaviors. The process of designing an educational game or exhibit, or any learning environment, is not an exact science. Accordingly, this article is not prescriptive. It is our hope that others will be able to use our findings to extract ideas about which design features will and will not work in their contexts, to reason about the potential impacts of other design features, and to identify leverage points where design tweaks can positively impact learners' experiences.

References

Allen, S. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education*, *88*(S1), S17-S33.

Anderman, E.M. & Wolters, C.A. (2006). Goals, values, and affect: Influences on student motivation. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 369-389). Mahwah, NJ: Erlbaum.

Bevan, B., Dillon, J., Hein, G.E., Macdonald, M., Michalchik, V., Miller, D. et al. (2010). *Making science matter: Collaborations between informal science education organizations and schools.* Washington, DC: Center for Advancement of Informal Science Education.

Bjørnebekk, G., & Diseth, Å. (2010). Approach and avoidance temperaments and achievement goals among children. *Personality and Individual Differences*, *49*(8), 938-943.

Conley, A.M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*(1), 32-47.

Dweck, C.S., & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Elliot, A.J., & Church, M.A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*(1), 218-232.

Elliot, A.J., & McGregor, H.A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501.

Harackiewicz, J.M., Barron, K.E., Carter, S.M., Lehto, A.T., & Elliot, A.J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*,*73*(6), 1284.

Harackiewicz, J.M., Barron, K.E., & Elliot, A.J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist*, *33*(1), 1-21.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundation and Practice. *Journal of the Learning Sciences*, *4*(1), 39–103.

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games.* Cambridge, MA: MIT Press.

Kaplan, A., & Maehr, M.L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19*(2), 141-184.

Karabenick, S.A., & Sharma, R. (1994). Perceived tutor support of student questioning in the college classroom: Its relation to the student characteristics and role in the classroom questioning process. *Journal of Educational Psychology, 86*(1), 90-103.

Latour, B. (1987). *Science in action*. Cambridge, MA: Harvard University Press.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory.* Oxford, UK: Oxford University Press.

Levenson, R.W., & Gottman, J.M. (1985). Physiological and affective predictors of change in relationship satisfaction. *Journal of Personality and Social Psychology*, *49*(1), 85-94.

Marchionini, G. (1995). *Information seeking in electronic environments.* Cambridge, UK: Cambridge University Press.

Newman, R.S., & Schwager, M.T. (1993). Student perceptions of the teacher and classmates in relation to reported help seeking in math class. *Elementary School Journal, 94*, 3-17.

Schenke, K., Lam, A.C., Conley, A.M., & Karabenick, S. (2015). Adolescents' help seeking in mathematics classrooms: Relations between achievement and perceived classroom

environmental influences over one school year. *Contemporary Educational Psychology, 41*, 133-146.

Schunk, D.H., Pintrich, P.R., & Meece, J. (2007). *Motivation in education: Theory, research and applications*. Upper Saddle River, NJ: Merrill Prentice-Hall.

Suchman, L.A. (1987). *Plans and situated actions: The problem of human-machine communication*. New York, NY: Cambridge University Press.

Urdan, T. (2004). Predictors of self-handicapping and achievement: Examining achievement goals, classroom goal structures, and culture. *Journal of Educational Psychology, 96,* 251-254.

Welsh, D. P., & Dickson, J. W. (2005). Video-recall procedures for examining subjective understanding in observational data. *Journal of Family Psychology*,*19*(1), 62.

Table 2.1

Typology of information-seeking behaviors

		←Social		Non-Social →
		Two-way interaction	**One-way interaction**	**No interaction**
More Directed →	**Creating opportunities to seek information**	Finding and speaking with a museum educator to ask for an explanation about how the heat pump works	Changing one's location to track different conversations to selectively eavesdrop on ones that are relevant to the information being sought	Testing personal hypotheses of how a heat pump works through connecting the visual with different speeds and directions of cranking
	Taking advantage of available opportunities with the intent to seek specific information	Asking the group why the colors in the scientific simulation changed while pointing to the heat pump chambers on the screen	Watching another student play to learn new techniques to apply them to the next cycle of the game	Reading the text on the exhibit walls and displaying an intention to use that information such as towards recording the iPod video
← Less Directed	**Exploratory behaviors**	Asking "what is going on?" without being specific about the point of confusion	Watching another student play without displaying an intention to gain or apply specific knowledge	Reading the text on the exhibit walls without displaying an intention to gain or apply that knowledge

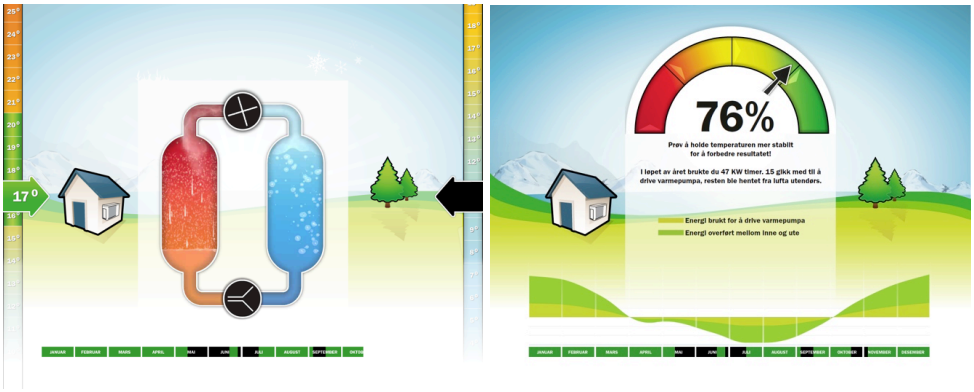

Figure 2.1. Heat pump game during play (left) and the display of the score at the end (right)

Figure 2.2. Heat pump simulation game setup with the crank and the heated handprints

Figure 2.3. Heat pump simulation game setup at the museum with two identical games back-to-back

CHAPTER 3

Study 2: Understanding Goal Change: Theoretical Advances and Leverage for Design

How do we account for the goals that students employ in a learning task to design for an effective learning experience? Part of this answer requires understanding the ways in which 1) goals shape students' interactions with the digital, physical, and social aspects of the task and 2) consequences from those interactions influence students' decisions to maintain or change their goals. Although a large body of literature exists regarding the initial adoption of goals, relatively little research has focused on how and why those goals are maintained or revised over time. Goal change (or lack thereof) deserves more attention because it is a marker of what students would like to gain from their experience and the role that the environment has in that pursuit. The aim of our paper is to first review Achievement Goal Theory (Dweck & Leggett, 1988), highlight the gaps in our understanding of goal change, and discuss methodological considerations for addressing those gaps. Second, we present findings from our study that used video and interview data to research goal change in a specific task and provide insights that extend current theoretical knowledge about achievement goals. Specifically, we present examples that highlight ways in which four students who endorsed different goals regulated those goals while playing a heat pump game during a science museum classroom field trip.

Our work is largely informed by Achievement Goal Theory (Dweck & Leggett, 1988), which focuses on the reasons individuals are motivated as well as the standard against which they judge their success. Researchers have distinguished two achievement goal orientations toward learning: mastery and performance. A mastery goal orientation focuses on developing and mastering skills and knowledge whereas a performance goal orientation focuses on appearing competent by either outperforming others (performance-approach) or by avoiding

appearing incompetent (performance-avoidance) (Urdan, 2011). The heat pump game that forms the context of our study is particularly appropriate for studying achievement goals because it provides affordances for both mastery goals and performance goals. For instance, students can pursue mastery goals of understanding the scientific mechanisms that underlie the functioning of the heat pump through the digital simulations of the inner workings of the heat pump (Figure 3.1, left). As players rotate a physical metal crank right or left, the simulation shows the resulting changes in the pressure and flow of heat in the depiction of the interior of the heat pump displayed on a digital screen. To pursue performance goals, students can focus on the game scores that reflect how successfully the players were able to operate the heat pump to keep a house at an optimal temperature throughout the year. Because there were two games occurring simultaneously and in close proximity to each other, students had opportunities to compare scores. These different game features allowed us to gain insight into how students use the exhibit in different ways to align available resources to their specific goals as well as how the consequences of their behaviors influence their subsequent goal pursuits and changes in goals during the game.

Students typically endorse multiple goals, though oftentimes one of those goals is more strongly endorsed than the others (Barron & Harackiewicz, 2001; Conley, 2012; Pintrich, 2000). Having stronger mastery goals has been linked to positive outcomes, including increased cognitive engagement, deeper cognitive strategies, and greater interest in the subject (Ames, 1992; Dweck & Leggett, 1988; Harackiewicz, Barron, Tauer, Carter, Elliot, 2000; Urdan, 2011). Having strong performance goals has varying consequences with an important distinction being whether students adopt a performance-approach focus on appearing highly competent or a performance-avoidance focus on not appearing dumb. For instance, performance-approach

students may want the highest score in the class whereas performance-avoidant students may not participate in discussions out of fear of embarrassing themselves with an incorrect answer. Research has generated mixed results on the associations between achievement and the performance-approach goal orientation, indicating that such orientations are linked to unrealistic high goals in some cases but also to high achievement in others (Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000; Kaplan & Maehr, 2007). Performance-avoidance goals, on the other hand, have robust associations with negative consequences such as procrastination, putting in little effort, and cheating (Elliot & Church, 1997; Harackiewicz, Barron, & Elliot, 1998; Urdan, 2004). These associations have been predominantly drawn from survey studies about motivation at a large grain size (e.g., for science class). As will be illustrated in our case studies, however, there are exceptions to these broader findings that can be explained by measuring motivation at a smaller grain size (e.g., at several points within a specific task) to allow for a more nuanced understanding of the relation between design, student-environment interactions, and goal change.

The context of our game—given its quick feedback cycle, replay opportunities, and multiple non-social and social resources—is particularly appropriate for studying the nature of goal change. Goals are dynamic and represent a form of self-regulation (Bandura, 1986; Locke & Latham, 1990), and optimal self-regulation entails not only endorsing goals prior to task engagement but also monitoring goal pursuit by evaluating goal progress and contemplating the need for goal revision (Shah, Kruglanksi, & Friedman, 2003; Wrosch, Scheier, Miller, Schulz, & Carver, 2003; Zimmerman, 1989). This revision process may be prompted by additional details about the task, available resources, and an evaluative environment. For example, information acquired from additional experience with the task (e.g., difficulty level, usefulness of feedback, competition level) may lead students to adjust their goal endorsement accordingly (Bong, 2005).

Researchers and designers alike need to better understand this process as it is expected that the design of a learning environment would interact with students' personal goals to influence behaviors and goal change.

Goal maintenance and change: What we do and do not know

A significant amount of theoretical and empirical work from various perspectives has focused on understanding the processes by which goals lead to behavior wherein individuals continuously adjust effort and strategies to maintain progress toward their goals (Carver & Scheier, 1998; Latham & Locke, 1991; Locke & Latham, 2002). Goals themselves are not static entities—for example, individuals switch from one goal to another over time (Fryer & Elliot, 2007) and revise their personal performance standards up and down as they are pursued (Bandura, 1991). Although a significant body of research shows that goals do change and that goals influence behavior, relatively little research has focused on why those goals are maintained or changed over time during task engagement. One exception is work by Tolli and Schmidt (2008) in which they conducted experiments to show that feedback, self-efficacy, and attribution influence students' self-set goals of the percentage of anagram tasks that they wanted to correctly answer each round. In particular, the researchers found that for those who attributed successful performance to internal factors, positive feedback resulted in a boost of self-efficacy, which in turn, resulted in upward goal revision. In contrast, those who attributed success to external factors did not experience this increase in efficacy, and self-set goal difficulty remained largely unchanged as a result.

