Copyright

by

Sa Liu

2018

The Dissertation Committee for Sa Liu Certifies that this is the approved version of the following dissertation:

The Impact of Learner Metacognition and Goal Orientation on Problem-Solving in a Serious Game Environment

Committee:

____

Min Liu, Supervisor

____

Lucas Horton

____

Paul Resta

____

Xiaofen Keating

The Impact of Learner Metacognition and Goal Orientation on Problem-Solving in a Serious Game Environment

by

Sa Liu

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2018

Dedication

This dissertation is dedicated to the memory of my grandpas, who were my inspirations to pursue my dreams and go for adventures. Because they were unable to see my graduation, this is for them.

Acknowledgements

This work would not have been possible without the support of my advisor, Dr. Min Liu. I am grateful for her valuable scholarly guidance, mentorship, and endless learning and research opportunities while I was pursuing the doctoral degree. I am also indebted to each of my committee members, Dr. Lucas Horton, Dr. Paul Resta, and Dr. Xiaofen Keating, who have been supportive of my career goals and have provided me with extensive guidance and persistent help with the dissertation.

In addition, I thank Dr. Diane Schallert and Hien Nguyen for their support on utilizing the Department of Educational Psychology participant pool. I would also like to thank my office members Dr. Karen French, Frank Escobedo, Dr. Rui Zeng, and Carmen Erdie for their warm support. I appreciate all the previous and current team members in the Alien Rescue team for their talent, motivation and enthusiasm on creating this environment. Special thanks to Dr. Jina Kang, Xin Pan, Wenting Zou and Zilong Pan on assisting the data analysis; Chenglu Li on maintaining the test version of Alien Rescue.

Thank you also to my friends Rulan Shangguan, Mihyun Lim, Morgan Jenatton, Christopher Brennan, Jason Harron, Patrick Parker, Sara Burnett, Ryan Myers, Yujung Ko, Emily Mckelroy, Hsu Hsiao Ping, and Sukanya Moudgalya, who made my experience in the graduate school exciting and fun. My sincere thanks also go to Dr. Matthew Mayer and Mushy Mayer for their encouragement, insightful comments and hard questions.

Last but not the least, I would like to thank my parents and sister, whose warm love, continued patience and endless support are always with me in whatever I pursue. I am also thankful for having two wonderful nephews, who have provided tremendous happiness to my life.

The Impact of Learner Metacognition and Goal Orientation on Problem-Solving in a Serious Game Environment

Sa Liu, Ph. D.

The University of Texas at Austin, 2018

Supervisor: Min Liu

Abstract: To understand the impact of two learner characteristics—metacognition and goal orientation—on problem-solving, this study investigated 159 undergraduate learners' metacognition, goal orientations, and problem-solving performances and processes in a laboratory setting using a Serious Game (SG) environment—Alien Rescue (AR)—that adopts Problem-based Learning (PBL) pedagogy for teaching space science. Utilizing multiple data sources, including computer log data and problem-solving solution scores within the SG, survey data, gameplay screencast videos, and interview data, this study combined a sequential mixed method design and serious games analytics techniques to answer the following two questions: (a) To what extent are learner problem-solving performance differences based on learner characteristics, and why? (b) To what extent are learner problem-solving process differences based on learner characteristics, and why?

The results indicated that (a) learner metacognition affected problem-solving. Specifically, there were statistically significant differences in learner problem-solving

performances based on metacognition, and learners also demonstrated different problem-solving processes based on metacognition. (b) Learner goal orientation impacted problem-solving. Particularly, learners in different goal orientation groups had different problem-solving processes. (c) The interaction between metacognition and goal orientations had an impact on learner problem-solving performances. Specifically, learners were clustered into three groups based on these two characteristics, including (a) high metacognition and high multiple goal orientations, (b) low metacognition and medium multiple goal orientations, and (c) medium metacognition and low multiple goal orientations. Learner problem-solving performances were statistically significant based on these three clusters. In addition, learner metacognition and goal orientations together could predict learner problem-solving performances. (d) The interaction between metacognition and goal orientations also had an impact on learner problem-solving processes. These differences in learner problem-solving performances and processes can be explained by learner characteristic differences, the problem complexity, SG design, and Dunning-Kruger effects (i.e., the cognitive bias that people of low metacognitive ability might mistakenly assess their metacognitive level as higher than it is). In addition, this study summarized 10 steps of how to be a successful and efficient problem solver in AR. These steps are as follows: 1) identify the problem correctly; 2) explore the 3D environment by visiting all rooms in AR and look over all tools; 3) discover what one alien species needs to survive in Alien Database; 4) search the Solar System Database for possible planets; 5) develop hypotheses about where this alien species can live; 6) figure out if there is any missing information needed for making a decision; 7) launch probes to gather information in the Probe Design room; 8) check the data from the probe in the Mission Control room; 9) decide whether the selected planet is a good choice for the selected alien species; 10) if so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4.

This research offers additional understanding of learner characteristic impacts on problem-solving in SG environments with PBL pedagogy. It can also contribute to future designs of these environments to benefit learners based on their metacognitive levels. In addition, the study limitations and further research in this area are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

SIGNIFICANCE OF THE STUDY

In recent years, educators have argued that students must be prepared for a vastly different working world from that of previous generation, which demands complex problem-solving skills (Duch, Groh, & Allen, 2001). According to the Program for International Student Assessment (PISA), U.S. student problem-solving performance was 24th out of 39 countries in 2003, and 24th out of 44 countries in 2012, which are both below average among OECD (Organization for Economic Cooperation and Development) countries (Lemke et al., 2004; OECD, 2014). Therefore, it is important to improve U.S. student problem-solving skills to prepare them for the future.

One educational approach to develop problem-solving skills is through facilitated problem-solving tasks utilizing a constructivist instructional method—problem-based learning (PBL)—which embeds student learning processes in real-life problems (Hmelo-Silver, 2004). Previous studies suggested PBL can facilitate long-term retention, skill development, and increase student and teacher satisfaction (Chowdhry, 2016; Oliveira, dos Santos, & Garcia, 2013; Strobel & van Barneveld, 2009). Meanwhile, with advances in computing technology and the gaming industry, scholars argued that Serious Games (SG), which were created for non-entertainment purposes (Abt, 1970, p. 9), can help student learning in educational settings (Boyle et al., 2016; Connolly, Boyle, Hainey, McArthur, & Boyle, 2012; Prensky, 2001). Therefore, by adopting PBL pedagogy in SG environments, instructional designers and scholars hope to increase learner motivation, enhance learning experiences, improve learning performances and develop critical thinking and problem-solving skills (Hou & Li, 2014; Lee & Chen, 2009; Sánchez & Olivares, 2011).

Despite many claimed advantages of using PBL, learners have faced some challenges, such as lack of guidance, distraction, and inadequate self-regulation and collaboration skills (Ertmer & Glazewski, 2015; Kirschner, Sweller, & Clark, 2006; Mayer, Griffith, Jurkowitz, & Rothman, 2008). Failing to overcome these challenges may cause learner frustration, disengagement, misconception, and eventually failure in SG environments that adopts PBL pedagogy. Thus, adopting PBL pedagogy in SG environments may not be an appropriate approach unless researchers and instructors acknowledge these challenges, incorporate necessary student support, and provide scaffolding to learner during the problem-solving processes. To decide when and how to provide scaffolding for learners, more research is needed to fully understand learner problem-solving—both problem-solving performances and processes—in SG environments.

Constructivism theory can help teachers and researchers understand learner problem-solving in SG environments. From a constructivist perspective, individual characteristics are important for understanding learning, such as personal characteristics including socioeconomic status, age, gender, and race/ethnicity (Lemke et al., 2004; OECD, 2014); academic characteristics such as prior knowledge and goal orientation (Hsieh et al., 2008; Liu, Kang, Lee, Winzeler, & Liu, 2015); social/emotional characteristics such as beliefs and self-efficacy (Hsieh et al., 2008; Jonassen, 2000; Liu, Cho, & Schallert, 2006); and cognitive characteristics such as memory and metacognition (Davidson & Sternberg, 1998; Shin et al., 2003). This study will focus on two learner characteristics—metacognition and goal orientation, because their importance for learning in SG environments with the PBL pedagogy.

Metacognition is an important learner characteristic, because it involves the process of thinking about thinking (Flavell, 1979), including knowing about one's own learning and memory capabilities, knowing what learning strategies are useful, and planning and monitoring one's own learning. Because metacognition is the "engine" that starts, regulates and evaluates the cognitive processes during learning (OECD, 2014, p. 121), learner metacognitive levels affect problem-solving in SG. Researchers have suggested that metacognition is necessary for students to succeed in PBL (Davidson & Sternberg, 1998; Marra et al., 2014; Shin et al., 2003). Without adequate metacognition, learners may have difficulty understanding complex topics in hypermedia environments (e.g., SG), because they may fail to plan, set goals, use effective strategies, and monitor and reflect learning processes, which would hinder deeper cognitive processing of core material during learning (Azevedo, Cromley, & Seibert, 2004; Mayer et al., 2008).

Goal orientation is another important learner characteristic, because preliminary studies have shown that student goal orientation critically influences their behavior in SG environments (Liu, 2005; Liu, et al., 2015; Hsieh et al., 2008). Goal orientation indicates individuals have different reasons or goals for learning (Elliot & Church, 1997; Elliot & McGregor, 2001; Elliot, Murayama, & Pekrun, 2011). Researchers have suggested that goal orientation plays an important role at the earliest stage of learner metacognitive regulation, which can further guide the entire metacognitive regulatory processes (Schraw & Dennison, 1994; Zimmerman, 2002, 2013; Moshman, 2017). Most recently, Elliot et al. (2011) constructed a goal orientation model that indicated learners could be grouped into six goal orientations groups, including task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance.

Although many important studies in the past four decades have revealed learner metacognition affects their learning (Flavell, 1979, 1987; Mihalca, Mengelkamp, & Schnotz, 2017; Schraw & Dennison, 1994; Shin et al., 2003), few studies described learning process differences based on learner metacognitive differences. Goal orientation has also been studied extensively for the past four decades (Locke & Latham, 2002; Middleton & Midgley, 1997; Ryan, 1970; Won, Wolters, & Mueller, 2017) and important advances have been made; however, the relationship between goal orientation and problem-solving processes in SG environments is still unclear. In addition, although scholars have suggested metacognition and goal orientation together could result in learner academic success, i.e., high GPA (Gul & Shehzad, 2012), there are few studies that have analyzed the impact of the interaction between metacognition and goal orientation on learner problem-solving performances and processes. Therefore, more research is needed on how metacognition and goal orientation differences would affect learner problem-solving in SG environments, including both problem-solving performances and processes.

To examine the impact of learner metacognition and goal orientation on problem-solving performances and processes in a SG environment, this study will employ Serious Game Analytics (SGA) technique, which could provide insights on learner game activities using various analysis techniques and software tools (Loh, Sheng & Ifenthaler, 2015). Specifically, as noted by

Zimmerman (2013), computers can be a valuable instrument for studying metacognition, because "students' learning processes and outcomes can be stored, analyzed, and graphed in various ways for students and researchers to uncover underlying strengths and deficiencies" (p. 165). Thus, in addition to use surveys, stimulated recall interviews and gameplay screencasts to gather data, this study will collect learner computer logged activity data, which will be utilized to analyze and visualize learner problem-solving processes in the SG.

Research findings will offer insights on the impact of learner characteristics on learner problem-solving performances and processes in SG environments. Particularly, this study will fill the gap in the lacking of research on metacognition and goal orientations in SG environments that adopt the PBL pedagogy. In addition, research results will elaborate learner problem-solving performance and process differences based on goal orientation and metacognition, which will help future design of these environments and help teachers to effectively provide support to learners based on learner characteristics while using SG environments.

PURPOSE OF THE STUDY

Based on constructivist theory, the purpose of this study is to examine the impact of learner characteristics (i.e., metacognition and goal orientation) on learner problem-solving (i.e., problem-solving performances and processes) in SG environments that adopt PBL pedagogy. To accomplish this purpose, this study will investigate learner metacognition, goal orientation, problem-solving performances and processes in a SG environment called Alien Rescue (AR), which adopts the PBL pedagogy for students to learn the space science in science subject matter.

RESEARCH QUESTIONS

This study will investigate the impact of learner characteristics (i.e., metacognition and goal orientation) on problem-solving (i.e., problem-solving performance and problem-solving process) in a SG environment for learning space science. The research questions for this study are:

1. To what extent are problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)? There are three sub-questions:

(a) Is there a statistically significant difference in learner problem-solving performances based upon metacognition (high metacognitive level, low metacognitive level) and goal orientation (task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation)?

(b) Can learner metacognition and goal orientation predict learner problem-solving performance?

(c) What are the reasons for any problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)?

2. To what extent are problem-solving process differences based on learner characteristics (i.e., metacognition and goal orientation)? There are five sub-questions:

(a) What are learner problem-solving process patterns?

(b) Are there any problem-solving process pattern differences among learners based on their metacognition?

(c) Are there any problem-solving process pattern differences among learners based on their goal orientation?

(d) Are there any problem-solving process pattern differences based on the interaction between learner metacognition and goal orientation?

(e) What are the reasons for problem-solving process pattern differences based on learner characteristics (i.e., metacognition and goal orientation)?

DEFINITIONS OF KEY TERMS

Constructivism

Constructivism is a learning theory that emphasizing knowledge is subjective, memory is constructed reality, and learner plays an important role in constructing own knowledge.

Learner Characteristics

The concept of learner characteristics is used in the sciences of learning and cognition to designate a target group of learners and define those aspects of the personal, academic, social, or cognitive self that may influence how and what the group learns.

Metacognition

Metacognition is learner knowledge and cognition about cognitive phenomena, including both metacognitive knowledge and metacognitive regulation.

Goal Orientation

Goal orientation is learner general purpose toward cognitive tasks. This study will use the 3 X 2 goal orientation model, which includes task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientations.

Problem-based Learning (PBL)

PBL is learner-centered constructivist instructional method that embeds student learning processes in solving ill-structured problems.

Problem-solving Process

Problem-solving process is a goal-oriented activity, in which the learner finds a solution for the problem during PBL. This study will identify learner problem-solving processes by examining learner computer activity log data, including both room visit and tool use activity.

Problem-solving Performance

One measurement of learner outcome during PBL is problem-solving performance. This study will evaluate learner problem solution quality in Alien Rescue, i.e. the rational for sending the alien to a corresponding planet. Learner solution score will be determined by how well he or she solved the problem of finding an appropriate relocation home for the alien Jakalay-Tay, which will be evaluated using an 8-point (0 to 7 points) rubric.

Serious Games (SG)

SG are digital games and simulation tools that are created for non-entertainment use, and with the primary purpose of improving skills and performance of learners through training and instruction.

Serious Games Analytics (SGA)

SGA are methods and techniques to collect and analyze user data to obtain actionable insights in order to improve the learning design of SG.

Chapter 2: Literature Review

This study is based on constructivism learning theory, integrating ideas from three aspects of this theory in educational literature including: (a) learner characteristics (i.e., metacognition and goal orientation), (b) problem-based learning (PBL), and (c) serious games (SG). Constructivism theory and each of the above topics are discussed in this chapter.

CONSTRUCTIVISM

In this section I will review constructivism learning theory in terms of its assumptions regarding knowledge, memory, learning, instructional design, and learning environments.

Knowledge

Constructivists believe that knowledge is subjective, which "arises from the active subject's activity, either physical or mental" (von Glasersfeld, 1995, p. 56). In addition, knowledge cannot be simply passed onto a learner from someone who has already known (Piaget, 1953). Rather, learners personalize the information into knowledge through observation, processing, and interpretation (Ertmer & Newby, 2013). Therefore, no two learners will possess the same prior knowledge and they will not construct the same understanding.

Memory

Constructivists believe that memory is a brain function which is distributed over the whole neuronal system and organizes itself based on its own history, which means memory does not necessarily represent but rather constructs reality (Schmidt, 2008). In addition, memory is always under construction as a cumulative history of interactions (Ertmer & Newby, 2013), and it is not isolated from previous learning but is a conscious process associated with past real-life experience (Bartlett, 1932). Constructivists suggested that the more conscious individuals are the less they guess or invent material during memory experiment recall (Gauld & Stephenson, 1967). Furthermore, in education, constructivists emphasize that memory is not for retrieving intact knowledge structures, but for updating learner knowledge of the world and providing learner with the means to use pre-existing knowledge to solve the problem at hand (Ertmer & Newby, 2013).

Learning

Constructivists consider learning as a mental activity. Learners play important roles in constructing their own knowledge rather than passively receiving it from someone who knows (Piaget, 1953). From the constructivist perspective, learning requires learner self-regulation; through observation, processing, and interpretation, learner constructs an independent reality based on personal experiences and perceptions of the world during the learning process (Cunningham & Duffy, 1996; von Glasersfeld, 1995). For example, instead of just listening, reading, and working through routine exercises, learners discuss, debate, hypothesize, investigate, and take viewpoints during learning (Perkins, 1999; Duffy & Jonassen, 2013).

Learner Characteristics

Based on the constructivist view of knowledge, memory, and learning, learner characteristics play critical roles while designing instruction in learning environments. According to Kirschner and Drachsler (2012), learner characteristics are used in learning sciences and cognition areas to describe learner "personal, academic, social/emotional, and/or cognitive characteristics that may influence how and what they learn" (p. 1743). Personal characteristics are related to learner demographic information such as age, gender, language, social economic status, cultural background and specific needs of a learner group such as disabilities and/or impairments in learning. Academic characteristics such as learning goals, prior knowledge, and educational level are more education/learning related. Social/emotional characteristics such as self-efficacy, mood, and sociability "relate to the group or individual with respect to the group" (Kirschner & Drachsier, 2012, p. 1743). As for cognitive characteristics, these refer to learner "attention span, memory, mental procedures, and intellectual skills, which determine how a learner perceives, remembers, thinks, solves problems, organizes and represents information in her/his brain" (Kirschner & Drachsier, 2012, p. 1743).

Researchers have investigated learner personal characteristics such as socioeconomic status, age, gender, and race/ethnicity (Lemke et al., 2004; OECD, 2014). Studies have also examined learner academic characteristics such as prior knowledge and goal orientation (Elliott & Dweck, 1988; Hsieh et al., 2008; Liu, Kang, Lee, Winzeler, & Liu, 2015); social/emotional characteristics such as beliefs and self-efficacy (Jonassen, 2000; Liu, Cho, & Schallert, 2006); and cognitive characteristics such as memory and metacognition (Davidson & Sternberg, 1998; Shin et al., 2003). Scholars suggested that there were often large characteristic differences between different groups of learners such as children, adults, students, professionals, and disabled people (Kirschner & Drachsier, 2012). These learners might differ in their motivation, prior knowledge, expertise level, study time, and physical abilities. Therefore, by taking account of learner characteristics, instructional designers are expected to design and develop "tailored instructions for a target group" (Kirschner & Drachsier, 2012, p. 1743), which can be more efficient, effective and/or motivating.

Instructional Design

Constructivist perspectives on knowledge, memory, learning, and learner characteristics have significant implications for instructional design. Constructivist designers encourage learners to construct their own understandings through solving real-world problems. They also validate new learner perspectives through his/her own learning experiences (Cunningham & Duffy, 1996; Jonassen, 1991; Lebow, 1993). Constructivists favor instructional methods and strategies where learners actively explore complex topics or environments and move closer to thinking as an expert might think in a given content area. Take learning computer programming as an example. A typical constructivist goal in teaching would not be to teach novice computer science students straight facts about programming languages, but rather to prepare students to use programming languages as a developer might use them. In this way, student performance is related to the processes of construction rather than to the content.

Some specific instructional strategies utilized by constructivists include PBL (solving a real-world problem) (Hmelo-Silver, 2004; Savery & Duffy, 1995), cognitive apprenticeship (coaching a student toward expert performance) (Brown, Collins, & Duguid, 1989; Resnick, 1987), collaborative learning (working together to develop and share alternative views) (Johnson, Johnson, & Smith, 1998), debate and discussion (Duffy & Jonassen, 2013), and scaffolding (providing guidance during the constructive processes) (Brush & Saye, 2001; Vygotsky, 1978).

Learning Environments

Constructivist perspectives on instructional design foster the design of learning environments, which are needed to help learners construct their own knowledge in an authentic context. Whether it is a computer-based environment or classroom-based environment, learning environments that make use of simulation, collaborative learning, problem-solving and inquiry, and apprenticeship are all consistent with constructivist views on learning (Wilson, 1996). For example, designers and developers can create SG environments that are based on constructivism, because they usually engage learners in exploring, discovering, and questioning in an authentic context, as well as encourage learners to construct their own knowledge during the game play (Rieber, 1996; Yang, 2012). In addition, learning environments adopting PBL pedagogy are also based on constructivism, because these environments (a) have an authentic task or problem to reflect the complexity of a real world situation; (b) anchor all learning activities to this task or problem; (c) support the learner to develop ownership of their own learning during the process; (d) challenge learner thinking; (e) encourage testing alternative views and ideas; and (f) provide refection opportunities on both the learning content and process (Savery & Duffy, 1995).

METACOGNITION AND GOAL ORIENTATION

Based on constructivist learning theory, learner characteristics are important for instructional designers to understand learners and design the learning environment. This section will review literature on two important learner characteristics for learning in SG environments—

metacognition and goal orientation. Particularly, to provide evidence for this study, I will mainly focus on the definition and measurements for each characteristic.

Metacognition

According to Kirchner and Drachsier (2012), metacognition is categorized as a learner cognitive characteristic. Since the late 1970s, researchers have concluded that learner metacognition plays an important role in learning (Flavell, 1979, 1987; Schraw & Dennison, 1994), because it "allows individuals to plan, sequence, and monitor their learning in a way that directly improves performance" (Schraw & Dennison, 1994, p. 460). In addition, for technology-based environments, researchers have suggested that enhanced metacognitive activities can lead to higher recall, comprehension, and deeper understanding (Bannert & Reimann, 2012; Lee, Lim, & Grabowski, 2010; Poitras, Lajoie, & Hong, 2012). The following paragraphs will mainly review the definition and measurements of metacognition.

Definition

Flavell (1979) defined metacognition as the "knowledge and cognition about cognitive phenomena" (p. 906). It is the awareness of one's own knowledge, of one's actions, and of one's current "cognitive or affective state" (Hacker, 1998, p. 3). Specifically, metacognition includes learner knowledge about how they learn new knowledge or skills, what strengths and weaknesses they have in learning, and what strategies they use in the area of study. Scholars further indicated that metacognition has two parts—metacognitive knowledge and metacognitive regulation (Brown, 1987; Flavell, 1979, 1987; Schraw & Dennison, 1994).

Metacognitive knowledge is the knowledge about cognition, which refers to knowledge about yourself, learning strategies, and knowledge about when, why, and how to use these strategies. Flavell and Wellman (1977) first proposed four types of metacognitive knowledge, including: (a) tasks—knowledge about how the nature of the task influences the task performance; (b) self—knowledge about one's own skills, strengths, and weaknesses; (c) strategies—knowledge regarding the alternative strategies for performing the task; and (d) interactions—knowledge about the preceding types of knowledge interact with one another to influence the outcome of cognitive performance. This taxonomy later evolved into three types: (a) declarative knowledge—knowing about self and strategies, (b) procedural knowledge—knowing how to do things or use strategies, and (c) conditional knowledge—knowing why and when to use strategies (Moshman, 2017; Schraw & Dennison, 1994).

As for metacognitive regulation, it refers to the control aspect of learning, such as planning (setting goals, predicting outcomes, scheduling strategies, and allocating resources), monitoring (testing, revising and re-scheduling during learning), and evaluation (appraising the effectiveness of regulation or learning). Metacognitive regulation is often referred to as self-regulation (Moshman, 2017; Schraw & Dennison, 1994; Sperling, Howard, Staley, & DuBois, 2004; Zimmerman, 2002, 2013). Zimmerman and Campillo (2003) suggested self-regulation ran through the entire problem-solving process in three phases, including forethought, performance, and self-reflection. In the forethought phase, there are two major categories: task analysis and self-motivation belief. During task analysis, problem-solvers would engage in goal setting (Latham & Locke, 1979, 1991; Locke & Latham, 2002, 2015) and strategic planning (Weinstein & Mayer, 1986; Zimmerman, 2002). Underlying forethought processes of goal setting and strategic planning there are several key factors, including self-efficacy, outcome expectations, intrinsic interest, and goal orientation (Zimmerman & Campillo, 2003). Particularly, for goal orientation, scholars have pointed out that metacognition and goal orientation have a strong correlation during learning (Gul & Shehzad, 2012). We will review the literature on goal orientation in further sections.

The second phase of self-regulation is the performance phase, which includes self-control and self-observation. Self-control can help problem-solvers to focus on the task while self-observation can help problem-solvers track their own performance (Zimmerman & Campillo, 2003; Zimmerman & Paulsen, 1995). The last phase of self-regulation is the self-reflection phase, which involves self-judgement and self-reaction. Self-judgement refers to evaluating one's own problem-solving performance and outcomes, while self-reaction includes self-satisfaction and adaptive inferences (Zimmerman & Campillo, 2003; Zimmerman & Martinez-Pons, 1992).

Zimmerman and Campillo (2003) gave an example in a medical practice setting. A family physician needs to treat an eight-year-old boy with a "breathing" problem. In this case, during the forethought process, the physician must define the medical problem, when and where breathing problems occur, and select an appropriate medication strategy. As for the physician's motivation, he or she might believe this case is beyond his/her level or that the treatment strategy might work. With respect to the performance phase, the physician will give drugs to the boy and monitor his reaction. In terms of the self-reflection phase, the boy might be cured, and the physician's self-efficacy in managing this type of case might be strengthened; or if the boy's condition gets worse, the physician might choose another treatment, or refer the boy to a specialist.

In addition, scholars suggested that self-regulated learners have three main characteristics: (a) intrinsically motivated—they find participating in learning activities to be its own reward and do not seek external incentives; (b) metacognitive active—they actively engage in planning, goal-setting and are able to monitor and evaluate learning effectiveness; and (c) behaviorally active—they select and use learning strategies to best suit their own learning needs (Ertmer & Newby, 1996; Zimmerman, 1990).

Measurements

Because of the importance of metacognition in learning process, researchers have employed both quantitative and qualitative methods to measure and study it. In quantitatively measuring learner metacognitive skills, the most influential instruments are the How Do You Solve Problems? (HSP) (Fortunato, Hecht, Title, & Alvarez, 1991), Inventory of Metacognitive Self-Regulation (IMSR) (Howard, McGee, Shia, & Hong, 2000), and the Metacognitive Awareness Inventory (MAI) (Schraw & Dennison, 1994).

HSP includes 21 statements to measure both metacognitive knowledge and metacognitive regulation (Fortunato et al., 1991). The statements are grouped into four sections: (a) interpreting the problem and planning the solution strategy before beginning to solve the problem; (b) monitoring the solution processes during problem-solving; (c) evaluating the execution after finishing solving the problem; (d) specific strategies of solving the problem. Learners can respond by selecting "YES - Yes, I did do this," "MAYBE - I may have done this," or "NO - No I didn't do this." Although Fortunato and his colleagues did not provide validity and reliability estimates for HSP, it was the first metacognition measurement and provided a theoretical foundation for future measurement development.

Based on HSP, Howard, McGee, Shia, and Hong (2000) developed the Inventory of Metacognitive Self-Regulation (IMSR). IMSR is a 32-item self-report inventory, which assesses student metacognitive skill (age 12 to 18) in the mathematical and scientific problem-solving context. It measures five dimensions: (a) knowledge of cognition—understanding one's cognitive abilities (e.g., "I can make myself memorize something", p.10), (b) objectivity—standing outside and thinking about one's learning (e. g., "When I am done with my schoolwork, I ask myself if I learned what I wanted to learn", p. 10), (c) problem representation—understanding the problem fully before proceeding (e. g., "I try to understand what the problem is asking me", p. 9), (d) subtask monitoring—breaking the problem down into subtasks and monitoring each subtask (e. g., "I ask myself if there are certain goals I want to accomplish", p. 10), and (e) evaluation—double-checking the problem-solving process (e.g., "I double-check to make sure I did it right", p. 10). The IMSR's reliability alpha is 0.93. Bulu and Pedersen (2012) used the IMSR to measure metacognitive skills of 322 students ranging from 11 to 12 years old in their study, and reported the reliability had a Cronbach's alpha of 0.89.

Another well-known measurement is MAI, which is a 52-item self-report inventory for measuring adult metacognitive knowledge and metacognitive regulation. There are 17 metacognitive knowledge items; a sample item is, "I understand my intellectual strengths and weaknesses." There are 35 metacognitive regulation items; a sample item is, "I ask myself periodically if I am meeting my goals." The measurement was originally scored on a 100-point, bipolar scale, with 0 being "totally untrue of me" and 100 being "totally true of me." The scale demonstrates high reliability ($a = .90$) and significant correlations between these two components in two studies ($r = .54$ and $r = .45$, respectively) (Schraw & Dennison, 1994). Many researchers directly adopted this measurement (Young & Fry, 2008) while some made modifications based on this measurement (Akin, Abaci & Çetin, 2007; Bendixen & Hartley, 2003; Corliss, 2005). Some even created new measurements for specific subject domains based on MAI (Mokhtari & Reichard, 2002; Sperling, Howard, Miller, & Murphy, 2002).

Young and Fry (2008) directly used the MAI to study 178 college students (45 undergraduate students and 133 graduate students) to examine how their metacognition related to academic achievement. The results suggested significant correlations between student metacognition and overall GPA ($r = .23$, $p < 0.01$) as well as course grades ($r = .19$, $p < 0.05$). They also found MAI scores are significantly different between graduate and undergraduate students on their metacognitive regulation $F\,(1,177) = 4.13$, $p < 0.05$.

As for researchers who modified the MAI, they mostly modified the MAI scoring or translated MAI items into another language, such as Turkish (Akin et al., 2007). Corliss (2005) modified the MAI scoring to a 5-point scale with 1 being "very untrue of me" to 5 being "very true of me" rather than 100 points for each item. She measured pre- and post-metacognitive awareness during student problem-solving processes using this modified scale. The results indicated the MAI was not the best measure of student metacognitive awareness for her study because individual scale scores were not correlated across time. She further suggested that metacognitive awareness might be better assessed through qualitative measures rather than using a quantitative instrument such as the MAI. For example, she suggested that observing students problem-solving behavior and interviewing them about their actions over time might be a better way to truly assess metacognitive awareness.

In addition to minor modification to the MAI, many researchers created new measurements based on MAI for a specific subject area. Mokhtari and Reichard (2002) developed the Metacognitive Awareness of Reading Strategies Inventory (MARSI), which was designed to assess adolescent and adult reader metacognitive awareness while reading academic or school-related materials. The MARSI has 30 statements, and participants can rate each statement on a scale from 1 (I never or almost never do this) to 5 (I always or almost always do this). Mokhtari and Reichard (2002) reported a high reliability for this measurement (*a* = .93). Sperling, Howard, Miller, and Murphy (2002) also developed a measurement called Junior Metacognitive Awareness Inventory (Jr. MAI) based on MAI, which has two versions. The Jr. MAI Version A includes 12 items with a 3-point Likert scale (1 = never; 2 = sometimes, 3 = always) and is intended for use in grades 3 through 5. The second version (Jr. MAI, Version B) has 6 additional items to Version A's 12 items and uses a 5-point Likert scale for using in grade 6 through 9. Sperling et al. (2002) did not report the validity and reliability for Jr. MAI, but their findings indicated that both versions had high correlations to HSP measurement (A: $r = .72$; B: $r = .68$).

Beside quantitative measurements, scholars have adopted qualitative methods to assess metacognition, such as think-aloud and interview. For think-aloud, Pressley and Afflerbach (1995) summarized 38 studies that had used this method to examine students cognitive and affective processes in their reading behavior. Hu and Gao (2017) also reviewed 29 papers that used think-aloud protocol in self-regulated reading research. They found that although researchers had relied on think-aloud to study self-regulated reading, this approach had been criticized for collecting inaccurate and incomplete reflections of learner thoughts. For example, studies reported participants (both junior students and adults) had difficulties in verbalizing their metacognition during think-aloud task, or only reported things they thought were important. Some participants were even unable to think aloud because they were occupied by the task (Barkaoui, 2011; Brach, 2001). Therefore, Hu and Gao (2017) suggested that although think-aloud was an important methodological tool to collect learner verbal data about metacognition, these data only indicated "a part of self-regulated reading process rather than a full picture of it" (p. 188).

Researchers also conducted structured interviews to study metacognition, especially on student metacognitive regulation (Zimmerman, 2013; Zimmerman & Martinez-Pons, 1986). Zimmerman and Martinez-Pons (1986) developed the Self-Regulated Learning Interview Scale (SRLIS), which provided a protocol for interviewing students on different strategies they might use during their metacognitive regulation, including: (a) self-evaluation, (b) organizing and transforming, (c) goal-setting and planning, (d) seeking information, (e) keeping records and monitoring, (f) environmental structuring, (g) self-consequences, (h) rehearsing and memorizing, (i) seeking social assistance from peers, teachers, and adults, (three separate categories) (j) reviewing records such as tests, notes, or textbooks (three separate categories), and (k) other. Zimmerman (2013) also noted that computers can be a valuable instrument to study metacognitive regulation in the future, because "students' learning processes and outcomes can be stored, analyzed, and graphed in various ways for students and researchers that uncover underlying strengths and deficiencies" (p. 145).

In summary, metacognition is the awareness of one's own knowledge, of one's actions, and of one's current "cognitive or affective state" (Hacker, 1998, p. 3), which includes metacognitive knowledge and metacognitive regulation. Researchers have employed both quantitative and qualitative methods to study learner metacognition. For quantitative measurement, MAI is the most popular measurement for assessing adult learner metacognition. With respect to qualitative methods, researchers suggested that think-aloud and interview were important ways to collect some parts of data on learner metacognition, and it is suggested that computers can be a valuable way to gather more information regarding learner metacognition.

Goal Orientation

Different from metacognition, goal orientation was categorized as a learner academic characteristic by Kirchner and Drachsier (2012). It is an important learner characteristic because having conscious goals will affect learner action (Locke & Latham, 2002; Pintrich, 2000a; Ryan, 1970) and learners display different behavior patterns and cognitive performances according to their goal orientation (Dweck, 1986; Elliott & Dweck, 1988; Meece & Holt, 1993). Goal orientation has been studied intensively as a factor to predict and understand learning outcomes during self-regulation (Locke & Latham, 2002; Middleton & Midgley, 1997; Porath & Bateman, 2006; Schmidt & Ford, 2003; Sideridis, 2008; Wolters, Yu, & Pintrich, 1996; Won et al., 2017). In addition, researchers have defined several goal orientation models and developed measurements (Ames, 1992; Dweck, 1986; Elliot & Church, 1997; Elliot et al., 2011; Elliot, Murayama, Kobeisy,

& Lichtenfeld, 2015; Pintrich, 2000a). This section will review these goal orientation models and their measurements.

Definition

Researchers suggested that individuals had different reasons or goals towards cognitive tasks, which are collectively defined as goal orientation (Ames, 1992; Dweck, 1986; Elliot & Church, 1997; Pintrich, 2000a; Shim & Ryan, 2005). Goal orientation models have evolved for the past three decades. At the very beginning, researchers suggested learners might have two different goal orientations: mastery and performance goal orientation (Ames, 1992; Dweck, 1986; Locke & Latham, 1990; Winters & Latham; 1996), which involved into three goal orientations: mastery, performance-approach and performance-avoidance orientation (Elliot, 1994; Elliot & Church, 1997; Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002; Middleton & Midgley, 1997). Through continuing exploration, researchers proposed a 2 X 2 goal orientation model (performance and mastery, approach and avoidance), which suggested there were four different goal orientations (i.e., mastery-approach, mastery-avoidance, performance-approach and performance-avoidance) (Elliot, 1999; Elliot & McGregor, 2001; Pintrich, 2000a; Linnenbrink & Pintrich, 2001), which then evolved into a 3 X 2 goal orientation model (three definitions: task, self, and other; two values: approach and avoidance) with six goal orientations (i.e., task-approach, self-approach, other-approach, task-avoidance, self-avoidance, and other-avoidance) (Elliot et al., 2011). Four years later, Elliot et al. (2015) proposed two more goal orientations (i.e., potential-approach and potential-avoidance) based on the 3 X 2 model. In addition, researchers started to collectively adopt a multiple-goal orientations perspective, which indicated individual can be motivated by endorsing more than one goal orientation (Barron & Harackiewicz, 2001; Button, Mathieu, & Zajac, 1996; Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008; Pintrich, 2000b; Zusho et al., 2005). Definitions on these goal orientation models are discussed in the following paragraphs.

Two goal orientations

Between the early-1980s and mid-1990s, scholars identified two goal orientations involving competence—mastery (or learning) and performance. These two goal orientations were considered at opposite ends for cognitive tasks (Ames, 1992; Dweck, 1986; Dweck & Leggett, 1988; Elliott & Dweck,1988; Locke & Latham, 1990). It is worth noting that in the human resources management field, Locke and Latham (1990) adopted "learning" goal orientation while Ames (1992) chose to label it as "mastery" goal orientation in the educational psychology field. "Mastery goal orientation" is used most prevalently in academic publication.

Mastery (or learning) goal orientation was considered as the orientation towards developing ability, and learners with this goal orientation pursued tasks because they wanted to develop competence, knowledge, understanding, skills and to achieve a sense of mastery (Ames, 1992; Dweck, 1986; Locke & Latham, 1990). On the other hand, performance goal orientation focused on demonstrating competence, and learners with this goal orientation pursued tasks because they wanted to manage the impression of their ability, especially through comparisons with others (Dweck, 1986; Locke & Latham, 1990; Nicholls, 1984).