For achievement goals in particular, the extant literature on goal stability and change has been limited in two important ways. First, many have examined shifts in goal endorsement for school across the elementary to middle school transition (e.g., Anderman & Anderman, 1999;

Anderman & Midgley, 1997) or shifts in goal endorsement for school within a school year (e.g., Bong, 2005; Conley, 2012). However, few have addressed stability and change in achievement goals across a specific task (for exceptions, see Bernacki, Aleven & Nokes-Malach, 2014; Fryer & Elliot, 2007; Senko & Harackiewicz, 2005). Second, achievement goal stability and change has been primarily investigated using changes in group means rather than individual differences (for exceptions, see Bernacki, Aleven & Nokes-Malach, 2014; Conley, 2012). Addressing those limitations, we continuously examined students' goal pursuit intensity and the type of goal endorsed for the full duration of their time spent with the heat pump game at the museum. This allowed us to assess ways in which students' interactions with the design of a specific task influence goal change, providing an understanding of the mechanisms that underlie goal change as well as an understanding of points of leverage that designers may have to positively influence students' goal regulation.

In our conceptualization of goal change, we view achievement goal adoption as not needing to be an all or none affair, and individuals can endorse many or few and can endorse each goal at varying levels of intensity. Goal change may therefore be construed in terms of a shift in degree of goal endorsement. Furthermore, as noted by Senko and Harackiewicz (2005), goal change may take place in two different ways: It may represent *goal intensification,* in which an individual increases or decreases commitment to a goal, or it may represent *goal switching,* in which an individual shifts commitment from one type of goal to another. Therefore, in analyzing goal maintenance and goal change, we consider both forms of goal change, acknowledge multiple goals, and conceptualize the endorsement of goals as a continuous rather than a discrete variable.

Aligning methods with goal revision theories

To effectively study goal maintenance and change, we must first align our methods of measuring goals with the goal revision theory. Current measures of achievement goals, which predominantly rely on self-reports administered a few times a year, are appropriate for measuring goals at a larger grain size. For measuring task goal change, however, it is necessary to use measures that are specific to the grain size of the task and that are more frequently employed. Therefore, it is necessary to adopt methods that better align with our theoretical assumptions about ways in which students self-regulate during complex learning tasks. Self-regulatory activities such as planning, monitoring, strategy use, and adaptation all form a dynamic system in which all components influence one another (Zimmerman, 2011). By using methods that allow us to measure goals and behaviors repeatedly or continuously, we are able to more precisely measure this dynamic process to understand how goals fluctuate in response to learner characteristics and the environment.

Several methodological trends have emerged for measuring goal change in ways that acknowledge that past experience may influence individuals' propensity to act a certain way but that the way behaviors manifest is also situated. For instance, experience sampling methods used to collect systematic data about what a person does, thinks, and feels during daily life can allow for continuous longitudinal data collection. Traditionally, participants carry a pager over the course of several days or weeks, and the pager beeps several times a day to elicit a report about their experience. With the omnipresence and popularity of cell phones, it is now possible to gather data over a longer period of time because participants are not required to carry an additional device. As another example, researchers have integrated frequent survey-item measures of achievement goals in educational software and used data mining to extract behavioral correlates (Bernacki, Aleven & Nokes-Malach, 2014). By logging students' behaviors

using learning software and frequently assessing goals in context, we can better understand learner's task-specific motivations, their influence on learning behaviors, and their relationship to task performance and goal change.

Our case study highlights a third method for continuously measuring goals: video data supplemented by interviews. Specifically, we coded information-seeking behaviors and traced the types of conversations that students engaged in or disengaged from as indicators of their achievement goals. For example, a student may seek information to understand the scientific mechanisms of the heat pump and disengage from a group conversation that is focused on comparing game scores, indicating a pursuit of a mastery goal. The advantage of measuring achievement goals through video data is that one does not interrupt the process of learning and that one is able to analyze the behaviors before and after goal pursuits to extract information on how the environment may have influenced goal maintenance and goal change. Fredricks and McColskey (2012) noted that a distinct advantage of observational methods is that they can provide rich description of different levels of engagement in the learning context. As illustrated in our case study, in addition to expanding our understanding of goal pursuit, this smaller grain size of analysis also provides details about specific instances in which designers may have leverage to influence learners' goal regulation.

Methods

Context of learning environment

Data presented were collected during a science museum visit during which high school students in Norway explored an interactive heat pump game booth designed by our university lab and the museum. Players learn about the general function of a heat pump, which moves heat from inside to outside and vice versa through processes of condensation and evaporation. Players

are challenged to keep the house temperature consistently warm throughout the year by operating the heat pump through physically rotating a metal crank underneath the screen in the appropriate direction using the appropriate speed (Figure 3.2). As students manipulate the heat pump compressor using the physical crank to heat up or cool down the house, the heat pump's inner workings dynamically move in real time in the middle of the screen (Figure 3.1, left). Those visualizations are linked to player's operation of the metal crank; the two chambers in the simulation react to clockwise and counter-clockwise operation of the crank, illustrating conditions of boiling and condensation. The result screen at the end of the game shows the percentage of time the house stayed within the desirable zone of warmth (Figure 3.1, right). Each game cycle is approximately one minute, and each student in this study visited the heat pump exhibit for about 15 minutes, as part of a group of five students.

Participant selection

The four key participants in our case study were chosen from 32 first-year students (50% male, ages 15 and 16) in one science classroom at a moderately selective high school in Oslo, Norway. We administered a nine-item Achievement Goals Questionnaire (Elliot & McGregor, 2001), adapted to focus on achievement goals during science class. Items assessing mastery, performance-approach, and performance-avoidance goal orientations were scored on a 7-point Likert scale ranging from 1 (*not at all true for me*) to 7 (*very true for me*). There was acceptable internal consistency for the performance-approach ($\alpha = 0.81$), performance-avoidance ($\alpha = 0.76$), and mastery scales ($\alpha = 0.83$). The self-reported responses allowed us to identify students who endorsed different levels and types of goals in order to assess how the design features of the exhibit related to interaction patterns across students with different motivational profiles.

Our selected data sample includes the interactions of five students in one group and of three students in a second group. Though both groups contained five members, we focused on the four target students selected based on their self-reported motivation profiles and those who interacted with them:

1) *Neil* reported **predominantly a mastery goal orientation** (average rating=6.7) and **low performance goal orientations for both approach** (average rating=2) **and avoidance** (average rating=1.5).
2) *Frank* reported **similarly high goal orientations** across mastery (6.67), performance-avoidance (5) and performance-approach (4.7) goal orientations.
3) *Carly* reported a **predominantly performance-approach goal orientation** (6) and **moderate performance-avoidance** (4.5) **and mastery** (4.7) goal orientations.
4) *Linda* reported **high performance-avoidance** (5.5) and **high mastery** (6) goal orientations and a **low performance-approach** (3.3) goal orientation.

Video analyses

Our analysis of video recordings of two groups of students interacting with the heat pump exhibit followed the principles of interaction analysis (Jordan & Henderson, 1995; Suchman, 1987). Interaction analysis emphasizes the patterns of the interaction among individuals and the objects of their environments, noting which resources and conversations get taken up and how. This includes talk, nonverbal interaction, and the use of physical (e.g., heat pump crank) and digital (e.g., heat pump screen animation) artifacts. Our aim was to uncover activity patterns that emerge as the interaction analysis proceeds to identify routine practices and problems within the game and the resources for their solution. We looked at groups of students (that included our target students) interacting with back-to-back heat pump games (Figure 3.3). The design of the

players facing each other allows for players on opposite sides to hear each other's conversations and move back and forth. Therefore, individuals can pursue their goals by weaving in and out of conversations that align and misalign with their goals. Each group spent about 15 minutes at the heat pump game and all interactions with our four target students were included in the analyses.

Project staff first prepared transcripts of utterances from the recorded footage. The first and second authors then viewed the videos and read the transcripts repeatedly to formulate tentative assertions. In subsequent viewings and transcript readings, we refined the transcriptions of information-seeking episodes to include other interaction factors including gestures, pauses, and overlaps of utterances. In cycles of interpretation of particular episodes and testing of categorical constructions of information seeking, we discarded, modified, or adopted emerging understandings. Interaction analysis aligns with our goals to identify how design features and the social environment interact to support or hinder the pursuit of different types of goals. Patterns of behaviors emerged as we identified key activity episodes that promoted or inhibited goal pursuit and analyzed the interactions prior to, during, and after those episodes to determine how particular game design elements facilitated information-seeking behaviors and how patterns of information-seeking behaviors relate to student goals.

Interviews

Interviews were conducted with students one-on-one with the second author about three weeks after the museum field trip, allowing sufficient time for the results from the video data to guide the interview structure. The semi-structured interviews lasted about half an hour per student and were video recorded. This session included an interviewer showing two short video clips, less than one minute each, of the interviewee's behaviors at the heat pump game. After each video episode, the interviewer asked the students to explain why they behaved as they did if

they are able to recall those reasons. Though video recall has not been used with high school students and in our context, previous research has supported the validity of video-recall procedures by simultaneously recording a variety of physiological measures during the original conversation and again during the video-recall procedure. The physiological data during the recall session were significantly related to participants' physiological data during the original interaction, suggesting that participants were "reliving" their experiences (Levenson & Gottman, 1985).

The video clips used during the interview were flagged as ones that lacked sufficient details, based on our interaction analyses, to include or rule out as information-seeking behaviors. For example, in one simulated video recall episode, a student left his game to watch another student play the same game for about five seconds before returning to his game. There were no actions prior to or after that episode that provided details for identifying if he was seeking specific information. There was no talk, as he simply went to the other side to watch when his game cycle ended and returned to his side to start a new game. When shown this video clip during the interview, the student indicated that he wanted to know the purpose of the physical handprints and if the other group had strategies for incorporating them into their gameplay. Because the other group also did not use the handprints, the student decided that they were not important and continued to not use them after returning to the game on his side. This supplementary data from the interview allowed us to document the clip as an information-seeking behavior.

Findings

We traced how the social environment and digital and physical resources of the exhibition influenced students' information-seeking behaviors and how the patterns of such

behaviors reflected the stability and change in their goals. In the case illustrations below, we highlight extensions to the achievement goal literature based on these analyses. Deriving these relations also provide insight on how game design elements function interactively with learners' goal orientations to better equip us with knowledge about the strengths and limitations of certain design elements with regard to allowing learners to adaptively regulate their goals.

Mastery goals do not always lead to adaptive outcomes

Although the endorsement of mastery goal orientations has been consistently linked to positive learning-related outcomes (Ames, 1992; Dweck & Leggett, 1988; Harackiewicz, Barron, Tauer, Carter, Elliot, 2000; Urdan, 2011), we found instances in which such orientations are not adaptive. In our case studies of two students who endorsed high mastery goal orientations for science class, we found that they either disengaged from their pursuit of their mastery goal or persisted in that pursuit and left unsatisfied because the exhibit did not provide them with the resources that they needed to understand the science behind the heat pump. Although both students entered with similar goals and were in the same group with the same resources, they left with very different levels of mastery goal orientations.