In addition, studies suggested that personal and contextual factors often made people have a dominant goal orientation (Meece & Holt, 1993; Van Yperen, 2006), and it can be identified at different levels (e.g., high, medium, or low on mastery goals and high, medium, or low on performance goals) (Ciani & Sheldon, 2010; Latham & Locke, 1991). For example, Meece and Holt (1993) used hierarchical cluster analysis to classify 261 students based on their mastery, ego, and work-avoidant goal orientations. They identified 3 clusters including high-mastery group, mastery-ego group and low mastery-ego group. Furthermore, Locke, Latham and their colleagues' 35-year studies on goals suggested that specific high-mastery goals lead to higher performance when people initially lacked the knowledge or skill to perform the task (Latham & Brown, 2006; Porter & Latham, 2013; Seijts & Latham, 2005, 2011; Seijts, Latham, Tasa, & Latham, 2004; Winters & Latham, 1996). So, it was suggested to set mastery goals when individuals lacked the ability to perform the task and set performance goals when they had the ability to attain a desired performance level (Seijts, Latham, & Woodwark, 2013).

Studies within the educational psychology field suggested the mastery goal orientation usually did not predict exam performance because learners might neglect boring topics and focus their efforts to preferred ones (Elliot & Church, 1997; Senko & Miles, 2008; Zusho et al., 2005). However, in previous studies, mastery goal orientation usually demonstrated a positive effect on intrinsic motivation, self-efficacy, self-regulated learning, persistence, preference for challenge, attitudes, and general well-being (see Ames & Archer, 1988; Dweck & Leggett, 1988; Elliot & Church, 1997; Harackiewicz, Barron, & Elliot, 1998; Kaplan & Flum, 2010; Sosik, Chun, & Koul, 2017; Pintrich, 2000a; Senko & Miles, 2008; Wolters et al., 1996). In addition, studies also reported that mastery goal orientation was positively related to age (Button et al., 1996; Utman, 1997), which might be because when people grow older they become more concerned with their own expectations rather than others'.

Opposite to mastery goal orientation, the results on performance goal orientation were not consistent (see reviews in Kaplan & Maehr, 2007). On the one hand, performance goal orientation was usually associated with negative learning processes and outcomes (Dweck & Leggett, 1988; Meece & Holt, 1993; Utman, 1997), such as using surface or defensive strategies rather than deep learning strategies (Ames, 1992; Dweck, 1986). Performance goal orientation also had negative effects in learning involving challenges or difficulty (Dweck, 1986), because individuals might sacrifice learning opportunities that involve challenges or risks to look smart. On the other hand, researchers have found performance goal orientation had more positive effects than mastery goal orientation on test scores (Harackiewicz, Barron, Carter, Lehto, & Elliot, 1997; Harackiewicz et al., 1998; Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000; Wolters et al., 1996). These findings were replicated across different academic subject areas (e.g., English, math, and social studies) for middle or high school students, and in different classroom settings for college students.

Three goal orientations

To resolve the inconsistent result pattern of performance goal orientation, researchers made a distinction between "approach" and "avoidance" orientations within performance goal orientation and proposed a three-goal orientation model—mastery, performance-approach and performance-avoidance goal orientations (Elliot, 1994, 1999; Harackiewicz et al., 2002; Middleton & Midgley, 1997; Vandewalle, 1997). According to Elliot (1994), an approach orientation referred to learners who focused on achieving success at a task, whereas an avoidance orientation referred to learners who focused on avoiding failure or other negative outcomes. Students with a performance-approach goal orientation strive to make a positive impression by outperforming others, while students with a performance-avoidance goal orientation avoid a negative impression by avoiding situations that would demonstrate their low ability. Elliot and Church (1997) suggested mastery goal orientation facilitated intrinsic motivation, performance-approach goal orientation enhanced graded performance (e.g., test scores), and performance-avoidance goal orientation harmed both intrinsic motivation and graded performance.

After distinguishing performance-approach and performance-avoidance goal orientations, several studies indicated performance-avoidance goal orientation was strongly related to negative outcomes, such as avoiding seeking help (Middleton & Midgley, 1997), diminished motivation (Elliot & Church, 1997; Shim & Ryan, 2005), anxiety (McGregor & Elliot, 2002; Middleton & Midgley, 1997; Zusho et al., 2005), and low grades (Elliot & Church, 1997). For performance-approach goal orientation, a few studies suggested it was also related to some negative outcomes such as disruptive behavior and only focusing on grades rather than deep processing (Elliot et al., 1999; Kaplan, Gheen, & Midgley, 2002; Midgley et al., 2001), however, it was mostly associated with positive outcomes, such as higher motivation, persistence, and higher graded performance (Elliot et al., 1999; Elliot & Church, 1997; McGregor & Elliot, 2002; Shim & Ryan, 2005; Zusho et al., 2005). Some studies also suggested performance-approach goal orientation could be beneficial in certain contexts (e.g., a competitive college setting) and for older students (Harackiewicz et al., 2002; Pintrich, 2000a).

2 X 2 goal orientations

After dividing performance goal orientation into performance-approach and performance-avoidance goal orientation, researchers proposed to divide mastery goal orientations (Elliot, 1999; Elliot & McGregor, 2001; Pintrich, 2000a). Elliot (1999) suggested that previous studies assumed mastery goal orientation only represents an approach form and overlooked the avoidance aspect. According to him, all goal orientations should have two dimensions based on learner competence: definition (mastery or performance) and valence (approach or avoidance). Elliot and McGregor (2001) proposed a 2 X 2 goal orientation framework, which contained mastery-approach (focused on attaining skills and competence), mastery-avoidance (focused on avoiding loss of skill and competence), performance-approach, and performance-avoidance goal orientation.

Researchers suggested outcome patterns for learners with mastery-avoidance goal orientation were more negative than those with mastery-approach goal orientation, but more positive than those with performance-avoidance goal orientation; however, there is little research on mastery-avoidance goal orientation (Elliot & McGregor, 2001; Elliot & Murayama, 2008). A few published studies suggested that mastery-avoidance goal orientation had positive effect on deep processing during learning (Elliot & McGregor, 2001) while others reported it was associated with anxiety (Sideridis, 2008), fear of failure (Elliot & McGregor, 2001; Sideridis, 2008), and fear of seeking help (Karabenick, 2003). One study of 333 college students (Age: M = 19.9) reported mastery-avoidance goal orientation was neither positively nor negatively related to graded performance or college student interest related to the class (Van Yperen, 2006). In addition, Ciani and Sheldon (2010) suggested mastery-avoidance goal orientation might be uncommon compared to other three types of goal orientations.

3 X 2 goal orientations

A decade after Elliot and Mcgregor (2001) proposed the 2 X 2 goal orientation model, Elliot, Murayama, and Pekrun (2011) proposed a 3 X 2 goal orientation model by re-constructing the goal orientation based on two competence dimensions—definition and valence. Specifically, definitions of competence include absolute (task), intrapersonal (self), and interpersonal (other), and two valences of competence include positive (approach) and negative (avoidance) (Elliot et al., 2011). Thus, the 3 X 2 model has six goal orientations: (a) task-approach (TAP: e.g., Do the task right), (b) task-avoidance (TAV: e.g., Avoid doing the task wrong), (c) self-approach (SAP: e.g., Do the task better than before), (d) self-avoidance (SAV: e.g., Avoid doing the task worse than before), (e) other-approach (OAP: e.g., Do better than others), and (f) other-avoidance goals (OAV: e.g., Avoid doing worse than others). See Table 1 for the model.

Table 1. The 3 X 2 Achievement Goal Orientation Model

		Definition		
		Absolute (task)	Intrapersonal (self)	Interpersonal (other)
Valence	Positive (approaching success)	Task-approach	Self-approach	Other-approach
	Negative (avoiding failure)	Task-avoidance	Self-avoidance	Other-avoidance

According to this model, task-based goal orientation uses task demand as the evaluative standard—getting an answer correct, understanding an idea or finishing a task. Self-based goal orientation uses one's own trajectory as the evaluative standard—doing well or poorly compared to how one has done in the past. Other-based goal orientation uses an interpersonal evaluative standard—doing well or poorly compared to others. Elliot et al. (2011) mentioned that they divided previous mastery goal orientation into task- and self-based goal orientation, because these two goal orientations were usually independent in daily life, and the self-based goal orientation required more cognitive capacity to recognize both the previous and present outcome of a learner (Elliot et al., 2011). Taking two learners who are working on ice skating for example, the first learner may simply want to glide one loop of the rink smoothly for 3 loops without stopping (task-based goal orientation), while the second learner may work on gliding one more loop than what he or she did yesterday (self-based goal orientation). According to Elliot et al. (2011), the second learner needs more cognitive capacity because he or she needs to recognize both the previous and present outcome compared with the first learner who only needs to focus on the current status.

Other goal orientations

Four years later, Elliot and his colleagues (2015) again proposed two more goal orientations: potential-approach and potential-avoidance goal orientations. They pointed out that the self-based goal orientations in the previous 3 X 2 framework were defined in terms of the past, which meant all the self-based goal orientation items were compared to what the learner did before. One example was "My goal is to do well on the exams in this class relative to how well I have done in the past on such exams," which did not mention the potential-based goal orientation (e.g., "My goal is to do as well as I can possibly do on the exams in this class"). Elliot and his colleagues did not simply add the new potential-based goal orientations to the 3 X 2 model to construct a 4 X 2 goal orientation model, as doing so would "disproportionately weight self-based standard over task-based and other-based" (p. 201) in the measurement.

Multiple goal orientations

Previous experimental research studied the impact of different goal orientations on student outcomes and usually suggested one orientation was more positive than the other. These research designs did not allow for multiple goal orientations or test the interaction between different goal orientations. Currently, many researchers endorse a multiple goal orientation perspective, suggesting students can adopt multiple goal orientations simultaneously and have different levels of both goal orientations to benefit their learning (Barron & Harackiewicz, 2001; Bereby-Meyer & Kaplan, 2005; Daniels et al., 2008; Harackiewicz et al., 2008; Jang & Liu, 2012; Meece & Holt, 1993; Pintrich, 2000b; Pintrich, Conley, & Kempler, 2003; Zusho et al., 2005).

Pintrich (2000b) argued that in some classrooms mastery and performance goal orientations were slightly positively related to each other. Moreover, he suggested that different patterns in the levels of the two goal orientations might lead to the same learning outcomes. This indicated a potential interaction between mastery and performance goal orientations, which might generate "multiple pathways" or "trajectories" (Pintrich, 2000b, p. 545). In addition, he proposed that learners with different goal orientations might follow different trajectories and have different experiences over time, but end up with the "same" achievement or performance (Pintrich, 2000b). He used the experience on a trip as a metaphor to demonstrate this idea. For example, mastery goal orientation learners might have a nicer experience (i.e., positive motivation or self-efficacy) along the way to their destination (i.e., performance or achievement). Performance goal orientation students could arrive at the same destination, but might have experienced a rocky road along the way (i.e., less interest, more anxiety, or negative affect).

Similar with Pintrich's (2000b) hypothesis, studies reported that learners did possess multiple goal orientations during learning and different goal orientations were related to different positive outcomes (Barron & Harackiewicz, 2001; Wolters et al., 1996). Barron and Harackiewicz (2001) studied 166 undergraduate student goal orientation and their math performance. The results indicated a positive correlation between mastery and performance goal orientation, $r\ (164) = .31$, $p < .05$, suggesting participants who adopted mastery goal orientation, rather than focusing on one goal orientation during the study session (45 minutes), were also likely to adopt performance goal orientation for the session. In addition, Barron and Harackiewicz (2001) highlighted that both types of goal orientation can be beneficial; specifically, mastery goal orientation was the only goal orientation positively related to participant interest, while performance goal orientation was the only goal orientation positively related to graded performance. Furthermore, it was reported that participants who adopted both goal orientations were more likely to become both interested and perform well in the learning session.

Harackiewicz and her colleagues (2008) offered further support for the multiple goal orientation perspective in a longitudinal study. They investigated 858 college students' continued interest and goal orientation during seven semesters. Findings indicated both mastery and performance-approach goal orientation were important in predicting continued interest. Daniels and his colleagues' study (2008) used *k*-means cluster analysis to classify 1002 undergraduate students according to their mastery and performance-approach goals. They identified four groups of students, including cluster 1, high mastery/performance (i.e., multiple goals); cluster 2, dominant mastery; cluster 3, dominant performance, and cluster 4, low mastery/performance (i.e., low motivation). They further investigated the relationship between these goal orientations and students' expected achievement, perceived success, achievement-related emotions (i.e., enjoyment, boredom, and anxiety), final grade in class, and GPA. The results showed that students in cluster 1, 2, and 3 showed same achievement level, while students in cluster 4 had lower achievement level. Interestingly, students in cluster 3 were more "psychologically and emotionally vulnerable" than students in cluster 1 and cluster 2 (p. 584), which indicates students who had dominant performance goal orientation tended to be afraid of failure, experience more anxiety and harder to adjust to the courses or university. Different from Daniels's study, Jang and Liu (2012) used hierarchical cluster analysis to group 480 students (aged between 13 and 14 years), and the resulted revealed five clusters of students based on their goal orientations, including (1) high multiple goal orientations, (2) high mastery-approach and low mastery-avoidance goal orientations, (3) low all goal orientations, (4) high mastery-avoidance goal orientations, (5) and low performance goal orientations.

Measurements

Researchers have developed surveys to indirectly or directly investigate learner goal orientation. With indirect measurements researchers asked about participant judgement on success or satisfaction rather than goal orientation (Archer, 1994; Button et al., 1996; Meece, Blumenfeld, & Hoyle, 1988; Nicholls, Patashnick, & Nolen, 1985). Nicholls, Patashnick, and Nolen (1985) created a popular survey to measure high school student goal orientation. This measurement had no clear distinction between mastery and performance goal orientation; rather, the goal orientations were operationalized as the individual's judgement of success on nine goals with each item beginning with the stem, "I feel most successfully if —." One example item was that "I feel most successfully if I score high without studying."

Compared to indirect approaches, other investigators have created more direct approaches to measure goal orientation, which have been used more in published studies (Elliot & Church, 1997; Midgley et al., 1998, Miller, Behrens, Greene, & Newman, 1993). The Patterns of Adaptive Learning Survey (PALS) directly asked elementary and middle school students about their reasons for doing schoolwork (Midgley et al., 1998). Miller, Behrens, Greene and Newman (1993) also directly asked college students about their goals in learning context, such as "One of my primary goals in this course was..."; and Elliot and Church's (1997) Achievement Goal Questionnaire (AGQ) asked college students to grade a list of 18 items consisting of their goals, what was important for them in the class, and to what extent they believed each item to be true of them on a scale of 1 to 7 (1 = not at all true of me to 7 = very true of me).

Among these direct measurements, The PALS (Midgley et al., 1998) and AGQ (Elliot & Church, 1997) were the most popular—PALS for elementary and middle school students; AGQ for college students and older. Although the PALS has been used widely on middle or high school students (Bereby-Meyer & Kaplan, 2005; Maltais, Duchesne, Ratelle, & Feng, 2015; Vedder-Weiss, 2017; Wolters et al., 1996) and been found to be reliable and valid with college students (Jagacinski & Duda, 2001; Shim & Ryan, 2005), it did not have a performance-avoidance dimension and has not evolved since 1997.

Unlike the PALS, the AGQ has evolved over the past two decades and has been applied in various studies (Elliot & Church, 1997; Hsieh et al., 2008; Schmidt & Ford, 2003; Senko & Miles, 2008). AGQ was an 18-item achievement goal questionnaire assessing college student adoption of mastery, performance-approach, and performance-avoidance goal orientation. Elliot and Church (1997) conducted a series of pilot studies to develop the questionnaire and tested each goal orientation via factor analysis and correlations with other relevant measures to make sure the AGQ had high internal consistency and validity (Cronbach's a = .89, .91, and .77 for mastery, performance approach, and performance-avoidance goal orientation, respectively). Within this measurement there were six items for each of the three achievement goal constructs. The scale for each item ranged from 1 (not at all true of me) to 7 (very true of me). Examples of items included: "I desire to completely master the material presented in this class" (mastery goal); "It is important for me to do well compared to others in this class" (performance-approach); and "I worry about the possibility of getting a bad grade in this class" (performance-avoidance).

Elliot (1999) modified the AGQ scale by replacing item 17 ("I wish my university classes were not graded") with the more face-valid item, "My goal for this class is to avoid performing poorly" to improve the internal consistency for the performance-avoidance goal dimension. Smith, Duda, Allen, and Hall (2002) tested the original AGQ with confirmatory factor analysis using data from 475 college students in the United Kingdom, which supported this modification. They also suggested that after deleting the item 17 all items in the survey had acceptable internal reliability (Cronbach's a = .81, .88, .69 for mastery, performance-approach, and performance-avoidance goral orientation, respectively). It also had acceptable validity after deleting the item 17 (i. e., factorial validity, χ^2 = 460.05, df = 134, CFI = .88, $RMSEA$ = .07; construct validity such as convergent validity, convergent correlation r = .68, .77, .53 for mastery, approach, and avoidance, respectively, and discriminant validity, discriminant correlation r = .11, .13, .06 for mastery, approach and avoidance, respectively).

Based on the modified AGQ, Elliot and McGregor (2001) then constructed a 2 X 2 goal orientation measurement, where mastery goal orientation was separated into approach and avoidance components. In this measurement three items were chosen to represent each goal orientation, which were graded on a 1 (not at all true of me) to 7 (very true of me) scale. They tested this measurement in an undergraduate classroom that suggested the measurement had high reliability: mastery-approach (a = .87), mastery-avoidance (a = .84), performance-approach (a = .96), and performance-avoidance (a = .82) (Elliot & McGregor, 2001). An example for mastery-avoidance item was "I worry that I may not learn all that I possibly could in this class" (p. 504).

Elliot and Murayama (2008) identified several specific problems with the modified AGQ measurement and revised this version. The first problem they pointed out was the previous measurement items were focused more on the student motive rather than goal. For example, the item “My fear of performing poorly in this class is often what motivates me” illustrated this problem. As a result, motive items were omitted in the revised version. The second problem regarded using the word “grade” in the measurement statement. For example, one performance-approach goal item stated: “My goal is to get a better grade than most of the other students.” It was suggested that “grade” could be applicable to either mastery-based or performance-based goals. “Grade” was omitted from this item in the revised measurement. In addition, the measurement scale was changed from 7 points to 5 points (1-strongly disagree to 5-strongly agree). By doing so, it was suggested that the updated measurement accurately represented the four goal orientations in this model and was superior to the previous measurement (Elliot & Murayama, 2008).

Based on the new 3 X 2 goal orientation model, Elliot et al. (2011) developed a new questionnaire to measure goal orientations. This measurement had three items for each goal orientation on a scale from 1 (not true of me) to 7 (extremely true of me). They tested this measurement with undergraduate students both in the United States and Germany and reported high reliabilities, including task-approach (one example is “To get a lot of questions right on the exams in this class”, $a = .84$), task-avoidance (one example is “To avoid incorrect answers on the exams in this class”, $a = .80$), self-approach (one example is “To perform better on the exams in this class than I have done in the past on these types of exams”, $a = .77$), self-avoidance (one example is “To avoid doing worse on the exams in this class than I normally do on these types of exams”, $a = .83$), other-approach (one example is “To outperform other students on the exams in this class”, $a = .93$), and other-avoidance goals (one example is “To avoid doing worse than other students on the exams in this class”, $a = .91$). Several published empirical studies used this measurement, demonstrating its structural validity and cross-cultural generalizability among college students in many countries, including Australia (Flanagan, Putwain, & Caltabiano, 2015), China (Ning, 2016), France (Gillet, Lafrenière, Huyghebaert, & Fouquereau, 2015), Germany (Elliot et al., 2011), Italy (Brondino, Raccanello, & Pasini, 2014), Norway (Diseth, 2015), Philippines (David, 2012, 2014), Singapore (Flanagan et al., 2015), U.K (Flanagan et al., 2015;

Stoeber, Haskew, & Scott, 2015) and the U.S (Elliot et al., 2011; Johnson & Kestler, 2013; Yang & Cao, 2013; Yang, Taylor, & Cao, 2016).

In Norway, Diseth (2015) investigated the Norwegian version's validity. He collected data from 217 Norwegian undergraduate psychology students (41 male, 176 female, age: *M* = 22.67). A confirmatory factor analysis (CFA) supported the expected 3 × 2 factor structure of this measurement (χ^2 = 247.14, *df* = 122, *CFI* = .96, *RMSEA* = .08). In Philippines, David (2012; 2014) tested this measurement in a university. In his 2014's study, he collected data from 487 first-year college students (93 male, 394 female, age: *M* = 16.48) to conduct the CFA. The results indicated that the 3 X 2 measurement met the criteria for a good fitting model (χ^2 = 124, *df* = 4, *CFI* = .98, *RMSEA* = .057). Furthermore, Brondino and her colleagues (2014) validated the Italian version with 466 Italian university students (33 male, 433 female, age: *M* = 23) with CFA and suggested that it was a good model (χ^2 = 352.81, *df* = 120, *CFI* = .93, *RMSEA* = .07). These studies are just a sampling of global wide studies that support the 3 X 2 measurement among college students; a summary of these research and additional studies is provided chronologically in Table 2.

Table 2: Published Studies on 3 X 2 Goal Orientation

Researchers	Region	Participants	Findings
Ning (2016)	China	384 first-year undergraduate students (133 male, 251 female, age: M =19, various disciplines)	TAP positively predicted the deep strategy of relating ideas and help-seeking. TAV negatively predicted the surface strategy of memorizing without understanding. SAP positively predicted the deep strategy of understanding for oneself. SAV positively predicted help-seeking. OAP positively predicted student overall GPA. OAV negatively predicted student overall GPA.
Yang et al. (2016)	U. S	209 online students in various distance education classes in a university (42 male, 167 female, age: M = 28.8)	OAV and SAP reported more help-seeking. OAP and SAV reported less help-seeking.
Diseth (2015)	Norway	217 undergraduate psychology students (41 male, 176 female, age: M = 22.67)	TAP and OAP were positively related to the motive for success. TAP and OAP positively predicted student self-efficacy. TAP positively predicted task value and strategic learning strategies TAV and SAV were positively related to the motive to avoid failure. SAP negatively predicted strategic learning strategies, but positively predicted surface learning strategies. SAP negatively predicted academic achievement. OAP and OAV positively predicted academic achievement.
Flanagan et al. (2015)	Australia/ Singapore/ U. K	286 undergraduate students	TAP was associated with lower test-irrelevant thinking and test anxiety. A higher OAP was related to higher worry and tension. Students tended to endorse multiple goals, and strongly endorsed self and task goals than other-related goal orientation.

Table 2 continued.

Researchers	Region	Participants	Findings
Gillet et al. (2015)	France	Sample 1: 278 undergraduates in an introductory level psychology class (48 male and 230 female age: *M* = 18.93) Sample 2: 327 undergraduates in an introductory level psychology class (56 male and 271 female age: *M* = 18.93) Sample 3: 169 workers were recruited via Amazon.com's Mechanical Turk online survey program (74 male and 92 female, three people did not report their gender, age: *M* = 32.48)	TAP was significantly and positively correlated to engagement and positive affect, and not significantly correlated to anxiety. SAV was not significantly correlated to satisfaction, engagement, positive affect, and anxiety. OAP was significantly and positively correlated to positive affect, and not significantly correlated to satisfaction and anxiety. SAP was significantly and positively correlated to satisfaction and engagement in sample 1 but not significantly correlated to these two variables in sample 2; significantly and positively correlated to engagement but not significantly correlated to satisfaction in sample 3. TAV and OAV were not significantly correlated to satisfaction, engagement, and positive affect in samples 1 and 3, while positively correlated in sample 2.
Stoeber et al. (2015)	U. K	100 undergraduate psychology students (11 male, 89 female, age: *M* = 19.9)	Only TAP predicted exam performance.
Brondino et al. (2014)	Italy	466 university students (33 male, 433 female, age: *M* = 23)	TAP predicted positive activity- and outcome-related emotions*. TAV predicted negative activity- and outcome-related emotions. SAP positively predicted one positive activity-related emotion. OAV positively predicted positive activity- and outcome-related emotions. * Positive activity-related emotions: enjoyment, relaxation; positive outcome-related emotions: hope, pride, relief; negative activity-related emotions: anger, boredom; negative outcome-related emotions: anxiety, shame, hopelessness.

Table 2 continued.

Researchers	Region	Participants	Findings
David (2014)	Philippines	487 first-year college students (93 male, 394 female, age: *M* = 16.48)	Participants reported higher level of SAP than TAP (t=8.68; p<. 001) and higher level of SAV than TAV (t=2.54; p<. 05). TAP and SAP positively predicted test hope. TAV negatively predicted test anxiety. SAV negatively predicted test performance.
Johnson & Kestler (2013)	U. S	123 traditional (37 male, 85 female, 1 not reported, age: *M* = 19.8) and 36 nontraditional (8 male, 28 female, age: *M* = 33.41) undergraduates in the School of Education.	Traditional students were statistically significantly more likely to be OAP and OAV than nontraditional students. OAV was negatively related to student grades (i.e., GPA).
Yang & Cao (2013)	U. S	93 university students in an educational psychology class (23 male, 70 female, 50 undergraduates, 43 graduates)	TAP had a detrimental influence on help seeking in e-learning environment. SAP had a positive impact on help-seeking in e-learning environment.
David (2012)	Philippines	350 first-year undergraduate students in math class (84 male, 266 female, age: *M* = 16.95)	Filipino students had relatively higher mean scores in self-based goals than task-based and other-based goals.
Elliot et al. (2011)	Germany/ U. S	Study 1: 126 undergraduate students in psychology class in Germany (22 male, 104 female). Study 2: 319 undergraduate students in psychology class in U.S (206 male, 113 female).	Participants had higher mean scores in task-based goals than self-based goals. TAP positively predicted intrinsic motivation, learning efficacy, and absorption in class. SAP positively predicted student energy in class. SAV negatively predicted student energy in class. OAP positively predicted exam performance and learning efficacy. OAV negatively predicted exam performance and learning efficacy. OAV positively predicted worry about exams.

According to Table 2, for the three approach goal orientations (i.e., TAP, OAP and SAP), published empirical studies usually suggested TAP and OAP had positive effects in learning context (Brondino et al., 2014; David, 2014; Diseth, 2015; Elliot et al., 2011; Johnson & Kestler, 2013; Stoeber et al., 2015). Stoeber et al. (2015) reported that TAP goal orientation positively predicted exam performance. Elliot et al. (2011) also reported that TAP was positively related to college student self-efficacy. With respects to OAP goal orientation, Diseth (2015) reported that it was positively correlated with graded performance, $r = .26, p < .01$. Elliot et al. (2011) also found it positively predicted exam performance and learning efficacy. For SAP goal orientation, Diseth (2015) suggested it was negatively correlated to graded performance, $r = .23, p < .05$, but it was positively related with more help-seeking behavior (Yang et al., 2016), and sometimes positively correlated to satisfaction and engagement (Gillet et al., 2015).

As for the three avoidance goal orientations (i.e., TAV, OAV, and SAV), studies usually reported they had negative effects. Although TAV was positively related to the motive to avoid failure (Diseth, 2015), David (2014) reported it negatively predicted test anxiety. Ning (2016) also suggested that TAV negatively predicted the surface strategy of memorizing without understanding. With respect to OAV goal orientation, Johnson and Kestler (2013) suggested it was negatively related to student grades (i.e., GPA). Elliot et al. (2011) also reported OAV goal orientation negatively predicted exam performance and learning efficacy, but positively predicted worry about exam. For SAV goal orientation, Elliot et al. (2011) suggested it was negatively related to college student energy in class. David (2014) also reported that SAV goal orientation was a negative predictor for college student performance.

Study also suggested other-approach and other-avoidance goal orientations were correlated with learner age. Johnson and Kestler (2013) assessed 157 college student goal orientations in the Midwestern region of the United States. They distinguished traditional student (i.e., undergraduates who continued into college education following high school graduation and were usually under 24 years old) from nontraditional college students (i.e., undergraduates who went to college after having other life experiences and were usually above 24 years old, $N = 36$). Their results indicated traditional undergraduate student goal orientation was more likely to be other-approach and other-avoidance oriented compared to nontraditional students.

It is worth noting that Elliot and his colleagues also proposed potential-approach and potential-avoidance goal orientations, but have not constructed a new measurement to date (Elliot et. al, 2015). They were concerned that lengthening the current 3 X 2 measurement, particularly through adding similarly worded items, may "heighten the already existing tendencies towards multicollinearity due to response sets and biases" (Elliot et. al, 2015, p. 201).

Relationship between Metacognition and Goal Orientation

Researchers have investigated the relationship between goal orientations and metacognition's two components (i.e., metacognitive knowledge, and metacognitive regulation/self-regulation) (Ning, 2016; Wolters, 1998; Yang et al., 2016), but few studies have been conducted on the relationship between learner specific metacognition levels and goal orientations (Gul & Shehzad, 2012).

For studying the relationship between metacognitive knowledge and goal orientations, Ning (2016) used the 3 × 2 goal orientation measurement to evaluate 384 freshmen from various disciplines in a university in Hong Kong on subjects such as arts, business, education, engineering, law, medicine and social science. The results showed SAP goal orientation can positively predict learner understanding for oneself (Ning, 2016), which indicated students who were SAP orientated were more likely have higher metacognitive knowledge about themselves.

For the relationship between metacognitive regulation and goal orientations, there are mixed results, including weak relationship, positive relationship and negative relationship. According to Wolters (1998), the relationship between learner mastery goal orientation and self-regulation is weak. He studied 115 college student self-regulated learning and motivational factors using an open-ended questionnaire and Likert scale survey, A small number of students (1%) indicated they adopted mastery goal orientation during self-regulated learning, while 15% students indicated they adopted performance goal orientation during learning. Some other studies indicated that there were both positive and negative relationships (Ning, 2016, Yang et al., 2016; Yang & Cao, 2013). Ning (2016) pointed out, compared to other goal orientations, SAV goal orientation positively predicted learner help-seeking in classroom setting. Yang et al. (2016) studied 209 online student goal orientation and online help-seeking behaviors. The results suggested learners with SAV and OAP goal orientation pursued less online help, while learners with OAV and SAP goal orientation sought more online help. Another study conducted by Yang and Cao (2013) indicated that TAP had a negative impact on help-seeking in e-learning environments, and SAP had a positive impact on help-seeking in e-learning environments.

As for the relationship between learner metacognitive level and goal orientation, Gul and Shehzad (2012) studied the relationship among 345 undergraduate student metacognition, mastery, performance goal orientation and GPA. They adopted the MAI (Schraw & Dennison, 1994) to measure student metacognition. The results showed mastery goal orientation was highly correlated with metacognition ($r = 0.53, p < 0.01$), while the performance goal orientation had slightly lower correlation with metacognition ($r = 0.49, p < 0.01$).

PROBLEM-BASED LEARNING

This section reviews the literatures on PBL. Firstly, this section identifies the definition and history of PBL. Then I will take a closer look at the two elements of PBL: the problem and problem-solving process, followed by the previous studies on the effects of PBL, especially those on the advantages and challenges of PBL. Lastly, this section will summarize the most current studies on the impact of learner characteristics on problem-solving.

Definition and brief history

PBL is a learner-centered constructivist instructional approach that embeds student learning processes in ill-structured problems (Barrows & Tamblyn, 1980; Hmelo-Silver, 2004; Savery, 2015). It was originally used for medical students. In 1979, the University of New Mexico was the first medical school to adopt PBL pedagogy (Neufeld & Barrows, 1974) in the United

States. The Mercer University School of Medicine in Georgia followed in 1982 and was the first medical school to employ PBL as its only curricular offering (Barrows, 1996).

Since its successful implementation in various medical education settings, PBL spread to many other disciplines throughout K–16 education all over the world. For instance, Dahlgren and Dahlgren (2002) described a Master program in Psychology, a Bachelor program in Physiotherapy, and a Master program in Computer Engineering which all adopted PBL in a Sweden university. In Australia, Maitland (1977) described where a four-year Architecture and Construction program had offered a problem-based architecture course at the fourth-year of the study for 11 years. In this course, students needed to solve a town-planning problem with attention to the environment, economy, and technical issues, and present their design to the client who owned the land. In computer sciences, Oliveira et al. (2013) reported at least 52 published studies used PBL between 1997 and 2011 to teach topics such as software engineering, robotics, embedded systems, operating systems, digital systems, and the development of software. Other disciplines such as economics (Maxwell, Bellisimo, & Mergendoller, 2001), languages (Lin, 2017), mechanical engineering (Yadav, Subedi, Lundeberg, & Bunting, 2011), music (Freer, 2017), political science (Maurer & Neuhold, 2014), and science (Liu, Bera, Corliss, Svinicki, & Beth, 2004; Liu et al., 2009) have also reported using PBL pedagogy.

Problem

According to Jonassen (2000), problems have two attributes. First, "a problem is an unknown entity in some situation" (p. 65), which indicates a difference between a goal state and a current state. If we consider designing a spaceship to explore the galaxy as a problem, then how to design the spaceship is the unknown entity. Second, there must be "someone [who] believes that it is worth finding the unknown" (Jonassen, 2000, p. 65). This means finding or solving for the unknown "must have some social, cultural, or intellectual value" (Jonassen, 2000, p. 65).

Additionally, there are two different structures for problems: well-structured and ill-structured problems (Jonassen, 2000). The well-structured problem has a limited number of rules, principles, and concrete solutions. For example, there is a math problem that asking students to find all the X for the following formula: X * 3 = 6. For this problem, there is only one answer, which is X = 2. An ill-structured problem, however, may have unknown elements, multiple solution criteria, and possess multiple solutions. For example, deciding how to travel from the United States to China. There are multiple solutions to this problem, such as the travel transportations, what airline does one choose if one flies, and to which city does one fly.

In PBL, there are different complexities for different problems, depending on the number of interacting elements within a problem. For example, it is more complex to solve a country's energy crisis than finding a single family's energy solution. Often, problems in PBL are designed to be ill-structured, complex, open-ended, and relevant to real life (Hmelo-Silver, 2004). This makes learning relevant and motivates learners to find solutions during problem-solving processes.

Problem-solving process

Gagne (1985) defined problem-solving as finding solutions to problems, and it was a goal-oriented activity to find the unknown between current state and the goal state. Similarly, according to Mayer (2013), problem-solving referred to cognitive processing directed at achieving a goal when the problem-solver did not initially know a solution method. Many well-known scholars agreed with these two definitions and studied learner problem-solving (Dominowski, 1998; Jonassen, 2000; Willingham, 2007). Cognitive scientists suggested that there were four general problem-solving strategies when facing an unfamiliar problem: random trial and error, hill climbing, working backward and means-ends (Mayer, 2013; Willingham, 2007).

"Random trial and error" is a fundamental strategy of problem solving, which means the problem-solver adopts repeated, varied random attempts until success. "Hill climbing" strategy means the problem-solver chooses the attempts that seem to lead most directly toward the goal state. As for the "working backward" strategy, this indicates the problem-solver starts with the result and applies the operations in reverse order until arriving the start point. According to Willingham (2007), by far the most thoroughly tested and probably the most broadly applicable strategy is means-ends analysis, which uses a combination of forward- and backward-moving strategies and dictates when it is effective to set sub-goals that should be completed before the main goal of the problem should be tackled. It can be summarized as following:

> *1. Compare the current state with the goal state. If there is no difference between them, the problem is solved.*
>
> *2. If there is a difference between the current state and the goal state, set as a goal to solve that difference. If there is more than one difference, set as a goal to solve the largest difference.*
>
> *3. Select an operator that will solve the difference identified in Step 2*
>
> *4. If the operator can be applied, apply it. If it cannot, set as a new goal to reach a state that would allow the application of the operator.*
>
> *5. Return to step 1 with the new goal set in step.* (p. 444)

Similar to the means-ends strategy, problem-solving has been described as an iterative cycle over the past four decades. One example is the classic problem-solver model (Newell & Simon, 1972), which indicated that problem-solving process included (a) internalizing the problem representation, (b) selecting a problem-solving method, (c) applying the selected method, which may be halted, (d) when a method is terminated, problem-solver may select another method, reformulate the problem presentation, or even abandon the problem, and (e) new problems may occur during the process, which the problem-solver may choose to solve one of.

In 1984, Bransford and Stein proposed the popular IDEAL model (Identify, Define, Explore, Act, and Look), which stated an ideal problem-solver should be able to *Identify* the problem, *Define* the problem, *Explore* alternative approaches, *Act* on a plan, and *Look* at the effects (Bransford & Stein, 1984). They suggested if an unknown obstructs the solution path, the problem-solver should have ability to initiate a new cycle and repeats the steps until reaching a solution.

Similarly, Mayer (2013) suggested that problem-solving process included "representing, planning, executing, and monitoring" four stages (Mayer, p.769). During the *representing*, the instructor presents a problem or scenario to the students, the students then start the *planning* process. During the *planning* stage, students first need to understand the current state of the problem and know what goal state they are seeking. Then they can discuss the problem with their peers, create their goals and determine how to proceed. Next, the students *execute* the plan either in groups or as individuals. Meantime, students also *monitor* their actions. When necessary, they can refine their approach to adjust the initial action until solving the problem. Recently, the Program for International Student Assessment (PISA) also suggested there are four stages during the problem-solving process: (a) exploring and understanding; (b) representing and formulating; (c) planning and executing; and (d) monitoring and reflecting (OECD, 2014).

These above models along with other similar models, indicated problem-solving follows a specific sequence, such as (a) problem presentation, (b) identification of problem-solving needs, (c) self-directed study, (d) group sharing, (e) solution development, (f) testing alternative solutions, and (g) post-problem debriefing/reflection (Hmelo-Silver, 2004; Pedersen, 2000; Savery & Duffy, 1995; Schmidt, Rotgans, & Yew, 2011; Yew & Schmidt, 2012).

Besides these general problem-solving process models, some scholars also specifically developed problem-solving models for ill-structured problems (Ge & Land, 2003, 2004). Ge and Land (2003, 2004) summarized four main processes during ill-structured problem-solving: (a) problem representation, (b) generating solutions, (c) constructing arguments, and (d) monitoring and evaluation. Some researchers studied problem-solving process specifically within technology-based PBL environments (Liu et al., 2004; Liu et al., 2009). For example, in Alien Rescue, a web-based PBL environment designed for sixth-grade middle school science subject, Liu et al. (2004) grouped student problem-solving processes into four conceptual stages to reflect student cognitive processes as a problem-solver: (a) understanding the problem (Stage 1—Understanding), (b) identifying, gathering, and organizing information (Stage 2—Researching), (c) integrating information and hypothesis testing (Stage 3—Hypothesis testing), and (d) evaluating the process and outcome (Stage 4—Evaluating).