Neil maintained the intensity of his mastery goal pursuit but left unsatisfied. Neil stated in his interview that his goal for the science museum visit was "to learn as much as possible. You are not going there to play, but it is to explore whether what I have learned is correct." This goal was clearly reflected in his interactions with the heat pump game and with group members. As his group focused on memorizing crank rotation patterns to achieve high scores while cheering and laughing, Neil stood on the side in between the two games and did not participate until he overheard the word "evaporation." He chimed in to ask who mentioned that word. Then as the game progressed, Neil pointed to the dynamic heat pump simulation in the middle to redirect his

group's attention and to steer the conversation towards scientific understanding. Another student responded, explaining what she thought caused the heat pump to generate cold temperatures. This conversation only retained its mastery orientation for about a minute, however, because the result screen appeared. The mastery conversation between Neil and his classmate came to an abrupt stop as the dynamic visual of the inner workings of the heat pump disappeared. This influenced the other members of the group to discuss how high of a score they received, and Neil disengaged. As the group was finishing their rotation at the heat pump game, Neil responded with, "I still don't understand how it worked," to which another student responded with, "We are done," as the team left.

Frank lowered his pursuit of mastery goals. In his interview, when asked about his goal at the museum, Frank stated that it "was about learning more" but, contrary to Neil, he also added that "often it is quite fun to visit the museum." In an attempt to understand how the heat pump works, Frank used physical, digital, and social resources to seek information. For example, he pointed to the simulation in the middle of the screen (Figure 3.1, left) and asked his group about what is going on and followed up with questions about specific parts of the simulation. Although his group ignored him and focused on the temperature meters on the side with the aim to crank for a high score, a student on the other side peered around the corner briefly to share her hypotheses of what the simulation was conveying. Frank then left his current game and moved to the other side to engage in the scientific conversation with the classmate during her game. However, that conversation ended when the game ended. After multiple unsuccessful attempts to understand the heat pump simulation, Frank disengaged from his mastery goal but did not switch to performance goals; rather, he chatted with others while laughing and singing. Even when a

museum staff came by to explain the heat pump at the end of the visit, Frank did not listen and continued to chat with other classmates.

Context shapes initial goals and influences goal change

Our interviews revealed that students do not always enter an achievement task with a goal. Furthermore, the context of the environment such as the type of activity and the way instructions are framed influenced changes in goals.

Carly switched between performance-approach and mastery goals. When asked during the interview about her goal for the museum visit, Carly said that she did not really have one and it was just to see what was there. Although Carly reported to be predominantly performance-approach oriented but also moderately mastery-oriented towards science class in the survey measure, there was no indication of a mastery goal orientation at the outset of her visit to the heat pump exhibit. This was evidenced by her lack of attention to the scientific simulations and by her conversations focused on comparing scores and generating high scores. In her interview, Carly explained that since the goal of a game is to get a high score, her focus was not on understanding science, but rather, she focused on any means to get a high score, specifying "I want to do well. Prefer to be better than others." Carly's group recognized a pattern between the digital temperature meter, digital calendar, and physical crank direction that allowed them to get high scores without paying attention to the scientific simulation of the inner workings of the heat pump that was at the center of the screen.

Towards the end of the heat pump exhibit visit, however, Carly switched to a mastery goal orientation. Students were asked to make a video using an iPod that displayed this set of instructions: "Try the heat pump game. Describe what you see and feel. Take a short video and some pictures where you try to explain why the house gets warmer and colder despite

temperature changes there." These instructions were intended by the design team to guide students towards adopting a goal of understanding how the heat pump works and making connections between the different elements of the game such as the crank, heat pump simulation screen, and handprints. After Carly read the iPod directions, she paused and reflected on her team's actions to figure out how to explain the simulation. She and her group also read the text on the walls to develop their iPod video. These interactions illustrate that even though students may endorse a particular goal at one point in the visit, the instructional text in environment can reshape these goals.

Performance-avoidance goals are not always damaging

Despite the robust findings in the achievement goal literature about the detrimental effects of performance-avoidance goals, we found that there are exceptions. As an example, we illustrate the interactions and thoughts of a student who adaptively aligned the environment's resources with her preference for both performance-avoidance and mastery goal orientations.

Linda maintained her performance-avoidance goal while vicariously participating. While watching a video of herself refusing to play, Linda explained during an interview that she declined when a classmate asked whether she wanted to try the heat pump game because she did not want others to see her get a low score. However, she vicariously participated with the aim to learn by watching her group interact with the game. The large, slanted screen at the heat pump game was adaptive for Linda's goal because it allowed for several visitors to simultaneously view the screen. This opportunity aligned with her high performance-avoidance orientation. Rather than participating by taking an active role in playing the game, she opted to learn through others' gameplay and their mistakes. Linda did, however, later try the game herself when no one was around. In her interview, she said she was thankful no one appeared because her score was

not very high. Linda, who endorsed both high performance-avoid and high mastery goal orientations, made the two seemingly incompatible goal orientations work for her as she clearly exhibited performance-avoidance behaviors but also had one of the best understandings of the heat pump mechanisms in her science class as evidenced by a test.

Allowing Linda to exercise her goal preference, as this exhibit did, may have been more beneficial than designing in a way that is intended to shift her goal to another orientation. It is quite possible that she would have otherwise disengaged entirely because it difficult to change avoidance-based goals (Fryer & Elliot, 2007). Since those goals are framed in terms of negative possibilities, students either successfully avoid that negative possibility or failure to avoid it. Neither of these outcomes is likely to shift the individual's focus from a negative possibility to a positive possibility to influence them to adopt an approach goal (Fryer & Elliot, 2007).

Discussion

Theoretical implications

Our study extends the theoretical framework on achievement goal revision by using video data to continuously measure behavior and supplementing that with interviews to provide examples of how mastery goal orientations can diminish over time. Much of the work on achievement goal theory shows consistent positive outcomes that result from mastery goal orientations such as better cognitive strategy use and greater persistence (Anderson & Wolters, 2006). Work on the precursors of such orientations have focused on elements such as the classroom environment, the types of praise that students receive from parents, and peer influence (e.g., Ames, 1992). What has been sparse in this work, however, are the factors that influence not just the adoption of mastery goal orientations and the behaviors that they initiate but also the *maintenance* of mastery goal orientations over time (Senko & Hulleman, 2013).

Within the goal revision realm, Dweck and Elliot (1983) posit that in light of the outcomes that students receive as they progress towards a goal, students may revise their expectancies for whether or not they can succeed based on changes in their perception of the stable, uncontrollable, and global nature of the outcomes, in line with the attribution theory (Weiner, 1972). Much of the work on goal revision, however, has looked at performance goals. Dweck and Elliot (1983) contended, "Such expectancy revision, particularly downward revision, is most likely to occur with performance goals" because they are "fostered by and, in turn, foster views of ability as a stable, uncontrollable, global factor." Students may, for example, switch between performance-approach and performance-avoidance goals as a result of information about their level of competence. Mastery goals, however, they argue, are more likely to reflect ability being viewed as specific and acquirable and encourage more effort to attain goals.

What is missing from this view, however, is the acknowledgement that effort does not always pay off and does not always lead to progression towards attaining goals for mastery-oriented students. The situated context can hinder the pursuit of mastery goal orientations when the resources available are not adequate for students to reach their learning goals or to gauge if they are progressing towards it. As illustrated in our study (Neil), it is not simply expectancy of one's ability to learn and perform that matters but also one's expectancy that the necessary resources to overcome obstacles are present. In the heat pump game, mastery goal pursuits were not successful because the resources did not help students progress towards mastering their understanding of how the heat pump works.

Our study shows that the mastery-oriented students employed great effort in learning the material and tried different strategies as they attempted to engage their peers in mastery-oriented conversations, experimented with the game, and participated in learning-related conversations.

However, each interaction left them unsatisfied, and their expectancies for the exhibit having the resources that they needed to attain their goals diminished. Dweck and Elliot (1983) state that, "Revision of expectancies plays a critical role in the maintenance or abandonment of achievement strivings. For example, the tendency to exhibit too-rapid expectancy decreases in the face of obstacles appears to be a key factor in certain maladaptive achievement patterns." As we have seen, technology and games provide a greater risk of producing "too-rapid expectancy decreases" given their quick interaction and feedback patterns, and in our heat pump game, this occurred. Focusing on goal attainment expectancies allows us to gain a deeper understanding of the ways in which students adjust their goal pursuits after setbacks.

Design implications

Our findings illustrate the need for considering students' goal orientations when designing educational exhibits, particularly for science. The two students in our study who reported high levels of mastery goal orientations for learning in science classrooms reported that their goal extended to their visit at the science museum, which was part of a school field trip. This motivational profile to want to develop knowledge and understand is extremely adaptive, yet it was these types of students that the exhibit managed to disappoint.

The museum exhibit was not designed to provide the resources the students needed to attain their goals of understanding the mechanisms that underlie how the heat pump works. The intent of the heat pump game, according to designers, was to provide students with the general understanding that the heat pump, as a single device, is able to both heat up and cool down a house while using significantly less energy than other alternatives. This was reflected in the game's purpose to operate the heat pump in a way that keeps the house at a comfortable temperature in both the winter and summer and the results screen showing a graph of the amount

of saved energy. The animation of the inner workings of the heat pump was intentionally made to be vague by the designers to introduce students the idea of energy transfer occurring when the heat pump is active without overwhelming them with scientific details. The assumption was that students would enjoy the game even without knowing all the scientific details as long as they were able to understand enough about the heat pump to perform well in the game. The other components of the heat pump were intended to catch their attention and be relatable for their later lessons on energy transfer. Additionally, the animation could be used by a museum educator or teacher to explain the details of the heat pump, but the lack of embedded explanations was not expected to be problematic.

As such, during iterative user testing sessions for the game, the focus was on leveling the difficulty of the game to make sure it was not too hard or too easy. What was missing however was ensuring that the game was appropriately scaffolded for students' learning goals to develop understanding. Had that been done, the team may have realized that some students wanted to know more details about what was happening in the animated inner workings of the heat pump and that was more important to them than attaining a high score. These goals were very important for the two mastery-oriented students in our study and could have been remedied before the launch of the exhibition at the museum. However, we should note that it is possible that the issues of not being able to learn how the heat pump works despite constant effort may have not emerged from user testing even if the team had been more cognizant about that issue because the user testing was done in a lab setting. Students were recruited to come into a university lab to test out the game, so they may have not brought with them the same expectations to learn that students did in our study when they took a school field trip to a museum. Regardless, this study shows that even in an informal setting, in the context of a school

field trip, there are students who are focused on learning and particularly want to understand the underlying science. It is concerning that the students who entered the exhibit with a strong desire to learn showed a diminished intensity of mastery goal pursuits over time and left disappointed.

Part of the confusion with the heat pump game was the students did not understand what was happening in the animation of the inner workings of the heat pump. There were several conversations, for instance, about what the changes in colors meant. Though the goal of the designers was not to explain the micro-level changes of the heat pump game, it was still an issue because the goal of the students was to understand that level of detail. When asked what they would have changed to make the game better, Frank and Neil said they would have included better explanations for what is going on such as by more text in the game itself. This conflict could have been resolved in two ways. The first is to provide more details. There could be different layers of the game in which students can zoom in for more details so that the details would not overwhelm the students who do not want that level of detail at the moment. Furthermore, there could be two modes of the heat pump—one as the currently operating game and another as simply a simulation in which students can experiment and discuss and pause as needed. This would help with facilitating mastery-oriented conversations, which were frequently interrupted as a result of the short game cycle, as discussed earlier. A second option would be to not include extensive visuals in which students are not able to derive more sophisticated meaning. Perhaps the evaporation and boiling in the heat pump animations as well as the changes in color within the chamber should not be there. The hope was that such animation would spark students' interest to later learn more about those topics, but what happened was that students expected to be able to resolve their confusion and develop that additional knowledge

immediately. As such, the graphics could have been simplified to not influence the construction of goals that are not attainable.

Our case illustrations show that students need to feel that they have control over obtaining their goals—goals that are refined as part of their interactions with the learning environment. Though students may enter the exhibit with a pre-established general goal to learn and understand science, their specific goals are influenced by the context at play. To increase students' expectancy to be able to attain their goals, designers must make sure that there are steps that learners can take towards their refined goals. This includes keeping in mind not only how design influences the construction of those goals but also the attainability of those goals.

References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.

Anderman, L.H., & Anderman, E.M. (1999). Social predictors of changes in students' achievement goal orientations. *Contemporary Educational Psychology*, *24*(1), 21-37.

Anderman, E.M., & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology*, *22*(3), 269-298.

Anderman, E.M. & Wolters, C.A. (2006). Goals, values, and affect: Influences on student motivation. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology* (pp. 369-389). Mahwah, NJ: Erlbaum.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248-287.

Barron, K., & Harackiewicz, J. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology, 80*, 706–722.

Bernacki, M.L., Aleven, V., & Nokes-Malach, T.J. (2014). Stability and change in adolescents' task-specific achievement goals and implications for learning mathematics with intelligent tutors. *Computers in Human Behavior*, *37*, 73-80.

Bong, M. (2005). Within-grade changes in Korean girls' motivation and perceptions of the learning environment across domains and achievement levels. *Journal of Educational Psychology*, *97*, 656-672.

Carver, C.S., & Scheier, M.F. (1998). On the self-regulation of behavior. New York: Cambridge University Press.

Conley, A.M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*(1), 32-47.

Dweck, C.S., & Elliott, E.S. (1983). Achievement motivation. In E. M. Hetherington (Ed.), *Socialization, Personality, and Social Development* (pp. 643-691). New York: Wiley.

Dweck, C.S., & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Elliot, A.J., & Church, M.A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*(1), 218-232.

Elliot, A.J., & McGregor, H.A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-519.

Fredricks, J.A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 763–782). New York, NY: Springer.

Fryer, J.W., & Elliot, A.J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology*, *99*, 700-714.

Harackiewicz, J., Barron, K., & Elliot, A. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist, 33,* 1-21.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., Carter, S.M. & Elliot, A.J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*, 316-330.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundation and Practice. *Journal of the Learning Sciences*, *4*(1), 39–103.

Kaplan, A., & Maehr, M.L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19,* 141–184.

Latham, G.P., & Locke, E.A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, *50*(2), 212-247.

Levenson, R.W., & Gottman, J.M. (1985). Physiological and affective predictors of change in relationship satisfaction. *Journal of Personality and Social Psychology*, *49*(1), 85-94.

Locke, E., & Latham, G. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Locke, E.A., & Latham, G.P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57,* 705–717.

Pintrich, P.R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and performance. *Journal of Educational Psychology, 92*, 544-555.

Senko, C., & Harackiewicz, J.M. (2005). Regulation of achievement goals: The role of competence feedback. *Journal of Educational Psychology*, *97*(3), 320.

Senko, C., & Hulleman, C.S. (2013). The role of goal attainment expectancies in achievement goal pursuit. *Journal of Educational Psychology*, *105*(2), 504.

Shah, J.Y., Kruglanski, A.W., & Friedman, R. (2003). Goal systems theory: Integrating the cognitive and motivational aspects of self-regulation. In S. Spencer, S. Fein, M. Zanna, & J. Olson (Eds.), *Motivated social perception: The Ontario symposium* (pp. 247-275). Mahwah, NJ: Lawrence Erlbaum Associates.

Suchman, L.A. (1987). *Plans and situated actions: The problem of human-machine communication*. New York, NY: Cambridge University Press.

Tolli, A.P., & Schmidt, A.M. (2008). The role of feedback, causal attributions, and self-efficacy in goal revision. *Journal of Applied Psychology, 93*, 692–701.

Urdan, T. (2004). Predictors of self-handicapping and achievement: Examining achievement goals, classroom goal structures, and culture. *Journal of Educational Psychology, 96,* 251-254.

Urdan, T. (2011). Achievement goal theory: Definitions, correlates, and unresolved questions. In. S. Jarvela (Ed.), *Social and Emotional Aspects of Learning* (pp. 21-27). Amsterdam: Elsevier.

Weiner, B. (1972). Attribution theory, achievement motivation, and the educational process. *Review of Educational Research, 42*(2), 203-215.

Wrosch, C., Scheier, M.F., Miller, G.E., Schulz, R., & Carver, C.S. (2003). Adaptive self-regulation of unattainable goals: Goal disengagement, goal reengagement, and subjective well-being. *Personality and Social Psychology Bulletin*, *29*(12), 1494-1508.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*, 329-339.

Zimmerman, B.J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B.J. Zimmerman & D.H. Schunk (Eds.), *Handbook of self-regulation of learning and performance.* (pp. 49-64). New York, NY: Routledge.

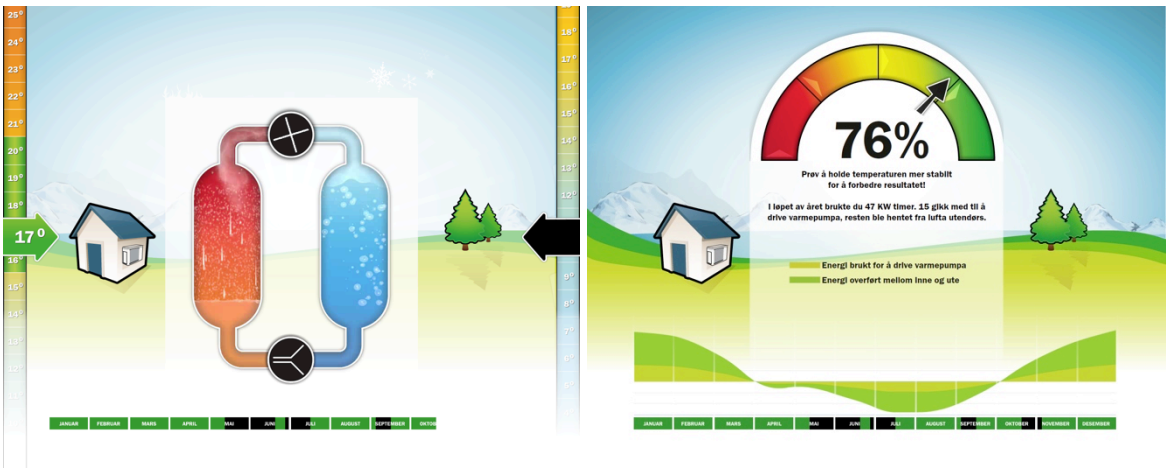

Figure 3.1. Heat pump game during play (left) and the display of the score at the end (right)

Figure 3.2. Heat pump simulation game setup with the crank and the heated handprints

Figure 3.3. Heat pump simulation game setup at the museum with two identical games back-to-back

CHAPTER 4

Study 3: The Effects of Miscalibration on Confusion, Learning Behaviors, and Learning Outcomes

Confusion is an affective state that occurs when learners confront breakdowns, contradictions, and anomalies and are uncertain about how to proceed (Carroll & Kay, 1988; VanLehn, et al., 2003). The state of confusion has been found to positively correlate with not just retention of knowledge but also deeper understanding and transfer of knowledge (Craig et al., 2004; D'Mello & Graesser, 2011; Pekrun, 2006). However, the mere experience of confusion does not lead to deeper learning; it is the effortful cognitive activities (e.g., problem solving, deliberation, reflection) that learners must engage in to resolve their confusion that leads to learning (D'Mello & Graesser, 2012; VanLehn et al., 2003). Capitalizing on such potential benefits of confusion for learning, we implemented a design intervention aimed at providing opportunities for confusion and confusion resolution during a math task. Specifically, a confidence judgment request was implemented which required students to rate their level of confidence after answering a math question before receiving feedback on the answer's correctness. The aim was to elicit confusion by raising awareness of cognitive conflicts in which when students reported high confidence for answers that turned out to be incorrect (high confidence errors) or low confidence for answers that turned out to be correct (low confidence corrects). The purpose of the study was to foster confusion in ways that promote the following learning behaviors: 1) increased attention to feedback; 2) increased information seeking; and 3) increased math knowledge.

Confusion: What causes it and what does it do?

Confusion is considered to be an epistemic or a knowledge emotion (Pekrun & Stephens, 2012; Silvia, 2010) because it arises out of appraisals of the degree to which incoming information aligns with existing knowledge structures. In sketching the antecedents and consequences of confusion, we draw from appraisal theories of emotions (for a review, see Moors, Ellsworth, Scherer, & Frijda, 2013) as well as cognitive theories linked to appraisals (Festinger, 1957; Mandler, 1976; Piaget, 1952). Appraisal theorists contend that emotions are adaptive responses that reflect one's assessment of how features of the environment are significant for one's well-being. These emotional experiences are associated with attending to and making sense out of incoming information and judging its relevance and congruency to one's goals as well as its association with one's expectancy, perceived levels of agency (i.e., cause of event by self, others, or impersonal factors), and abilities to cope with the event (Moors et al., 2013). Appraisals have motivational implications in that they provide information about what one can or cannot do, thereby influencing one's emotions and the behavioral actions that follow (Clore & Ortony, 2000; Lazarus, 1991). These theories therefore account for differences in individuals' responses to the same situation.

To understand how appraisals may elicit confusion, we turn to theories of cognitive dissonance (Festinger, 1957), cognitive disequilibrium (Piaget, 1952), and interruption (Mandler, 1976). Cognitive dissonance or disequilibrium results when there is inconsistency between expected and observed outcomes. This inconsistency then, according to the interruption theory, produces arousal of the autonomic nervous system. Empirical evidence suggests that this arousal influences individuals to strive to make sense of the unexpected information. For example, individuals remember more expectancy-violating behavior than expectancy-consistent behavior (Srull & Wyer, 1989), implicating increased working memory and more elaboration and effortful

processing (Bartholow, Fabiana, Gratton, & Bettencourt, 2001; Stangor & Duan, 1991). When cognitive inconsistency is not quickly resolved, affective states of confusion and frustration often accompany cognitive disequilibrium (Kort, Reilly, & Picard, 2001).

With regard to complex learning, such as problem solving in math and science, cognitive inconsistency is often beneficial because such a conflict can arise from active learning. Deep comprehension occurs when learners confront contradictions, obstacles to goals, salient contrasts, anomalies, and other experiences that fail to match expectations (Carroll & Kay, 1988; Mandler, 1976; VanLehn, et al., 2003). Cognitive conflict can activate one's efforts to deliberate and reflect on the situation in an attempt to restore cognitive equilibrium. However, in line with appraisal theories, these efforts may only be employed if other factors such as goal relevance and coping potential are favorable. As such, a crucial component of the present study is to also provide supports for resolving confusion.