In summary, problem-solving is a goal-oriented activity to achieve a goal by finding the unknown between current state and the goal state. In addition, researchers have summarized several similar problem-solving process models that indicated leaner problem-solving was an iterative cycle until solutions were found. Besides developing these problem-solving models, studies have been conducted to investigate the effects of PBL. I will review the relevant literature in the following section.

Effects of Problem-based Learning

PBL has been utilized in a variety of fields for over 40 years, and scholars have studied extensively its effectiveness. This section will review the research findings, which indicate there are both advantages and challenges when applying PBL pedagogy.

Advantages

With respect to the positive effects while using PBL, studies have suggested that PBL surpassed the traditional instructional method in long-term retention (Dochy, Segers, Van den Bossche, & Gijbels, 2003; Strobel & van Barneveld, 2009), knowledge and skill development (Albanese & Mitchell, 1993; Hmelo-Silver & Barrows, 2015; Lin, 2017; Oliveira et al., 2013), and student and teacher satisfaction (Berkson, 1993; Colliver, 2000; Lou, Shih, Diez, & Tseng, 2011; Newman, 2003).

With respect to learner long-term retention, Strobel and van Barneveld (2009) compared PBL to conventional classroom by using a qualitative meta-synthesis approach. They analyzed eight meta-analyses about medical education effectiveness and found that PBL was more effective when it was relevant to long-term retention, while traditional approaches were better for short-term retention as measured by standardized test. Another example is that Dochy et al. (2003) compared knowledge outcomes based on a retention period or no retention period, and suggested that long-term knowledge retention favored PBL.

Besides long-term retention, PBL was reported to improve knowledge and various skills across multiple disciplines. In medical education, Albanese and Mitchell (1993) looked at PBL efficacy research from 1972 to 1992. They suggested PBL students in medical education felt better prepared in self-directed learning skills, problem-solving, information gathering, and self-evaluation techniques. Kasim's (1999) research on PBL in medical education was consistent with previous findings that PBL tended to produce better results for clinical knowledge and skills. In computer science, Oliveira et al. (2013) conducted a systematic mapping study on using PBL in Computing between 1997 and 2011, and they suggested that students were ready to deal with the demands of actual software development projects after the PBL instruction. In addition, other researchers also found PBL promoted "the development of the students' reasoning ability" (Masek, & Yamin, 2011, p. 217), reading comprehension ability (Lin, 2017), and collaboration skills (Segrelles, Martinez, Castilla, & Moltó, 2017). Furthermore, Hmelo-Silver studied PBL for more than two decades, and she suggested PBL helped students develop understanding, self-directed learning skills, effective collaboration skills, and lifelong learning skills (Hmelo, 1998; Hmelo-Silver, 2004; Hmelo-Silver & Barrows, 2015).

Besides helping students with knowledge and skill development, PBL also enhanced student and teacher satisfaction in learning and teaching. Berkson's (1993) review of ten pre-1992 studies indicated students and faculty in medical education favored PBL. Colliver's (2000) review of literature published from 1992 to 1998 also acknowledged PBL provided a more challenging, motivating, and enjoyable approach to medical education. In addition, Newman's (2003) review of PBL in medical education was consistent with previous findings whereby students were more satisfied with PBL. Besides the medical education field, students also showed a higher degree of active leaning attitude and desire to learn English more when using PBL (Lin, 2017). Lou et al. (2011) studied 40 females tenth-grade student attitude toward science, technology, engineering, and mathematic (STEM) learning using PBL. The results suggested that through solving ill-structured problems, students actively engaged in the learning process and took more responsibility of their own learning. In addition, after one month of STEM learning using PBL, these students recognized that learning STEM knowledge can be "very effective and interesting" (p. 203). Furthermore, after students realized that they could participate in PBL activities to increase their exposure to STEM knowledge, their career attitudes also changed. Specifically, these female students showed increased interests in STEM-related jobs.

Challenges

While the benefits of PBL have been widely accepted, some researchers argued PBL was challenging for both teachers and students, especially when it was applied in K-12 settings (Ertmer & Simons, 2006). For teachers, many researchers suggested K-12 teachers lacked pedagogical knowledge on how to effectively engage students in ill-structured problem-solving (Bransford, Brown, & Cocking, 1999; Bransford & Stein, 1984; deChambeau & Ramlo, 2017). In addition, curricular and assessment demands may constrain teacher PBL adoption in the classroom (Ertmer & Glazewski, 2015). Even when teachers started using PBL in the K-12 classroom, it was difficult for them to provide all the support that learners might need (Pedersen, 2000). In the medical school environment, an expert facilitator or teacher works closely with a small group of students. This kind of close small group facilitation usually is not possible in K-12 settings where one teacher is often in charge of several groups. Hmelo-Silver (2004) called this as “wandering facilitation” whereby one facilitator moves from group to group. Since students in K-12 settings are not used to this kind of independent learning that is required in PBL, and a single teacher may not be able to provide needed assistance to groups when needed, students would have trouble moving forward in their problem-solving processes.

In addition, due to the minimally guided instruction in PBL, researchers suggested students also faced several challenges, such as frustration, disengagement, and misconceptions unless they possessed the required collaborative learning skills and self-regulation skills (Clark, Kirschner, & Sweller, 2012; Kirschner et al., 2006; Norman & Schmidt, 1992). Some researchers indicated students needed to collaborate with group members in PBL to solve problems (Norman & Schmidt, 1992); however, most PBL groups needed some help to collaborate effectively, because both young and adult learners might be able to work in a group but do not know how to work as a group (Evensen, Salisbusry-Glennon, & Glenn, 2001; Mercer & Littleon, 2007). Learners might also need facilitators’ help to ensure that all learners were involved in the discussion (Barrows, 2000), and had assigned collaborative roles in the group to make sure all group members are cognitively engaged (Hmelo-Silver, 2002).

PBL also requires learners to monitor their own learning as skilled self-regulated learners. They must be able to continually set learning goals, develop strategies for reaching the goals, monitor their understanding, and change strategies as needed (Azevedo et al., 2004). Research suggested this process was a challenge even for adult learners (Chowdhry, 2016; Evensen et al., 2001). Chowdhry (2016) conducted a case study on the PBL environment of a mechanical engineering course called “Engineering Design and CAD.” The studies lasted 13 weeks where 79 university students met once a week for two hours within this PBL environment. This study indicated some students were “having trouble understanding what to do” (p. 25), because the PBL environment’s open-endedness and no frequent task-focused feedback.

Furthermore, studies suggested that students were often not prepared for the cognitive tasks in environments adopting PBL. Land’s (2000) meta-analysis on open-ended learning environments found students usually failed to engage in reflective thinking and metacognition. In addition, students routinely attached meaning to irrelevant information, made biased, incomplete, and unreliable observations, failed to refine problem-solving strategies over time, and attempted to apply incomplete and potentially inaccurate knowledge of the domain (Land, 2000).

In summary, the previous studies have identified both the advantages and challenges for applying PBL pedagogy in K-16 settings. Specifically, PBL can benefit learners on their long-term retention, knowledge and skill development, and learning satisfaction. However, successful PBL requires learner understand the problem then engage in metacognition or have instructor support during the problem-solving processes.

Learner Characteristics affect problem-solving

Besides the advantages and challenges of using PBL in K-16 settings, scholars have also investigated the impact of different learner characteristics on problem-solving. This section will review previous studies on metacognition, goal orientation and other learner characteristics’ impact on problem-solving.

Metacognition

As a cognitive characteristic, metacognition has been linked to successful problem-solving in PBL (Davidson & Sternberg, 1998; Gourgey, 1998; Marra et al., 2014; Mihalca et al., 2017; Shin et al., 2003). Three decades ago, Gourgey (1998) described his observation during student mathematical problem-solving processes, “effective problem solvers seek to understand concepts and relationships, monitor their understanding, and choose and evaluate their actions based on whether the actions are leading toward their goals” (p. 89). Davidson and Sternberg (1998) also suggested that during the problem-solving, metacognitive skills can help learners strategically encode the nature of the problem, form a mental model or representation of its elements, select appropriate plans and strategies for reaching the goal, and identify and overcome obstacles that may impede progress. Specifically, by monitoring their progress in reaching a solution, learners can adjust their plan and strategies if needed to successfully solve the problem (Davidson & Sternberg, 1998). A recent study conducted on undergraduate students also found that students who had better monitoring accuracy on their own learning had better post-test performances with the problem-solving tasks (Mihalca et al., 2017). This study also suggested that learner monitoring accuracy was more important when they had no instruction during the problem-solving processes compared to having instruction.

Since students must “exercise metacognitive skills” in PBL (Marra et al., 2014, p. 233), researchers described how they helped with learner metacognition in a PBL program called *Iron Range Engineering* (IRE), in which learners can attain technical and professional competencies by solving authentic industrial engineering problem (Marra et al., 2014). The researchers asked students to complete a “metacog” memo that documents their reflections on the learning processes, the judgments they made on the quality of the learning, and the regulative changes made based on these judgments. They observed that students really valued the metacognition memos, and found it “extremely useful” (p. 234). In addition to the “metacog”, IRE students were also asked to produce a “metachron” to reflect on their time spent on learning activities, and student interview data indicated that the “metachron” helped them with time management.

Shin, Jonassen, and McGee (2003) compared the metacognition required for solving well-structured problems and ill-structured problems in the context of a multimedia environment called *Astronomy Village* (AV) for high school astronomy. They recruited 124 students in ninth-grade, then measured their metacognition using an instrument adapted from Fortunato et al.'s (1991) How Do You Solve Problems? (HSP). They also developed two instruments to measure student well-structured and ill-structured problem-solving results after a three-week investigation in AV. The results indicated that student metacognitive skills were more important to solving ill-structured problems than well-structured problems. Their further analysis suggested that the second component of metacognition (i.e., metacognitive regulation) is a significantly predictor for solving ill-structured problems both in familiar and unfamiliar contexts.

In addition, Ge, Law, and Huang (2016) developed a conceptual framework to describe the interrelationships between learner self-regulation and ill-structured problem-solving. In this framework, once given a problem, the problem-solver would start two self-regulation cycles—*Problem Representation* and *Solution Generation* (see the two boxes in Figure 1)—to find the *Final Solution*. Within each cycle, there were three self-regulation phases: *Planning, Execution* and *Reflection* to generate either a plausible problem representation or solution. The two cycles were affected by problem-solver's *Motivation & Beliefs*, and the relationship between these two cycles is circular—a *Problem Representation* cycle can lead to a cycle of *Solution Generation*, and vice versa (see the red dot arrow in Figure 1).

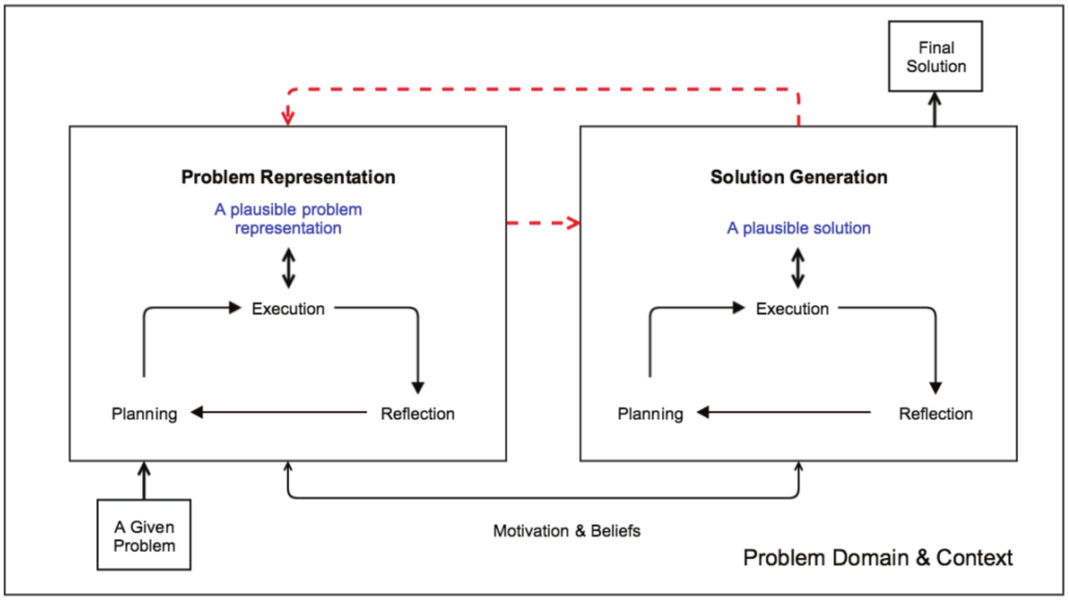


Figure 1. A conceptual framework of self-regulated ill-structured problem-solving

Adapted from "Detangling the Interrelationships Between Self-Regulation and Ill-Structured Problem Solving in Problem-Based Learning," by X. Ge, V. Law, & K. Huang, 2016, *Interdisciplinary Journal of Problem-Based Learning, 10*(2).

Goal orientation

Besides metacognition, a few scholars have explored the relationship between learner goal orientation and problem-solving. Bereby-Meyer and Kaplan (2005) investigated the effect of children goal orientation on a problem-solving strategy transfer with different task. They recruited second-grade (*N* = 60) and sixth-grade students (*N* = 60) to work on the task. The task was about matching singers with their respective bands, which required using a matrix (1–2, 1–3, 1–4, 1–5; 2–3, 2–4, 2–5, etc.); and the analogous problem was about finding all pairs of shapes that complete a rectangle in five packs of geometrical shapes, which required the same strategy with previous task. The result suggested, regardless of age and/or perceived ability, when the participants all had high level of mastery goal orientation, participants with higher performance-approach goal orientation were less likely to transfer the strategy compared to those with lower performance-approach goal orientation. Although this study did not use the latest 3 X 2 model to study goal orientation, it indicated that learner goal orientation affected problem-solving.

Other learner characteristics

In addition to metacognition and goal orientation, previous studies investigated other learner characteristics, including (a) personal characteristics such as socio-economic status, gender difference, and race/ethnicity (Lemke et al., 2004; OECD, 2014); (b) academic characteristics such as familiarity with the problem (Brand-Gruwel, Wopereis, & Vermetten, 2005; Jonassen, 2000); and (c) social/emotional characteristics such as beliefs and self-efficacy (Jonasseen, 2000; Phillips, 2001; Liu et al., 2006; Pajares & Kranzler, 1995; Schommer-Aikins et al., 2005).

Personal Characteristics

For learner personal characteristics, researchers compared problem-solving performances based on learner socio-economic status. The PISA included a problem-solving assessment, which was first implemented in 2000 and is carried out every three years in all the countries that joined the Organization for Economic Cooperation and Development (OECD). It used a variety of contexts, disciplines, and problem types to measure 15-year-old student problem-solving skills and described seven problem-solving proficiency levels based on student scores. These scores included Level 6 (Equal to or higher than 683 points), Level 5 (618 to less than 683 points), Level 4 (553 to less than 618 points), Level 3 (488 to less than 553 points), Level 2 (423 to less than 488 points), Level 1 (358 to less than 424 points) and Below 1 (Below 358 points) (OECD, 2014; p. 50). PISA indicated that among all the OECD countries, the higher the student socioeconomic status, the higher their problem-solving scores. In the United States, the relationship between these two variables was even stronger; a one standard deviation change in the student socioeconomic status affected 31 points on the problem-solving proficiency scale in 2004 (Lemke et al., 2004), and 35 points in 2012 (OECD, 2014). However, their further analyses pointed out that the socio-economic status of students did not appear to have a direct causation with problem-solving performances. Instead, it reflected the unequal access to good teachers and schools.

Studies also explored the impact of learner gender differences on problem-solving. It was reported in PISA 2003 that boys scored higher than girls in the majority of tested countries (32 out of 39 countries) including the United States. In 2012, boys also scored seven points higher on average than girls in problem-solving across OECD countries (OECD, 2014). However, there were no statistically measurable differences in problem-solving scores by gender.

With the respect to race/ethnicity variable, Black and Hispanic students problem-solving skills average scores were below the respective OECD average scores while scores for White students were above the OECD average scores. Students who were White and Asian scored higher, on average, than Black, Hispanic, and students of more than one race (Lemke et al., 2004). However, Lemke et al. (2004) did not report this was statistically measurable differences.

Academics Characteristics

For learner academic characteristics, Jonasseen (2000) suggested that the problem-solver's familiarity with the problem was a predicator for their problem-solving skills, because experienced problem-solvers had better developed problem schemas and they could utilize these schemas more automatically. Dominowski (1998) also noticed that novice problem-solvers would simply leap into problem-solving action rather than taking the time to understand the problem. They also rarely engaged in solution monitoring and reflection. In addition, Brand-Gruwel, Wopereis, and Vermetten (2005) studied the problem-solving process differences between experts and novices. They recruited five doctoral students (experts) and five freshmen students (novices) in the study and asked them to write a 400-words argument on the topic of food consumption for a consumer magazine. The researchers adopted think-aloud method, and coded the data later. The results revealed that both experts and novices invested a lot of time in writing the argumentation, but experts spent more time on defining the problem and more often activated their prior knowledge, elaborated on the content, and regulated problem-solving processes.

Social/Emotional Characteristics

With respect to learner social/emotional characteristics, researchers suggested that learner epistemic beliefs about the nature of problem-solving affected their problem-solving (Jonasseen, 2000; Phillips, 2001; Schommer-Aikins et al., 2005). Schommer-Aikins et al. (2005) studied 1269 middle school student epistemological beliefs and problem-solving performances on mathematical problems. Based on regression analyses, the results indicated that the less students believed in quick/fixed learning (i.e., that learning is fast and instinctual) and the more they believed in math is useful, the better they were at mathematical problem-solving. In addition, student beliefs might affect their problem-solving strategies during the problem-solving processes (Phillips, 2001). For example, in solving an ill-structured problem, depending on whether they considered knowledge was complex or not, one student might insist on simple answer and the other might be open to complex and alternative solutions.

Another studied social/emotional characteristic was self-efficacy, which referred to the beliefs people had about whether they could successfully complete a task (Bandura, 1986). Many studies have indicated that problem-solver self-efficacy would significantly affect their problem-solving (Liu et al., 2006; Pajares & Kranzler, 1995). Pajares and Kranzler (1995) tested the influence of math self-efficacy on the math problem-solving performances of 329 high school students. They found most (86%) high school students were overconfident on their problem-solving abilities, a few of them (9%) were under-confidence, and only 4% of them successfully predicated their results. For all 18 math problems, students in the overconfidence group erred more often (6.2 problems on average) than did those in the under-confidence group (3.5 problems on average). Moreover, students in the under-confidence group had higher performance scores. Another self-efficacy study conducted by Liu et al. (2005) in a PBL environment Alien Rescue indicated that student self-efficacy was a statistically significant predictor of science achievement scores. That is, students with high self-efficacy scored higher on the post science achievement test than students with low self-efficacy.

In summary, learner characteristics, especially metacognition played an important role in problem-solving, including both problem-solving performance and problem-solving process, because it could help learner to understand the problem, make appropriate plan, select appropriate strategy, manage time, overcome obstacles, monitor their problem-solving processes, and eventually link to successful problem-solving in PBL. Although scholars have developed the conceptual framework between learner self-regulation and ill-structured problem-solving, little research illustrated the problem-solving process differences based on individual metacognition differences. In addition, only a few studies investigated goal orientation on problem-solving. More research is needed to understand the impact of goal orientation on learner problem-solving.

SERIOUS GAMES

This section will review the literatures on Serious Games (SG) including its definition and current trends, the previous reported advantages and challenges of using SG for learning, and the impact of learner characteristics on using SG.

Definition and current trends

Abt (1970) pointed out Serious Games (SG) "have an explicit and carefully thought-out educational purpose and are not intended to be played for amusement" (p. 9). While Abt considered the educational value of SG, Zyda (2005) emphasized SG entertainment value and suggested SG are "mental contests played with a computer in accordance with specific rules that use entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives" (p. 26). Their definitions have provided some insights on SG; however, Abt's definition neglected SG's entertainment purpose and Zyda's definition was too broad, because it included all the games that can be used for educational purpose no matter the initial goal for which they were created. Therefore, I agree with Loh, Sheng and Ifenthaler's (2015) definition, "SG are digital games and simulation tools that are created for non-entertainment use, but with the primary purpose to improve skills and performance of play-learners through training and instruction" (p. 7).

Currently, SG has been applied to many fields, such as energy (Morganti et al., 2017), language (Alyaz, Spaniel-Weise, & Gursoy, 2017), healthcare (Graafland, Schraagen, & Schijven, 2012), military (IBM, 2017a), government (IBM, 2017b), religion (Nazry & Romano, 2017) and K-16 educational settings (Ke, 2008; Liu et al., 2015; Michael & Chen, 2006). Morganti and his colleagues summarized ten SG that had been used in three different areas related to energy efficiency: environmental education, consumption awareness, and pro-environmental behaviors (Morganti et al., 2017). They suggested the appealing and motivating feature of SG could foster energy-saving behaviors among consumers. Graafland et al. (2012) reviewed thirty SG that were developed for healthcare purpose, such as team training in critical care, triage training, and laparoscopic psychomotor skills training. They suggested SG were beneficial to train both technical and non-technical skills relevant to healthcare. Alyaz et al. (2017) also investigated two German language SG in Turkey (Adventure German—The Mystery of the Nebra Sky Disc and A Mysterious Mission), which were found useful for developing language skills.

As for the K-16 setting, SG have been applied to different subject areas such as engineering (Perini, Margoudi, Oliveira, & Taisch, 2017), language (Barendregt & Bekker, 2011; Johnson, Vilhjalmsson, & Marsella, 2005; Liu & Chu, 2010), mathematics (Ke, 2008; Trespalacios & Chamberlin, 2012; Winburg, Chamberlain, Valdez, Trujillo, & Stanford, 2016) and science (Liu et al., 2015; Mayo, 2007; Rowe, Shores, Mott, & Lester, 2011; Tsai, Huang, Hou, Hsu, & Chiou, 2016). A SG called EcoFactory was designed for middle school students to learn about the modern manufacturing industry (Perini et al., 2017). An English language SG called Hello You consisted of several quests and 21 mini games for children between 10 and 12 years old to improve their English vocabulary. The learner must walk around in the SG and gather information to fulfil the quests by finding objects, people, or words hidden in conversations with non-player characters (Barendregt & Bekker, 2011). As for math subject, the suite of Math Snacks products was a SG designed to help middle school learners understand traditionally misunderstood mathematics concepts. As of 2015, it included six animations and five games, all of which were available in English and Spanish and freely available to play online (http://mathsnacks.org) (Winburg et al., 2016). Another SG called Alien Rescue was created for middle school students to improve their problem-solving skills and science knowledge about the solar system (Liu, 2005).

Besides the above single SG products, there are also organizations dedicated to design and develop SG, such as The Education Arcade (TEA) and Muzzy Lane (http://www.muzzylane.com/). TEA is based at the Massachusetts Institute of Technology

(Jenkins, Klopfer, Squire, & Tan, 2003; Klopfer & Yoon, 2005; Klopfer & Haas, 2012), and has developed more than fifteen SG to support teaching across math, science, humanities, arts, and social sciences at high school or the early college level. For example, TEA developed Revolution (Jenkins et al., 2003), which was a multiplayer historical role-playing game. In this SG, students played a role of a townsperson in a colonial Virginia community confronting the events leading up to the American revolution. TEA also developed four different UbiqBio games that focused on four subjects of a high school biology curriculum including classical genetics, protein synthesis, evolution, and food webs (Perry & Klopfer, 2014). Muzzy Lane is a technology company founded in 2002 that specifically develops SG. They have developed more than nine SG covering area such as language (Practice Spanish: Next Best Thing to Being There), history (Making History), mathematics (Algeburst), biology (Hungry Birds), healthy eating and cooking (StudenTopia), computer skills (Tech Town), and even tobacco cessation (QuitIT). They have priced all the SG for sale, and as a technology company, there were few research publications about these SG. One conference proceeding mentioned that they used Making History for teaching World War II in a teacher training program to help pre-service teachers try out teaching the curriculum (Williams, Lai, Ma, & Prejean, 2008). Another journal paper described the development of QuitIT, a SG promoting skills for coping with smoking urges (Krebs et al., 2013).

In summary, this study agrees with Loh et al.'s (2015) definition that SG are "digital games and simulation tools created for non-entertainment use, but with the primary purpose to improve skills and performance of play-learners through training and instruction" (p. 7). In addition, literature shows SG have been applied in different fields including healthcare, defense, arts, government, industrial engineering, and educational institution. Furthermore, not only academic scholars investigated SG's value in different subject fields, but technology companies also worked on the development and tried to commercialize their SG products.

Serious Games with Problem-based Learning pedagogy

To take advantage of PBL, researchers have adopted PBL pedagogy in designing and developing SG in both K-12 and university settings. These SG were also developed for different subject matter, such as science, mathematics, computer skills, and so on (Hou & Li, 2014; Lee & Chen, 2009; Sánchez & Olivares, 2011; Liu et al., 2014; Spires, Rowe, Mott, & Lester, 2011). For example, Crystal Land was a SG adopting PBL pedagogy (Spires et al., 2011), which was aligned with North Carolina's standard for eighth-grade microbiology. Learners played the role of the protagonist, and needed to solve a problem regarding the cause of an outbreak on a tropical island.

Alien Rescue team also adopted PBL pedagogy to design the SG for middle school science subject (Liu, 2005; Liu et al., 2014). In this SG, students faced a problem—finding appropriate planets for six displaced alien species in the solar system—which required them to learn space science knowledge and practice problem-solving skills. Lee and Chen (2009) investigated a SG with PBL pedagogy—Frog Leaping Problem. This SG required ninth-grade learners to solve a non-routine mathematical problem — "If there are 'n' frogs in the left (right) group and 'n' frogs in the right (left) group, how many times do you move the frogs to finish the game?" Sánchez and Olivares (2011) also described two SG with PBL pedagogy—Museum and BuinZoo—that guided eighth-grade student visits to a museum and zoo to learn about the "Evolution of species" topic. Both SG required students to resolve a problem while visiting the museum or zoo. For Museum, students needed to figure out how to explain fossilization and the evolution of species. For Buinzoo, students needed to explain three species' adaptions during the evolution.

Besides middle school students, literature also indicated SG with PBL pedagogy had been used in university settings. For example, university students were situated in a locked room in the SG Boom Room. They had to escape the room by finding all computer components in the room and assembling them correctly within 10 minutes (Hou & Li, 2014). Hou and Li (2014) designed and developed this SG to help university students to learn personal computer assembly through problem-solving process.

In summary, there have been many SG that have adopted PBL pedagogy in different subject areas to enhance learning. Whether SG adopted PBL pedagogy or not, SG stakeholders such as teachers, researchers and designers wanted to improve SG design. The following paragraphs detail techniques that designers, developers and researchers have used to study and improve SG.

Serious Games Analytics

Commercial games owners want to know which games players enjoy and are interested in what kind of content players are willing to pay for in the future. Similarly, SG stakeholders are interested in understanding what learners might do in certain learning scenarios. Even though profits are important to some SG developers, the primary goal for SG is to enhance learning or improve skills (Loh et al., 2015). To achieve this goal, serious games analytics (SGA) have been used to obtain valuable, actionable insights to improve the learning design of SG, which helps to improve learner skills and performances and convince stakeholders of SG's effectiveness.

Firstly, SGA need to collect learner gameplay data including both "Ex-situ" and "In-situ" data (Loh et al., 2015, p. 16). Ex-situ data are collected from "outside the system," such as user survey data—demographics, feedback (Kang, Liu, & Liu, 2017; Ke, 2008; Liu, Toprac, & Yuen, 2009), pretest/posttest (Kang et al., 2017; Ke, 2008), think-aloud data (Ke, 2008), focus-group interviews (Liu et al., 2009), videotapes of student game-play (Perkins, 2016) and eye tracking technique (Kickmeier-Rust, Hillemann, & Albert, 2011; Tsai et al., 2016). Reese et al. (2015) used a survey to collect player's age, gender, and ethnic group information to understand different learner learning progress in a SG called Selene. Ke (2008) used think-aloud technique to ask participants to report whatever went through their mind as they were playing a SG. They recorded the participants' verbal self-report for further analyzes. Perkins (2016) video-recorded learner gameplay and transcribed the video-recorded data into text format for further analyses. Kickmeier-Rust et al. (2011) used Tobii 1750 eye tracker equipment to capture nine Austrian learners' (age: M = 13) eye movement while playing a geography SG for learning about earth.

Different from Ex-situ data, In-situ data are collected from "within the system," such as player in-game actions and behaviors that are recorded by the computer (Kang et al., 2017; Reese et al., 2015; Reese & Tabachnick, 2010). By collecting "In-situ" computer log data, Reese and Tabachnick (2010) investigated learner "learning moments" in the SG Selene, which was designed for standard-based science about fundamental geology and space science concepts. The goal of the game was to slingshot particles to build the Earth's Moon (accretion), and then change it over time by peppering its surface with impact craters and flooding it with lava. The researchers calculated player progress score every 10 seconds on their gameplay then generated a timed report based on player progress scores toward the game goal. They interpreted the scoring as continuous data, calculated for interpretation as "-1" (away from goal), "0" (no progress), or "1" (toward goal). They also generated a gesture report, which was based on a player- or game-initiated event (behavior) that changed the game state. By using these data, they identified the "ah ha!" moment—"moment of learning"—for players. They named this moment as accretionLM, which was the instant when a learner transitions his/her behavior from initiating very high velocity slingshot gestures to sustained low velocity slingshots in the game. This transition indicated player successfully learned how to correctly execute accretion in the game to complete the task. In their follow-up study (Reese et al., 2015), they collected the same type of "In-situ" data.

After the data collection, SGA adopt different data analyzing techniques to generate actionable insights. These techniques include coding, descriptive analyses, drawing analyses, creating statistical models, developing metrics, data visualization, and so on (Kang et al., 2017; Ke, 2008; Liu et al., 2009; Liu et al., 2015; Loh & Sheng, 2014; Loh, Li, & Sheng, 2016; Perini et al., 2017; Reese et al., 2015; Reese & Tabachnick, 2010). Loh et al.'s (2016) exploratory study adopted similarity measures of learner in-game actions (i.e., navigational sequences) as a performance metric to differentiate learners. In this study, they created a SG using *Unity3D* game engine. Learners needed to escape a maze from an exit portal as quickly as possible in the SG. They recruited 31 university students to solve the maze puzzle and traced all participant Gameplay Action-Decision (GAD) in the SG. As a result, 19,960 action data points were collected for all 31 players. Based on learner in-game actions, they grouped these 31 learners into 3 groups such as *Explorer* (i.e., learners who were not just satisfied for one working solution but looking for better solutions), *Fulfiller* (i.e., learners who were only fulfilling the goal of SG, not searching for alternative solutions), and *Quitter* (i.e., learners who tended to give up too early or too easily). Then, they examined five most commonly used similarity measures to see if these measures, or combinations of these measures would identify different learner GAD. These measures include *Dice*, *Jaccard* (*Jac*), *Overlap* (*OVL*), *Cosine* (*Cos*), and the *Longest Common Substring* (*LCS*) coefficients. They further used two metrics such as Average Similarity Index (ASI), and Maximum Similarity Index (MSI) to process similarity measure coefficients. The findings indicated that similarity measures on GAD (i.e., navigational sequences) can differentiate players in SG. In addition, their finding revealed that combined similarity measures could gain more strength on understanding learner actions. Specifically, combined ASI and MSI had 93.55% correct rate to identify player profile compared to MSI alone (77.42%), and ASI alone (67.74%). Since there are still 6.45% rate of identifying incorrect learner profile, they also suggested more research is needed to "create or develop new metrics and methods" for studying learner action and behavior in SGA.

In addition, several software tools have been used to help with SGA, such as *Nvivo/Atlas* for coding, SPSS/R for statistical analyses, and Tableau/Processing/R for data visualization (Kang et al., 2017; Ke, 2008; Loh et al., 2016). Ke (2008) used Nvivo to code participant think-aloud data. Reese and Tabachnick (2010) used SPSS to conduct repeated-measure ANOVA analyses on learner pre- and post- time reports. Loh et al. (2016) used *R* and "stringdist package" for their similarly analyses. Using *Tableau* software, Kang et al. (2017) combined learner activity log data with visualization techniques to understand relationship between learner in-game activities and their performances. They generated different visualization of learning paths based on the log data then analyzed the differences with respect to learner performances.

In summary, researchers have used various SGA techniques and software tools to provide insights on learner game activities. These insights can further assist SG stakeholders in making future decisions including enhancing the design of SG, improving learner skills and performance, and increasing return of investment (Loh et al., 2016).

Effects of Serious Games

In the past two decades, many researchers conducted studies on the effects of SG. Studies have shown that SG had positive effects on learner motivation, knowledge acquisition, and skill development including problem-solving, collaboration, and social skills. Scholars also indicated that there were challenges in using SG for learning and teaching. The following section will briefly review studies on these positive effects and challenges.

Positive Effects

Researchers suggested that SG had positive effects on learner motivation and learning outcomes. Prensky (2001) created more than 50 software games for learning. He claimed putting learning into game context was much more motivating than formal academic education for today's students. This claim was supported by many studies that demonstrated a significant relationship between learner motivation and SG. Particularly, learners that used games tend to be more engaged and intrinsically motivated when they actively solved problems (Barendregt & Bekker, 2011; Huang, Huang, & Tschopp, 2010; Liu & Chu, 2010; Liu et al., 2009). Liu and Chu (2010) conducted a quasi-experimental study on a SG environment HELLO for seventh-grade student English learning. The results revealed that learners who used HELLO demonstrated increased attention and motivation for learning. Interviews with students in the experimental group suggested that this SG provided an enjoyable experience for assisting listening and speaking English.

Besides SG's motivational power, researchers also argued that there were positive effects on student learning outcomes through SG. Connolly and his colleagues reviewed 272 empirical computer games and SG studies from January 2004 to February 2014, and found the most frequently occurring outcome was knowledge acquisition/content understanding, followed by physiological and social skills outcomes (Boyle et al., 2016; Connolly et al., 2012). While some studies suggested that SG had no impact on student knowledge acquisition (Annetta, Minogue, Holmes, & Cheng, 2009; Ke, 2008; Spires et al., 2011), researchers tended to report that playing SG led to better knowledge gain compared to other conditions, such as traditional lectures (Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013), project-based instruction (Huizenga, Admiraal, Akkerman, & Dam, 2009), and web-based content learning experiences (Cheng, Su, Huang, & Chen, 2014; Papastergiou, 2009).

Wouters et al. (2013) used meta-analytic techniques to investigate whether SG were more effective for learning than conventional instruction methods, such as lectures, reading, drill and practice, or hypertext learning environments. They located 38 empirical studies published between 1990 and 2012, and found SG learners learned more than those taught with conventional instruction methods (weighted mean effective size in knowledge acquisition: $d = 0.27$, $p < .05$). Scholars from the Netherlands also compared the history knowledge acquisition between students (age from 12 to 16 years) who played a SG called Frequency 1550 and those who followed the regular project-based lesson series (Huizenga et.al, 2009). They suggested students gained significantly more knowledge about medieval Amsterdam than those who received regular project-based instruction ($d = 0.62$; $p < 0.001$). Another study compared SG with web-based content also revealed that ninth-grade students who learned by playing a biology SG called Humunology (adjusted $M = 1.51$) significantly outperformed those who learned by using web-based content (adjusted $M = 1.31$) on procedural knowledge (Cheng et al., 2014).

Besides knowledge acquisition, researchers also proposed that SG had positive effects on skill development such as problem-solving, collaboration, and social skills. For problem-solving skills, Cheng et al. (2014) suggested the SG Humunology, which was developed for learning body defense system, could provide a new way to learn scientific concepts and "develop the mental resources for problem solving" (p. 831). Several empirical studies also reported that SG could improve learner problem-solving skills (Lester et al., 2014; Sánchez & Olivares, 2011). Sánchez and Olivares (2011) compared the self-reported problem-solving skills between eighth-grade students who used SG Museum and BuinZoo for the class and student who did not use these SG. The result indicated students in the experimental group (i.e. students who used these SG) had higher scores in problem solving skills. Lester et al. (2014) also reported that after playing Crystal Land, 400 learners in fifth-grade gained significant problem-solving skills based on pre (M = 1.83, SD = 1.47) and post (M = 2.03, SD = 1.49) problem-solving question scores.

As for collaboration skills, Sánchez and Olivares (2011) asked students to work collaboratively in SG Museum and BuinZoo. Result indicated students in the experimental group (i.e. students who used these SG) had a higher score in collaboration skills. With respect to student social skills, Thomas and Vlacic (2102) constructed a SG scenario called TeamMATE, and argued that this SG had the potential for learners to interact with both machines and others, which could help them develop collaboration skills and decision-making.

In addition, although lacking empirical support, scholars suggested SG had positive effects on innovativeness and creativity, due to the instant feedback and risk-free environments, which invited exploration and experimentation, curiosity and discovery learning (Gee, 2007; NESTA Futurelab, 2004; Prensky, 2001).

Challenges

Besides the positive effects, researchers have identified some challenges while adopting SG for learning, specifically on sustaining learner motivation, utilizing SG, and employing effective SG design strategies (Perrotta, Featherstone, Aston, & Houghton, 2013; Yang, 2012).

Although SG could enhance learner motivation and engagement, it was unclear whether the impact would be sustainable (Perrotta et al., 2013). One study asserted that motivation was sustained when using a commercially available game, SimCity Societies, in a ninth-grade civics and society class, but the motivation was only measured over a single semester (Yang, 2012). Therefore, it was unknown whether the study was long enough to be confident of SG's continued impact on motivation.