Previous research on confusion induction during learning

Although work on the association between confusion and learning has a long history, research on proactively inducing confusion during learning has only recently emerged. In this line of work, confusion has been strategically elicited through contradictory information (D'Mello, Lehman, Pekrun, & Graesser, 2014) and system breakdowns (D'Mello & Graesser, 2014). The presentation of contradictory information as a method of prompting confusion has been investigated in experiments in which two disagreeing computer agents asked the learner which opinion had more scientific merit. Learners reported more confusion when the agents presented contradicting opinions rather than similar opinions, and those who were successfully confused performed better on retention and near transfer learning measures when in the contradictory information conditions than in the no contradiction condition.

In another study, D'Mello and Graesser (2014) investigated the effectiveness of presenting system breakdowns as a method of inducing confusion. Participants were asked to study illustrated texts of everyday devices and then were presented with the same illustrated text and an additional prompt. In the control condition, participants received an additional prompt that instructed them to read the same illustrated text again. In the experimental condition, the additional prompt described a situation in which the device did not function properly. Participants were then asked to determine why the device did not function properly, and based on self-reports, indicated more confusion when presented with a breakdown scenario compared to the control condition. Furthermore, participants who were able to partially-resolve their confusion performed better on the device comprehension task than their counterparts who could not resolve their confusion, highlighting the important role of confusion resolution in learning.

Present study: Mechanisms through which cognitive conflict influences learning

In the present work, we introduce a scalable design intervention to naturally provide opportunities for cognitive conflict and examine the mechanism through which it may impact learning. Specifically, we asked students to report their level of confidence after answering a math question on a computer. Feedback was then provided indicating whether or not their answer was correct. The intent was to induce confusion through raising awareness of cognitive conflicts that arise when students report high confidence for answers that turn out to be incorrect (high confidence errors) or low confidence for answers that turn out to be correct (low confidence corrects) (see Table 4.1). Through provision of feedback, students are made aware of instances in which they miscalibrated what they did and did not understand.

This particular design intervention was chosen in part because literature on high confidence errors and low confidence corrects have shown that such conflicts can be beneficial

for learning (Metcalfe, Butterfield, Habeck, & Stern, 2012). Research on the effects of high confidence errors has shown that both children and adults are more likely to correct errors made with high confidence than errors made with low confidence in an immediate post-test of factual recall of trivia questions (Metcalfe & Finn, 2012). Unexpected feedback is thought to elicit feelings of surprise and has been found to lead to a greater expenditure of effort to encode that feedback, with positive consequences for memory (Fazio & Marsh, 2009). Given the similarities between surprise and confusion, as both are emotions relating to an uncertainty about new information (Pekrun, 2010), this evidence supports our hypothesized effect of high confidence errors inducing confusion. Furthermore, research using a test of general knowledge facts has shown that when correct responses are made with low confidence, feedback serves to correct this initial metacognitive error, enhancing retention of low confidence correct responses (Butler, Karpicke, & Roediger, 2008).

Little is known, however, about whether these miscalibration effects of high confidence errors and low confidence corrects apply beyond problems of factual recall and memorization to problems that require more complex thinking. Our work examined that extension by using math problems. Unlike the previously mentioned studies, which used trivia questions and retested participants on identical questions, our retest used similarly structured math problems that were not identical, preventing the straight memorization of correct answers. Further building on previous research, we analyzed the mechanisms by which miscalibration may influence increased knowledge through measuring students' learning behavior and knowledge outcomes. Our specific research questions are as follow:

1. What are the effects of miscalibration (e.g., high confidence errors, low confidence corrects) on one's level of confusion?

2. What are the effects of miscalibration on attention to feedback, information seeking, and increased math knowledge?

Method

Participants

A group of 42 fourth- and 28 fifth-grade students, 54% girls, participated in this study. The students were from the same public elementary school in California. Data about the ethnic and socioeconomic status breakdown of our sample was unavailable, but the most recent publicly available school-wide records show that the 2012-2013 student population included 97% Hispanic students with 84% of all students eligible for free- or reduced price lunch, which serves as a proxy of low income. On the 2013 California Standardized Test, 66% of the fourth grade students and 45% of the fifth grade students scored at or above the state-indicated level of proficiency for math. The study took place over two days. We had full participation from two fourth-grade classrooms and one fifth-grade classroom on Day 1 of data collection. On Day 2, 68 of 70 students returned. One data file did not properly save, leaving us with Day 2 data for 67 students.

Procedure

Students participated in a two-day research session that was carried out in group sessions of 21 to 28 students in the school computer lab. Each class came as a group to participate in a one-hour session each of the two days. On Day 1, students were briefed as a group on how to use a 1 to 5 Likert rating scale and asked to complete and discuss example problems that illustrated all points on the scale. This scale was used for rating confusion and the helpfulness of the explanation boxes presented in the math software. Following rating scale training, 58 students individually completed math questions on the computer for 30 minutes. Due to technological

issues, the other 12 students completed math questions for 20-29 minutes. Finally, students filled out a rating scale to indicate the degree of helpfulness of the explanation boxes presented in the software. On Day 2, students individually completed math problems on the computer. Each student answered a set of problems that was similar to the ones they previously incorrectly answered on Day 1. Following that, 30 randomly selected students were briefly interviewed one-on-one by a researcher about their perception of and experience with confusion. Those who were not interviewed—or who had finished the interview—drew on scratch paper or completed mazes and puzzles.

Math software

Math questions. Multiple-choice math questions for this study were taken from the Spatial Temporal Math program (ST Math; MIND Research Institute, Irvine, CA). In a sample of 1,522 fourth grade students who answered an average of 155 ST Math multiple-choice math problems over a school year, the students incorrectly answered 39% of the items and indicated high confidence for 65% of those errors. Similarly, in a sample of 1,032 fifth grade students who answered an average of 137 problems, the students incorrectly answered 35% of the items and indicated high confidence for 67% percent of those errors. Thus, for all fourth and fifth grade students, incidences of high confidence errors occur for about a quarter of all questions.

We used a pool of 40 general math questions from ST Math and divided them into three batches based on the problem's difficulty and its likelihood to elicit high confidence errors. Batch 1 included questions incorrectly answered by more than half the students, and of those students who were highly confident about their answer, 35% or fewer correctly answered. Batch 2 included questions also incorrectly answered by more than half the students, and of those who were highly confident about their answer, between 36 and 50% of students correctly answered.

Batch 3 included questions that were incorrectly answered by less than half of the students. The software started with four questions from Batch 2 and was programmed to select a question from a harder batch when students answered more than half of their previous four problems correctly, select from an easier batch when students answer fewer than half of their previous four problems correctly, and select from the same batch if students answered exactly half of their previous four problems correctly. If there were not problems available from the targeted batch, students received problems from the next closest batch.

Math software structure. Students answered all math questions on the computer and rated their confidence and level of confusion during each problem. The software recorded all mouse click locations and times between those clicks for analytic purposes. The sequence of computer events during the math task on Day 1 was as follows:

1) The computer displayed a math question.
2) Students selected one of five provided answer alternatives.
3) Students responded to "How sure are you?" by selecting one of three icons that represent varying levels of confidence (see Figure 4.1).
4) The screen displayed "This is correct" or "This is not correct" (see Figure 4.2).
5) Students answered "How confused are you right now?" by selecting a number on a scale of 1 to 5 anchored by 1 as being "Not at all confused", 3 as being "Somewhat confused", and 5 as being "Very confused" (see Figure 4.3).
6) If students reported a confusion rating of 2 or higher, they were asked "Why are you confused?" and given two options. If their answer was correct, the options were "I thought I picked the wrong answer but it was right" and "A different reason." If their

answer was incorrect, the options were "I thought I picked the right answer but it was wrong" and "A different reason."

7) The screen displayed a green check mark next to the correct answer and, if applicable, a red "x" next to the incorrect answer. In addition, "Explain More" and "Next" buttons appeared at the bottom of the screen (see Figure 4.4).
8) Students had the option to click on the "Explain More" button to view details of the problem being worked out as well as the option to click on "Next" to advance to the next problem. Students had full autonomy over whether to request an explanation, how long to look at the explanation, and how long to wait before clicking "Next."
9) After clicking "Next," students rated "How confused are you right now?" on a scale of 1 to 5 with the same previous anchors.
10) If students reported a confusion rating of 2 or higher, they were asked "Why are you confused?" and given two options: "I still don't understand how to solve this problem" and "A different reason."
11) The next math problem was displayed.

On the Day 2, each student completed a posttest with questions that were similar to the ones he or she incorrectly answered on the first day. The questions' structures were identical but the numbers changed so that students could not simply choose the correct answer through memorization. For example, if the question "4,292 + 3,794 " was incorrectly answered on Day 1, the posttest question on Day 2 was "5,257 + 3,149". During the posttest, students only answered the questions and were not asked to rate their confidence or confusion levels, were not given feedback on whether their answers were correct, and were not given explanation boxes.

Interview

Given that students' perceptions of different emotional states impact their behavior and endorsement of particular emotions, we conducted brief interviews with a subset of students to gain insight into how they viewed confusion. Following the posttest, students selected by a random number generator were interviewed regarding their thoughts about and experience with confusion. Interviews were about five minutes long and were one-on-one with one of three university researchers. Specifically, we asked students to define confusion, talk about an incident during which they were confused, discuss how they respond to confusion, and explain if they think confusion is a good or bad thing and if certain types of people are more often confused. Anecdotes from these interviews complement our statistical results in the discussion section.

Analysis plan

We begin by examining the descriptive statistics of students' math performance and calibration patterns. To investigate the degree to which accuracy of calibration predicts students' likelihood to correctly answer the problem, seek information, and correct a previous error, we ran a series of logistic regressions. In seeking to understand the role of students' accuracy of calibration in predicting their confusion level as well as time spent on feedback, we used ordinary least squares regression methods. In all models, time-invariant student characteristics were controlled with student fixed effects, which amounts to including dummy variables for all but one student. In effect, the fixed effects adjustment ensures that the estimated links are based solely on within-student variation in dependent and independent variables. Finally, we discuss results about students' perceptions of confusion and how such perceptions can influence the relation between confusion and learning by sharing excerpts from our interviews.

Results

Descriptive statistics

Performance. On the initial test, 70 students answered an average of 27.7 (SD=8.15; range=11-40) questions, of which 35% (SD=13%) were correct answers. The next day, on a posttest with only the incorrectly answered questions, the 67 students who returned and had valid data answered on average 18.6 questions (SD=6.7; range=7-34), of which 25% (SD=13%) were correct answers. Our numbers are in line with a previous experiment on the effects of high confidence errors among grades 3-6 children in which researchers administered an average of 27.6 problems per child to research a criterion of 16 incorrect answers (Metcalfe & Finn, 2012).

Calibration accuracy. Our sample of 70 students answered a total of 1,936 problems and reported high confidence for 44% of their answers, medium confidence for 24%, and low confidence for 31%. Examining calibration accuracy, we found that students experienced high confidence errors for 24% of the problems and low confidence corrects for 7% of the problems. As such, 31% of their answers were miscalibrated. Descriptive statistics illustrating additional relations between students' confidence and accuracy are in Table 4.2.

Helpfulness of explanation box. Students mean rating for the helpfulness of the explanation boxes provided for the math problems was a 3.92 (SD=1.33; range=1-5) on a scale of 1 to 5 in which 1 was "not at all helpful," 3 was "somewhat helpful," and 5 was "very helpful."