Researchers indicated another challenge was SG might not generate learning outcomes without instructional facilitation (Erhel & Jamet, 2013; Harr, Buch, & Hanghøj, 2008). Harr et al. (2008) studied eight ninth-grade student critical thinking while playing a SG called Global

Conflicts: Palestine (GC:P). This SG offered students an experience of exploring the Israeli-Palestinian conflict through the eyes of a journalist. They conducted observations and post-game interviews with teachers and students and suggested this SG was unlikely to teach critical thinking merely by letting students interact with the game environment without actively engaging in meaningful conversation. Erhel and Jamet's (2013) study supported Harr et al.'s finding. They conducted two experiments on a SG called ASTRA (Appréhender par la Simulation les TRoubles liés à l'Age), which was designed for undergraduate students to learn about four different medical diseases. In their first experiment, they designed two different types of instructions (learning instruction vs. entertainment instruction) in the SG. Results showed the learning instructions elicited deeper learning than the entertainment instructions. In the second experiment, they added regular feedback about learner performance in entertainment instructions. The results showed that with feedback students also engaged in deeper learning. Therefore, these two experiments demonstrated a SG environment can promote learning despite the instructional method if the SG provided learners a way to actively process the educational content.

In addition to motivation and utilization, many studies highlighted the challenges of game design (Foster & Mishra, 2009; Michael & Chen, 2006; Pedersen, Liu, & Williams, 2002). Scholars indicated that compared to video game players, SG players might have little or no experience with video games, therefore, designers needed to take this into consideration and make SG accessible for first-time players (Michael & Chen, 2006). Besides the accessibility for the first-time game player, Lee and Chen (2009) suggested it was important to add more scaffolding techniques in a mathematical SG to help students with low-prior-content-knowledge to enhance their performances of mathematical problem-solving.

Researchers also identified a gap between disciplinary knowledge and SG design and suggested designers and researchers need to think more deeply about how content (disciplinary knowledge) can be fruitfully integrated within the SG design (Foster & Mishra, 2009; Pedersen et al., 2002). Specifically, Foster and Mishra (2009) suggested designers should deal with the interaction among content, pedagogy, and technology (Technological Pedagogical Content knowledge [TPACK] framework) while designing for SG. They argued SG design should consider the pedagogical affordance of different game genres (pedagogy) and how to integrate disciplinary knowledge (content) with the SG (technology).

Another group of scholars also described their challenges while designing and developing the SG Alien Rescue (Liu et al., 2014; Pedersen et al., 2002). According to Liu et al. (2014), the SG employed a design based research (DBR) approach, which followed four phases: (a) concept; identify the goals, objectives, and desired outcomes of project related tasks, (b) design; identify components to be produced or refined for the SG, (c) development; develop functional components for feedback and revision, and (d) implementation; test the SG in real classroom setting. Each phase was coordinated through continuous and iterative processes of planning and evaluation. They mentioned that there were at least three challenges they had faced including always needing research collaborators during the design phase, highly demanding translation between research evidence and design modifications, and needing multiple iterations.

In summary, through SGA, researchers have identified benefits and challenges in SG. To move this field forward and better understand SG, many researchers have started to investigate learner role in SG. Particularly, the impact of learner characteristics on SG have been insensitively studied. I will review the relevant literature in the following paragraphs.

Learner characteristics affect learning in Serious Games

Besides studies on the effects of SG, researchers have investigated the impact of learner characteristics in SG. Particularly, this section reviews studies on learner metacognition, goal orientation in SG and other studied learner characteristics in SG, such as age, race/identity, gender, in-game behaviors, fantasy proneness, and self-efficacy.

Metacognition

A few studies indicated that learner metacognition affected their learning in SG (Tsai et al., 2016), and SG affected learner metacognition (Ke, 2008; Perkins, 2016). Tsai et al. (2016) utilized eye-tracking technology to study 22 college student eye movements in a physics game called Escape the Lab, which was a role-play problem-solving game for teaching electromagnetism. The researchers grouped students into high- and low-conceptual-comprehension—two groups based on their comprehension post-test score—and measured student prior knowledge and flow experiences. The results indicated that students in the high-comprehension group had better metacognitive controls of visual attention in the SG, while the low-comprehension group students had difficulties in the SG, particularly on decoding the conceptual representations of the SG.

Ke (2008) used mixed methods to investigate 487 fifth-grade student metacognitive awareness and learning in a quasi-experimental study. In this study, students were divided into paper-and-pencil drills group and game group who played a mathematics SG called ASTRA EAGLE. Ke (2008) used Jr. MAI to measure student metacognition. In addition, field observation and participant think aloud data were also collected in the study. The results indicated there was no statistically significant difference on improving student metacognitive awareness between the paper-and-pencil drills group and SG group. However, the qualitative data indicated that SG group participants seemed to more frequently engage in metacognitive regulation processes and exhibit more self-regulative manners than those in paper-and-drills group. In addition, Perkins's (2016) doctoral dissertation suggested playing SG resulted in student metacognitive ability gain.

Although it was not clear the clear the causal relationship between metacognition and learning in SG, these preliminary studies suggested that there was a connection between them. Therefore, more research is needed to explore the role of metacognition for learning in SG.

Goal orientation

Goal orientation affected student achievement and behavior within SG (Hsieh et al., 2008; Liu, 2005; Liu et al., 2015; Tran, Smordal, & Conley, 2016). Liu (2005) studied student goal orientation from a motivational perspective. She adopted *Motivated Strategies for Learning Questionnaire* (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991) to measure student intrinsic and extrinsic goal orientation (N = 437) in Alien Rescue. The results suggested student science knowledge scores were positively related to their intrinsic goal orientation. Hsieh et al. (2008) studied the same SG but used a larger sample size (N = 549), and adopted the *Achievement Goal Questionnaire* (Elliot & Church, 1997) to measure goal orientation. Results suggested student performance-avoidance goal orientation moderated the relation between self-efficacy and science achievement, indicating self-efficacy had positive influences on achievement when students were not performance-avoidance oriented.

In 2015, Liu and her colleagues studied student goal orientation in the same SG again, but they adopted a different measurement called the *Patterns of Adaptive Learning Scales* (PALS; Midgley et al., 2000) to assess student goal orientations including mastery, performance-approach, and performance-avoidance goal orientations (Liu et al., 2015). Using data visualization of different goal orientation student group log data in the SG, they analyzed student (N = 38) behavior pattern. Using "natural groupings of the goal orientation scores" (p. 190), the findings suggested students with high mastery-oriented scores (N = 9) tended to behave more appropriately during the problem-solving processes, and were more productive in tool use than students in other groups. Students with high scores in performance approach (N = 3) and avoidance goal orientation (N = 3) showed inappropriate behavior patterns, such as exploring more fun tools rather than gathering information to solve the problem.

Another study conducted by Tran and her colleagues in Norway had a slightly different finding on student goal orientation effects in a SG, which was designed for high school students to learn about condensation and evaporation science (Tran et al., 2016). They selected four participants—ages 15 and 16—who represented a diverse goal orientation profile in a high school including (a) predominantly mastery oriented; (b) predominantly performance-approach oriented; (c) predominantly performance-avoid and mastery oriented; and (d) similarly performance-approach, performance-avoid and mastery oriented. They closely studied student behavior and analyzed data from self-reports about goal orientations using *2 x 2 Achievement Goal Questionnaire* (Elliot & McGregor, 2001), video analyses of the museum visit, retrospective think-aloud on the videos and semi-structured interviews after the visit. The results suggested students who had mastery goal orientation did not always have adaptive outcomes. For example, one student who reported a high level of mastery goal orientation showed diminished engagement with the game and activities because the game started to disappoint this student over time. In addition, students who had performance-avoidance goal orientation did not always had negative outcomes. For example, one student who had a higher level of performance-avoidance goal orientation adaptively participated in playing the game, because she also had a higher level of mastery goal orientation.

Despite the meaningful preliminary finding on student behavior difference based on their goal orientation, both Liu et al. and Tran et al.'s studies had small sample size, which created a significant limitation for generalization. In addition, both studies did not include the task goal orientation dimension. Therefore, to fully understand learner learning process in SG based on goal orientation, future studies need to include more participants. Measurements of task-approach and task-avoidance goal orientation dimensions also need to be added.

Other learner characteristics

Besides goal orientation, researchers also examined other leaner characteristics related to SG including (a) personal characteristics such as age, race/ethnicity and gender (Alyaz et al., 2017; Nazry & Romano, 2017; Reese et al., 2015); (b) academic characteristics such as prior knowledge (Lee & Chen, 2009), and (c) social/emotional characteristics such as in-game behaviors, fantasy proneness, and self-efficacy (Ketelhut, 2007; Lee, 2015; Loh et al., 2016).

Personal Characteristics

Studies indicated play-learner age was a significant factor in some SG, but it was not clear which age group benefitted most from SG in general (Nazry & Romano, 2017; Reese et al., 2015). Nazry and Romano (2017) suggested younger learners (18-24 years old) achieved a significant higher performance score (score $M = 18.16$) compared to older users (25-34 years old, score $M = 13.44$) in a SG for learning about religion, $F(2, 49) = 3.755$, $p = 0.030$; however, Reese et al. (2015) suggested that younger learners (12-year-olds) experienced a slightly greater learning challenge in SG *Selene* than did the older learners (15-year-olds). In another study about a language SG, the participants were all undergraduate students. It was found that the higher age group participants (above 23 years old) had a statistically significant higher test results compared to younger participants (age 18-19, and age 20-22), $F\ (2, 57) = 25.15$, $p = .000$ (Alyaz et al., 2017).

With respect to the ethnic group, in SG Selene, Reese et al. (2015) found that game completers reported a diverse background: White, African, African-American, Asian, Caribbean, Hispanic, mixed, Native American and others, but a slightly higher proportion (3%) of white (non-Hispanic) players than the US population average completed the game.

Researchers have also looked at the gender difference in SG, and the results were inconsistent. On the one hand, some researchers found gender was a significant factor for SG (Nazry & Romano, 2017; Reese et al., 2015). Reese et al. (2015) found out, although an equal number of males and females registered to play Selene, three male players completed the game for every two females, $\chi^2(1, n = 999) = 12.6$, $p < .001$. They also found females took longer to finish the SG than males. Nazry and Romano (2017) studied sixteen adults' (8 women, 8 men, Age: M = 25.88) performance and mood in a SG for learning about religion. Although it was a small sample, the results suggested that women performed better when they were happier.

On the other hand, some studies suggested there were no significant gender effects on learning or in-game performance (Lester et al., 2014; Mavridis, Katmada, & Tsiatsos, 2017; Papastergiou, 2009; Rowe et al., 2011; Spires et al., 2011). Papastergiou (2009) studied 88 students (46 boys, 42 girls, Age M = 16.58) utilizing a SG called LearnMem in computer science class at a Greek high school. She found that there was no significantly difference in learning performance based on gender, and the game was found to be equally motivational for both genders. Mavridis et al. (2017) studied 79 student gender difference (46 boys, 33 girls, age 12–14) and the attitude towards a math SG. They found that student gender did not influence the improved test performances and positive attitudes towards the SG. Although Rowe et al. (2011) found that males tended to report significantly greater presence in a SG called Crystal Island, which was designed for an eighth-grade microbiology curriculum, however, the analyses suggested that differences in presence might be more strongly associated with previous game play experience than gender. The results from other studies on Crystal Island consistently suggested there was no gender differences on student test scores and in-game performances (Lester et al., 2014; Rowe et al., 2011).

Academic Characteristics

With respect to learner academic characteristics, Lee and Chen (2009) suggested that student prior knowledge on reasoning affected their problem-solving performances in SG. They studied seventy-eight ninth-graders in a mathematical SG game for 6 weeks. Based on student Pattern Reasoning Test scores, they grouped 38 students into a high-prior knowledge group, and 40 students into a low-prior-knowledge group. The results suggested students with high prior knowledge in reasoning got higher scores in the problem-solving performances than those with low prior knowledge did (M = 4.99, SD = .57; M = 3.89, SD = .80).

Social/Emotional Characteristics

For learner social/emotional characteristics, researchers studied learner in-game behavior characteristics, fantasy proneness, and self-efficacy. Loh et al. (2016) categorized learners into three categories based on in-game behavior sequences characteristics: (a) *Explorer*—players who are not satisfied with just one working solution, (b) *Fulfiller*—players who are single-minded about fulfilling the goal of the SG rather than exploring the environment, and (c) *Quieter*—players who give up too early or too easily.

Another studied social/emotional learner characteristics was fantasy proneness, which was defined as a characteristic established from early childhood through exposure and engagement with imaginative activities (Merckelbach, Horselenberg, & Muris, 2001). A few studies have investigated individual fantasy proneness effects in SG (Lee, 2015; Liu et al., 2016). Lee's doctoral dissertation investigated the effects of learner fantasy proneness in Alien Rescue (Lee, 2015). The findings showed that students with higher fantasy proneness scores showed better game engagement in the SG; however, too deep an involvement in fantasy also resulted in ineffective and inefficient learning outcomes. Liu et al. (2016) used the same SG, but looked at students' various tool utilization in the SG based on their fantasy proneness level. The results indicated students used game-provided tools differently according to their fantasy proneness level. Specifically, students with higher fantasy proneness spent significantly more time on a tool that embodied more fantasy elements than students with low fantasy proneness. Since this same tool also had more information relevant to the content subject, it was augured that students with high fantasy proneness also spent more time on learning about the information.

Learner self-efficacy has also been studied, and scholars suggested learner self-efficacy affected their learning performance in SG (Hsieh et al., 2008; Liu et al., 2006; Yang, Quadir, & Chen, 2016). Liu et al. (2006) examined the relationship between sixth-grade student self-efficacy and achievement in Alien Rescue, and suggested self-efficacy was a significant predictor of science achievement scores. Yang et al. (2016) also studied how third-grade student self-efficacy affected their English learning performance in a SG. The results revealed self-efficacy had a significant positive influence on student learning performance in the SG—students with higher self-efficacy performed better than those with lower self-efficacy.

Besides learning performance, researchers also suggested self-efficacy impacted learner in-game behavior. Ketelhut (2007) studied the relationship between 100 seventh-grade student self-efficacy and their data-gathering behaviors while engaged in a SG called River City, which was designed to engage students in a collaborative scientific inquiry-based learning. She examined student moment-by-moment data-gathering behavior in the SG, and found self-efficacy played different roles in behavior during the process. For example, high self-efficacy students engaged in more data gathering than students with low self-efficacy when they first entered the SG; then, the impact of student self-efficacy changed; by the end of the study, student self-efficacy did not impact data-gathering behavior. In addition, student self-efficacy level did not affect how many sources students chose to gather data from.

Summary

To provide a theoretical foundation and research evidence for this study, this chapter reviews constructivism learning theory, two important learner characteristics (i.e., metacognition and goal orientation), PBL, and SG. From a constructivist perspective, learner plays an important role in the learning process; therefore, learner characteristics are important for understanding learning. According to literature, both characteristics have been proved to affect learner academic performance and other outcomes. Particularly, learner metacognition is the "engine" that starts, regulates and evaluates the cognitive processes during learning. For learner goal orientation, it plays an important role at the very early stage of metacognitive regulation, which can further guide the entire metacognitive regulatory process.

As a constructivist instructional method, PBL has several advantages compared to traditional classroom instruction. These advantages include engaging students, maintaining long-term retention, and developing skills. As for SG, previous studies argued that there was a positive relationship between SG and the learning process, learning outcomes, and engagement. To take advantage of PBL, some SG adopt PBL pedagogy. However, there are also challenges while adopting PBL, especially for young learners and teachers in the K-12 setting. These challenges include a lack of guidance, demanding self-regulation skills, and collaboration skills. Failing to overcome these challenges may cause student frustration, disengagement, misconception, and eventually failure in PBL. Meanwhile, there are also challenges while using SG, such as how to sustain learner motivation, utilize SG, and better design SG.

According to literature, researchers have used various SGA techniques (e.g., drawing analyses, creating statistical models, developing metrics, and data visualization) and software tools (e.g, R, Processing, Tableau, and so on) to provide insights on learner game activities, which may assist in enhancing the design of SG, improving learner skill and performance, and eventually increasing return on investment for all parties. Particularly, learner metacognition and goal orientation have been proved to affect academic performance, problem-solving, and other outcomes in PBL and SG. However, little research has looked at how metacognition and goal orientation together would affect learner problem-solving (both problem-solving performances and problem-solving processes) in a SG environment that adopts PBL pedagogy. In addition, literature shows the needs to utilize both quantitative measurement and qualitative methods to better understand learner metacognition, especially the needs for analyzing computer tracked data for researching metacognition, because it can record the entire learning processes and outcomes of learner. Furthermore, literature indicates that the goal orientation model has evolved multiple times through four decades, and using the 3 X 2 questionnaire (Elliot et al., 2011)—the most updated and widely used measurement for college students—is needed to understand the impact of learner goal orientation in SG.

Therefore, to fill the gap in understanding the interaction among learner metacognition, goal orientation, and problem-solving in SG environments, and to advance current study on metacognition and goal orientation using computer tracked data and the most updated measurement, this study proposes to examine learner problem-solving (both problem-solving performances and processes) based on learner characteristics (i.e., metacognition and goal orientation) using the most updated measurements in a laboratory setting. The next chapter will describe in detail the specific methodology I used for this study.

Chapter 3: Methodology

To understand learner problem-solving (problem-solving performances and processes) based on metacognition and goal orientation, this study employed sequential mixed research design, SGA and multiple data sources to analyze learner problem-solving in a SG environment in a laboratory setting. This chapter will describe the research context, participants, data sources, research procedure and data analysis to address the research questions.

RESEARCH QUESTIONS

This study investigated learner problem-solving (both problem-solving performances and processes) based on their two characteristics (i.e., metacognition and goal orientation) in a SG environment designed for learning space science. The research questions are:

1. To what extent are problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)? There are three sub-questions:

(a) Is there a statistically significant difference in learner problem-solving performances based upon metacognition (high metacognitive level, low metacognitive level) and goal orientation (task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation)?

(b) Can learner metacognition and goal orientation predict problem-solving performances?

(c) What are the reasons for any problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)?

2. To what extent are problem-solving process differences based on learner characteristics (i.e., metacognition and goal orientation)? There are five sub-questions:

(a) What are learner problem-solving process patterns?

(b) Are there any problem-solving process pattern differences among students, based on their metacognition?

(c) Are there any problem-solving process pattern differences among students, based on their goal orientation?

(d) Are there any problem-solving process pattern differences based on the interaction between learner metacognition and goal orientation?

(e) What are the reasons for any problem-solving process pattern differences based on learner characteristics (i.e., metacognition and goal orientation)?

RESEARCH CONTEXT

Alien Rescue as a Serious Game environment

This study utilized Alien Rescue (AR, http://alienrescue.edb.utexas.edu; Liu et al., 2016) as the SG environment. AR adopts PBL pedagogy, and focuses on teaching knowledge about our solar system and complex problem-solving skills for middle school students. In this environment (see Figure 2), learners face an ill-structured problem—to save six displaced alien species (i.e., Akona, Eolani, Jakala-Tay, Kaylid, Sylcari, and Wroft) due to the destruction of their home planets. Learners need to utilize the information provided within this environment to find the suitable planets for these aliens and explain their rationale in the problem solution form.

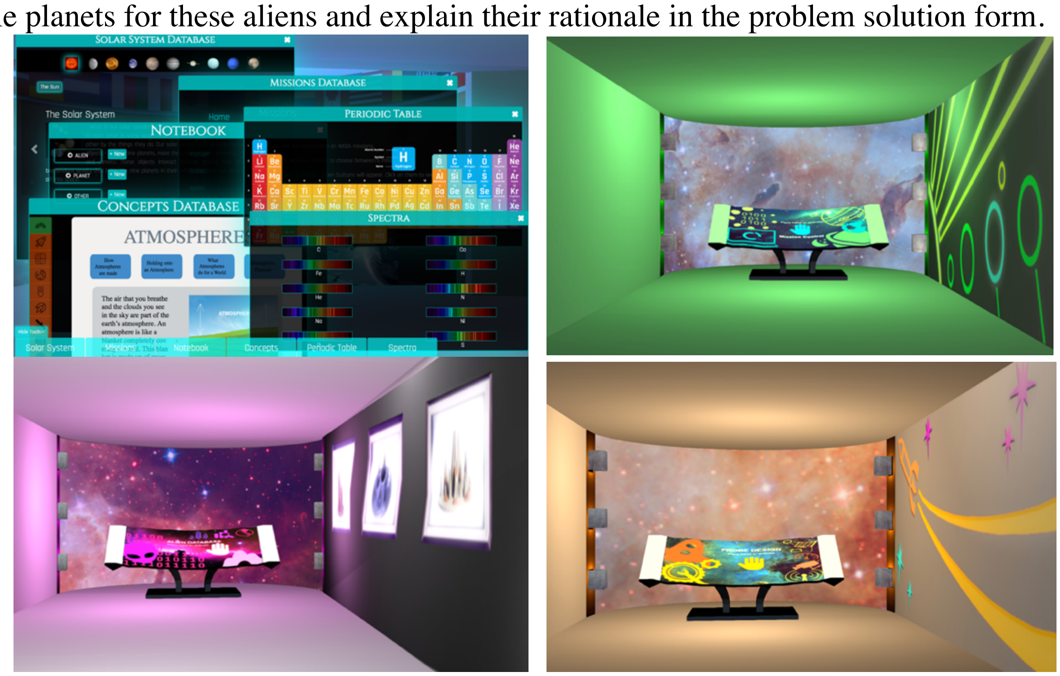


Figure 2. Alien Rescue environment

There are four rooms in the AR environment including the Main Room, Probe Design Room (Room P), Alien Information Room (Room A), and Mission Control Room (Room C). In addition, there is a game console in each room (i.e., Communication Center Console, Probe Design Console, Alien Database Console, and Mission Control Console respectively). To enter the rooms, the learner must go to the Main Room then go through the different sliding gates (i.e., Probe Design Room gate, Alien Information Room gate and Mission Control Room gate) to each room (see Figure 3).

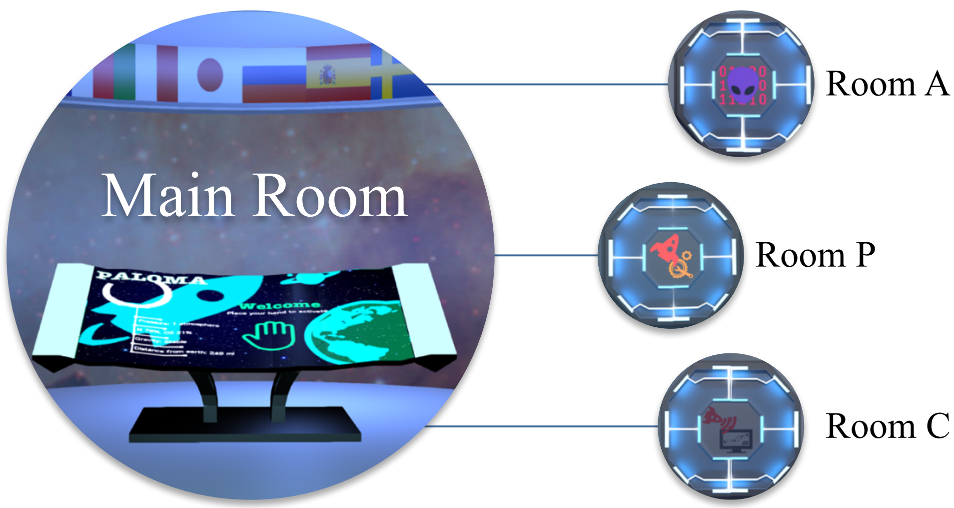


Figure 3. Alien Rescue room layout

Furthermore, there is a toolbar in AR, which can be accessed in all four rooms. Overall, AR provides 10 cognitive tools that can be categorized into four types including (a) sharing cognitive load, (b) supporting cognitive process, (c) supporting otherwise out-of-reach activities, and (d) supporting hypothesis testing (Liu & Bera, 2005; Liu et al., 2014, 2016). See Table 3 for tool descriptions and locations.

Table 3: Descriptions of Cognitive Tools Provided in Alien Rescue

Tool Types	Name	Tool Location	Tool Functions
Share cognitive load	Alien Database	Room A	Presents textual descriptions and 3D visuals of the alien information including alien home solar system, journey to Earth, and the characteristics and needs of each species.
	Solar System Database	Toolbar	Provides information on the planets and moons in our solar system. Intentionally incomplete data ensures learner must design and send probes to test hypotheses.
	Missions Database	Toolbar	Presents information on the mission, technology and findings of historical NASA probe launches.
	Concepts Database	Toolbar	Provides supplemental instruction on scientific concepts presented elsewhere in AR
	Spectra	Toolbar	Provides students spectral information to interpret spectral data encountered in AR.
	Periodic Table	Toolbar	Provides a periodic table of all the elements for reference.
Support cognitive process	Notebook	Toolbar	Provides a place for students to take notes as they engage in solving the central problem.
Support otherwise out-of-reach activities	Probe Design Center	Room P	Allows students to design, build and send probes to gather data on worlds in our solar system.
Support hypothesis testing	Mission Control Center	Room C	Provides an interface to view data from launched probes.
	Message Tool/ Problem Solution Form	Main Room	Allows students to read messages regarding the background story. Provides the Solution Form, which allows students to submit their planets recommendations and rationale for review by teachers.

In this study, AR was used as a SG environment to understand the impact of learner metacognition and goal orientation on their problem-solving. Specifically, undergraduate students were recruited to solve the problem—helping one displaced alien species (i.e., Jakala-Tay) to find a suitable home within 60 minutes. By studying undergraduate students using AR in a laboratory setting, this study hope to control variables that might affect learner problem-solving processes (e.g., teacher guidance and peer influences in the real classroom).

Previous Alien Rescue studies

This study is informed by previous AR studies in four main areas including two studies on metacognition (Bogard et al., 2013; Liu et al., 2004), three studies on goal orientation (Hsieh et al., 2008; Liu, 2005; Liu et al., 2015), six studies on problem-solving processes (Bogard et al., 2013; Kang, 2017; Liu et al., 2004; Liu & Bera, 2005; Liu et al., 2009; Liu et al., 2015), and six studies on data collection and analysis in AR (Kang, 2017; Liu et al., 2004, 2009, 2013, 2015; Liu & Bera, 2005). These foundation studies will be discussed in the following paragraphs, See Table 4 for a summary of these studies organized in a chronological order.

Metacognition in Alien Rescue

This study is informed by two AR studies about learner metacognition (Bogard et al., 2013; Liu et al., 2004). In 2004, Liu et al.'s (2004) exploratory study on sixth graders indicated learner tool use patterns reflected learner characteristics (i.e. information processing and metacognition orientated). Specifically, students who were more metacognitive oriented were more thoughtful and consistent in their tool selection, while students who were more information processing oriented were more active on their tool use and spent more time on action-related tasks. They suggested that more research is needed to understand the connection between learner characteristics and tool use patterns.

Bogard et al.'s (2013) study took a closer look on 15 advanced learners (i.e., graduate students) cognitive processes while solving the problem in AR. They conducted a cross cluster analysis, and the results showed that there were 4 clusters emerged based on learner prior knowledge and problem solution performance including: (a) Cluster 1 (N = 4): low prior knowledge, unsuccessful; (b) Cluster 2 (N = 3): medium prior knowledge, successful; (c) Cluster

3 (N = 4): low prior knowledge, highly successful; (d) Cluster 4 (N = 4): high prior knowledge, highly successful. The authors identified learners in cluster 1 as “inadequate self-regulation learners” (p. 482), who focused more on finding shortcuts than building knowledge and identifying the factors impacting the problem. Cluster 2 learners were identified as “delayed self-regulation” (p. 484), who engaged in planning and strategizing processes, but did not reevaluate their approach until they were lacking resources of time and money in AR. Cluster 3 learners were identified as “high self-regulation” (p. 486), who focused on self-regulating the procedures for problem-solving, and were highly successful. Although this study only studied 15 graduate students, and did not connect cluster 4 student success with their metacognitive regulation, the Cluster 1-3 indicated that leaner metacognition, particularly the self-regulation element, impacted advanced learner problem-solving.

Informed by these two studies, this current study looked at learner tool use patterns during problem-solving, and used a larger sample size to further understand the connection between learner metacognition and problem-solving.

Goal orientation in Alien Rescue

This study is also informed by three AR studies on learner goal orientation (Hsieh et al., 2008; Liu, 2005; Liu et al., 2015). The first study examined 437 sixth-grader goal orientation from a motivational perspective (Liu, 2005). It adopted Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991) to measure student intrinsic and extrinsic goal orientations in AR. The results suggested student science knowledge scores were positively related to their intrinsic goal orientation. The second study used a slightly larger sample size (N = 549), and used the *Achievement Goal Questionnaire* (Elliot & Church, 1997) to measure interactions between sixth-grader goal orientation and self-efficacy (Hsieh et al., 2008). Results suggested student performance-avoidance goal orientation “moderated the relation between self-efficacy and science achievement” (p. 34). Specifically, it showed self-efficacy had positive influences on student science achievement when students were not performance-avoidance oriented.

The most recent study investigated student goal orientation using a measurement called the *Patterns of Adaptive Learning Scales* (PALS; Midgley et al., 2000) to assess sixth grader mastery, performance-approach, and performance-avoidance goal orientations (Liu et al., 2015).

Table 4: Previous Alien Rescue Studies

Researchers	Informed Area	Participants	Data Sources	Data Analysis	Research Questions (Q) and Findings (F)
Liu, Bera, Corliss, Svinicki, & Beth (2004)	Problem-solving process; Metacognition	161 sixth graders	Log data, frequency; Science Knowledge Test (SKT); Cognitive Task Questionnaire	Chi-Square; Exploratory factor analysis; One-way MANOVA	Q: Examine the connection between sixth-grader cognitive tool use and the cognitive processes as they solve problem in AR. F: Different cognitive tools were used for different cognitive processes in problem-solving, and the degree of engagement was positively related to the tool use frequency. In addition, tool use patterns reflected learner characteristics (i.e. information processing and metacognition orientated). Students who were more metacognitive oriented were more consistent in their tool selection, while students who were more information processing oriented were more action oriented in performing the tasks.
Liu, 2005	Goal Orientation	437 sixth graders	Motivated Strategies for Learning Questionnaire (MSLQ); SKT	ANOVA	Q: Examine the effect of AR on sixth-grader science knowledge, attitude toward learning science, and motivation toward learning. F: Student science knowledge scores were positively related to their intrinsic goal orientation. Student science knowledge, attitudes toward science and intrinsic goal orientation have significantly increased after using AR.

Table 4 continued.

Researchers	Informed Area	Participants	Data Sources	Data Analysis	Research Questions (Q) and Findings (F)
Liu & Bera (2005)	Problem-solving process	110 sixth graders	Log data: frequency; SKT	Cluster analysis	Q: How the built-in tools were used and if tool use was associated with different problem-solving stages. F: In the early stage of problem-solving processes, students primarily used tools supporting cognitive processing and tools sharing cognitive load. In the later stages, students used multiple tools more often, especially tools supporting hypothesis generation and testing. There was a positive correlation between SKT scores and productive use of the tools: higher score students used the tools more productively than lower score ones.
Corliss, 2005	Metacognition	298 female college students	Metacognitive Awareness Inventory (MAI)	ANOVA	Q: what are the effects of reflective prompts and collaborative learning on problem-solving and metacognitive skills in AR? F: there was no significant effect of reflective prompting, collaborative learning, or an interaction effect on problem-solving performance, near transfer task performance, far transfer task performance, and metacognitive skill performance.
Hsieh, Cho, Liu, & Schallert, 2008	Goal Orientation	549 sixth graders	Achievement Goal Questionnaire; SKT; Eight items in MSLQ for self-efficacy.	ANOVA	Q: What are the interactions between sixth-grader goal orientation and self-efficacy. F: Student performance-avoidance goals moderated the relation between self-efficacy and science achievement, indicating self-efficacy has positive influences on achievement when students are not performance-avoidance oriented.

Table 4 continued.

Researchers	Informed Area	Participants	Data Sources	Data Analysis	Research Questions (Q) and Findings (F)
Liu, Horton, Corliss, Svinicki, Bogard, Kim, & Chang (2009)	Problem-solving process	61 undergraduate students	Log data (*N*=59): frequency, duration; Cognitive Task Questionnaire; 11 Stimulated recall interviews; Problem solutions scores	Chi-Square MANOVA Qualitative analyses	Q: Examine the tool use patterns, and understand what tools were used and why they were used. F: Confirmed the findings from previous two studies: strong connections between cognitive processes and cognitive tool use. No tool use differences based on performance.
Bogard, Liu, & Chiang (2013)	Problem-solving process; Metacognition	15 graduate students	Observation; Think aloud and stimulated recall protocol; Problem solution scores	A cross cluster analysis Qualitative analyses: grounded theory	Q: what are advanced learner cognitive processes and problem-solving performances while solving a complex problem? F: Thresholds of knowledge development: mastering problem-solving operations within each threshold enhanced learner conceptual awareness of where to apply cognitive processes and increased the combinations of cognitive processes they activated at higher thresholds of knowledge development.
Liu, Kang, Lee, Winzeler, & Liu (2015)	Problem-solving process; Goal orientation	47 sixth graders	Log data: 3 weeks of 47 students' tool use frequency, duration; 38 students' Problem solution scores; 16 Learners' goal orientation: PALS	Data visualization: action shapes ANOVA	Q: How do learners access tool differently based on their goal orientation, and problem solution score? F: Learners in high performance and mastery-oriented groups tended to use the tools more appropriately relative to the stage they were in the problem-solving processes, and were more productive than students in low performance groups.

Table 4 continued.

Researchers	Informed Area	Participants	Data Sources	Data Analysis	Research Questions (Q) and Findings (F)
Kang, 2017	Problem-solving process	237 sixth graders	Log data: tool use frequency, duration sequence, probe design center; Problem solution scores.	Lag sequential analysis Sequential pattern mining Cluster analysis Data visualization	Q: Identify student navigation behavior patterns in cognitive processes between at-risk and non-at-risk students; Examine the relationship between student learning performance and scientific inquiry behaviors. F: The problem-solving processes were different between non-at-risk and at-risk students. The game metrics developed in AR Probe Design Center improved the predictions of student in-game and after-game performance.

Using data visualization of different group student computer log data, the researchers analyzed student behavior patterns (*N* = 38). The findings suggested students with high mastery-oriented scores (*N* = 9) tended to behave more appropriately in each problem-solving process stage, and were more productive than students in other groups. Students with high scores in performance approach (*N* = 3) and avoidance goal orientation (*N* = 3) showed an inappropriate behavior pattern, such as exploring more fun tools rather than gathering information to solve the problem.

These three studies indicated learner goal orientation affected their problem-solving in AR, but all three studies were conducted in real classroom setting; thus, it is unclear whether teachers had affected learner problem-solving during these studies. Therefore, this study was conducted in a laboratory setting by recruiting college students, and this method is supported by two previous studies conducted in laboratory settings (Bogard et al., 2013; Liu et al., 2009). In these two studies, rather than studying sixth graders in a classroom, researchers studied more advanced learners in laboratory setting—61 undergraduate students and 15 graduate students. In both studies, researchers asked the participants to find a home for one alien species (i.e., Akona in 2009, and Jakalay-Tay in 2013) within 90 minutes. By studying advanced learners in a laboratory setting, this study eliminated teacher influence during learner problem-solving processes, and collected more accurate computer log data on learner problem-solving behavior.

In addition, the goal orientation measurement used in previous studies has been further developed; therefore, this study used the most recent 3 X 2 goal orientation measurement (Elliot et al., 2011), which includes six types of goal orientations (i.e., task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation) to better understand the impact of learner goal orientation on problem-solving.

Problem-solving in Alien Rescue

Besides pervious AR studies on learner metacognition and goal orientation, this study is also based on a line of research about problem-solving in AR (Bogard et al., 2013; Liu & Bera, 2005; Liu et al., 2004; Liu et al., 2009; Liu et al., 2015). These studies provided the theoretical foundation for this study—successful problem-solver will go through all four conceptual stages while solving a problem in AR: (a) understanding the problem, (b) identifying, gathering, and organizing information, (c) integrating information, and (d) evaluating the process and outcome (Liu et al., 2004). In addition, previous studies indicated learners had different tool use patterns (based on tool use frequency, duration, and sequences) during their problem-solving processes (Kang, 2017; Liu et al., 2004; Liu et al., 2009). Particularly, in the early stage of problem-solving processes, students primarily used tools supporting cognitive processing and tools sharing cognitive load; in the later stages, students increasingly used multiple cognitive tools, especially tools supporting hypothesis generation and testing (Liu & Bera, 2005). This study continued to examine learner problem-solving processes using computer log data on tool use patterns in AR.

Data collection methods and analysis in Alien Rescue

The data collection and analysis in this study is also informed by previous AR research. Firstly, same as Liu et al. (2009) and Bogard et al. (2013), this study used stimulated recall to collect participant thoughts on the problem-solving processes. Stimulated recall interview is "a valuable tool for investigating cognitive processes in a naturalistic context" (Lyle, 2003, p. 861), because it involves an observer making notes while participant working through a problem, then asking probing questions using the observational notes as stimulus (Calderhead, 1981; Lyle, 2003). In addition, to reduce human error during the observation and note taking in stimulated recall, this study recorded learner gameplay processes using a screencast tool, then used these gameplay screencast recordings as stimulus to ask probe questions during interview.

Besides stimulated recall data, previous AR study also suggested using data visualization with log data is a promising technique to interpret the complex data set (Liu et al., 2015). Therefore, this study collected learner computer log data to capture learner actions in AR, which is similar to five previous AR studies (Kang, 2017; Liu et al., 2004, 2009, 2015; Liu & Bera, 2005). In addition, previous AR studies have analyzed various log data patterns, including tool use frequency (Liu et al., 2004; Liu & Bera, 2005), duration (Liu et al., 2009; Liu et al., 2015), and sequences (Kang, 2017). To fully understand the data, this study analyzed all these three log data patterns (i.e., frequency, duration and sequences). Furthermore, similar to Kang (2017) and Liu et al. (2015), this study utilized data visualization to interpret the analysis.