Examining the influence of calibration accuracy on confusion and learning behaviors

Effect of confidence rating on accuracy of answer. Conducting a logistic regression, we found a significant predictive effect of confidence on the accuracy of students' answers during Day 1, ($\beta = .68$, SE=.07, $Z = 9.76$, $p < .01$). Significant positive β estimates indicate an increase in the log odds, and hence an increase in the likelihood of occurrence of the dependent variable with the predictor considered (calculated using the inverse logit function). We computed

the Wald's Z statistic, testing whether the estimates are significantly different from 0. The β coefficient of .68 indicates that for every unit increase in confidence, participants were almost twice as likely to be accurate (logit^{-1} (.68) = 1.97 log odds).

Effect of miscalibration on confusion rating. We analyzed whether miscalibrations between students' confidence in their answer choices and the correctness of the answers predicted higher self-reported confusion. Descriptive measures in Table 4.3 show no apparent relation between confidence of errors and level of confusion; however, lower confidence in correct answers appears to be associated with higher levels of confusion. To confirm this, we ran two sets of ordinary least squares regressions. First, we restricted our data to the 1,275 incorrectly answered problems. A regression analysis with the predictor, confidence level (low, medium, high), and outcome variable, confusion rating (1, 2, 3, 4, 5), showed no significant influence of confidence level on confusion rating (see Table 4.4). Next, we restricted our dataset to the 661 correctly answered problems. A regression analysis showed a significant relation between confidence level and confusion rating such that every unit decrease in confidence of a correct answer corresponded to a .24-unit increase in confusion on a 1-to-5 scale (see Table 4.4). Therefore, the inverse association between confidence of answer and confusion was only apparent for incidences of correct responses.

Effect of miscalibration on information seeking. We analyzed whether miscalibration resulted in greater likelihood to seek more information. In particular, we assessed whether students requested an explanation after being shown whether their answer was correct or incorrect. Descriptive statistics in Table 4.5 show no apparent relation between confidence of errors and likelihood to request an explanation; however, lower confidence in correct answers appears to be associated with higher likelihood of requesting an explanation. To confirm this, we

restricted the data to correctly answered problems. A logistic regression model with the predictor of confidence level (low, medium, high) and dichotomous outcome variable of information seeking (1=requested explanation box; 0=did not request explanation box) revealed a significant influence of confidence level on whether or not students requested an explanation box (see Table 4.6). The odds ratio was 2.37 and 2.63 for low and medium confidence relative to high confidence corrects, respectively, meaning that the likelihood that students requested an explanation box for low and medium confidence corrects is more than two times higher than the likelihood that students requested an explanation box for high confidence corrects.

We also conducted a second set of analyses with a subset of the data after observations that some students did not request an explanation box because it did not take them much time to understand their error (e.g., "oh I get it now!" right after seeing the correct answer noted). Therefore, to analyze if students requested an explanation box only during instances in which that may have be helpful, we restricted the sample **t**o problems in which students reported that: 1) they were still confused (rating above 1) after deciding to move to the next problem, and 2) reported that the reason for the confusion was they still did not understand how to do the problem. This filter eliminated 740 problems, leaving 470 problems. Furthermore, 31 participants (180 problems) were dropped because those students either always or never pressed the explanation box, leaving 290 problems. Results revealed that there was still not a significant relation between confidence of incorrect answers and information seeking ($\beta = .01$, SE=.20, $Z = .03$, $p > .05$).

Effect of miscalibration on total time spent on feedback. We assessed whether miscalibration was related to more time spent on feedback. Time spent on feedback was calculated as the total number of seconds that students spent between being told whether their

answer was correct and clicking on the button to go to the next problem. This includes time spent looking at the problem again, reworking the problem on scratch paper, and viewing the explanation box. Descriptive statistics in Table 4.7 show trends between miscalibration and time spent on feedback such that lower confidence related to more time spent on feedback for both correct and incorrect responses. Regression analyses confirmed the pattern for incorrectly answered problems such that students spent on average 0.7 seconds (inverse log transformation of $10^{-.132}$) more time on feedback for low confidence errors compared to high confidence errors (see Table 4.8). To put this in perspective, students on average spent 6.88 seconds on feedback thus 0.7 second equates to approximately 10% more time. The pattern for correctly answered problems, however, was not significant.

To test if students spent more time on feedback only during instances that may have be helpful, we restricted the sample **t**o problems in which students reported that: 1) they were still confused (rating above 1) after deciding to move to the next problem, and 2) reported that the reason for the confusion was they still did not understand how to do the problem. This filter eliminated 1,403 problems, leaving 483 problems for the analysis with the outcome variable of time on feedback. Results revealed a similarly significant relation between calibration accuracy and time on feedback for incorrect answers ($\beta = .13$, SE=.05, $t = -2.14$, $p < .05$) but not for correct answers ($\beta = -.08$, SE=.21, $t = -.38$, $p > .05$).

Mediation. We tested a mediation model of whether confusion mediated the influence of students' confidence of correct answers on their likelihood to request an explanation box. Previous regressions illustrated that lower confidence for correct answers were related to both higher confusion and higher likelihood to request an explanation box. However, an additional

regression revealed that confusion did not predict likelihood to request an explanation ($\beta = .094$, SE=.142, Z =.66, p >.05) and therefore did not support the possibility of mediation.

Examining the influence of calibration accuracy on achievement

Effect of miscalibration on correcting previous errors. A logistic regression model with the predictor of confidence level (low, medium, high) and dichotomous outcome variable of accuracy on posttest (accurate, not accurate) showed no significant influence of confidence level (see Table 4.9). However, given that it would be unreasonable to expect students to correct errors in which they were unable to resolve, we then restricted our sample to only problems in which students resolved their confusion. Therefore, we re-ran our statistical analyses while only including the problems in which students reported that they were not at all confused (rating of 1) before moving onto the next problem. This eliminated 696 problems, leaving 514 problems for the analyses. The relation between calibration accuracy and the correction of previous errors remained insignificant.

Discussion

Investigating the impact of miscalibration on confusion and learning, we found that miscalibration by the way of low confidence correct responses had a significant influence on confusion and learning behaviors. Previous studies on feedback have generally conceptualized feedback as a method for correcting errors rather than as a method for learning the reason behind correct responses. A notable exception, however, is a set of experimental studies that showed that feedback doubled the retention of correct low-confidence responses relative to providing no feedback (Butler, Karpicke, & Roediger, 2008). These studies were conducted with undergraduate college students and used trivia questions (e.g., What is the longest river in the world?). Extending those findings, our study provides insights about mechanisms that underlie

why correct low-confidence responses may be beneficial for learning by exploring affective and metacognitive components. Specifically, we found that students reported higher confusion and were more than twice as likely to request an explanation box for correct answers made with low confidence than correct answers made with medium or high confidence.

Contrary to previous studies that showed that high confidence errors are more likely to be corrected on a posttest than low confidence errors, we did not find that relation and our findings suggest that the correction of high confidence errors (also known as the hypercorrection effect) may not be a general phenomenon. A distinct difference between our study and previous studies on the hypercorrection effect is our use of math problems rather than trivia questions, which implicate different kinds of cognitive processes. On the posttest, our study used math questions that were similar in structure to the ones students incorrectly answered but replaced with different numbers. As such, students could not simply memorize the correct answer for the posttest. This was not the case for the trivia questions because students were given the exact same questions on the posttest. Students can quickly accept that a particular trivia answer is correct (or least memorize it to reproduce on a posttest even if they do not believe the answer is correct). Resolving mathematical misunderstanding, however, is typically more involved and a large body of literature shows that conceptual change in math is difficult to facilitate (e.g., Chi & Roscoe, 2002). Therefore, feedback that merely shows the correct answer is likely not enough to correct math errors.

A limitation that made it difficult for us to assess the implications of our design intervention on math achievement, as measured by the correct of errors on a posttest, was the fact that many students were unable to resolve their confusion. Even though our study coupled correct answer feedback with the option for students to request an explanation box that showed

the math problem worked out step-by-step, that was not enough to resolve confusion for almost half the problems in which students reported initial confusion when their answer was shown to be incorrect. In that sense, confusion was neither productive nor desirable, which aligns with the perspectives that students shared in interviews. Many stated that confusion is a negative state when confusion cannot be resolved, especially when the stakes are high (e.g., "when you are taking a timed test and you don't know what to do") or when one is not progressing despite putting in time and effort (e.g., "you might have been doing it for an entire month… the same thing over and over again"). We recommend that future work take great care in providing sufficient resources for students to resolve their confusion with greater success than students were able to do in our study. It is unclear, however, whether students in our study did not resolve their confusion because they were unable to with the provided resources or if they did not want to put forth additional effort for resolving their confusion. While it takes seconds to process a trivia answer and memorize it as being the correct answer, it typically takes longer to process why a particular math answer is correct so persistence plays a larger role in that context.

Our study proposed that the design intervention of prompting students for confidence judgments provides opportunities for confusion, which we found to be partly successful. While lower confidence for correct answers was related to higher confusion, higher confidence for incorrect answers was not. A possible explanation for this is that for the incorrectly answered problems, students interpreted the "How sure are you?" question that they were asked right after selecting an answer in different ways. For instance, we observed one student click on "very sure" and then said to himself, "I am very sure that I got this wrong" and then reported that he was "not at all" confused as he said to himself, "I am not confused. I knew that I would get this

wrong." Future studies extending this work should be more clear about the context of the confidence judgment such as by asking, "How sure are you that your answer is correct?"

Through our interviews with students, we also explored whether our results could be explained by a social desirability bias resulting from positive or negative connotations that students attach to the state of confusion. Several students perceived confusion as being an affect that is experienced by those who are not "good students" which may have influenced students to report low or no confusion. Interviewed students explained that classmates are often confused when they are not paying attention (e.g., "students that are playing around" and "they don't listen"). Of particular note was one student who stated that when he is confused, he "keeps quiet" rather than seek help. This not only explains the potential of confusion being underrated in our study but also underscores the importance of creating a culture in which confusion is valued as part of the process for learning in order to facilitate productive regulatory learning behaviors for resolving cognitive conflicts.

In conclusion, promising implications emerged from our findings regarding learning behaviors. Miscalibration prompted students to seek more information (in the case of low confidence corrects) and spend more time on feedback (in the case of high confidence errors). These findings align with research that cognitive conflict can prompt learning behaviors that are beneficial for learning, especially in math and science domains. Prompting students to reflect on their confidence for their answer provided opportunities for their miscalibration and cognitive conflict to be salient, and for those problems, students exhibited behaviors that indicated that they employed greater efforts in understanding the mathematics.

References

Bartholow, B. D., Fabiana, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197–204.

Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918.

Carroll, J., & Kay, D. (1988). Prompting, feedback and error correction in the design of a scenario machine. *International Journal of Man-Machine Studies*, *28*, 11-27.

Chi, M. T., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 3-27). Netherlands: Kluwer Academic.

Clore, G. L., & Ortony, A. (2000). Cognition in emotion: Always, some- times, or never? In L. Nadel, R. Lane & G. L. Ahern (Eds.), *The cognitive neuroscience of emotion* (pp. 24–61). New York, NY: Oxford University Press.

Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media, 29,* 241-250.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29,* 153-170.

D'Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition and Emotion, 25*, 1299-1308.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*, 145-157.

D'Mello, S., & Graesser, A. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychologica*, *151*, 106-116.