Summary

The previous AR studies have shown that learner metacognition affected tool use in AR, and goal orientation affected problem-solving in AR. These studies also built a foundation and pointed the direction for future research. Particularly, the connection between learner metacognition and problem-solving in AR is still not clear. In addition, the latest goal orientation measurement might provide more information on the impact of goal orientation on problem-solving. Furthermore, examining larger sample size of learner problem-solving in a laboratory setting to eliminate teacher influence are also needed to expand previous studies.

Therefore, this study draws insights from previous research to investigate the impact of learner characteristics (i.e., goal orientation and metacognition) on the problem-solving using computer log data, gameplay recordings, and stimulated recall interviews in a laboratory setting. In addition, this study expanded the data analysis by looking at all three log data patterns (i.e., frequency, duration and sequences) and used data visualization to further interpret the data.

RESEARCH PARTICIPANTS

The participants were 159 undergraduate students mainly from the Department of Educational Psychology participant pool at a large public university in the southwestern United States ($N = 116$). In addition, participants included other undergraduate students ($N = 43$) that were recruited from seven undergraduate courses that were offered by the College of Education, College of Natural Sciences, College of Liberal Arts, School of Nursing and College of Fine Arts. Students from the participant pool were also majored in various disciplines, but taking at least one of the following courses in Education Psychology department at the university, including EDP 304 (Strategic Learning for the 21st Century), EDP 306 (Human Sexuality & Relationships), EDP 350E (Introduction to Life Span Development), EDP 350G (Adolescent Development), EDP 350L (Human Sexuality), and EDP 371 (Introduction to Statistics). All participants had not played AR prior to participating in the study. At the time they were participating, they were all over 18 years of age, and voluntary to join in the study.

An a-priori analysis was conducted in G-Power to determine the necessary sample size N of this study (Mayr, Erdfelder, Buchner & Faul, 2007). The input parameter for conducting regression included one tail, Slope $H_1 = .15$, $\alpha = .20$, $1-\beta = .80$, Slope $H_2 = 0$, Std dev $\sigma\_x = 1$, Std dev $\sigma\_y = 1$. The suggested sample size is $N = 124$. Although there is no real rule for conducting cluster analysis, it has been suggested that 2^k can be used, preferably 5×2^k, where k = number of clustering variables (Dolnicar, 2002). This study used multiple types of cluster analysis, and the maximum k is the different goal orientation variables, which are six. Therefore, the suggested sample size is 64.

RESEARCH DESIGN AND PROCEDURE

This study used sequential mixed design, which indicates "the research questions and procedures for one strand depend on the previous strand" (Teddlie & Tashakkori, 2009, p. 143). In this study, qualitative data collection depended on quantitative data analysis. Firstly, in a laboratory setting, participants were given an online survey (online survey link: https://tinyurl.com/gamesurvey2018) on their demographic information, metacognition and goal orientation before playing AR. After the survey, the researcher played the opening video scenario, which provided the context and described the problem. The participants were asked to find the home planet for one alien named Jakala-Tay. Then each participant was given a username and password to login to the AR environment and had 60 minutes to work independently on the problem. At the end of the playthrough, the participants were asked to submit home solutions for Jakala-Tay. Their solution scores and activity logs in the SG environment were collected during the problem-solving processes in AR. Then stimulated recall interviews were conducted on 12 selected participants immediately after they played AR. The selection was based on participant metacognitive level (i.e., high and low) and goal orientation (i.e., task-approach, task-avoidance, other-approach, other-avoidance, self-approach, and self-avoidance). See the procedure illustrated in Figure 4.

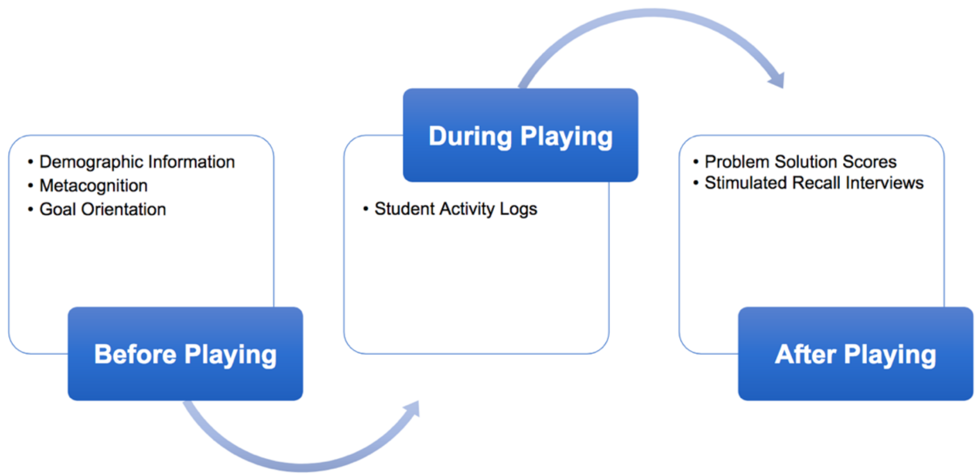


Figure 4. Sequential design procedure

DATA SOURCES

As a mixed-methods research, this study collected both quantitative and qualitative data. The quantitative data included student activity logs, problem-solving solution scores in the SG, demographic information, metacognition measurements, and goal orientation measurements. The qualitative data included gameplay screencasts and stimulated recall interviews.

Student activity logs

All student actions performed while using AR were logged to a data file. Each log file contained student ID, timestamp including start time (recorded to the precise minute), end time, tool name, gate access, tool use action—open or close, gate access action—go through, and problem solution texts. Because no direct teaching or guidance was provided for the participants, the computer log data indicates individual participant problem-solving processes in AR.

Problem solution scores

Learner problem solution scores were used to indicate problem-solving performances, which were evaluated by the quality of their solutions to the problem, i.e. the answer and rationale for sending the alien to a corresponding planet. The solution scores were determined by how well the learner solved the problem of finding an appropriate relocation home for the alien Jakalay-Tay, which were evaluated using an 8-point (0 to 7 points) grading rubric. This rubric has been used in multiple previous AR studies (Bogard et al., 2013; Liu et al., 2009; Liu et al., 2015). See Table 5 for the grading rubric.

In this study, a few learners have submitted multiple solutions to Jakala-Tay, or even other Aliens. In these cases, similar to Liu et al.'s (2015) study, the data were filtered to ensure that only the last score for the Jakalay-Tay was included, which "assumed the quality of solutions would increase as a student gained more experience in solving the problem" (p. 190). In addition, similar to Horton's (2014) approach, the solution scores were scored using the rubric by a panel of three trained raters including the author. Each rater graded half of the solutions and reached 100% agreement. Following the same rating standard, the author graded the other half of the solutions.

Table 5: Problem Solution Grading Rubric

Description	Score
The student recommends an unsuitable home for the alien species.	0
The student recommends a suitable home, but does not provide any reasons to substantiate their choice.	1
The student recommends a suitable home and is awarded one additional point for each reason provided to substantiate their choice.	2-7

Demographic information

Participant demographic information was collected, including gender, age, ethnicity, and college affiliation within the university. Specifically, the student demographics were as follows: 47.2% female students (*N* = 75), 52.8% male students (*N* = 84); 5% African American (*N* = 8), 26.4% Asian (*N* = 42), 18.9% Hispanic (*N* = 30), 39.6% White (*N* = 63), 9.4% Two or more races (*N* = 15), and 0.6% student (*N* = 1) chose to not answer. Most of these students were at age 20 (*N* = 33, 20.8%), followed by age 19 (*N* = 30, 18.9%), 21 (*N* = 29, 18.2%), 22 (*N* = 25, 15.7%), 23 (*N* = 16, 10.1%), older than 23 (*N* = 14, 8.8%), and 18 (*N* = 12, 7.5%).

With regarding the year at college, most of them were at the senior year (*N* = 56, 35.2%), followed by freshman (*N* = 45, 28.3%), junior (*N* = 28, 17.6%), sophomore (*N* = 27, 17%). There were also 1.9% students (*N* = 3) at their fifth year at the university. Students were from ten different colleges in the university: 25.2% of these students (*N* = 40) were from College of Natural Sciences, 20.8% (N = 33) were from College of Liberal Arts, and 17.6% (*N* = 28) were from McCombs School of Business (see Table 6).

Table 6: Participants Demographic Information

Demographic Information	Count (*N*)	Percentage (%)
Gender		
Female	75	47.2
Male	84	52.8
Ethnicity		
African American 18.9%	8	5%
Asian	42	26.4%
Hispanic	30	18.9%

Table 6 continued.

Demographic Information	Count (*N*)	Percentage (%)
White	63	39.6%
Two or more races	15	9.4%
I don't want to answer	1	0.6%
Age		
18	12	7.5%
19	30	18.9%
20	33	20.8%
21	29	18.2%
22	25	15.7%
23	16	10.1%
Older than 23 years old	14	8.8%
Year at the College		
Freshmen	45	28.3%
Sophomore	27	17.0%
Junior	28	17.6%
Senior	56	35.2%
Fifth year	3	1.9%
College Affiliation		
Cockrell School of Engineering	7	4.4%
College of Education	21	13.2%
College of Fine Arts	7	4.4%
College of Liberal Arts	33	20.8%

Table 6 continued.

Demographic Information	Count (*N*)	Percentage (%)
College of Natural Sciences	40	25.2%
McCombs School of Business	28	17.6%
Moody College of Communication	16	10.1%
School of Social Work	1	0.6%
School of Undergraduate Studies	4	2.5%
Other	2	1.3%

Metacognition measurement

Learner metacognition was measured using the Metacognitive Awareness Inventory (MAI) (Schraw & Dennison, 1994). This is a 52-item self-reported survey to measure adult metacognition, which consists of metacognitive knowledge and regulation. There are 17 metacognitive knowledge items; a sample item is, "I understand my intellectual strengths and weaknesses." There are 35 metacognitive regulation items; a sample item is, "I ask myself periodically if I am meeting my goals." The measurement is scored on a 100-point, bipolar scale, with 0 being "totally untrue of me" and 100 being "totally true of me." The scale demonstrates high reliability (a = .90) and significant correlations between these two components in previous studies (r = .54 and r = .45 respectively) (Schraw & Dennison, 1994). The measurement also had a high reliability (a = .95) using the sample data from this study. The complete measurement can be found in Appendix A.

Goal orientation measurement

Learner goal orientation was measured using the 3 X 2 Achievement Goal Orientation Inventory (Elliot et a., 2011). This has six subscales with high reliability, which are task-approach (*a* = .84), task-avoidance (*a* = .80), self-approach (*a* = .77), self-avoidance (*a* = .83), other-approach (*a* = .93), and other-avoidance goals (*a* = .91). Using data from this study, the reliability numbers were task-approach (*a* = .85), task-avoidance (*a* = .90), self-approach (*a* = .81), self-avoidance (*a* = .81), other-approach (*a* = .92), and other-avoidance goals (*a* = .92). For each subscale, there are 3 items for a total of 18 items. These items were randomized in the online survey based on Elliot's suggestion. Participants can rate these statements on a 7-point scale (1 = not true of me, 7 = extremely true of me). See Table 7 for the example items. The complete measurement can be found in Appendix B.

Table 7. 3 X 2 Goal Orientation Example Items

Goal Orientation	Example
Task-approach (TAP)	To get a lot of questions right on the exams in this class.
Task-avoidance (TAV)	To avoid incorrect answers on the exams in this class.
Self-approach (SAP)	To perform better on the exams in this class than I have done in the past on these types of exams.
Self-avoidance (SAV)	To avoid doing worse on the exams in this class than I normally do on these types of exams.
Other-approach (OAP)	To outperform other students on the exams in this class.
Other-avoidance (OAV)	To avoid doing worse than other students on the exams in this class.

Gameplay recording and stimulated recall responses

All participant gameplay screens were recorded using software called *QuickTime* player while they were playing AR. The screencast videos served as a recall tool for stimulated recall interviews. In other words, the researcher conducted the interviews while watching the gameplay screencast videos with the participants. Interview questions focused on participant problem-solving processes and metacognition, which were guided by participant actions in AR. The example questions were:

1. Can you tell me about your overall impression about this game?

2. Can you describe the whole process how you find the home for Jakala-Tay?

3. Has it ever occur to you that there might be other planets they can go to rather than the one you chose?

4. Have you look at the screen to check on the timer? Or Have you checked the time?

5. Did you feel pressure from your classmates while you are solving the problem?

6. If you are going to help the second alien, what the procedure would be like?

7. How did you go through all the 22 planets to pick a solution?

8. What tool is the most helpful one for you in this environment?

9. How did you know you need to send the probe, and check mission control room?

10. The video showed you clicked the alien database. What were you thinking that made you want to click here?

11. The video showed you stayed in the probe design room for 5 minutes. What do you want to find out there?

All the responses were audio recorded and transcribed for further analysis. Interviews were conducted for twelve selected participants based on their approximate metacognitive level (i.e., high and low) and goal orientation (i.e., task-approach, task-avoidance, other-approach, other-avoidance, self-approach, and self-avoidance). Therefore, there were six participants with higher metacognitive levels and six participants with lower. In addition, each of the six participants had one dominate goal orientation that is different from other participants within the same metacognitive level group.

RELIABILITY AND VALIDITY

This study ensured the reliability and validity of the quantitative data. As for achieving the trustworthiness of the qualitative data, this study used data triangulation, member checking, peer debriefing and peer coding techniques.

Quantitative

To ensure the validity of quantitative data, the researcher considered both the internal and external validity (Onwuegbuzie & Johnson, 2006). For internal validity, during research design and data colletion, this study was conducted in the laboratory envrioment, which could reduce participants' behavior bias. In addition, survey measurements was used to reduce the research observational bias. Furthermore, during the data analysis, the study checked data multicollinearity. During the data interpretation, this study also reported the effect size to reduce the threats to internal validity. As for external validity, this study conducted power analysis to determine an appropriate sample size to ensure the polupaliton validity. The researcher was aware of the self-reported survey and order bias that might affect the external validty of the stuty, which will be discussed in the study limitation section (Onwuegbuzie & Johnson, 2006). Finally, the goal orientation and metacognition measurements have been developed by experts in the field; therefore, these measurements all have high content validity and construct validity.

To address reliability, the game log files were consistent over time and samples. Specifically, for each participant, computer log file recorded his/her every mouse clicks and timestamps in a cloud-based database consistently. The solution scores were graded by three trained graders in the AR team to ensure reliability. The goal orientation and metacognition measures were all reliable according to their reliability alpha number.

Qualitative

To ensure the trustworthiness of the study, this study employed data triangulation, member checking, peer debriefing and peer coding techniques.

Data triangulation

This study combined different data sources to build a coherent justification of themes (Creswell, 2014). Specifically, this study used interview transcripts, student activity logs, and survey results during analysis to ensure data triangulation.

Member checking

Member checking was conducted with participants, because this is “the most important strategy for determining the credibility of the researcher’s interpretation of the participants’ perceptions” (Teddlie & Tashakkori, 2009, p. 213). Specifically, the researcher asked for participant feedback and corrections regarding interview transcripts. The researcher also asked all participants, via email, for feedback on the data, themes, and major findings discovered through the interviews. Three of these participants provided feedback on the interviews. This process had provided an opportunity for the participants to verify the information they shared was accurately transcribed and interpreted.

Peer debriefing and coding

To "clarify interpretations and identify possible sources of bias" (Teddlie & Tashakkori, 2009, p. 295), the study used peer debriefing to ask colleagues to comment on findings as they emerge. The researcher also shared the initial coding with two researchers in the AR research team to ask for feedback. These colleagues acted as second coders to confirm or question the codes, interpretations and themes of the study.

DATA ANALYSIS

Question 1: problem-solving performance differences based on learner characteristics

There are three sub-questions for question 1 including: (a) is there a statistically significant difference in learner problem-solving performances based upon metacognition (high metacognitive level, low metacognitive level) and goal orientation (task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation); (b) can learner metacognition and goal orientation predict problem-solving performances? and (c) what are the reasons for any problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation);

For question (a), cluster analyses in SPSS were used to show if there is a statistically significant difference among learner groups based on their metacognition and goal orientation, because it has been used in multiple studies to generate groups based student goal orientation (Daniels et al., 2008; Jiang & Liu, 2012; Meece & Holt, 1993). It has also been used in several AR studies (Bogard et al., 2013; Kang, 2017; Liu & Bera, 2005).

STATA—a statistical software package used mostly in the fields of economics, sociology, political science, biomedicine and epidemiology—indicated that "Although some have said that there are as many cluster-analysis methods as there are people performing cluster analysis. This is a gross understatement! There exist infinitely more ways to perform a cluster analysis than people who perform them" (STATA, 2018, p. 3). Several prominent researchers suggested that k-means cluster analysis is the most popular technique when exploring participant homogeneous groups (Dolnicar, 2003; Jain, 2010). Dolnicar (2003) reviewed 243 studies that used cluster analysis in business setting at the time, and 76% of them used k-means cluster analysis. The k-means cluster analysis is typically used with the Euclidean metric for calculating the distance between points and cluster centers (Jain, 2010). In SPSS, researcher needs to identify a *k*, which indicates the hypothesized group number. Therefore, based on literature and the collected data, this study used multiple *k* numbers (i.e., $k = 2 \ldots 12$) in k-means cluster analysis to explore student groups based on their characteristics.

For question (b), a multiple regression was used to see if learner metacognition (MC) and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, OAV) could predict problem-solving performance differences (i.e., Solution Scores, SS). Specifically, the regression model is as follows:

$$Y_{ss}=\beta_1 MC+\beta_2 TAP+\beta_3 TAV+ \beta_4 SAP+ \beta_5 SAV+ \beta_6 OAP+ \beta_7 OAV + u$$

As for questions (c), based on Glesne's (2015) suggestion, 12 stimulated recall interview transcripts were coded using "line-by-line coding" (p. 195), which requires the researcher to generate codes by going through transcripts line-by-line and extract the code words. It helps the researcher to "immerse [themselves] in the data and discover what concepts they have to offer" (Glesne, 2015, p. 195). A codebook was developed based on the emerging coding scheme, metacognition and goal orientation definition in the literature

(Brown, 1987; Flavell, 1979, 1987; Schraw & Dennison, 1994; Sperling et al., 2004; Zimmerman, 2002, 2013). The grounded theory (Glesne, 2015) were used to analyze codes, and explain the reasons for any possible problem-solving differences among learners. Three categories of codes were emerged based on the interviews, including problem-solving process, metacognition, and goal orientation. See Appendix C for the codebook.

Question 2: problem-solving process differences based on learner characteristics

There are five sub-questions for question 2: (a) what are learner problem-solving process patterns; (b) are there any problem-solving process pattern differences among learners based on their metacognition; (c) are there any problem-solving process pattern differences among learners based on their goal orientation; (d) are there any problem-solving process pattern differences based on the interaction between learner metacognition and goal orientation; and (e) what are the reasons for problem-solving process pattern differences based on learner characteristics (i.e., metacognition and goal orientation)? The following paragraphs will describe data analysis for Question 2 using collected data.

Visualizing learner problem-solving processes

For question (a), (b), (c), (d), *Tableau* and *R* were used to visualize learner problem-solving processes, based on their activity log, to identify if there were any existing patterns. In AR, learner activity log data can be grouped into two types of actions during problem-solving in AR including room visit action (sequences) and tool use action (frequency, duration). Chord Diagrams (Flajolet & Noy, 2000) and the *R circlize* package (Gu, Gu, Eils, Schlesner, & Brors, 2014) were used to visualize learner tool use frequency and duration action in AR, since these diagrams provide a compact way of representing information (Wei et al., 2016).

Furthermore, besides "eyeballing" at the visualization to decide the action (i.e., frequency and duration) differences during problem-solving processes based on learner characteristics, similarity measurements (Loh et al., 2016; van der Loo, 2014) were used to analyze learner action sequences during problem-solving based on learner characteristics, which can indicate the problem-solving process differences among learners who have different characteristics. In addition, two-proportion z-test was used to answer question (b), (c), and (d)—whether the differences between two groups of learners who have different characteristics are statistically significantly different. The following paragraphs will describe similarity measure and how to use it with two-proportion z-test to identify learner problem-solving process differences in AR.

Similarity Measure

A similarity measure is a statistical function to quantify the (dis)similarity of two objects, such as text strings, documents, audio files, digital photographic images, DNA sequences, and other digitized objects for pattern recognition (Dengfeng & Chuntian, 2002; Loh & Sheng, 2014; Loh et al., 2016; Van der Loo, 2014). Mathematically, the value of (dis)similarity is between 0 and 1, which indicates completely different to identical, respectfully. In addition, the dissimilarity of two object is defined as distance. Thus, the relationship between similarity and dissimilarity of two objects X and Y is:

$$\text{Similarity (X, Y)} = 1 - \text{Distance (X, Y)} \quad (1)$$

Based on Loh et al.'s (2016) finding that combined similarity measures bolster understanding of learner action, this study used three different similarity measures including *Cosine (Cos), Jaccard (Jac)*, and *Longest Common Substring (LCS)* coefficients to analyze learner action sequential strings in AR. *Cosine* and *Jaccard* are *q-grams* based on distances, which slice the string by number *q*, then count the number of *q-grams* that are not shared between two strings (Van der Loo, 2014). It is suggested that bi-gram (q = 2) is sufficient for slicing medium corpora (thousands of words) (Loh et al., 2016).

The *Cosine* distance equals 0 when two strings are identical and 1 when two strings have no *q-gram* in common. The definition of *Cosine* distance is

$$\mathrm{d_{cos}(s, t{:}\, q) = 1 - \frac{v(s;q) \cdot v(t;q)}{||v(s;q)||_2 ||v(t;q)||_2}} \tag{2}$$

Here, s and t are the two strings, $v(t;q)$ is a nonnegative integer vector whose coefficients represent the number of occurrences of every possible q-gram in string t, and $|| \cdot ||_2$ indicates the standard Euclidean norm, which indicates the magnitude of the vector. For example, we could assign stringA = "Alien Rescue", stringB = "Alen Resc", and stringC = "Book Club". Since these three strings are short, instead of using 2 as the q to slice the string, we use q = 1. We can "eyeball" the distance between stringA and stringB (disAB) as being smaller compared to the distance between stringA and stringC (disAC). We can also calculate disAB and disAC by using the stringdist package in R (Van der Loo, 2014) as follows:

```
> stringdist (stringA, stringB, method='cosine', q=1)
[1] 0.07613023
> stringdist (stringA, stringC, method='cosine', q=1)
[1] 0.7867993
```

Therefore, the calculated *Cosine* distance indicates disAB (.076) is much closer compared to disAC (.786), which means compared to stringC, stringB is more like stringA.

Similar to *Cosine* distance, the *Jaccard* distance varies from 0 to 1, where 0 corresponds to two strings full overlap and 1 to no overlap. The *Jaccard* distance is defined as

$$d_{jaccard}(s, t{:}\, q) = 1 - \frac{|Q(s;q) \cap Q(t;q)|}{|Q(s;q) \cup Q(t;q)|} \tag{3}$$

where *s*, *t* are the two strings, *q* is the number that the two strings are sliced by, and *Q(t:q)* indicates the unique set of *q*-grams occurring in string *t*. Take stringA (“Alien Rescue”), stringB (“Alen Resc”), and stringC (“Book Club”) for example; the distance between stringA and stringB (disAB) can be computed as follow:

> stringdist("Alien Rescue", "Alen Resc", method='jaccard', q =1)

[1] 0.2

The distance between stringA and stringC (disAC) can be computed as follow:

> stringdist("Alien Rescue", "Book Club", method='jaccard', q =1)

[1] 0.8

Therefore, the *Jaccard* distance indicates stringA and stringB have more overlap (0.2) compared to stringA and string C (0.8), which means compared to stringC, stringB is more like stringA.

Different from *q-grams* based distances such as *Cosine* distance and *Jaccard* distance, *LCS* is an edit-based distance, which counts the number of deletions and insertions necessary to transform one string into another (Loh et al., 2016; Van der Loo, 2014). It is recursively defined as

$$d_{lcs}(s,t)\begin{cases} 0\ if\ s = t = \varepsilon, \\ d_{lcs}\left(s_{1:|s|-1}, t_{1:|s|-1}\right)\ if\ s_{|s|}\ =\ t_{|t|} \\ 1 + \min\{d_{lcs}\left(s_{1:|s|-1}, t\right), d_{lcs}\left(s, t_{1:|t|-1}\right)\}\ otherwise. \end{cases} \tag{4}$$

Take stringA (“Alien Rescue”), stringB (“Alen Resc”), and stringC (“Book Club”) for example, the *LCS* distance between stringA and stringB (disAB) can be computed as follow:

> stringdist("Alien Rescue", "Alen Resc", method='lcs')

[1] 3

The distance between stringA and stringC (disAC) can be computed as follow:

```
> stringdist("Alien Rescue", "Book Club", method='lcs')
[1] 17
```

This indicates that to transform stringB to stringA, at least 3 edits are required, while it takes 17 edits to transform stringC to stringA. In addition, to calculate the *LCS* distance coefficient, the formula 5 will be used (Loh et al., 2016), where d_{LCS} is calculated using formula 4, and d_{max} is the maximum number of insertions or deletions for the string transformation to occur.

$$\frac{d_{LCS}(X \cdot Y)}{d_{max}(X \cdot Y)} \qquad (5)$$

disAB coefficient is .142 and disAC coefficient is .809, which indicates string A and string B are similar compared to string C. Similar to *Cosine* and *Jaccard* distance, the *LCS* coefficient also varies from 0 to 1, where 0 shows two strings are identical and 1 shows no substrings in common.

Identifying Problem-solving process differences using Similarity Measure

Previous SGA studies indicated that learner action sequence is important for understanding learning processes in SG (Kang, 2017; Loh et al., 2016). Therefore, this study analyzed learner log data sequences during the problem-solving process. Specifically, to identify whether there were existing problem-solving process differences among learners based on their metacognition and goal orientation, this study used similarity measure to analyze both room visit sequences and tool use sequences during problem-solving.

To conduct similarity measure, firstly, learner room visit sequences in AR during the problem-solving processes were converted into strings to facilitate similarity measure analysis. Then the data analysis tool *R* and stringdist package were used to calculate the similarity of learner action sequential strings (van der Loo, 2014). According to Loh et al.

’s (2016) suggestion, three different similarity measures were conducted together to make sense the behavior differences including *Cosine (Cos), Jaccard (Jac)*, and *Longest Common Substring (LCS)* coefficients.

For example, a learner named Matt started from the Main Room (Room M), then went to the Alien Information Room (Room A), followed by Problem Design Room (Room P), and Mission Control Room (Room C) during the first 5 minutes. He might go back to Room A again. Since in the current environment, the learner must go back to the Main Room to go to any other room, the location sequence for Matt is MAMPMCMA. While another learner, Emily, might have a different sequence such as MPMAMCM, because she started from Room M, then visited Room P first, then went to Room A and C. Therefore, stringMatt = MAMPMCMA; stringEmily = MPMAMC. In addition, this study adopted bigram ($q = 2$) to slice the string, because it has proved to be sufficient for medium corpora (thousands of words) (Loh et al., 2016). Using stringMatt [MAMPMCMA] and stringEmily [MPMAMC] as examples, the sliced strings are:

stringMatt (bigram): [MA, AM, MP, PM, MC, CM, MA];

stringEmily (bigram): [MP, PM, MA, AM, MC];

Let X = stringMatt, and Y = stringEmily; q-grams in this analysis are: [MA, AM, MP, PM, MC, CM], and $v(X;2)$ is (2, 1, 1, 1, 1, 1), $v(Y;2)$ is (1, 1, 1, 1, 1, 0). According to the *Jaccard*, *cosine*, and *LCS* distance, Matt and Emily have 83.3%, 89.4%, and 57.1% similarity (i.e., 16.7%, 10.6%, and 42.9% differences) in their room visit sequences during the problem-solving in AR. See Table 8 for respective formulas and coefficients.

Table 8: Coefficient Formula for Different Similarity Measures

Similarity Coefficient	Formula	Calculation	Coefficients
Jaccard	$\frac{\lvert Q(X;q) \cap Q(Y;q) \rvert}{\lvert Q(X;q) \cup Q(Y;q) \rvert}$	$5/6$	0.833
Cosine	$\frac{v(X;q) \cdot v(Y;q)}{\lvert\lvert v(X;q) \rvert\rvert_2 \lvert\lvert v(Y;q) \rvert\rvert_2}$	$6/(\sqrt{9} \cdot \sqrt{5})$	0.894
Longest Common Substring	$1 - \frac{d_{LCS}(X \cdot Y)}{d_{max}(X \cdot Y)}$	1-6/14	0.571

Use the above method, this study calculated the similarity coefficients in learner action sequences among all students based on their clustered groups. This study compared all learners to learner who had the highest solution score (i.e., 7 points). To describe how to use similarity measure to identify problem-solving process differences, hypothesized room visit sequences of 12 selected participants are used. See Table 9 for the similarity measure result using these 12 examples.

Table 9: Hypothesized Room Visit Sequences and Similarity Measure Coefficients

Compared Students Group	Hypothesized Room Visit Sequences	Similarity Measure Coefficients		
		Jaccard	Cosine	LCS
HMC + TAP vs HMC + TAP	MAMPMA	0	0	0
HMC + TAP vs HMC + TAV	MCMPMC	0.667	0.714	0.4
HMC + TAP vs HMC + SAP	MCMP	0.833	0.782	0.5
HMC + TAP vs HMC + SAV	MPMAMC	0.2	0.155	0.4
HMC + TAP vs HMC + OAP	MPMAMAM	0	0.044	0.417
HMC + TAP vs HMC + OAV	MPMPMPMP	0.5	0.471	0.429
HMC + TAP vs LMC + TAP	MAMCMC	0.667	0.571	0.4
HMC + TAP vs LMC + TAV	MAMCMPMPMA	0.333	0.122	0.286
HMC + TAP vs LMC + SAP	MCMAMCMCMP	0.5	0.67	0.444
HMC + TAP vs LMC + SAV	MAMCMCMC	0.667	0.707	0.429
HMC + TAP vs LMC + OAP	MPMAMCMMPMAMCM	0.429	0.244	0.364
HMC + TAP vs LMC + OAV	MCMPMAMCMPMA	0.333	0.258	0.333

Based on the similarity coefficient generated in Table 10, the similarity of 12 selected participant room visit sequences during their problem-solving processes were visualized (see Figure 5). According to Figure 5, the three different similarity measures showed consistent similarity among all three similarity measures based on 12 learner room visit sequences, which is consistent with the previous Loh et al. (2016) study that combined similarity measure could help with better understanding sequence similarity. Therefore, using similarity measure and visualization techniques to examine both room visit sequences and tool use sequences, this study hopes to identify the differences among different student groups based on goal orientation and metacognition.

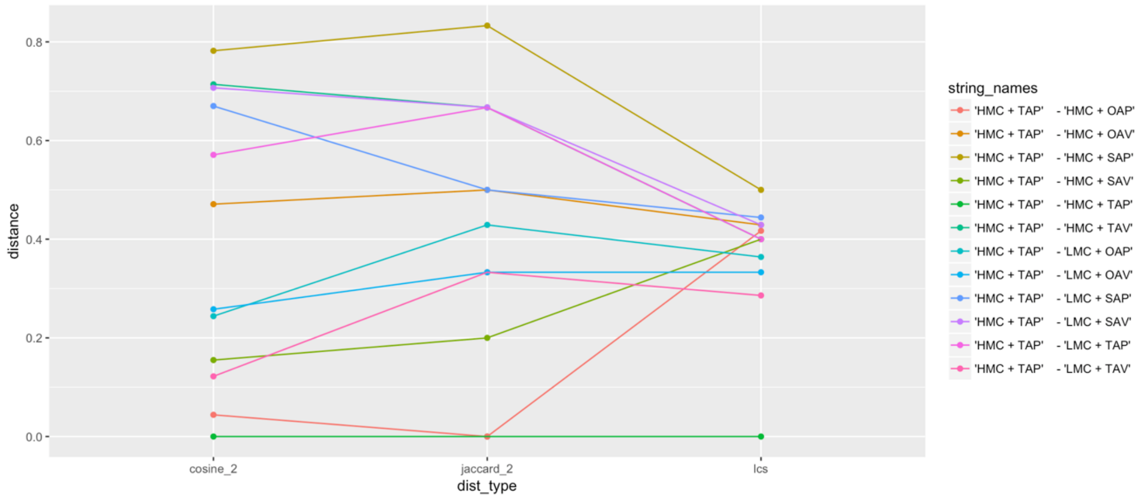


Figure 5. Hypothesized similarity based on learner characteristics

Are the differences significant?

Based on Figure 5 and Figure 6, we could "eyeball" the differences among different learner groups, but it is not clear whether these differences are statistically significant or not. Therefore, to answer research question 2(b), 2(c), and 2(d), two-proportion z-test were used be conducted to decide whether any of the two group of student similarity measures are statistically significant during the problem-solving process.

At last, to answer research question 2(e), the same as answering question 1(b), participant stimulated recall interview transcripts were coded using "line-by-line coding" (Glesne, 2015, p. 195). Grounded theory (Glesne, 2015) was used to analyze the codes, and explain the reasons for any possible problem-solving process differences among learners. See Appendix C for the codebook.

Chapter 4: Results

This chapter will present the results for question one (i.e., problem-solving performance differences based on learner characteristics) and question two (i.e., problem-solving process differences based on learner characteristics) based on both quantitative and qualitative data.

QUESTION 1: PROBLEM-SOLVING PERFORMANCE DIFFERENCES BASED ON LEARNER CHARACTERISTICS

There are three sub-questions for question one including: (a) is there a statistically significant difference in learner problem-solving performances based upon metacognition (high, medium, and low metacognitive levels) and goal orientation (task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation); (b) can learner metacognition and goal orientation predict problem-solving performances; and (c) what are the reasons for any problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)? Results for the three sub-questions are discussed in the following sections.

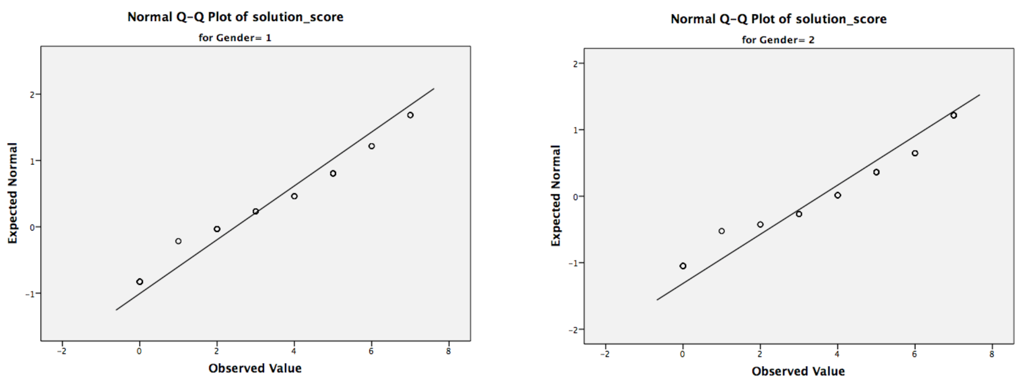
1.a. learner performance differences based on learner characteristics

To examine if there was a statistically significant difference among learners based on their metacognitive levels and goal orientations, k-means cluster analyses were used. Specifically, k-means cluster analyses were used to cluster each of the three variables, including learner metacognitive levels (k = 3), goal orientations (k = 6), problem performances (k = 3). In addition, all 159 participants were clustered based on all three variables together (k = 3). The following section will describe learner problem-solving performance differences based on demographics and all four cluster analyses.

Demographics

Using one-way ANOVA in SPSS, this study analyzed learner problem-solving performances based on age, year at the university, college affiliation and ethnic groups. The Independent Sample *t* test was conducted to analyze learner gender differences.

According to SPSS, the average problem-solving performance (i.e., solution score) was 3.04 points on an 8-points scale. An Independent Sample *t* test was conducted to compare problem-solving performance based on gender difference, which met assumptions of normality (see Figure 6) and homogeneity, Levene's Test: $F(1, 157) = 1.056$, $p = .306$. The test yielded a significant result ($t = -2.592$, $p < .01$). Specifically, female participants had significantly lower problem-solving performance scores ($M = 2.48$, $SD = 2.462$, $N = 75$) compared to male participants ($M = 3.55$, $SD = 2.704$, $N = 84$).



Note: Gender = 1 indicates female, while Gender = 2 indicates male.

Figure 6. Normal Q-Q Plot of solution score based on gender

After meeting the assumptions for one-way ANOVA, including homogeneity of variances and normality, there were statistically significant differences in learner problem-solving performance based on participant ethnic groups, $F(4, 153) = 3.774$, $p = .006$, $\eta^2 = .097$; Levene's Test: $F(4, 153) = .944$, $p = .440$. Specifically, Asian students had the highest solution score—3.74 points, followed by White students—3.49 points, African American students—2.75 points, Multi-Ethnic students—2.33 points and Hispanic students—1.67 points. The post hoc analysis using Tukey's HSD test indicated that Hispanic participants scored significantly lower than Asian and White participants, $p < 0.05$. However, there was no significant difference among African America, Asian, and White participants.

There were no statistically significant differences in learner problem-solving performances based on age, college affiliation, and subject area (i.e., natural sciences or social science). There was also no significant differences in learner problem-solving performance based on year at the university, but learner solution scores did show an increasing trend based on years of university study (see Figure 7).

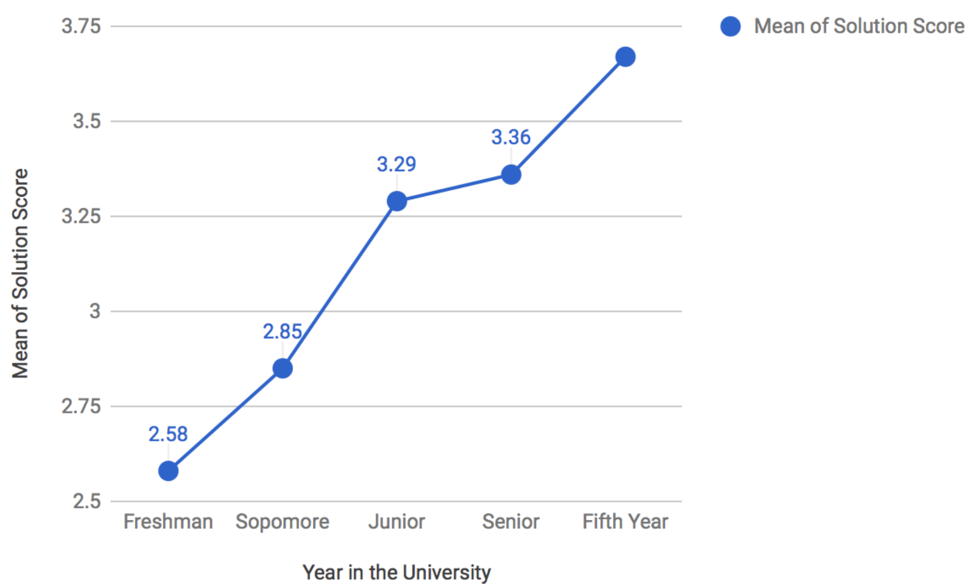


Figure 7. Learner solution scores based on their year at the university

Learner Metacognitive Levels

Learner metacognitive levels were clustered into 3 groups using k-means cluster analysis, including high (*Mean* = 85.17, *N* = 24), medium (*Mean* = 67.46, *N* = 79), and low metacognitive levels (*Mean* = 51.77, *N* = 56) (see Figure 8).

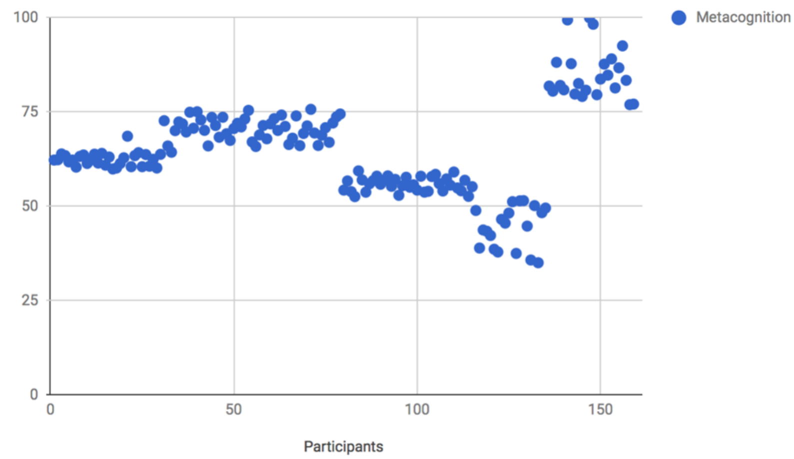


Figure 8. Learner metacognitive levels using cluster analysis

One-way ANOVA showed there were statistically significant differences in learner problem-solving performances based on these three groups, $F(2, 156) = 4.848$, $p = .009$, $\eta^2 = .058$; Levene's Test: $F\ (2, 156) = .458$, $p = .633$. Interestingly, learners in the lowest metacognitive level group had the highest solution score (*Mscore* = 3.86), followed by learners in high metacognitive level group (*Mscore* = 3.08), and medium metacognitive group (*Mscore* = 2.46). The post hoc analysis using Tukey's HSD test indicated that participants in the medium metacognitive group scored significantly lower than those in the low metacognitive group, $p < 0.05$. However, there was no significant difference between the high and low metacognitive level groups, or high and medium metacognitive groups.

In addition, since there were statistically significant differences in learner problem-solving performance based on participant ethnic groups and gender, One-way ANCOVA analyses were further conducted to determine a statistically significant difference among high, medium, and low metacognitive level learners on problem-solving performances controlling for ethnic groups and gender. The SPSS results indicated that there were significant effects of learner metacognitive levels on problem-solving performances after controlling for ethnic groups, $F(2, 156) = 4.999$, $p = .008$, $\eta^2 = .061$; Levene's Test: F (2, 156) = .801, $p = .451$ and gender $F(2, 156) = 4.677$, $p = .011$, $\eta^2 = .057$; Levene's Test: F (2, 156) = .281, $p = .756$.

Goal Orientations

After identifying one outlier using SPSS, other learners were clustered into five groups based on their goal orientation results, including 1) Cluster 1: medium in all six goal orientations ($N = 41$); 2) Cluster 2: low in all six goal orientations ($N = 5$); 3) Cluster 3: high in all six goal orientations ($N = 85$); 4) Cluster 4: high in TAP, TAV, SAP, SAV, but low in OAP and OAV ($N = 23$); and 5) Cluster 5: high in TAP and OAP, Medium in SAP, but low in TAV, SAV and OAV ($N = 4$). See Table 10 and Figure 9 for the cluster results. It is interesting that students in Cluster 2 (i.e. low in all six goal orientations) had the highest problem-solving performances (*Mscore* = 4.60), while students in Cluster 5 (i.e., high in TAP and OAP, Medium in SAP, low in TAV, SAV and OAV) had the lowest performances (*Mscore* = 2.25). However, these two clusters only had nine learners, and it is worth noting that using ANOVA in SPSS, there was no significant difference on problem-solving performances based on goal orientation groups among learners, $F(4, 153) = 1.520$, $p = .199$, $\eta^2 = .038$; Levene's Test: F (4, 153) = .884, $p = .475$.

Table 10: Cluster Centers for Learner Goal Orientation Groups.

Goal Orientations	Clusters				
	1 ($N = 41$)	2 ($N = 5$)	3 ($N = 85$)	4 ($N = 23$)	5 ($N = 4$)
TAP	16	11	19	18	19
TAV	16	10	19	19	10
SAP	15	9	19	17	14
SAV	15	9	19	17	11
OAP	14	10	18	9	18
OAV	15	10	19	12	9
Solution Score	3.71	4.60	2.72	3	2.25

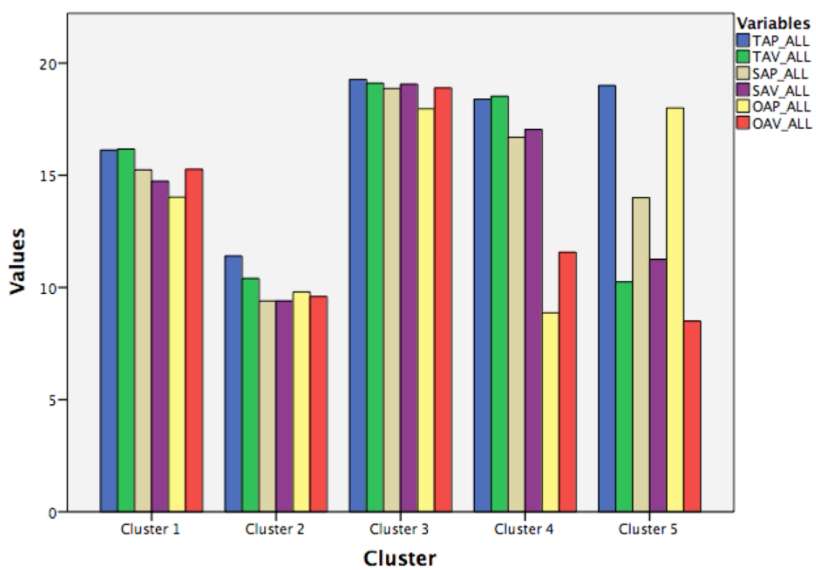


Figure 9. Learner goal orientation groups using cluster analysis

Final Cluster

To prepare for the final cluster, learner performances were also clustered into 3 groups, including 1) high performance ($6 <= Mscore <= 7$, $N = 35$), 2) medium performance (2 <= *Mscore* <= 5, N = 66), and low performance (0 <= *Mscore* <= 1, N = 58) (see Figure 10).

Figure 10. Learner solution scores using cluster analysis

Using k-means cluster analysis, learner problem-solving performance, goal orientation, and metacognition were grouped into three clusters, including 1) Cluster 1: high metacognition and high multiple goal orientations, 2) Cluster 2: low metacognition and medium multiple goal orientations, and 3) Cluster 3: medium metacognition and low multiple goal orientations. According to one-way ANOVA, learner problem-solving performances were statistically significant based on these three clusters, $F(2, 155) = 11.208$, $p = .000$, $\eta^2 = .126$; Levene's Test: $F(2, 155) = .989$, $p = .374$. Specifically, learners in Cluster 2 (i.e. low metacognition and medium multiple goal orientations) and Cluster 3 (i.e., medium metacognition and low goal orientation group) had nearly 2-point higher scores—in an 8-point scale system—compared to learners in Cluster 1. The post hoc analysis using Tukey's HSD test indicated that participants in Cluster 1 scored significantly lower than those in Cluster 2 and 3, $p < 0.05$. The cluster analysis results are presented in Table 11.

Table 11. Final Cluster Analysis Results

Mean	Clusters		
	1 ($N = 61$)	2 ($N = 51$)	3 ($N = 46$)
Solution Scores	1.89 (Low)	3.80 (High)	3.80 (High)
Metacognition	74.28 (High)	56.64 (Low)	60.59 (Medium)
TAP	19.38 (High)	18.73 (Medium)	15.61 (Low)
TAV	18.80 (High)	18.51 (Medium)	15.54 (Low)
SAP	18.77 (High)	17.63 (Medium)	14.61 (Low)
SAV	18.79 (High)	17.86 (Medium)	14.15 (Low)
OAP	17 (High)	15.02 (Medium)	13.57 (Low)
OAV	17.56 (High)	16.37 (Medium)	14.65 (Low)

In addition, since there were statistically significant differences in learner problem-solving performance based on participant ethnic groups and gender, One-way ANCOVA analyses were further conducted to determine statistically significant differences among Cluster 1, Cluster 2, and Cluster 3 learners on problem-solving performances controlling for ethnic groups and gender. The SPSS results indicated that there were significant effects of learner final clusters on problem-solving performances after controlling for ethnic groups, $F(2, 155) = 9.726$, $p = .000$, $\eta^2 = .112$; Levene's Test: $F\ (2, 155) = .711$, $p = .493$ and gender $F(2, 155) = 11.148$, $p = .000$, $\eta^2 = .126$; Levene's Test: $F\ (2, 155) = 1.092$, $p = .338$.

1.b. Can learner characteristics predict problem-solving performance differences?

Multiple regression was conducted in SPSS to identify significant predictors of learner problem-solving performance differences (i.e., Solution Scores, SS). The predictors included learner metacognition (MC) and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, OAV). Specifically, the regression model is as follows:

$$Y_{ss}=\beta_1 MC+\beta_2 TAP+\beta_3 TAV+\beta_4 SAP+\beta_5 SAV+\beta_6 OAP+\beta_7 OAV + u$$

Assumptions of linearity, reliability of measurement, homoscedasticity, multicollinearity and normality for multiple regression were tested using the data (Osborne & Waters, 2002). According to Cook's and leverage values that were generated using SPSS (Hutcheson & Sofroniou, 1999), this study identified three outliers. After eliminating these three outliers, scatterplots of the residuals showed linear relationships with residuals between the independent and dependent variables (see Figure 11). For reliability, as stated in the previous chapter, all the measurements are reliable based on the literature. The Cronbach's Alpha for this study was also tested in SPSS, and had high reliabilities. Specifically, the reliability numbers were task-approach (a = .85), task-avoidance (a = .90), self-approach (a = .81), self-avoidance (a = .81), other-approach (a = .92), other-avoidance goals (a = .92), and metacognition (a = .95).

In addition, the results of the residual graphs for independent variables indicated that the samples of this study met the assumption of homoscedasticity—the points equally distributed above and below zero on the X axis and to the left and right of zero on the Y axis (see Figure 11). Furthermore, the Variance Inflation Factor (VIF) values for the predictors ranged from 1.108 to 3.967, while the Tolerance values ranged from .252 to .902, indicating that multicollinearity was not a problem in the data (see Table 12). Finally, the P-P plot showed the sample data approximated a normal probability line (diagonal), which indicated that the residuals were normally distributed (see Figure 11).

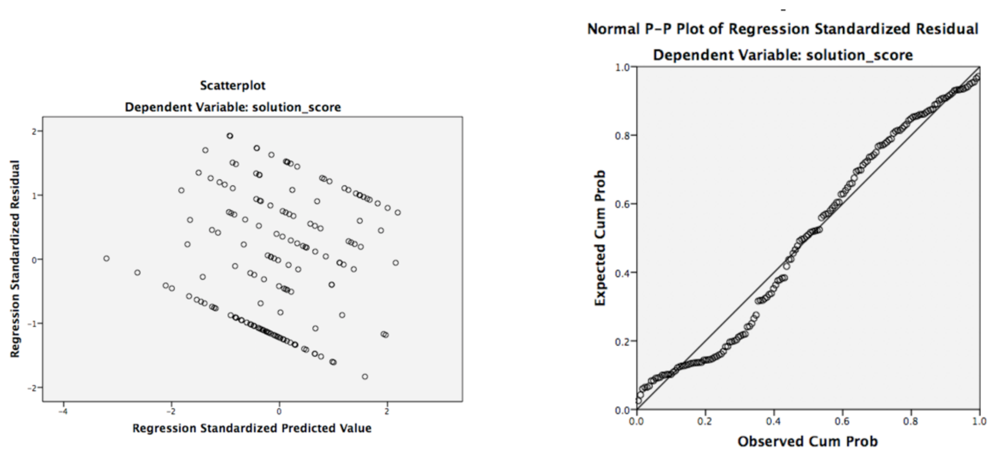


Figure 11. Scatterplot and P-P plot in Multiple Regression

The results showed that the model as a whole was significant ($p < .01$), which indicated that learner goal orientation and metacognition were significant predictors for problem-solving performance, $R^2 = .134$, $F(7, 155) = 3.283$, $p < .01$. The regression equation was:

$Y_{ss} = -.025*MC - 0.34*TAP + 0.252*TAV + 0.006*SAP - 0.229*SAV - 0.003*OAP + 0.065*OAV$

In addition, learner TAP, TAV, and SAV were significant predicators of performance during problem-solving. Specifically, for every point increase in TAP, a 0.34-point decrease in learner problem-solving performance was predicted; for every point increase in TAV, a 0.252-point increase in learner problem-solving performance was predicted; and for every point increase in SAV, a 0.229-point decrease in learner problem-solving performance was predicted. Furthermore, there was a weak relationship between learner metacognition and problem-solving performance ($r = -0.19$, $p = 0.009$). SAP and problem-solving performance also showed a weak relationship ($r = -0.211$, $p = 0.004$). See Table 12 for the regression results.

Table 12. Summary of Multiple Regression Results

Predictor	Solution Score (N = 156)					
	Pearson Correlation	B	R^2	β	Tolerance	VIF
Model			.134**			
Metacognition	-0.19**	-0.025		-0.122	0.902	1.108
TAP	-0.277***	-0.34		-0.299**	0.457	2.186
TAV	-0.073	0.252		0.275*	0.366	2.731
SAP	-0.211**	0.006		0.006	0.346	2.893
SAV	-0.207**	-0.229		-0.259	0.252	3.967
OAP	-0.112	-0.003		-0.004	0.48	2.083
OAV	-0.045	0.065		0.097	0.406	2.465

Note: $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

Although the regression model was considered significant, it only had a small R^2 (R^2 = .134), which indicated this model could only predict 13.4% of the data. Therefore, based on the final cluster result, this study proposed a new regression model, which used the three final cluster groups as variables to predict learner problem-solving performances, as follows:

$$Y_{ss} = \beta_1 Final\_GO + \beta_2 Final\_MC + u$$

In this model, Final_GO indicated learners in high, medium or low goal orientation groups, while Final_MC indicated learners of high, medium and metacognitive levels.

The multiple regression results showed that the new model as a whole was significant ($p < .000$), which indicated that high metacognitive level-high goal orientation, low metacognitive level-medium goal orientation and medium metacognition level-low goal orientation were significant predictors for problem-solving performance, R^2= .445, $F(2, 129) = 51.7$, $p < .000$. The regression equation was as follows:

$$Y_{ss} = -0.318 * Final\_GO - 0.428 * Final\_MC + u$$

Both Final_GO and Final_MC were significant predicators of performance during problem-solving. Specifically, with every point increase in Final_GO, a 0.928-point decrease in learner problem-solving performance is predicted; and with every point increase in Final_MC, a 1.314-point decrease in learner problem-solving performance is predicted. In addition, there was a strong correlation between Final_GO and problem-solving performance (r = -0.571, p = 0.000). Final_MC and problem-solving performance also showed a weak relationship (r = -0.616, p = 0.000). See Table 13 for the regression results.

Table 13. Summary of Multiple Regression Results Using a New Model

Predictor	Solution Score (N = 132)					
	Pearson Correlation	B	R^2	β	Tolerance	VIF
Model			.445***			
Final_MC	-0.571***	-.928***		-0.318	0.648	1.542
Final_GO	-0.616***	- 1.314***		0.428	0.648	1.542

Note: *p < 0.05; **p < 0.01; ***p < 0.001.

1.c. Reasons for differences in learner performance

Learner problem-solving performances were measured using their solution scores, which were between 0 to 7 points. Specifically, for Jakala-Tay, there are two appropriate planets—Venus and IO. The learner can get one point if they chose either of them. In addition, each piece of evidence the learner provided for the solution was worth one point.

In the Alien Database tool, AR provided four types of hints for finding Jakala-Tay a new home, including 1) habitat: sulfur in atmosphere, no hydrogen; 2) food: nitrogen; 3) dwellings: metals needed, and earthquakes is okay; and 4) inhabited world: temperature is 200-500K, gravity is one-third of Earth.

To understand the reasons for any possible problem-solving differences among learners based on metacognition and goal orientation, the 12 selected participants were divided into four groups based on the final cluster analysis result, including, 1) high metacognition, high multiple goal orientations, and low performance group (*N* = 3); 2) low metacognition, medium multiple goal orientations, and high performance group (*N* = 3); 3) medium metacognition, low multiple goal orientations, and high performance group (*N* = 2); and 4) outliers (*N* = 4). See the detailed information about these 12 interviewees in Table 14, including pseudonyms using the Greek alphabet, major, ethnicity, gender, age, metacognition score (MC), goal orientation Group (GO), solution score (SS), and clustered groups.

Table 14. Information for 12 Stimulated Recall Interviewees

Learner	Major	Ethnicity*	Gender**	Age	MC	GO	SS	Group
Alpha	College of Liberal Arts	White	F	20	61.34	TAP TAV	0	Outlier
Beta	College of Liberal Arts	African American	M	20	71	TAP	0	1
Chi	Moody College of Communication	White	M	21	57.88	TAP TAV	0	Outlier
Delta	College of Liberal Arts	Asian	M	19	66.38	TAV	4	3

Table 14 continued.

Learner	Major	Ethnicity*	Gender**	Age	MC	GO	SS	Group
Epsilon	College of Natural Sciences	White	F	20	50.17	SAP	4	2
Gamma	College of Natural Sciences	Asian	M	18	71.84	SAP	5	1
Kappa	School of Undergraduate Studies	African American	F	19	52.9	TAP TAV SAV OAP OAV	2	2
Lota	College of Liberal Arts	White	F	23	82.5	OAV SAV	7	Outlier
Mu	College of Natural Sciences	White	F	23	43	OAV	6	2
Omega	Engineering and Natural Sciences	White	M	22	72.36	TAP TAV OAV	4	Outlier
Theta	Moody College of Communication	White	M	22	58.85	OAP	7	3
Zeta	College of Fine Arts	Hispanic	F	18	75.1	OAP OAV SAV SAP	0	1

Note: * White indicates White/Caucasian/European American; Asian indicates Asian American/South Asian American; Hispanic indicates Hispanic American/Latino/Chicano; African American indicates African American/Black; Multi indicates Multi-Ethnic, such as both Hispanic and Caucasian.

** F indicates Female, and M indicates Male.

Group 1: Identifying the wrong problem

Three interviewees were clustered into group 1, including Beta, Gamma, and Zeta. Among these three students, Beta and Zeta did not solve the problem although they both were considered as having high metacognitive levels and high goal orientations in all categories compared to their peers. Gamma successfully solved the problem, and had a high solution score (i.e., 5 points). See Table 15 for the detailed information on these three interviewees.

Table 15. Information for 3 Interviewees in Group 1

Learner	Major	Ethnicity *	Gender **	Age/ Year	MC	GO	SS
Beta	College of Liberal Arts	African American	M	20 Junior	71	TAP:21 TAV:20 SAP:19 SAV:19 OAP:17 OAV:18	0
Gamma	College of Natural Sciences	Asian	M	18 Fresh man	71.84	TAP:19 TAV:18 SAP:20 SAV:19 OAP:15 OAV:16	5
Zeta	College of Fine Arts	Hispanic	F	18 Fresh man	75.1	TAP:21 TAV:21 SAP:21 SAV:21 OAP:12 OAV:13	0

Beta was a Junior student in the College of Liberal of Arts. He thought AR was cool, and he pointed out that AR was a problem-solving game. During problem solving, he sent one probe to Mars, but did not take any notes on the planets or aliens. He chose Mars as the final solution, and suggested that,

i recommend Jakala-Tay for mars. i believe it would be a good fit. i think think that its the closes to earth and they have things very similar to where they were before. if they dig, which they like to do they may be able to fins frozen water. They chemicals found on mars aren't harmful so all in all i think they will be okay. [Quoted without editing]

Based on his solution message, it seems he had a false concept—the planet had to be similar to earth—before he started to learn about the alien. He also did not identify the problem correctly, so he wasted some time on investigating a different alien—Akona—for 20 minutes before realizing he needed to help Jakala-Tay. He did have help-seeking behavior by asking the researcher some clarification questions during problem-solving, which was consistent with literature about learners who had a higher TAP score (Ning, 2016). Since he had slightly lower OAV and OAP score, it made sense that he also was not affected by his peers during the study, which indicated he was more task goal orientated. For example, when asked if he was checking out another participant's computer screen on her progress, he answered, "I didn't look at her screen just because… like… I was kind of, like, focused on trying to make sure that I could figure it out, because I know that…that was a big thing for it." Despite his high metacognitive level, Beta's failure to identify the problem correctly at the initial stage caused he was ultimately unable to solve the problem.

Likewise, Zeta was not able to solve the problem. She was a freshman in the College of Fine Arts. She also chose Mars as the answer, and the reason she wrote in the recommendation was "Mars has been recently speculated as a planet that is capable of housing life. It's frozen ice caps prove that water can exist on the planet." Zeta mentioned she felt lost at the beginning, and she even sent a recommendation before doing any research. After 20 minutes of playing AR, she was still exploring the environment, wandering around, and trying to identify the problem. Therefore, she did not have time to learn about other planets. Rather, she used her previous knowledge of the solar system,

I feel like since Mars was expected to have traces of life, they could probably, like, find a way to grow, like, the plants and stuff, they weren't there. Even though it'll be hard, I mean, if they're aliens, they'll have additional something, I guess.

Then she explained more:

I'm not really know too much about, like, the other planets, They are not as commonly talked about, so I was just, like, I'm just gonna talk about Mars, because I feel like that's, like, what I am familiar with, like, I feel like he'd be fine with Mars.

Since Mars is the most familiar planet to her, she decided to send Jakala-Tay to Mars. According to the interview, Zeta did not really understand the problem in the game. She also did not regulate her problem-solving behavior despite her self-identifying as a high metacognitive level student.

For Beta, Zeta, and some other students in Group 1 who reported a higher metacognition level but had the lowest problem-solving performance, one possible explanation is the Dunning-Kruger Effect (Kruger & Dunning, 1999), which indicates a cognitive bias that people of low metacognitive ability usually mistakenly assess their ability as greater than it is. Likewise, it is possible that Beta and Zeta overestimated their metacognitive levels during the study.

Different from Beta and Zeta, Gamma had a high solution score. He was a freshman in the College of Natural Sciences. He was selected as an interviewee because he had a high SAP goal orientation score compared to other goal orientations. From the interview, Gamma also indicated that he was not clear about the problem he needed to solve, so he spent 10 more minutes on the wrong task:

First I was just checking all the details. I probably for the first 10 minutes doing that. Then I figured out you supposed to come here to... I was looking for the all the information for Akona. I thought it was the alien.

Although he did not send a probe, he did draw a correct conclusion and explained the rational:

Io contains lots of volcanic activity, like the Jakala-Tay are used to. It also has a relatively similar gravity to that of their home planet. It contains sulfur and has a temperature range that is close to what the Jakala-Tay are used to. [Quoted without editing]

Since he mentioned four more pieces of evidence to support his choice, such as the volcanic activity, gravity, sulfur, and temperature range, he got 5 points for the solution score.

In summary, these three participants had trouble identifying the problem—helping Jakala-Tay find a home. Beta spent 20 minutes on the wrong alien, Zeta spent over 20 minutes exploring the environment, and Gamma spent 10 minutes helping alien Akona. In addition, although Beta did ask for help identifying the problem, it was too late, while Gamma was able to correct himself in the early stage of the problem-solving. Furthermore, Beta and Zeta might have overestimated their metacognitive levels during the study.

Group 2: Measure twice, cut once

Three interviewees were clustered into group 2, including Epsilon, Kappa and Mu. They were all considered to have low metacognition, medium goal orientation, and high solution scores. See Table 16 for the detailed information for these three participants.

Table 16. Information for 3 Interviewees in Group 2

Learner	Major	Ethnicity *	Gender **	Age/ Year	MC	GO	SS
Epsilon	College of Natural Sciences	White	F	20 Junior	50.17	TAP:18 TAV:20 SAP:21 SAV:20 OAP:16 OAV:14	4
Kappa	School of Undergraduate Studies	African American	F	19 Sophomore	52.9	TAP:21 TAV:21 SAP:16 SAV:21 OAP:21 OAV:21	3
Mu	College of Natural Sciences	White	F	23 Senior	43	TAP:17 TAV:17 SAP:17 SAV:16 OAP:18 OAV:20	6

Epsilon was a Junior student in the College of Natural Sciences. She sent Jakala-Tay to Venus with the rationale of "Venus has good atmosphere, and an appropriate gravitational field. Additionally: if the Jakala-Tay live in tunnels, they would be able to cope with the high temperatures. Thank,s for coming to my TED talk, have a good day [Quoted without editing]." Her funny rationale provided three pieces of evidence including atmosphere, gravitational field, and temperature, which was graded at 4 points. The positive tone in her explanation was consistent with literature on learners who have a high SAP. These learners usually possess positive activity-related emotion (Brondino et al., 2014). During the interview, when asked whether she was familiar with the solar system, she said, "uh-emm, technically, but not really; like, I don't know anything about the planets' moons, but I know, like, that Jupiter has…it is a gas planet, you know—very basic information."

Although she was not a subject matter expert, after some exploration in AR, she identified the problem and started to take some notes. Initially, she thought Titan, IO, and Venus were the possible planets, because she "ruled out everything that wasn't compatible," and then "sent a probe to the ones that, like, didn't right away get scratched off," then "looked at them, and then [she] just kind of, like, chose from there, that one [she] thought would be [correct]" after double-checking her notes and the results from sending the probes,

I went back, because I was, like… hmm… let me see if there's anything that I forgot about them, like, I wanted to double-check with the notes I'd taken… uh… I'd like to see it compared to the planet, you know.

She realized she had the wrong hypothesis at first time then made the correct final decision. Epsilon had a slightly lower score on her OAP and OAV, which indicated she was not easily affected by her peers. For example, when being asked whether she felt pressure from her peers, especially by those who finished and left early, she said, "not really, cause, like, when I saw people were leaving, I was, like, I still have, like, 17 minutes, you know; so, and then I think, after this, I went and did a probe." It seems she did not care much about people's leaving, and she was just focusing on her own problem-solving.

Different from Epsilon, Kappa—a sophomore student in Undergraduate Studies—sent Jakala-Tay to another appropriate planet named IO, and the reason was "It is the right colors, has the right things for them to breathe and vegitation?" Since she provided two pieces of evidence regarding atmosphere and vegetation, she was given 3 points for the solution score. Different from the short answer she submitted to the system, Kappa articulated her problem-solving process and rationales for selecting IO very well while the researcher conducting the interview.

I open, like, all the information to try to see what I was doing. Oh no… I was in the Mission Control Center and I was, like… what's this?… I realized that I didn't have any of the stuff that I needed to look at the Mission Control Center, so I went to look at the aliens and I read up about, like, the specific alien that we were doing, and then I went…um… I went back and forth between making my probe and looking what happens when every your probe actually lands, and then I realized that, oh then… It finally clicked that I am supposed to be looking for, like, whether the planets had a certain, like, things that the Jakalay-Tay they needed. So, then, I went through each one, and I was, like, reading… oh that's what I felt, like, that's what took the most time… cause I read one now, okay, does this one have sulfur? Does it have hydrogen? Does it have nitrogen? And so, that's when I went back. So after that, I just went back and forth between making the probes and sending them to different places to check the spectrum.

Only if she could write down the above information, could she have a higher solution score. For testing her hypothesis on the appropriate planet, she sent two probes to Mars, followed by one to IO, and one to Ganymede. She also explained the reason for doing so as follows,

I sent the probe to Mars, because Mars said that, well, first of all, Mars was like the certain colors; I don't know why I was looking for colors, but Mars had certain colors that, mmm-hmm, the Jakala-Tay had… and then it had volcanoes, and Mars had something… of which was a reason… Oh it had water for the vegetation yeah. So I was, like…okay, maybe Mars is right, but then when I checked, Mars didn't have what it needed to breathe; it didn't have hydrogen, which would kill it, but it didn't have what needed to be something… I knew that. That one was wrong.

Once she found Mars was unsuitable for the alien, she sent one probe to IO and wrote a recommendation. In addition, she kept looking for a better planet, even there were only 12 minutes left:

Yeah, so that's when I started looking for other ones, because, like, even though IO was very, like, ideal, I noticed it didn't have the nitrogen, so I was trying to look for the one that was, like, perfect, but I didn't do… I did not have time to verify…

According to the log data, Kappa sent a probe to planet Ganymede, but just like she said, there was no time for her to verify other planets. Eventually, she only submitted one solution to IO with limited rationale. Regarding not providing more evidence for the rationale, it might be relevant with Kappa's information gathering behavior while using the notebook tool in AR. She suggested that

I didn't really use the notebook tool. Because, well, for one, I didn't really know how to use it, and then, for two, the only thing that I really just needed to remember was that, like, was the three things… about, like, you know, what it needs to be, what will kill it, and what it needs to eat and so.

Since she thought it was unnecessary to take notes, and she could not remember all the details regarding the planet and alien, it was inevitable that she could not write more on the rationales. She also admitted that she should have taken more notes:

Yeah, which is kind of bad, because sometimes I had to keep going back, so maybe I should've written it down, but sometimes I didn't, so I have to keep going back and, like, make sure… but I felt like before I even found the spectrum thing, I was trying to figure out what exactly the spectrum meant, so that's what I wanted to write down in my notes.

However, AR did not provide her an effective way to take notes—when she wanted to take notes on spectrum, "there wasn't a way to, like, take a picture of this guy." Therefore, "after my probe was sent and when I checked the control center, I recently, yeah, when it was telling me which elements were on that planet, I didn't know what spectrum went with the… you know what I mean?"

Different from Kappa's short answer in solution message, Mu (Senior student in the College of Natural Sciences) found two possible solutions—Venus and IO—and submitted one of the correct answers (i.e., Venus) with elaborated evidences:

Venus and Io are both close matches, but Venus has the N needed in the atmosphere needed for plant growth and has a heavy atmosphere while Io has a thin atmosphere. Venus is quite hot but there was no specific specification that the Jakala-Tay require a certain temperature, especially because they can live in tunnels. The colors of the environments are also good for both. [Quoted without editing]

Since she provided sounding evidence regarding elements in atmosphere, atmosphere thickness, temperature, living conditions, and coloring of the planet, she was granted 6 points for the solution score. Despite having a low metacognition score (i.e., 43 points), Mu had a higher OAV score compared to other five goal orientations, which was consistent with literature that learners who have high OAV have a higher academic achievement (Diesth, 2015).

She also had a positive attitude towards solving the problem. She indicated that "I was actually having a pretty good time, like, trying to figure out, like…oh, well, obviously, none of these are gonna be, like, perfect matches but, like, let's try to find something that will work and then, like, kind of going beyond the logic of the game." Mu behaved the same as Epsilon—she also double-checked many times before the final submission. She admitted that she was "that kind of person, check their guess themselves (laugh)." For example, when asked the reason for coming back to the Alien Database after taking notes, she answered "just cuz I had so much extra time, I was, like, I should, yeah; so, I guess if I have extra time, I would usually go back and check things." She even said that "just double-checking, a lot of my time." Due to her double-checking behavior and positive learning attitude, she solved the problem and gained a higher score.

Group 3: Double check and time management

Delta and Theta were clustered into group 3, and they were considered as having medium metacognitive levels, low goal orientation scores, and high problem-solving performance (see Table 17).

Table 17. Information for 2 Interviewees in Group 3

Learner	Major	Ethnicity *	Gender **	Age/ Year	MC	GO	SS
Delta	College of Liberal Arts	Asian	M	19 Sophomore	66.38	TAP:15 TAV:17 SAP:16 SAV:16 OAP:14 OAV:16	4
Theta	Moody College of Communication	White	M	22 Senior	58.85	TAP:13 TAV:12 SAP:14 SAV:14 OAP:15 OAV:12	7

Delta is a Sophomore student in the College of Liberal Arts. He had a medium metacognition score and low goal orientation scores compared to his peers. He sent Jakala-Tay to IO, and provided the rationale, "Jupiter's moon of IO is a good candidate for the Jakala-Tay's new homeworld. The moon has a similar geography in that volcanoes cover the moon, and have frequent eruptions that contain sulphur, which is the air they breathe. The atmosphere is 30% Sulphur and contains no trace of Hydrogen [Quoted without editing].” Since Delta pointed out three critical elements such as volcanoes, Sulphur, and Hydrogen, he earned four points for the solution score.

Compared to students in Group 1, who spent time on gathering unnecessary information for solving the problem, Delta directly identified the problem:

I think it's been a good, like, a good amount of my time was spent in that room [Alien Database] that was the most. It was just most… I think the thought process was just finding information about, like, this species and then… about, like, just both the species themselves and their home planet. Afterwards it kind of, just was, just using the tools, just to find, like, information about that or just find planets or moons that, like, was similar to the description.

Delta was also good at time management and filtering information. When asked whether he checked at the timer, he said,

Yeah, I was… I was periodically looking at it, just so I could pace myself and make sure that I wasn't wasting time, you know? Don't lean on something that was irrelevant, because I think in the beginning… like, I was just looking through, like, just stuff and there's some information that, like, was not, like, really helpful, like… I think there's, like, in the beginning, like, that the first, like, monitor [Message Tool], there was, like, information that was, like, it wasn't really helpful… is just more, like, just introduction.

During the study, Delta stayed until the last minute, while five other participants in his session had finished and left early. However, he was not affected by their leaving or feeling rushed; rather he just focused on his own task and double-checking the answers,

It really didn't bother me that much, cause with me that… like, in how I operate whether it comes to working on assignments or taking tests, I would take, like, the whole time even if, like… I finish, like… I guess even, like, an hour or 30 minutes before the time is up, I would still, like, use all of it just to double-check and make sure that, or like, I'm *sure my answers, because I don't know… I just really don't like having the thought of missing answers or being incorrect on my responses just because I didn't use the whole time.*

Theta was a Senior student in the College of Communication; he had an even lower score on metacognition and goal orientations compared to Delta, but he had a top score for problem-solving performance. The same as Delta, he also chose IO as the home planet for Jakala-Tay, and the reasons were:

IO has a similar size and rough gravity to the home planet of Tay. Their is presence of an atmosphere with sulfur which is what the Tay breathe. The chemical composition and coloring of the landscape is reminiscent to their home world. In addition, the low temperature is balanced by a similar high from the planet Tay. There is presence of ice as well dirt and hard minerals to build with. Furthemore, frequent volcanic activity leads to earthquakes. The volcanic activity allows nitrogen to be released which will help plant species grow there. The Tay are used to digging and building tunnels in case of volcanic activity or earthquakes. [Quoted without editing]

Because he mentioned more than six pieces of evidence for selecting IO, such as gravity, sulfur, chemical composition, low temperature, building materials, volcanic activity, nitrogen, and tunnels, he scored 7 points. He also had similar double-checking behavior as Delta; specifically, he mentioned that,

I liked the fact that you had to really cross-check all the facts—make sure something was correct before you were spending. Double-check everything. It was kind of intriguing, saying I think, this is because I initially had two options, and I ended up going with a different thing than my initial plan.

And his initial options were Venus and Titan, because he "read about their home world using lots of sulfur." Then, after reading more, he realized "Venus was a little too hot, and it didn't have all the correct composition of things, and it was a little too similar in size to earth." As for Saturn's Titan, "the ground seemed muddy, which wouldn't be good composition-wise for building things, and there wasn't a sense of the other necessary requirements for Jakala-Tay to survive besides the sulfur and atmosphere, really." Therefore, he did more research; then he "came across IO, which had the same things, but it also had the same color scheme that was reminiscent and lots of volcanic activity, which they're used to… which would lead to nitrogen, which allows them to use the nitrogen for the gases or growing the plants; so it had the two main gases they needed, the rocky composition. Also the temperature highs and lows were similar, and it was smaller than Earth—closer in size to our moon, in fact."

It is worth noting that Theta drew the conclusion without sending any probes to verify the hypothesis. He only depended on the information provided in the solar system and alien database to make the decision and told the researcher that "I didn't know how to use the probes." He then explained more:

(I) tried, and that's it. There are no probes, so I didn't send any probes out. Well you were, like, thinking about… we were, like, trying to send something… I was thinking it could be useful to test my theories. but I never sent any, so I could be totally wrong.

Theta also figured out a way to "take notes" as compared with Kappa—the students in Group 2, who did not take notes regarding the spectrum because she said there was no way she can take a picture. Theta did exactly what Kappa did not do—take a 'picture' of the spectrum—he took some screenshots.