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*(1), 88-92.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Kort, B., Reilly, R. & Picard, R. W. (2001) External representation of learning process and domain knowledge: affective state as a determinate of its structure and function. *Proceedings of the Artificial Intelligence in Education Workshops* (San Antonio, TX, AIED), 64–69.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Mandler, G. (1976). *Mind and emotion*. New York: Wiley.

Metcalfe, J., Butterfield, B., Habeck, C., & Stern, Y. (2012). Neural correlates of people's hypercorrection of their false beliefs. *Journal of Cognitive Neuroscience*, *24*(7), 1571-1583.

Metcalfe, J., & Finn, B. (2012). Hypercorrection of high confidence errors in children. *Learning and Instruction, 22*, 253–261.

Moors, A, Ellsworth, P., Scherer, K. R., Frijda, N. H. (2013) Appraisal theories of emotion: State of the art and future development. *Emotion Review, 5,* 119-124.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341.

Pekrun, R. (2010). Academic emotions. In T. Urdan (Ed.), *APA Educational Psychology Handbook*. Washington, DC: American Psychological Association.

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. Harris, S. Graham, T. Urdan, S. Graham, J. Royer & M. Zeidner (Eds.), *APA Educational Psychology Handbook: Individual differences and cultural and contextual factors* (pp. 3-31). Washington, DC: American Psychological Association.

Piaget, J. (1952). *The origins of intelligence.* New York, NY: International University Press.

Silvia, P. J. (2010). Confusion and interest: The role of knowledge emotions in aesthetic experience. *Psychology of Aesthetics Creativity and the Arts, 4*, 75-80.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review, 96*, 58-83.

Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27,* 357–378.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*, 209-249.

Table 4.1
Six Possible Calibration Categories

Confidence judgment	Accuracy of student answer	
	Correct	Incorrect
Low	Low confidence correct	Low confidence error
Medium	Medium confidence correct	Medium confidence error
High	High confidence correct	High confidence error

Table 4.2
Number of Answers for Different Calibration Categories

Confidence judgment	Accuracy of student answer	
	Correct	Incorrect
Low	Low confidence correct	Low confidence error
Medium	Medium confidence correct	Medium confidence error
High	High confidence correct	High confidence error

Table 4.3
Mean Confusion Level for Different Calibration Categories

Confidence Rating	Correct Mean (SD)	Incorrect Mean (SD)
Low	1.94 (1.46)	3.30 (1.62)
Medium	1.65 (1.10)	2.96 (1.59)
High	1.23 (0.75)	2.72 (1.73)

Note. Numbers represent the means of confusion ratings. For example, 1.94 was the mean confusion rating for correct answers rated with low confidence. Confusion level was rated on a scale of 1 (not at all confused) to 5 (very confused).

Table 4.4
Ordinary Least Squares Regressions Predicting Confusion Rating from Calibration Category

Dependent variable = Confusion rating				
	β	*SE*	*t*	*p*
Incorrect Answer				
Confidence	.08	.05	1.63	.10
Constant	2.84	.10	27.11	< .01
Correct Answer				
Confidence	-.24	.04	-5.66	< .01
Constant	2.04	.11	18.81	< .01

Note. Model was run with student fixed effect to account for the within-subjects sample. Coefficients are unstandardized. Analyses with regressions treating confidence as a categorical variable resulted in similar findings.

Table 4.5
Number of Problems for which an Explanation Box was Requested for Different Calibration Categories

Confidence	Correct		Incorrect	
	Yes (%)	No (%)	Yes (%)	No (%)
Low	45 (36)	80 (64)	178 (37)	304 (63)
Medium	38 (29)	95 (71)	127 (38)	210 (62)
High	72 (18)	331 (82)	120 (26)	336 (74)

Note. Numbers represent the frequency for which the explanation box was requested (yes) and not requested (no). For example, 45 was the number of times that an explanation box was requested for correct answers rated with low confidence.

Table 4.6
Logistic Regressions Predicting Likelihood to Request an Explanation from Calibration Category

Dependent variable = Requested explanation (1=yes, 0=no)				
	Odds Ratio	*SE*	*Z*	*p*
Incorrect answer				
Medium confidence	1.42	.30	1.64	.10
High confidence	1.04	.22	.20	.84
Correct answer				
Medium confidence	2.69	.98	2.71	.01
High confidence	2.34	.87	2.29	.02

Note. The reference category was low confidence. Model was run with student fixed effect to account for the within-subjects sample. Coefficients are unstandardized. Analyses with regressions treating confidence as a continuous variable resulted in similar findings.

Table 4.7
Mean Time Spent on Feedback for Different Categories of Calibration

Confidence Rating	Correct Mean (SD)	Incorrect Mean (SD)
Low	7.15 (11.51)	8.21 (13.00)
Medium	6.24 (8.56)	7.87 (10.41)
High	4.91 (7.36)	6.58 (7.69)

Note. The numbers represent the mean number of seconds spent on feedback. For example, 7.15 was the average number of seconds spent on feedback for correct answers rated with low confidence.

Table 4.8
Ordinary Least Squares Regressions Predicting Time Spent on Feedback from Calibration Category

Dependent variable = Log of feedback time (seconds)				
	β	*SE*	*t*	*p*
Incorrect answer				
Medium confidence	.02	.06	.39	.70
High confidence	-.13	.06	-2.19	.03
Constant	1.53	.04	37.16	< .01
Correct answer				
Medium confidence	.17	.09	1.84	.07
High confidence	.13	.09	1.43	.15
Constant	1.14	.04	27.03	< .01

Note. The reference category was low confidence. Model was run with student fixed effect to account for the within-subjects sample. Coefficients are unstandardized. Analyses with regressions treating confidence as a continuous variable resulted in similar findings.

Table 4.9
Regression of Confidence Rating of Incorrect Answers Predicting Correction of Errors on a Posttest

Dependent variable = Error corrected (1=yes, 0=no)				
	Odds Ratio	*SE*	*Z*	*p*
Medium confidence	.92	.18	-.44	.66
High confidence	.91	.17	-.50	.62

Note. The reference category was low confidence. Model was run with student fixed effect to account for the within-subjects sample. Coefficients are unstandardized. Analyses with gamma correlations and regressions treating confidence as a continuous variable resulted in similar findings.

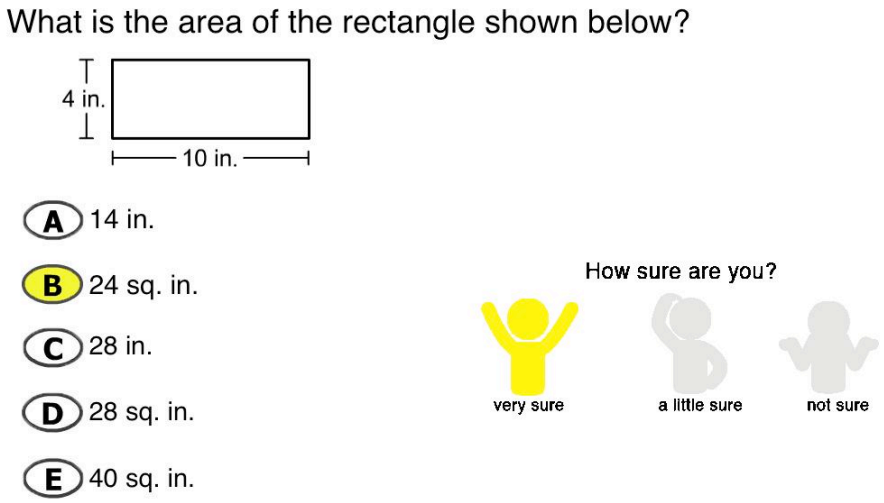

Figure 4.1. Screenshot of icons used by students to rate their confidence immediately after selecting an answer.

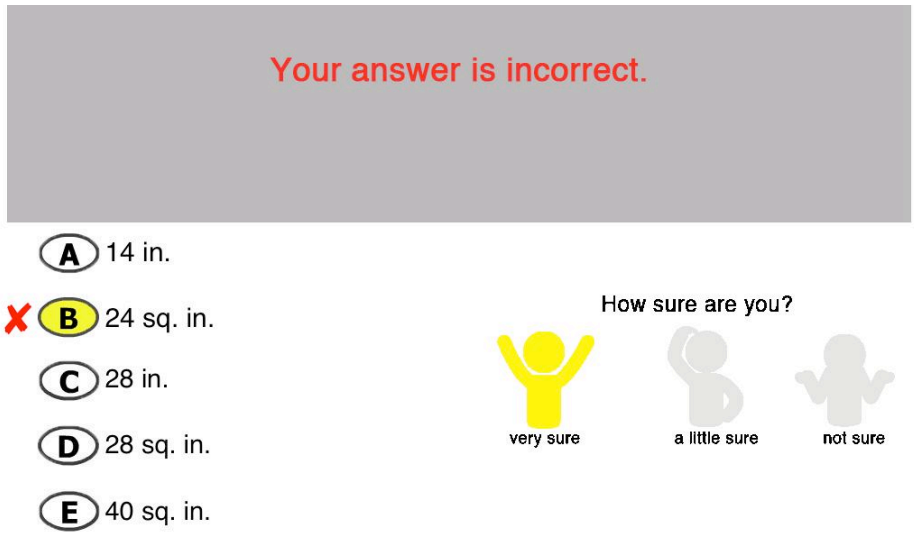

Figure 4.2. Screenshot of feedback interface for incorrect answer shown immediately after students' confidence rating.

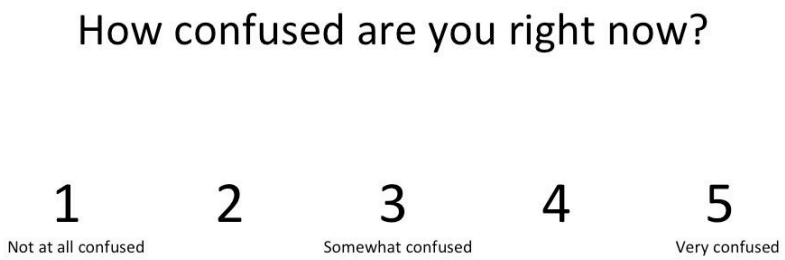

Figure 4.3. Screenshot of confusion rating scale shown immediately after students rate their confusion.

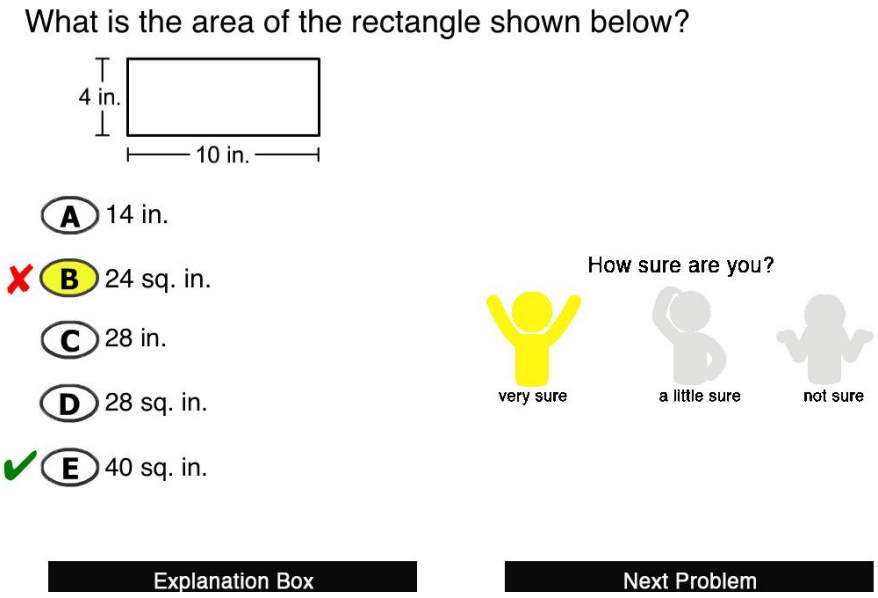

Figure 4.4. Screenshot of feedback symbols representing correct and incorrect answers and option buttons to request an explanation or move onto the next problem immediately. This was shown immediately after students rated their confusion and reason for confusion (if applicable).