Mostly, I did a few things in a notebook, and then I checked mission results and the spectra; so here I find out information about the Tay food, and I took screenshots. So I can look at them again.

He then explained that "that's before I realized I could just click that right overlap it. I thought it was only one menu at a time at first."

In summary, both Zeta and Theta demonstrated double-checking behaviors and efficient problem-solving process. They also provided detailed evidences for their solution, which granted them higher solution scores than their peers.

Group 4: Outliers

Four interviewees were in Group 4, including Alpha, Chi, Lota, and Omega. They were clustered in the above three groups initially but were considered as outliers in the multiple regression analysis, because the regression model had a significantly higher predication rate when excluding these four learners (see Table 18)

Table 18. Information for 4 Interviewees in Group 4

Learner	Major	Ethnicity *	Gender **	Age/ Year	MC	GO	SS
Alpha	College of Liberal Arts	White	F	20 Junior	61.34	TAP:21 TAV:21 SAP:17 SAV:14 OAP:14 OAV:16	0
Chi	Moody College of Communication	White	M	21 Senior	57.88	TAP:18 TAV:18 SAP:16 SAV:16 OAP:12 OAV:12	0
Lota	College of Liberal Arts	White	F	23 Senior	82.5	TAP:18 TAV:18 SAP:20 SAV:21 OAP:19 OAV:21	7
Omega	Engineering and Natural Sciences	White	M	22 Fifth Year	72.36	TAP:21 TAV:21 SAP:16 SAV:17 OAP:19 OAV:21	4

Alpha and Chi both got zero points for the solution score. Alpha was a Junior student in the College of Liberal Arts. She had high TAP and TAV scores, and a medium metacognition score. The solution she came up with was Titan, and reason was “this was the best one that I can find and recommend in the given time. The temperature isn't perfect, but it's more consistent, and the magnetic field isn't overwhelming like Jupiter [Quoted without editing].” According to Alpha’s answer, she only had enough time for considering the temperature factor. However, the time-constraint might be due to her distracted problem-solving behavior in AR. For example, she spent over 5 minutes checking out the alien pictures on the wall in the Alien Database room,

Then I was, like, wandering around, and I felt like this is a good place to start [Alien Database]. And then I was, like, looking at the picture here [Alien pictures on the wall]… I usually, like… people don't pay much attention to how it works. I was, like, look at these! How did they put this in? It was just awesome!

She also did not filter the information efficiently. Rather, she checked on everything in the environment—"yeah, I definitely clicked on everything just to see how it works…," because she wanted to "*make sure that I was thorough in trying to find what I could*." She also spent a lot of time on taking notes. When asked the reason, she admitted that she was weak in science knowledge,

"I did [take lots of notes], just because, like… I knew I was gonna be really weak as far as my natural planet knowledge. So I was, like, maybe if I have, like, really, really extensive notes, I could finally match things up."

Until she saw other participants were leaving, and realized that the time was running out, "*I probably started to get, like, I start… that's probably around the time… I realized, like, time was running out, so I started to get more impatient."*

She did seek help during the problem-solving by asking some clarify questions about the game controls. Different from many participants, after the study, she also showed concern about her learning outcome. She asked two questions regarding the outcome — one on the right answer for the solution—"Yeah…what was the planet?" and the other was about how to design a probe to test the atmosphere—"*okay, just the last question with the whole… like how it was like, hey, this atmosphere will kill them… this atmosphere is what they breathe… it's, like, what instrument could I have used to have found that?*"

However, due to poor time management and inefficient information filtering behavior, Alpha did not select an appropriate planet for Jakala-Tay.

As for Chi, he is a Senior student in the College of Communication. Same as Alpha, he had a medium metacognition score but had slightly lower scores on goal orientation. He wrote two recommendations for Jakala-Tay. One is Earth, because "Earth is an accommodating place, and considering that the human race knows more about it than any of the other planets, I'd feel more comfortable with Jakala-Tay living there [Quoted without editing]." The other is Pluto, because "I recommend Pluto. It has the gas you need to produce your food and does not have the gas that is poisonous to your species [Quoted without editing]." During the interview, he admitted that "I think the earth was a mistake," and he selected Pluto because "when I was looking at the different types of gases they thought existed at the other planets, I saw a lot of them contained levels of hydrogen which was poisonous for this species." So, he only investigated the gas factor that affected Jakala-Tay, but ignored other factors, which led him to draw a wrong conclusion.

In addition, Chi was "not really good at time management." When asked whether he had checked on the timer, he said, "I was so focused on learning and trying to figure things out, I didn't really pay attention at the Time." He also admitted that,

hmm…I'm not really good at time management. I know that, I know when I'm doing homework, I kind of just start working on it, and then whenever I finished, I finish, and I probably should pay more attention to time, obviously.

He also suggested that he would solve the problem more efficiently if he were given a second time, because

I think initially it took me a while just to kind of figure out what it is that I was doing, because I had a hard time exactly figuring out where to begin, but it's after some trial and error, I feel like I'd be able to help the other aliens a lot faster.

Therefore, lacking time management and information filtering skills, Chi had an unsuccessfully problem-solving session.

The other two participants—Lota and Omega—were also considered as outliers, because they had a higher metacognition score, higher goal orientation, and higher solution score. Lota is a Senior student in the College of Liberal of Arts. She had 82.5 points in metacognition, higher SAV and OAV score, and earned 7 points for her solution. She sent Jakala-Tay to IO, because

After doing some research, I had reason to believe the atmosphere if Io would be suitable for Jakala-Tay. It frequently has volcanic eruptions that shoots out sulfur putting sulfur in the air. I then sent a land probe to Io and evaluated the percentages in the atmosphere. It has no hydrogen, which is perfect since hydrogen is deadly to Jakala-Tay. It also has Sulfur which is ideal and is made up of 10% Nitrogen which means with some alterations, they may be able to grow the plant life from their home planet. The volcanic eruptions and earthquakes make Io scary to us but Jakala-Tay are used to this from home. I believe that Io is the best option. [Quoted without editing]

Omega was a fifth-year student double-majored in Engineering and Natural Sciences. There were only 3 students who were beyond the Senior level in this study. Literature suggested that as a nontraditional student, these three students were statistically significantly more likely to be OAP and OAV than nontraditional students. Omega did have a high OAV score (i.e., 21 points). He also had a high metacognition score, TAP score, TAV score, and earned 4 points for his solution. Different from Lota, Omega sent Jakala-Tay to Venus, because it "Fits most of the specifications of gases temperature and seismic activity [Quoted without editing]." Although his rationale was just one sentence, he mentioned gases, temperature, and seismic activity, which were counted as three pieces of evidence.

Although Lota and Omega had different majors, genders, and ages, they had at least three things in common. Firstly, they both mentioned they like to double check their answers or relevant information before making a final decision. Lota said that,

yeah, en-hmm. and I always, like, if I'm working on an assignment, I always double-check, like, the key requirements and when it's due, and I look at the grading rubric, and, like, the checklist.

Omega also double checked the Solar System Database multiple times:

I just wanted to see if there was, for example, something else… one of the moons or planets that I missed and that would be more close… that perhaps, like, what I thought of what was a good planet or moon, or maybe there's something better that I just missed. I'm trying to find mainly the atmosphere—a good qualities on the atmosphere, and temperature.

In addition, when asked if he had similar behavior in class, he admitted that,

Yes, I'd have to reread something several times. You know that the type of people who like to read something, and then they almost forget what they read? Like the minute they read it? It's kind of like that. It's not—I guess it wouldn't be so much that you completely forget it, but it's, like, let me just look at it again—let me think about it.

Secondly, both participants fully utilized the 60-minute study time. Lota checked on her watch periodically and knew when she started and when to finish, "I check my watch and I knew I started at 9:20 [am]." When asked if she had pressure compared to other participants who might have finished the task early, she said that, "um…I just, I didn't really think about that. Yeah, not really. I was just dong my thing, be like, it is not a race. Everyone would have 60 minutes." As for Omega, when asked if he was hurried to finish the study, he said, "well, because, I mean, we're here for 60 minutes and you're gonna use the whole time"; and he explained further that, "I don't get anything…I don't know…submitting the answer earlier, it's not like I hit a bonus."

Finally, different from other interviewees, they both showed concerns regarding their learning outcomes. For Lota, at the end of the interview, she was eager to know, "Was I wrong?" When being told her decision would be considered as correct as long as she explained her rationale for selecting the planet, she said, "I would like to know why." Omega had a slightly different concern about the outcome. When asked if he had any questions for the interviewer, he said, "no…only…not about this…but just, just like the credits and all that?" He would like to know when he could get the extra credit for participating in the study.

Because of their time management skills, double-check on problem-solving, and care for outcomes, these two participants had high performance scores.

QUESTION 2: PROBLEM-SOLVING PROCESS DIFFERENCES BASED ON LEARNER CHARACTERISTICS

Different from research question one that focused on learner problem-solving performances, question two focused on learner problem-solving processes, five sub-questions included (a) what are learner problem-solving process patterns; (b) are there any problem-solving process pattern differences among students based on their metacognition; (c) are there any problem-solving process pattern differences among students based on their goal orientations; (d) are there any problem-solving process pattern differences based on the interaction between learner metacognition and goal orientation; and (e) what are the reasons for any problem-solving process pattern differences based on learner characteristics (i.e., metacognition and goal orientation)? The following sections report the results for the above five sub-questions using both quantitative and qualitative data.

2. a. Visualizing learner problem-solving process patterns

This study visualized three types of learner problem-solving process patterns including learner tool use frequency (Liu et al., 2004; Liu & Bera, 2005), duration (Liu et al., 2009; Liu et al., 2015), and sequences (Kang, 2017), particularly room visit sequences.

Tool use frequency

In this study, 159 participants used 10 tools for a total of 79,998 times (see Figure 12). Among all the 10 tools, Solar System Database was the most popular one—used for 17,815 times, followed by Probe Design (16,205 times) and Mission Control (12,395 times). The three least frequently used tools were Periodic Table (732 times), Spectra (1, 668 times), and Concepts Database (2,361 times).

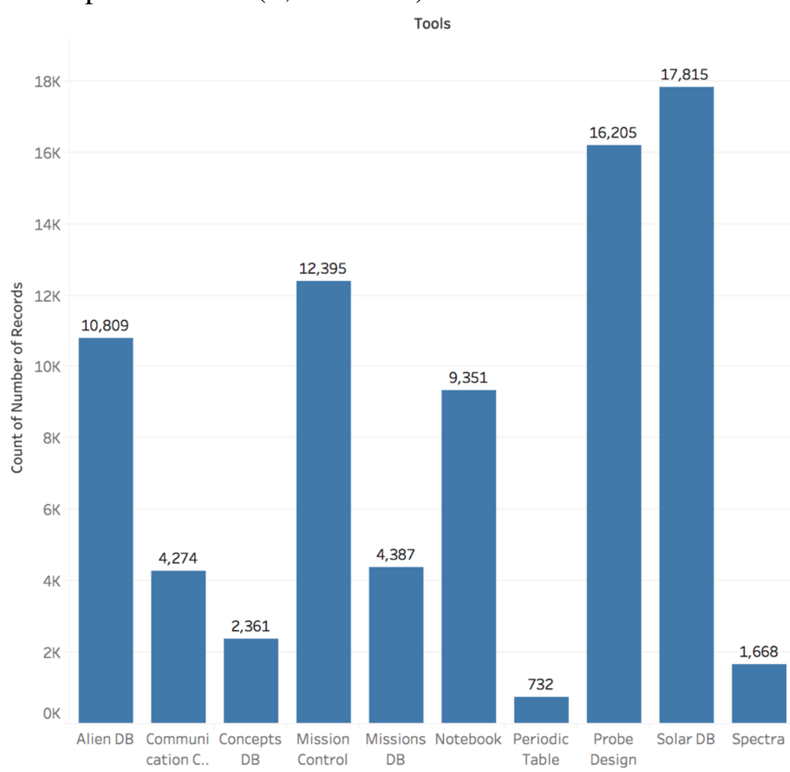


Figure 12. Tool use frequency among all learners

Previous quantitative analysis indicated that there was a significantly difference in solution scores between male and female students. There was also a difference in tool use average frequency among learners based on their gender—female students used all the tools less frequently. Particularly, for Solar System Database, on average, each female student used the tool 101 times, while male student average used it 121.9 times during 60-minute of problem-solving in AR. Female student average also used Probe Design 28.4 times fewer than male student (see Figure 13).

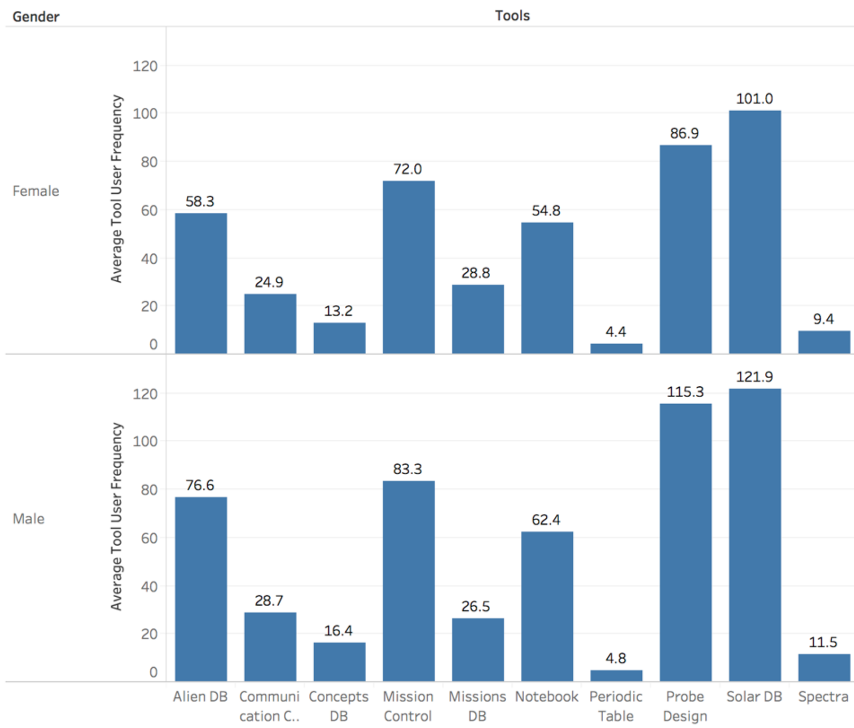


Figure 13. Tool use average frequency based on gender

To further understand the tool use frequency pattern, Tableau was used to visualize learner average tool use frequency based on learner solution score (see Figure 14). According to the visualization, the highest-scoring student used Solar System Database more frequently than the lowest scoring students, while the lowest scoring students used

Alien Database, Communication Center, Mission Control, and Probe Design more frequently compared to the highest-scoring students.

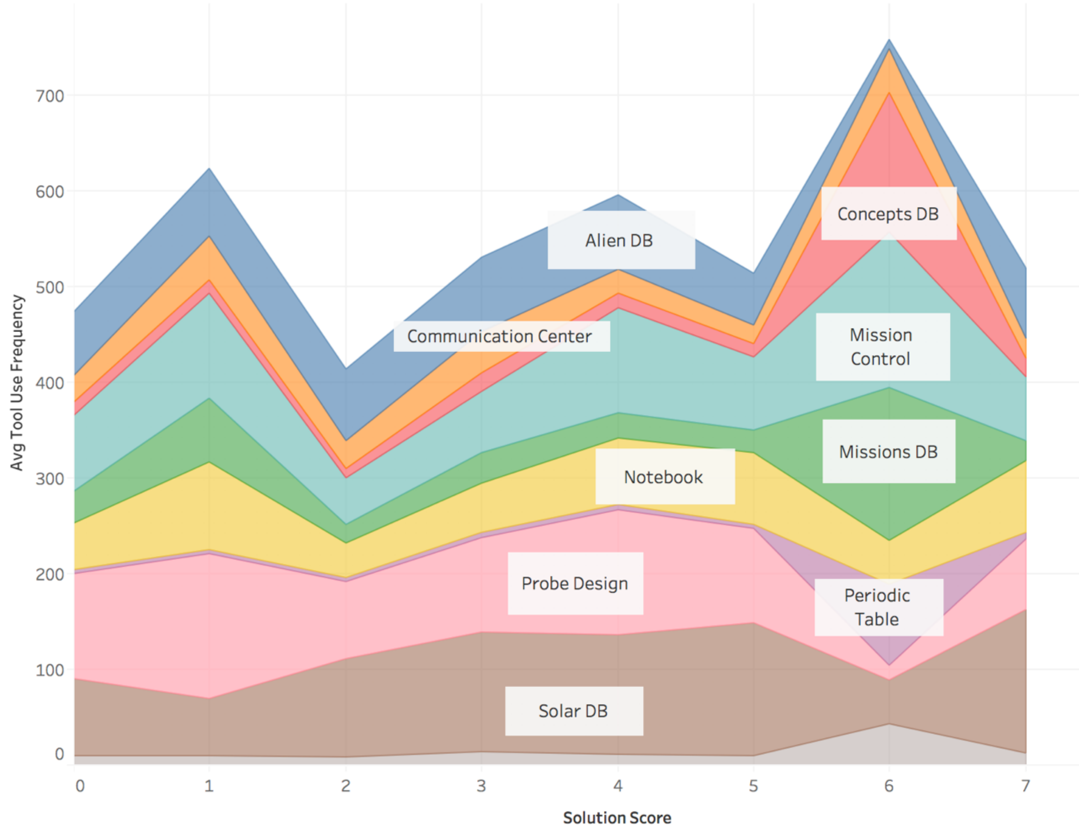


Figure 14. Tool use frequency based on solution scores

Tool use duration

Besides tool use frequency among all learners, Tableau was also used to visualize tool use duration for all learners. During 60-minute of problem-solving in AR, the top three tools that learners spent time on were Alien Database (M = 19.67), Probe Design (M = 13.93), and Solar System Database (M = 9.06). On average, learners spent the least time on tools such as Concepts Database, Missions Database and Spectra—less than 4 minutes (see Figure 15). It is interesting that learners spent the most time on Alien Database, while they did not open it often (ranked 4th place in tool use frequency). Learners also spent more time on Period Table, although this tool was used the least frequently.

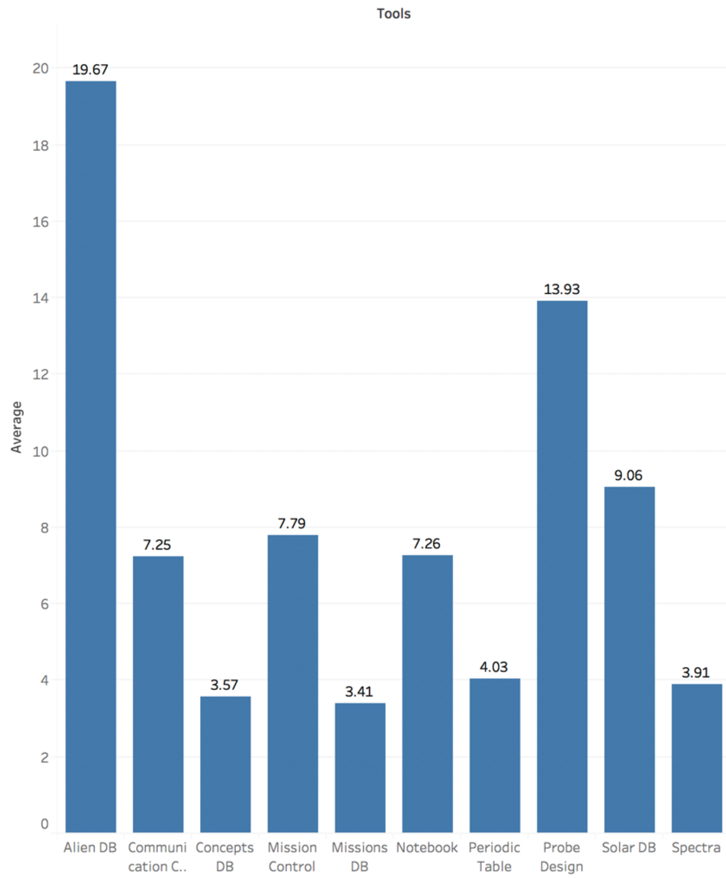


Figure 15. Tool use duration averages for all learners

Considering the gender factor, data visualization showed male students stayed in Probe Design longer (M = 14.63) than female students (M = 13.15). In addition, male students stayed in Mission Control, Notebook, and Spectra longer than female students (M = 8.69, 7.77, 4.56 for male students; M = 6.77, 6.69, 3.17 for female students). Both genders stayed in Communication Center and Periodic Table about the same amount of time—about 7 and 4 minutes (see Figure 16).

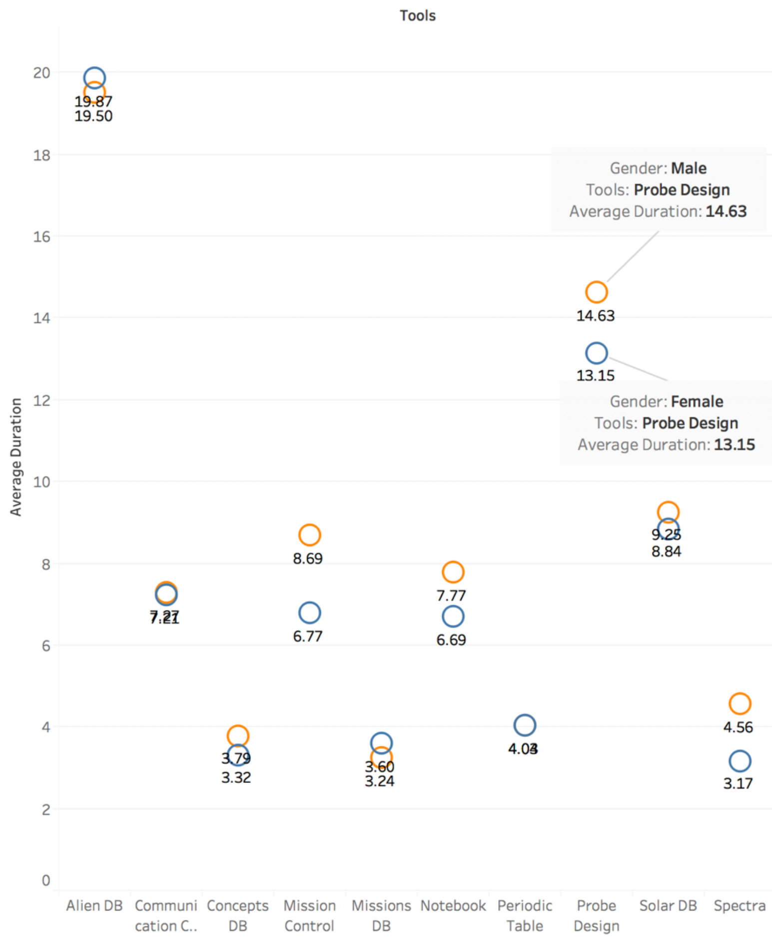


Figure 16. Tool use duration averages based on gender

This study also examined the tool use duration differences based on learner problem-solving performances (i.e., solution scores), see Figure 17. The horizontal axis indicated learner solution score groups, which contained 8 groups (0 to 7 points). The vertical axis indicated the average minutes of each tool use duration. According to Figure 17, learners who had the highest scores used Mission Control, Concepts Database, Notebook, Probe Design, and Solar System Database longer than students who had the lowest scores ($M_{MissionControl}$ = 8.292, $M_{ConceptsDatabase}$ = 4.583, $M_{Notebook}$ = 8.292, $M_{ProbeDesign}$ = 15.167, $M_{SolarSystemDatabase}$ = 8.917 for the highest-scoring students; $M_{MissionControl}$ = 6.963,

$M_{ConceptsDatabase} = 3.907$, $M_{Notebook} = 7.204$, $M_{ProbeDesign} = 13.074$, $M_{SolarSystemDatabase} = 7.963$ for the lowest scoring students). Learners who had the lowest scores and highest scores spent similar amounts of time using Alien Database and Periodic Table (around 18 and 4 minutes). Learners who had the lowest scores stayed Missions Database and Communication Center longer than learners who had the highest scores ($M_{MissionsDatabase}$ = 3.722, $M_{CommunicationCenter}$ = 7.278 for the lowest scoring students; $M_{MissionsDatabase}$ = 3.041, $M_{CommunicationCenter}$ = 6.417 for the highest-scoring students).

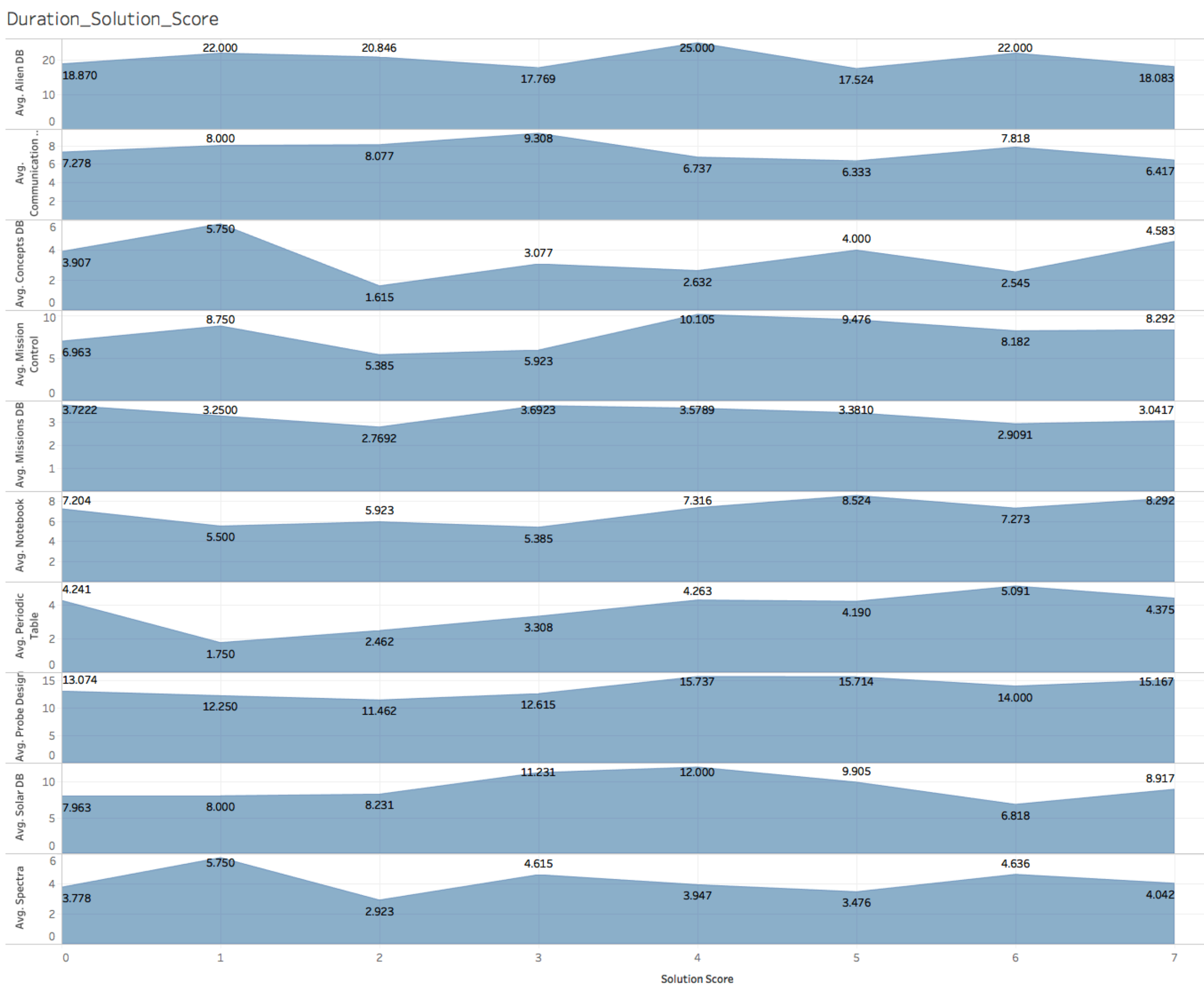


Figure 17. Tool use duration averages based on solution scores

Room visit sequences

Besides analyzing tool use duration and frequency, this study adopted software R to conduct similarity measure. Particularly, to visualize the differences, three similarity measures (i.e., Cosine, Jaccard, and LCS) were used to compare room visit sequences of learner groups who had different solution scores with the highest-scoring student group. See Table 19 for the average similarity measure coefficients based on learner groups.

Table 19. Similarity Measures Based on Learner Solution Score

Solution Score	Cosine	Jaccard	LCS	String Length	N (159)
7	0.30	0.07	0.39	74.21	24
6	0.36	0.08	0.42	82.82	11
5	0.30	0.07	0.40	82.38	21
4	0.28	0.09	0.45	79.53	19
3	0.33	0.13	0.44	71.69	13
2	0.41	0.10	0.41	90.23	13
1	0.42	0.07	0.39	84.00	4
0	0.36	0.12	0.44	78.22	54

Using a simple line graph in R, this study visualized the similarity among learners (see Figure 18). Based on the visualization, learners in Group 0, Group 2, and Group 3 had larger distances from learners in Group 7, which indicated learners in these three groups had much more different room visit sequences compared to learners in Group 7. In addition, learners in Group 4, Group 5, and Group 6 had smaller distances from learners in Group 7, which indicated learners in these three groups had much more similar room visit sequences compared to learners in Group 7. As for Group 1, there were only 4 participants, and the data showed that they had larger distances from Group 7 based on Cosine measure, but had the same coefficients based on LCS and Jaccard measure.

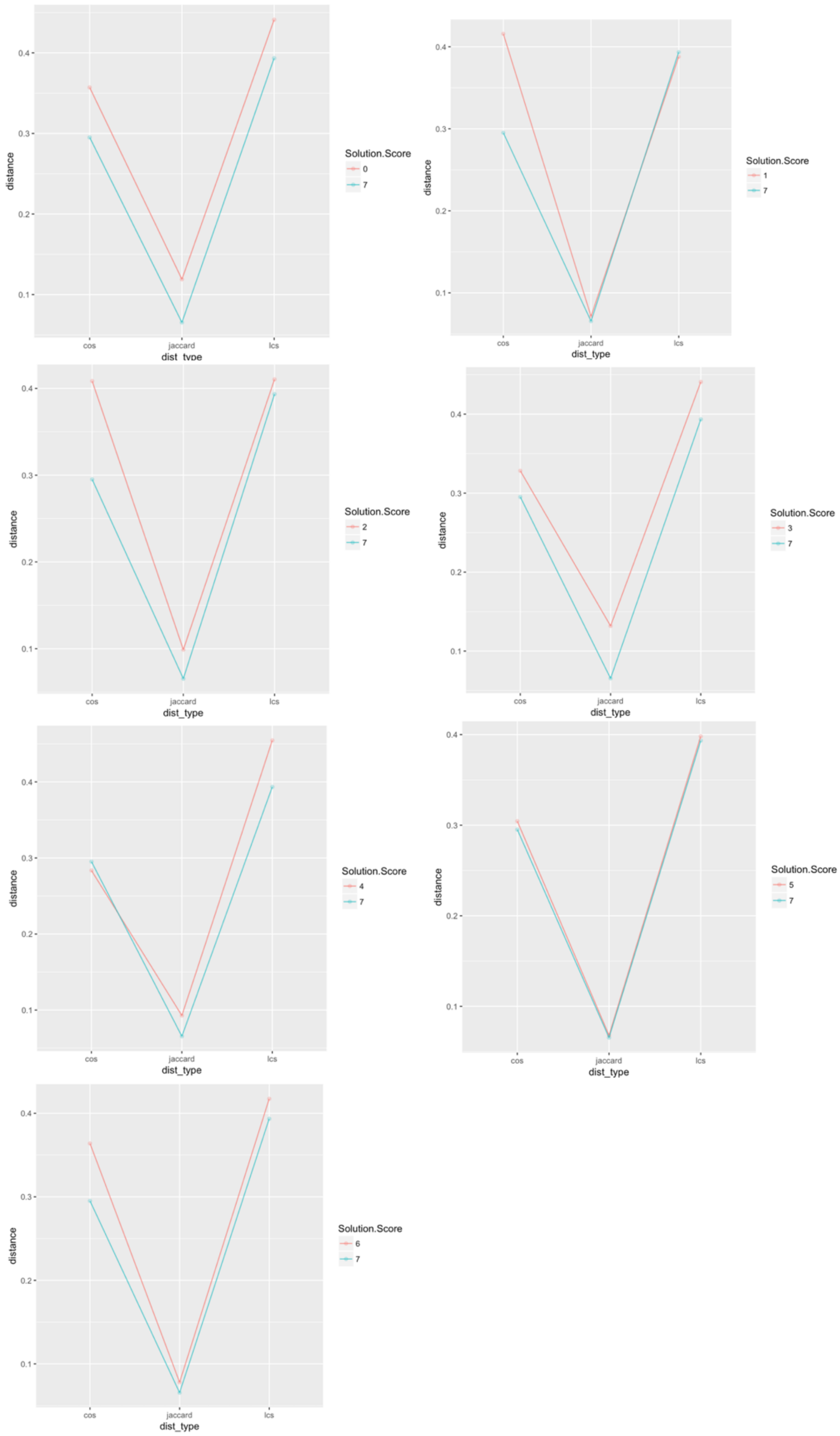


Figure 18. Visualization for similarity measures based on solution scores

Beyond eye-balling the visualization to decide the similarity of the room visit sequences, this study used two-proportion z-test to further investigate whether the similarity of the room visit sequences were significantly different based on the solution score groups. Specifically, the coefficient score of group 7 was used to compare with all the other groups. The results indicated that there were no significant differences in learner room visit sequences between learners in Group 7 and any other of the 7 groups. See Table 20 for the two-proportion z-test results.

Table 20. Two-proportion z-test of Similarity Measure Based on Learner Solution Score

Solution Score Group	Z_lcs	Z_cosine	Z_jaccard
6	0.43	0.90	0.27
5	0.14	0	0
4	0.86	0.31	0.52
3	0.72	0.46	1.41
2	0.28	1.62	0.76
1	0	1.78	0
0	0.72	0.90	1.20

2. b. Problem-solving process patterns based on metacognition

Using the metacognition cluster groups generated from research question one, this study analyzed learner problem-solving process pattern, including the tool use frequency, duration, and room visit sequences based on metacognition. The three groups were generated using k-means cluster analysis, including high (*Mean* = 85.17, *N* = 24), medium (*Mean* = 67.46, *N* = 79), and low metacognitive levels (*Mean* = 51.77, *N* = 56).

Tool use frequency

Previous one-way ANOVA analysis indicated that there were a significantly different on solution scores based on learner metacognition. This study further examined the tool use frequency based on learner metacognition. See Table 21 for learner average tool use frequency in each metacognition group.

Table 21. Tool Use Frequency Based on Metacognition

Tools Name	Metacognition			
	Low	Medium	High	Total
Spectra	11.64	10.82	6.71	29.17
Solar DB	127.46	105.48	97.67	330.61
Probe Design	106.66	106.94	74.33	287.93
Periodic Table	5.27	4.22	4.33	13.82
Notebook	61.80	58.33	53.42	173.55
Missions DB	23.84	29.19	31.08	84.11
Mission Control	85.45	77.30	62.63	225.38
Concepts DB	16.46	15.09	10.29	41.84
Alien DB	64.27	74.97	53.63	192.87
Communication Center	23.43	31.05	21.21	75.69
Total	526.29	513.39	415.29	1454.97

To help better understand the data, Chord Diagram was used to visualize the data in Table 21 (see Figure 19).

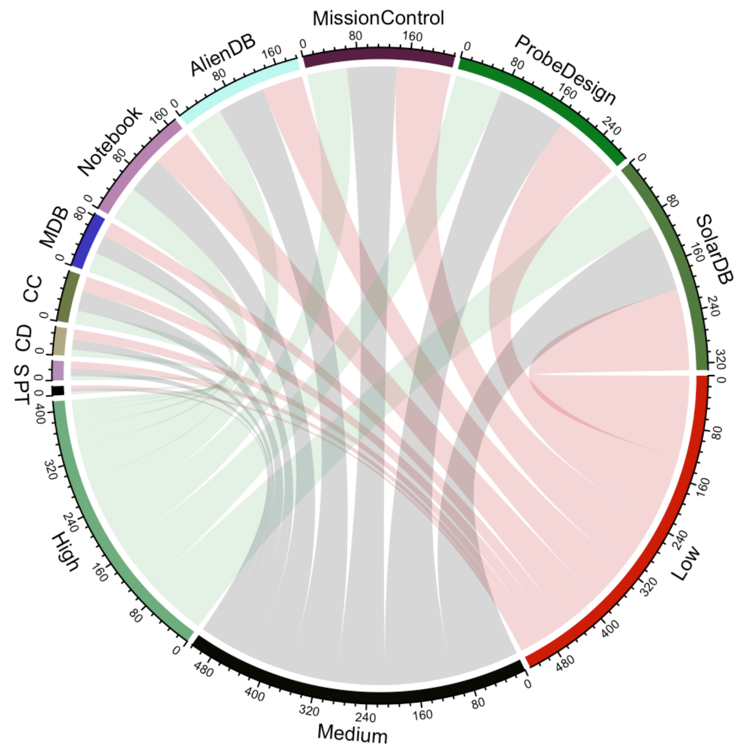


Figure 19. Tool use frequency based on metacognition

This Chord Diagram consists of two major sectors, namely the metacognition sector (bottom) and average tool use frequency sector (top). Learner metacognition was labeled as High, Medium, and Low. Learner average tool use frequency was labeled using the 10 tool names in AR. The width of tool use sector showed the average frequency of use by all learners during the problem-solving process. The width of learner metacognition sector visualized all tool use frequencies during learner problem-solving processes. The thickness of directional links from the learner metacognition sector to the tool frequency sector represented the behavior inclination toward using a tool. Specifically, in Figure 19, the Chord Diagram showed that learners possessing a low metacognition level had higher tool use frequency, followed by learners with medium and high metacognition. Among all learners, Solar System Database, Probe Design, Mission Control, Alien Database, and

Notebook were used the most. Among all 10 tools, it is worth noting that learners in the high metacognition group used nine of the tools less often than learners who had lower metacognition. Missions Database was the only tool that learners in the high metacognition group used more frequently than those who were in the low metacognition group.