CHAPTER 5

Conclusion

When designing learning environments, often the focus is on how to design in ways that allow for certain types of interactions to occur with the thought that those interactions lead to learning. However, there is another part of this equation: it is equally as important to consider that learners have different motivations and goals and will accordingly differ in how they perceive what the design affords and how they interact with the environment. My dissertation examined ways in which design affords different types of information-seeking behaviors while accounting for those behaviors being reciprocally influenced by perceptions that are colored by students' personal traits, such as their goals and level of background knowledge. My focus on unraveling these interactive patterns to better understand learners' reactions to failure and ways to design for productive persistence was largely inspired by my observations of the desire for challenge and the persistence that players exhibit in games. One quote that remained at the forefront of my mind as I worked on my dissertation vividly demonstrates how challenge, confusion and even frustration often bind players to a game:

> It's easy to tell what games my husband enjoys the most. If he screams 'I hate it. I hate it. I hate it,' then I know he will finish it and buy version two. If he doesn't say this, he'll put it down in an hour. (Lazzaro, 2008, p. 686)

The driving force behind my dissertation was the pondering of how wonderful that would be if we could encourage students to have the same desire for challenge and persistence after failure for more formal learning such as in understanding math and science. How do we motivate students to view failure as desirable and as progress towards learning and how do we promote productive responses to failure?

To tackle these questions, I first sought a better understanding of how students react to failure in the first study of my dissertation. The heat pump game was an appropriate context for doing so given the quick feedback cycle, the digital, physical and social resources, and the ample opportunities to retry. Video recordings of high school students' experiences with the game allowed for second-by-second analyses, with replays and pauses, to closely examine interactions among the entire social, digital, and physical system of the learning environment. Information seeking emerged as a critical behavior that reflected learners' goals and motivation, and accordingly, my colleagues and I developed a typology for information-seeking behaviors based on two dimensions: social to non-social and goal-directed to not goal-directed. We then detailed how certain design features of the heat pump game provided affordances for specific types of information-seeking behaviors. This micro-level interaction analysis provided insight towards specific leverage points that designers can change to impact learners' experiences.

Data from the heat pump game also informed my second study which explored how students' interactions with the exhibit influenced changes in their goals. The digital, physical, and social resources of a learning environment can shape students' task goals, influence the level of intensity of their goal pursuit, and impact their decision to switch goals. To better understand this process, I focused on four students who endorsed different types of goals for science learning and traced their interactions at the heat pump exhibit using video analyses as well as supplementary interviews to better understand the rationale behind their actions. These case studies illustrated ways in which the design of a science museum exhibit provided affordances for some types of goals but not others, resulting in shifts in goal pursuit intensity and goal type. These findings not only have implications for the design of learning environments but also

underscore the importance of aligning research methods with the dynamic nature of task goal change to extend theoretical knowledge about achievement goals.

The final study in my dissertation was conducted in a different context with fourth- and fifth-grade students and a computerized math task. It focused on eliciting confusion, which is an affective state that has been found to be an ingredient in enjoyable games as well as a key emotion that is related deeper, conceptual understanding of math and science. Specifically, I explored a scalable design intervention aimed at providing opportunities for confusion and the productive resolution of confusion. The design intervention required students to rate their confidence level after selecting an answer to a math problem. Our findings showed that the cognitive conflict that arises from the discrepancy between what one thinks one knows and what one actually knows can influence students to invest more effort towards better understanding math problems but that these efforts can dampen when feedback does not successfully resolve students' confusion.

These studies illustrate how one can draw from different disciplines and perspectives to inform the design of learning environments in an effort to promote productive persistence after failure. However, what also emerged was a lesson that one needs to be cautious and mindful when applying design features from one context to another. It is important to not blindly incorporate "best practices" or specific design elements without understanding the underlying mechanisms of how those elements work in their current context and system. As I draw from the insights from my research—both from the studies themselves but also from the collaborations and discussions connected to that work—to conclude my dissertation, I find two themes as especially worth elaborating upon. First, I discuss what makes failure appealing in games and the difference between persistence in video games and persistence in school-related subjects. Several

of the design elements related to game persistence may not be applicable to or appropriate for learning in more formal environments; I discuss why that is so and how to address those limitations. Second, I wrap up this chapter with a discussion of improvement science and how future research and design of learning environments could benefit from improvement science, which differs from yet still complements research for theoretical advancements.

Complexities of Adapting Design Features from Video Games

Contextual differences between video games and academic learning environments have implications for how failure is perceived in those environments and how productive persistence can be promoted. Several of the design elements related to game persistence may not be applicable to learning school subjects. Video games generate an expectancy for success, allow for choice, and is part of a low stakes environment in ways that are typically not aligned with the learning culture in schools (Tran, Lehman, Dockterman, Juul, D'Mello, & Graesser, 2013).

An important feature of games is the implicit promise that one can tackle the challenge with sufficient time and effort (Juul, 2013). Players often are engaged in games in which they expect to eventually win and are given constant feedback and unlimited retries. This is very different from learning in more formal environments in which students are required to attend classes even if they have low expectancy for success in those classes. Students who fall behind in math may not be able to catch up as the class continues moving forward. Video games, however, patiently wait for players to master a level before moving on, whether that takes minutes, hours, days or weeks. There is an expectancy for success inherent in the design of the video game itself but also from the player who chooses to engage in the game. Rarely, if ever, would a player engage in a game in which he or she knew there was no chance of progressing. Therefore,

applying game design features to more formal learning to promote persistence requires special care in constructing the feedback loop in ways that promote high expectancy for success.

Another distinction of video games is that players often get to choose whether or not they play a video game whereas students do not often get a choice of which subjects they are taught in school, especially for core subjects such as math and science. Therefore, in choosing games, players are able to self-regulate and select games that are of interest to them and games in which they believe they can eventually master. While it may not be appropriate to give students a choice of whether or not to learn math, educators can design learning environments such that students are likely to choose to engage in the task if they had the choice. The expectancy-value model (Eccles, 1983) proposes that students' achievement-related choices can be understood as a function of their expectancy for success across tasks ("Can I do this?") and the values they attach to various tasks ("Do I want to do this?"). Value refers to interest, utility, importance of identity, and cost related to the task. Research has shown that value, especially interest and utility, are powerful predictors of choice (e.g., Hidi, 1990; Hidi & Harackiewicz, 2000; Hidi & Renninger, 2006). Therefore, although one may not be able to give students choice in learning math or science, designers and educators can provide learning experiences that students would want to pursue by being aware of students' values and integrating those values into the instructional practices of both game and non-game learning environments.

Finally, a notable distinction between failure in video games and failure in school-related subjects is that failure in a game is low-stakes and that games therefore provide us with the option to deny that we care about failing in a particular game (Juul, 2013). The game is generally an artificial construction that creates a new, arbitrary measure of success (e.g., get the ball in the goal as many times as possible). However, when a teacher describes a classroom assignment as a

"game" that students are graded on, the grading makes it impossible to deny that failure matters. A possible avenue towards reproducing the appeal of failure in education is then to design ways to keep the freedom to learn and experiment that is often characteristic of games, such as by integrating more low-stakes formative assessment that is focused on progress and does not have implications for grades (Tran, Lehman, Dockterman, Juul, D'Mello, & Graesser, 2013).

Next Steps: Integrating Improvement Science with Design Research

Given that my projects were research-practice partnerships that focused on the interplay between design features and learners' behaviors, it was necessary to adopt an analytical method that allowed for both extending learning theory and providing insights on leverage points for designers to improve their products. A prevailing theme during my dissertation research was the importance of aligning the research methods with the aims of the study. In the project at the Norwegian Museum of Science and Technology, using video data supplemented with simulated video-recall interviews allowed me to assess interactions and derive clear design implications for promoting specific learning behaviors. In collaboration with the MIND Research Institute, the design intervention in math software to provide opportunities for confusion was scalable and therefore also had wide practical implications.

The cycle of research informing design and then using the process to evolve design is a core component of improvement science. Improvement science is distinctive in that it focuses on improving learning environments rather than on theoretical advancement or research for accountability (Solberg, Mosser & McDonald, 1997). For example, when the purpose of research is for theoretical development about the relationships between conceptual variables, assessment typically occurs once or twice per study using a large test that captures a large sample size "just in case" to allow for statistical significance. Research for improvement, however, aims to

develop and evaluate changes in practice. This is conducted more frequently and consists of sequential tests that gather "just enough" data in small samples. Improvement research can be shared in a low-stakes, safe environment conducive to change to revise design and other teaching practices while accounting for knowledge about the variation in different contexts (Lewis, 2015). My dissertation takes a step towards providing insights on how the design of a learning environment interacts with learners' personal characteristics to influence their behaviors, and the next step would be to implement the suggested changes derived from the research. The improvement science field, however, can provide insights as how one can establish a more efficient cycle of research and design to impact learning.

In future work on researching ways to design for productive persistence, I plan to develop and incorporate practical measurements which provides data against which teams can test their theory of practice improvement (Takahashi, 2014). As practitioners make changes in the design of learning environments, data from the measurement system can provide insight about whether the changes result in the hypothesized outcomes. Practical measures are specific to the work processes and related outcomes that are the object of change; produce data accessible in a timely manner; have formative value signaling subsequent action useful to consider; and are framed in a language that is meaningful to those engaged in the work (Takahashi, 2014). These criteria are especially crucial for research that is intended to be in collaboration with practitioners and to have impact outside of the academic circle, which is particularly of relevance to design research.

References

Eccles, J. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: Freeman.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*, 549–571.

Hidi, S., & Harackiewicz, J. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*, 151–179.

Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41* (2), 111-127.

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games.* Cambridge, MA: MIT Press.

Lazzaro, N. (2008) Why we play: Affect and the fun of games. In S. Sears & J. Jacko (Eds.), *The Human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* (pp. 679–700). New York, NY: Lawrence Erlbaum Associates.

Lewis, C. (2015). What Is Improvement Science? Do We Need It in Education?. *Educational Researcher*, *44*(1), 54-61.

Solberg, L., Mosser, G., & McDonald, S. (1997). Why are you measuring? *Journal on Quality Improvement*, *23*(3), 135-147.

Takahashi, S. (2014, August 21). Revisiting the purposes of practical measurement for improvement. Retrieved May 28, 2015, from http://www.carnegiefoundation.org/blog/revisiting-purposes-practical-measurement-improvement-learning-bten-measurement-system

Tran, C., Lehman, B., Dockterman, D., Juul, J., D'Mello, S., & Graesser, A. (2013). *Designing for productive failure*. Proceedings of the Games, Learning, and Society Conference, Madison, WI.