Tool use duration

In addition to analyzing tool use frequency based on learner metacognition, this study also examined tool use duration based on learner metacognition (see Table 22).

Table 22. Tool Use Duration Based on Metacognition

Tools Name	Metacognition			
	Low	Medium	High	Total
Spectra	4.50	3.82	2.79	11.11
Solar DB	9.79	8.28	9.92	27.98
Probe Design	14.91	13.37	13.50	41.78
Periodic Table	3.48	4.91	2.42	10.81
Notebook	8.59	6.72	5.96	21.27
Missions DB	3.20	3.62	3.21	10.03
Mission Control	8.64	7.59	6.42	22.65
Concepts DB	3.86	3.73	2.33	9.92
Alien DB	18.57	20.27	20.29	59.13
Communication Center	7.23	7.18	7.50	21.91
Total	82.77	79.49	74.33	236.59

To help better understand the data, the Chord Diagram was also used to visualize tool use duration based on learner metacognition (see Figure 20).

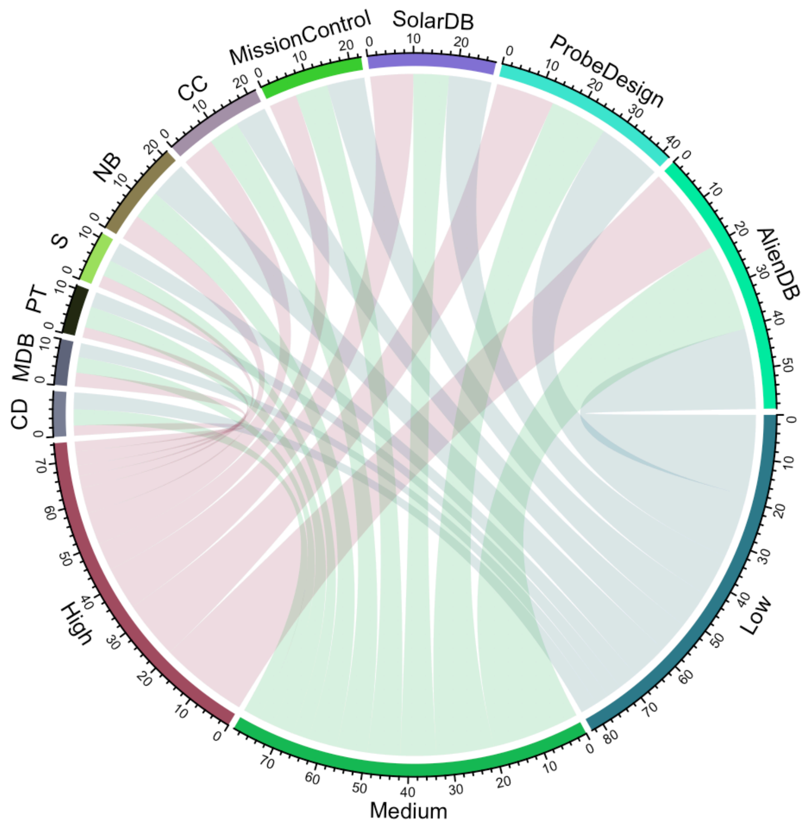


Figure 20. Tool use duration based on metacognition

Considering learner metacognition and tool use duration, data visualization showed that learners in the low metacognition group used all the tools longer (82.77 minutes) compared to learners in the higher metacognition groups (79.49 minutes and 74.33 minutes). In addition, for individual tools, learners in the low metacognition group stayed in Spectra (4.50 minutes), Probe Design (14.91 minutes), Notebook (8.59 minutes), and Mission Control (8.64 minutes) longer than learners in the high metacognition group (Spectra = 2.79, Probe Design = 13.50, Notebook = 5.96, and Mission Control = 6.42 minutes, respectively). However, they had a shorter duration in Alien Database (18.57 minutes) compared to the other two groups (20.27 minutes and 20.29 minutes). They also stayed a slightly shorter time in Solar System Database (9.79 minutes) compared to the high metacognition group (9.92 minutes) but a longer time than students in the medium metacognition group (8.28 minutes). Furthermore, these three groups stayed in Communication Center about the same time—around 7 minutes.

Room visit sequences

Similarity measure was also conducted for learner room visit sequences based on metacognition groups. Particularly, to visualize the differences, three similarity measures (i.e., Cosine, Jaccard, and LCS) were used to compare room visit sequences of learner groups based on their metacognition groups (low, medium, and high). All learner room visit sequences were compared to the highest metacognition learner groups (see Table 23).

Table 23. Similarity Measures of Room Visit Sequences Based on Metacognition

Metacognition	Cosine	Jaccard	LCS	String Length	N (159)
Low	0.28	0.07	0.45	78.00	56
Medium	0.36	0.12	0.45	82.15	79
High	0.36	0.07	0.47	72.50	24

Using a simple line graph in R, this study visualized the similarity among learners (see Figure 21). Based on the visualization, the distances among all learners are similar, which indicated learners in these 3 groups had a similar room visit sequences.

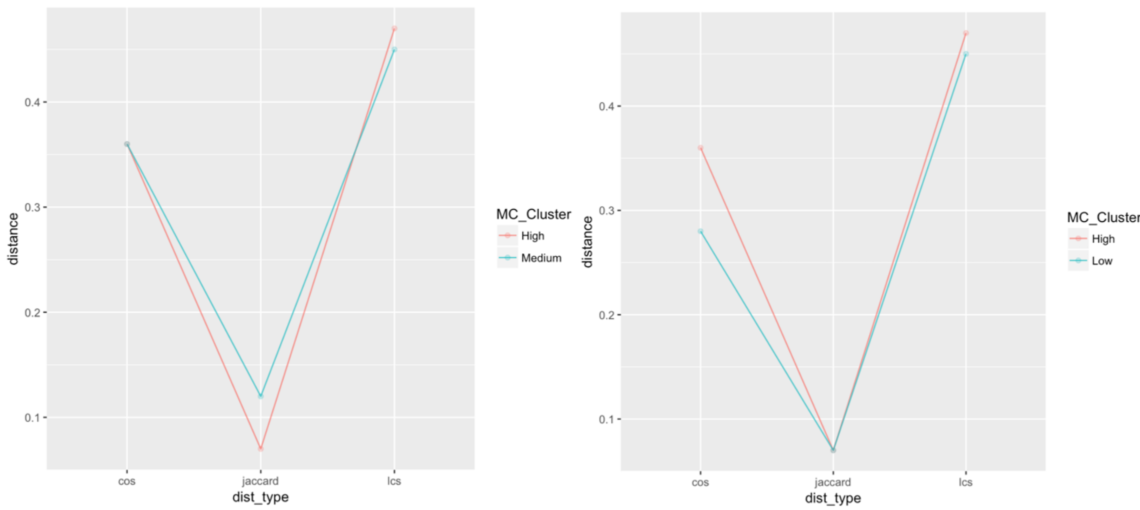


Figure 21. Visualization for similarity measures based on metacognition

2. c. Problem-solving process patterns based on goal orientation

In addition to metacognition, this study analyzed learner problem-solving process patterns, including tool use frequency, duration, and room visit sequences based on the five goal orientation cluster groups generated from research question one. These five groups were generated using k-means cluster analysis including: 1) medium in all six goal orientations (*N* = 41); 2) low in all six goal orientations (*N* = 5); 3) high in all six goal orientations (*N* = 85); 4) high in TAP, TAV, SAP, SAV, but low in OAP and OAV (*N* = 23); and 5) high in TAP and OAP, medium in SAP, but low in TAV, SAV and OAV (*N* = 4).

Tool use frequency

This study examined averages of learner tool use frequency based on the five goal orientation groups. See Table 24 for detailed learner average tool use frequency in each group.

Table 24. Tool Use Frequency Based on Goal Orientation

	Goal Orientation Group					
Tools Name	1($N = 41$)	2($N = 5$)	3($N = 85$)	4($N = 23$)	5($N = 4$)	Total
Spectra	8.73	10.80	10.19	13.26	13.25	56.23
Solar DB	126.51	60.40	102.18	116.22	220.00	625.31
Probe Design	77.95	89.40	106.12	129.13	141.00	543.6
Periodic Table	5.49	4.40	4.35	4.35	3.25	21.84
Notebook	65.66	38.60	55.26	62.83	80.00	302.35
Missions DB	27.76	40.40	24.82	36.96	18.25	148.19
Mission Control	64.41	71.20	80.14	96.30	91.00	403.05
Concepts DB	17.05	20.40	13.75	14.17	13.00	78.37
Alien DB	66.46	81.60	66.42	63.52	113.25	391.25
Communication Center	23.98	27.40	26.60	29.17	45.25	152.4
Total	484.00	444.60	489.84	565.91	738.25	2722.59

To help better understand the data, Chord Diagram was used to visualize the data in Table 24 (see Figure 22).

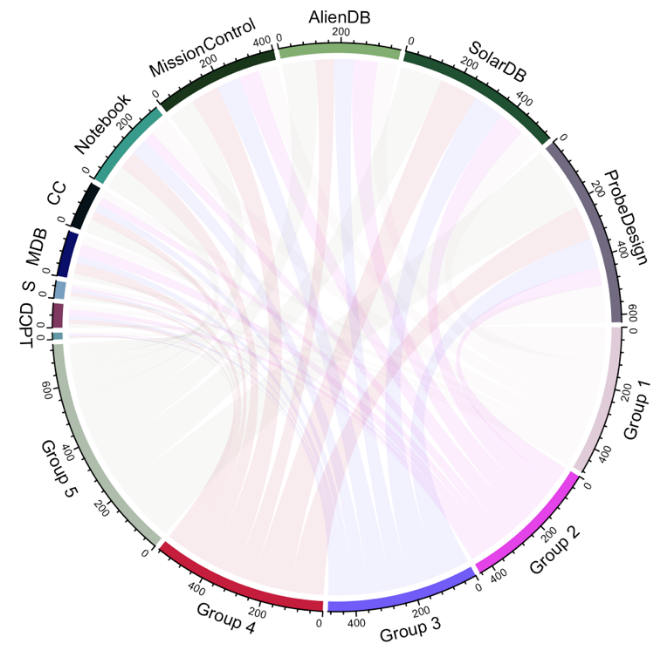


Figure 22. Tool use frequency based on goal orientation

This Chord Diagram consists of a goal orientation sector (bottom) and an average tool use frequency sector (top). Learner goal orientations were labeled from Group 1 to Group 5. Learner average tool use frequencies were labeled using the 10 tool names in AR. The width of tool use sector showed the average frequency of use by all learners during the problem-solving process. The width of learner goal orientation sector visualized all tool use frequency during the learner problem-solving process.

According to the visualization, the Chord Diagram showed that learners in Group 5 (i.e., high in TAP and OAP, medium in SAP, but low in TAV, SAV and OAV) had the highest tool use frequency, followed by Groups 4, 3, and 1. Learners in Group 2 (i.e., low in all six goal orientations) had the lowest tool use frequency. Particularly, learners in Group 5 visited Solar System Database, Probe Design, Notebook, and Alien Database most frequently, while learners in Group 1 (i.e., medium in all six goal orientations) visited

Periodic Table the most. Learners in Group 2 visited Missions Database and Concepts Database the most. Group 4 learners (i.e., high in TAP, TAV, SAP, SAV, but low in OAP and OAV) visited Spectra and Mission Control the most.

Tool use duration

Besides analyzing tool use frequency based on goal orientation, this study also examined learner tool use duration (see Table 25). The same as previously analyzed, the data was based on the average tool use duration in minutes of each goal orientation group.

Table 25. Tool Use Duration Based on Goal Orientation

	Goal Orientation Group					
Tools Name	1($N = 41$)	2($N = 5$)	3($N = 85$)	4($N = 23$)	5($N = 4$)	Total
Spectra	3.41	4.40	4.12	3.57	5.25	20.75
Solar DB	9.22	4.60	8.64	11.39	9.50	43.35
Probe Design	12.20	9.40	14.11	17.78	14.50	67.99
Periodic Table	5.12	6.40	3.62	3.65	1.50	20.29
Notebook	7.29	4.40	7.12	8.17	9.75	36.73
Missions DB	3.17	2.20	3.39	4.43	2.25	15.44
Mission Control	8.66	5.80	7.12	9.09	9.50	40.17
Concepts DB	4.46	2.80	3.08	2.87	9.50	22.71
Alien DB	19.12	15.00	20.39	17.91	24.00	96.42
Communication Center	7.68	9.40	6.67	7.91	9.50	41.16
Total	80.34	64.40	78.25	86.78	95.25	405.02

To help better understand the data, the Chord Diagram was used to visualize tool use duration based on goal orientation (see Figure 23).

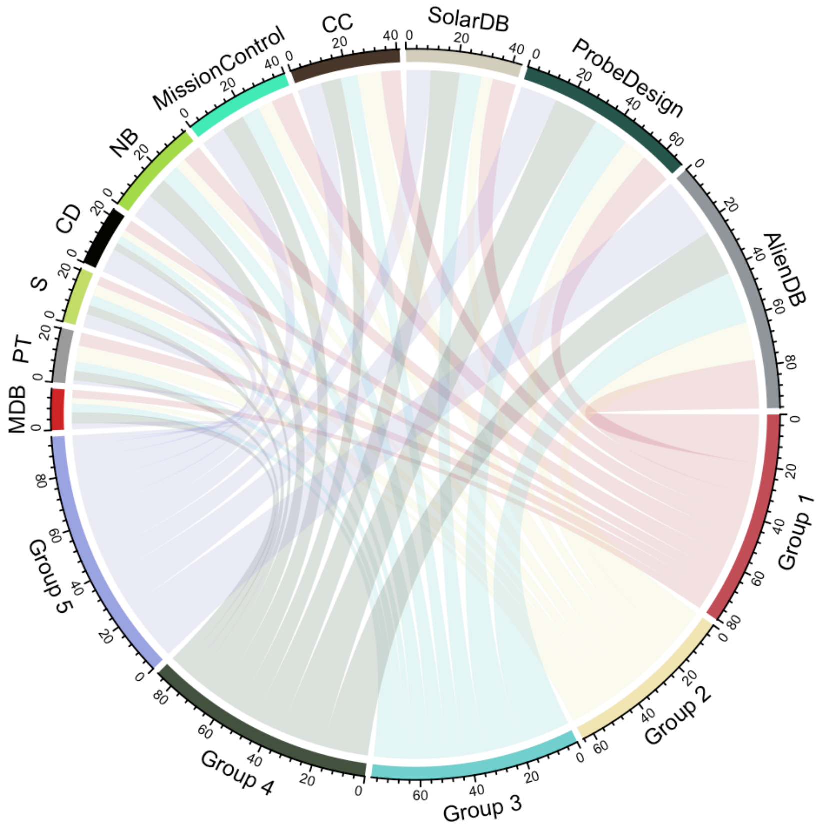


Figure 23. Tool use duration based on goal orientation

The above data visualization of learner tool use duration based on goal orientation showed that learners in Group 5 (i.e., high in TAP and OAP, medium in SAP, but low in TAV, SAV and OAV) used all the tools longer (95.25 minutes) compared to learners in the other four groups; and learners in Group 2 (i.e., low in all six goal orientations) used all the tools for the shortest duration (64.40 minutes). In addition, learners in Group 5 used Spectra, Notebook, Mission Control, Concepts Database, Alien Database, and Communication Center longer compared to the other four groups. Group 4 learners (i.e., high in TAP, TAV, SAP, SAV, but low in OAP and OAV) used Probe Design, Solar System Database, and Missions Database the longest compared to other learners. Group 2 (i.e., low in all six goal orientations) used Period Table the longest compared to their peers.

Room visit sequences

Similarity measure was also conducted for learner room visit sequences based on goal orientation groups. Specifically, all learner room visit sequences were compared to the highest performance learner group. See Table 26 for the average similarity measure coefficients using three methods (i.e., LCS, Cosine, and Jaccard).

Table 26. Similarity Measures Based on Goal Orientation

Goal Orientation Group	Cosine	Jaccard	LCS	String Length
1	0.38	0.12	0.44	78.22
2	0.32	0.09	0.39	86.40
3	0.35	0.09	0.42	79.07
4	0.27	0.09	0.41	82.13
5	0.14	0.12	0.47	79.00

Using simple line graphs in R, this study visualized the similarity among learners based on their distance relative to learners in Group 1 (see Figure 24). Based on the visualization, the room visit sequences were slightly different between Group 1 and Group 3, but there were bigger differences between Group 1 and Group 2. Cosine measure also indicated there were bigger differences between Group 1 and Group 4, as well as Group 1 and Group 5.

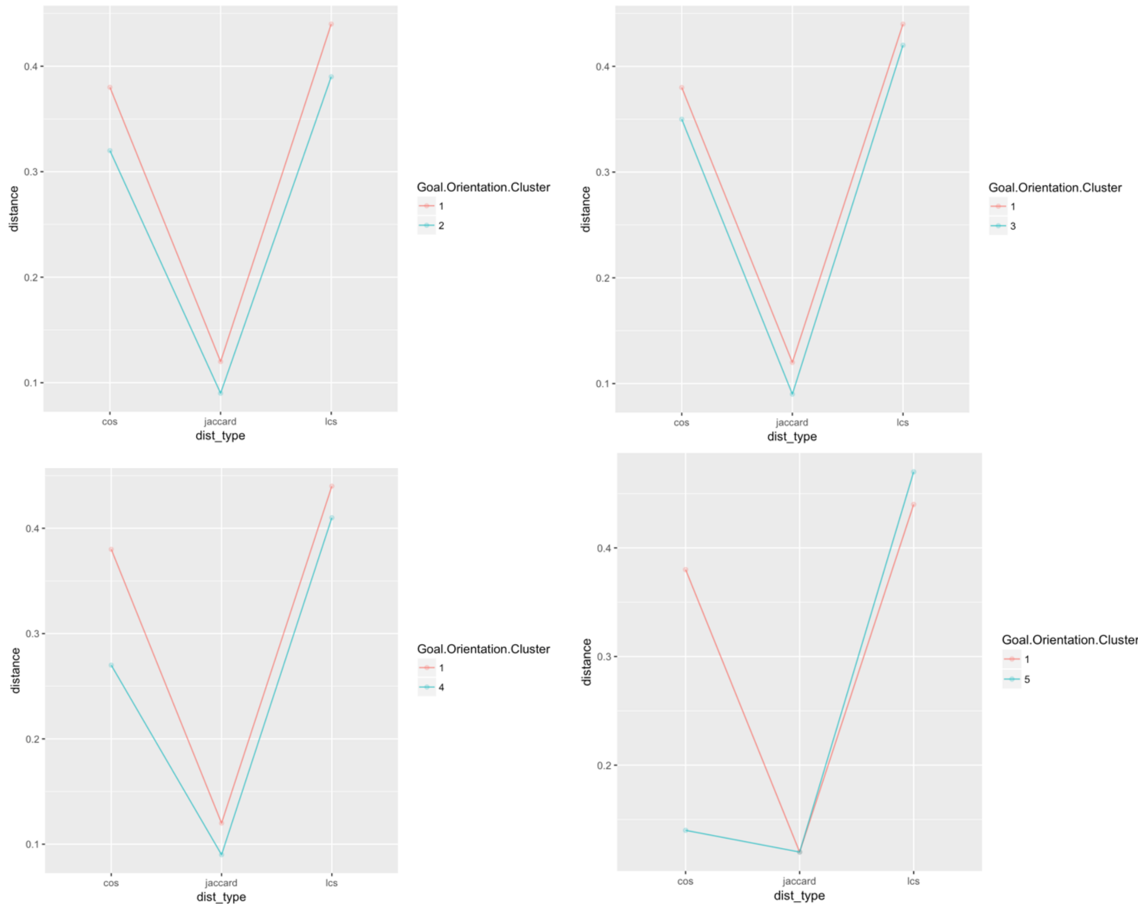


Figure 24. Visualization for similarity measures based on goal orientation

To further investigate whether there were significant differences between groups, two-proportion z-test were used in this study. Specifically, the coefficient score of Group 1 was used to compare with the other 4 groups. The results indicated that there were no significantly differences on learner room visit sequences between learners in Group 1 and Group 2, 3, 4. There were also no significantly different on learner room visit sequences between Group 5 and Group 1 using LCS and Jaccard methods. However, Cosine method suggested that Group 5 and Group 1 learners had significantly different room visit sequences, see Table 27 for the two-proportion z-test results.

Table 27. Two-proportion z-test of Similarity Measures Based on Goal Orientation

Goal Orientation Group	Z_lcs	Z_cosine	Z_jaccard
2	0.72	0.89	0.69
3	0.29	0.44	0.69
4	0.43	1.66	0.69
5	0.42	3.87***	0

Note: ***$p < 0.001$.

2. d. Problem-solving process patterns based on the interation between metacogniton and goal orientaion

This study further analyzed learner problem-solving process patterns based on the interaction between learner metacognition and goal orientation. Particularly, learner tool use frequency, duration and room visit sequences were analyzed based on the three cluster groups generated from research question one, including: 1) high metacognition and high multiple goal orientations group, 2) low metacognition and medium multiple goal orientations group, and 3) medium metacognition and low multiple goal orientations group.

Tool use frequency

This study examined the average of learner tool use frequencies based on the three cluster groups. See Table 28 for detailed learner average tool use frequency in each group.

Table 28. Tool Use Frequency Based on Final Cluster

	Final Cluster Group			
Tools Name	1(N = 61)	2(N = 51)	3(N = 46)	Total
Spectra	9.57	12.16	9.30	31.03
Solar DB	91.49	125.71	122.33	339.53
Probe Design	102.95	122.47	79.37	304.79
Periodic Table	3.61	4.69	5.67	13.97
Notebook	55.57	57.53	63.98	177.08
Missions DB	32.46	18.67	30.22	81.35
Mission Control	71.54	98.18	65.43	235.15
Concepts DB	12.57	14.67	17.72	44.96
Alien DB	68.05	65.65	68.65	202.35
Communication Center	29.84	24.86	24.74	79.44
Total	477.66	544.57	487.41	1509.64

To help better understand the data, Chord Diagram was used to visualize the data in Table 28. This Chord Diagram consists of three cluster sectors (bottom) and average tool use frequency sector (top). Cluster groups were labeled from Cluster 1 to Cluster 3. Learner average tool use frequencies were labeled using the 10 tool names in AR (see Figure 25). According to the visualization, the Chord Diagram showed that learners in Cluster 2 (i.e., low metacognition and medium multiple goal orientations) had the highest tool use frequency (i.e., 544.57 times), followed by Cluster 3 (i.e., medium metacognition and low multiple goal orientations, 487.41 times) and 1 (high metacognition and high multiple goal orientations, 477.66 times).

In addition, learners in Cluster 2 used Probe Design, and Mission Control tools the most; learners in Cluster 3 used Periodic Table, Notebook, Concepts Database, and Alien

Database tools most frequently; and learners in Cluster 1 used Missions Database and Communication Center most frequently.

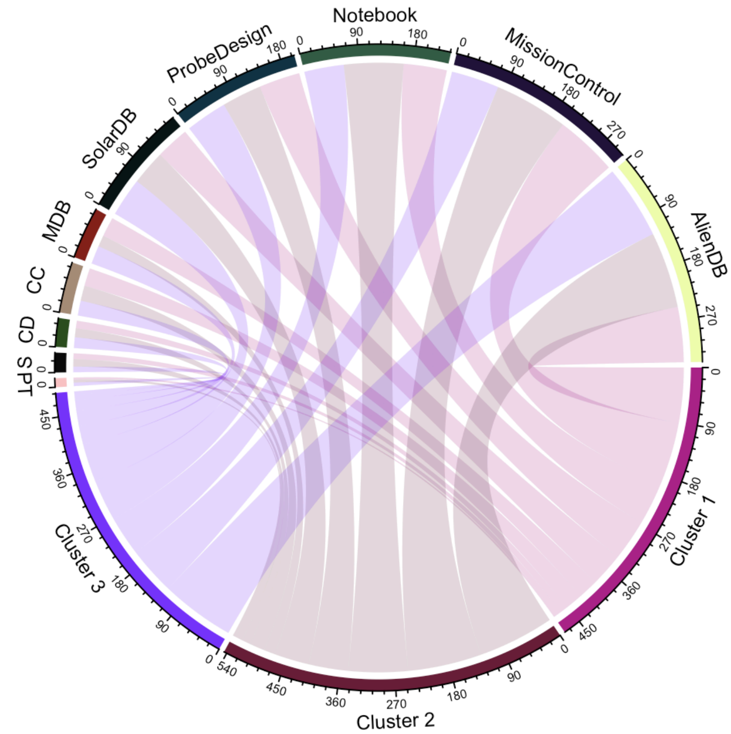


Figure 25. Tool use frequency based on final cluster

Tool use duration

Beside analyzing tool use frequency based on the interaction of learner metacognition and goal orientation cluster, this study also examined learner tool use duration (see Table 29). Just as previously analyzed, the data was based on the average tool use duration minutes of each cluster group.

Table 29. Tool Use Duration Based on Final Cluster

	Final Cluster Group			
Tools Name	1(N = 61)	2(N = 51)	3(N = 46)	Total
Spectra	3.67	4.49	3.52	11.68
Solar DB	8.93	9.59	8.72	27.24
Probe Design	13.44	16.59	11.89	41.92
Periodic Table	3.67	3.41	5.26	12.34
Notebook	6.64	8.37	6.98	21.99
Missions DB	3.89	3.18	3.07	10.14
Mission Control	6.59	8.82	8.35	23.76
Concepts DB	3.18	3.37	4.28	10.83
Alien DB	20.84	19.02	18.67	58.53
Communication Center	6.87	7.22	7.87	21.96
Total	77.72	84.06	78.61	240.39

To help better understand the data, Chord Diagram was used to visualize tool use duration based on final cluster (see Figure 26). According to the data visualization of learner tool use duration based on final cluster, learners in Cluster 2 (i.e., low metacognition and medium multiple goal orientations) had the longest tool use duration (i.e., 84.06 minutes), while learners in Cluster 1 (i.e., high metacognition and high multiple goal orientations) had the shortest tool use duration (i.e., 77.72 minutes). In addition, learners in Cluster 2 stayed in Spectra, Solar System Database, Probe Design, Notebook, and Mission Control longer compared to the other four groups. Cluster 3 learners (i.e., medium metacognition and low multiple goal orientations) stayed in Periodic Table, Mission Control, Concepts Database, and Communication Center the longest compared to other groups. Cluster 1 learners stayed in Missions Database and Alien Database the longest relative to their peers.

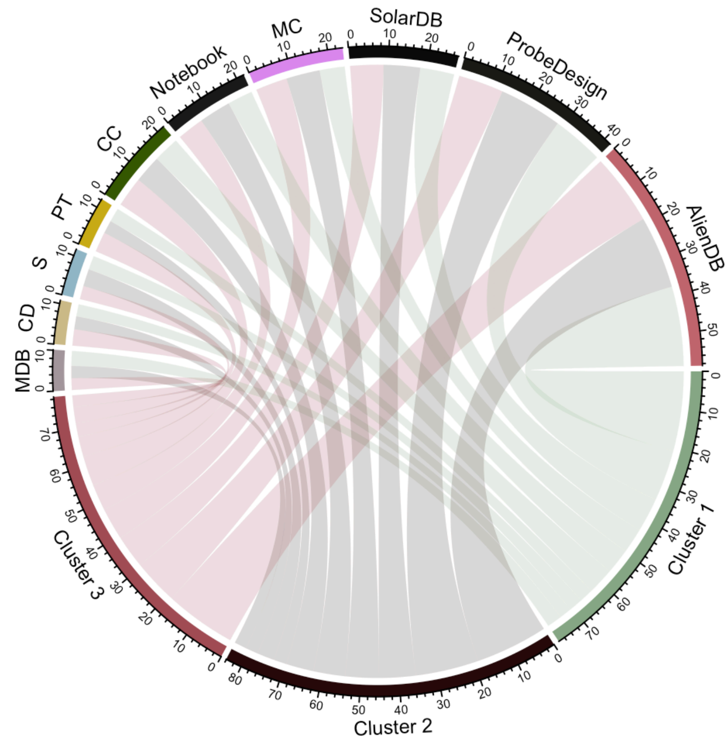


Figure 26. Tool use duration based on final cluster

Room visit sequences

Similarity measure was also conducted for learner room visit sequences based on the interaction between learner metacognition and goal orientation. Specifically, all learner room visit sequences were compared to the highest performance learner group in cluster 3. See Table 30 for the average similarity measure coefficients using three methods (i.e., LCS, Cosine, and Jaccard).

Table 30. Similarity Measures Based on Final Cluster

Final Cluster Group	Cosine	Jaccard	LCS	String Length
1	0.39	0.11	0.43	81.36
2	0.25	0.07	0.41	77.71
3	0.37	0.12	0.43	79.11

Using simple line graphs in R, this study visualized the similarity among learners based on similarity coefficient values (see Figure 27). Based on the visualization, the room visit sequences were very similar between Group 1 and Group 3, but there were bigger differences between Group 2 and Group 3.

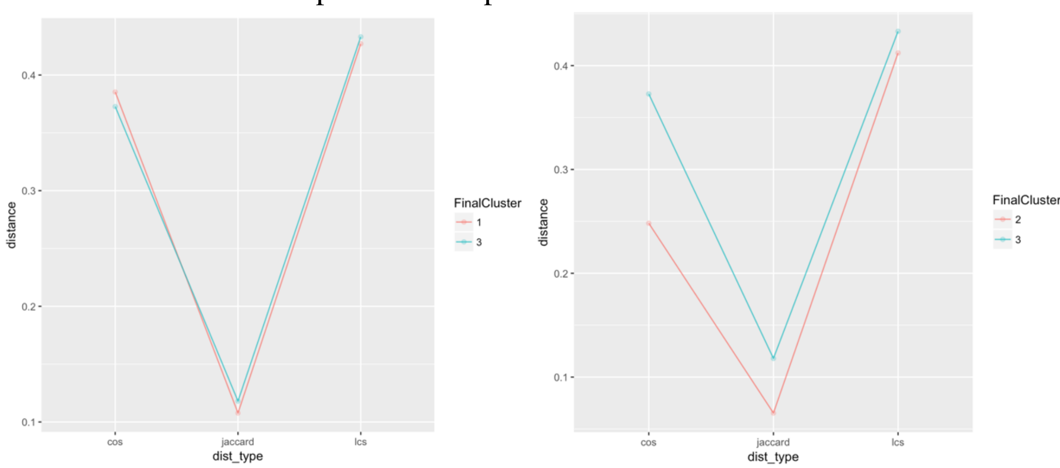


Figure 27. Visualization for similarity measures based on final cluster

To further investigate whether there were significant differences between groups, two-proportion z-test was used in this study. Specifically, the Cluster 3 group was used to compare with the other two cluster groups. Results indicated that there were no significant differences in learner room visit sequences among learners in Cluster 3 and Cluster 1. There were also no differences between learners in Cluster 3 and Cluster 1. See Table 31 for the two-proportion z-test results.

Table 31. Two-proportion z-test of Similarity Measures Based on Final Cluster Group

Final Cluster Group	Z_lcs	Z_cosine	Z_jaccard
1	0	0.29	0.99
2	0.29	1.83	0

2. e. Reasons for the learner problem-solving process differences

Previous AR studies suggested that a successful problem-solver will go through all four conceptual stages while solving a problem in AR: (a) understanding the problem, (b) identifying, gathering, and organizing information, (c) integrating information, and (d) evaluating the process and outcome (Liu et al., 2004). In this study, learner problem-solving processes were analyzed based on learner behavior patterns in AR, including tool use frequency, tool use duration, and room visit sequences.

To understand the reasons for problem-solving process differences among learners based on metacognition and goal orientation, 12 selected participant interviews were analyzed using grounded theory. Similarly as the previous section, these participants were grouped into four groups based on the final cluster analysis result, including: 1) high metacognition, high multiple goal orientations, and low performance group, including Beta, Gamma, and Zeta; 2) low metacognition, medium multiple goal orientations, and high performance group, including Epsilon, Kappa, and Mu; 3) medium metacognition, low multiple goal orientations, and high performance group, including Delta and Theta; and 4) outliers, including Alpha, Chi, Lota and Omega.

Based on 12 stimulated recall interviews, the afore mentioned four conceptual stages were integrated with 10 steps that a successful problem-solver usually performs during this study, including: 1) identify the problem correctly; 2) explore the 3D environment by visiting all rooms in AR and look over all tools; 3) discover what Jakala-

Tay needs to survive in Alien Database; 4) search the Solar System Database for possible planets; 5) develop hypotheses about where Jakala-Tay can live; 6) figure out if there is any missing information needed for making a decision; 7) launch probes to gather the information in the Probe Design room; 8) check the data from the probe in the Mission Control room; 9) decide whether the selected planet is a good choice for Jakala-Tay; 10) if so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4 (see Figure 28). In addition, the example quotes from participants describing these 10 steps are presented in the codebook. See Appendix C for the codebook.

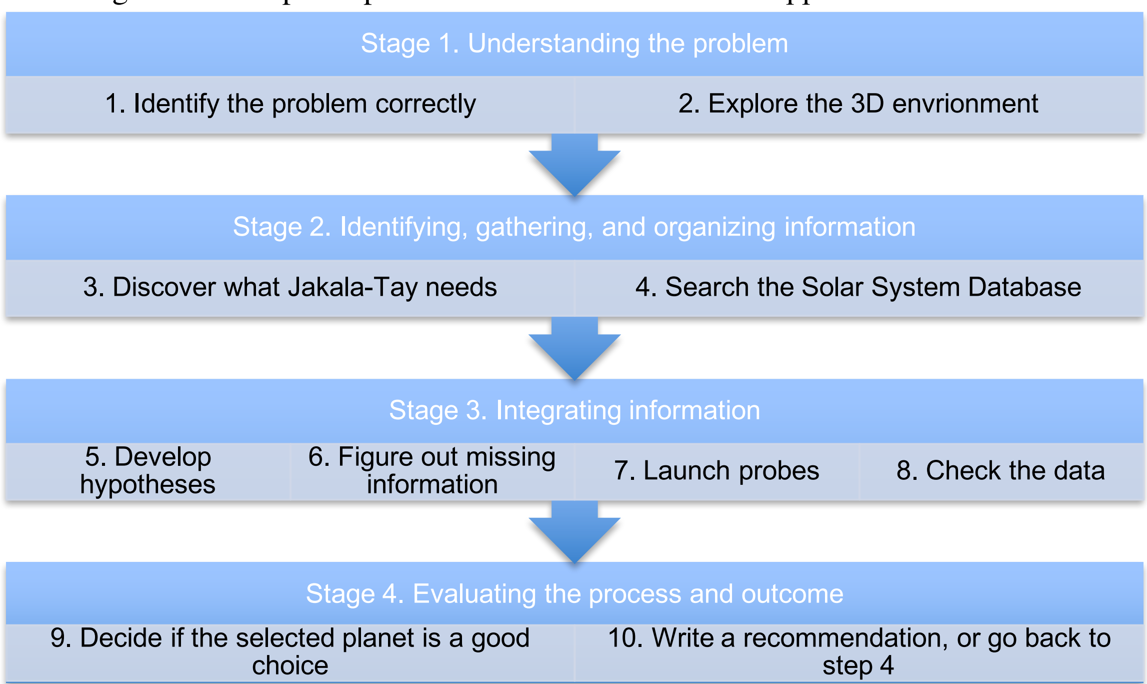


Figure 28. Problem-solving processes of successful learners in AR

Group 1: Self-correction matters

Based on previous findings, learners in Group 1 had a higher metacognition and goal orientations compared to the other two clusters, but they all failed at a very early stage—understanding the problem—during the problem-solving process.

According to Beta, he started AR by exploring the environment, "I wanted to look around first, to see where I was—to see what I can do, or where I would go." Then he started to identify the problem, "I didn't know at first if we were with, like, focusing on a specific alien"; meanwhile, he started to form a hypothesis even though he had not identified the problem yet. "I was, like, okay, well, in my mind, I was, like, the closest thing to Earth is Mars. That's, like, just my growing up and hearing different things about how, like, they sent the rover to Mars"; then he made a quick decision; "So I was, like, okay… I can send them there," which led to a score of 0 for his solution score.

As for Zeta, she sent a recommendation without having conducted any research, and her reason was, "I just, like… for any the alien, like to live on a planet…, I don't know there was, like, specifically features yet. So just, like, a basic idea, I just, like, oh… Mars, cause mars is, like, maybe be able to have life." In addition, she was reading about Akona for more than 20 minutes when she finally realized the problem was to help Jakala-Tay, which made her fall behind. Eventually, she was not able to go through the Solar System Database and chose Mars as the destination for Jakala-Tay.

Both Beta and Zeta thought they possessed what Flavell and Wellman (1977) called tasks-metacognitive knowledge. Specifically, they thought they had knowledge about Mars, which turned out to hinder their problem-solving during the study. Beta's quick decision also demonstrated he did not really have a high metacognition knowledge, because he was not able to do things or use strategies during the problem-solving (Weinstein & Mayer, 1986; Zimmerman, 2002). Zeta's poor time management indicated she had low metacognitive level, because she failed to focus on the task while tracking her own performance (Zimmerman & Campillo, 2003). Therefore, both Beta and Zeta might have overestimated their metacognitive levels during the study (Kruger & Dunning, 1999).

Gamma was at the same group with Beta and Zeta, and he also failed to identify the problem in the first place—reading about Akona for about 10 minutes. However, he had a high self-approach goal orientation, and self-corrected at an early stage. He efficiently skipped steps 6 and 7 regarding sending probes and read the data to make a correct selection. He suggested that, "I was probably gonna send a probe, but when I got the information on IO, I thought there are not many options for me except IO." Therefore, although Gamma only successfully finished step 1 through 5 and step 8 through 9, he still scored 5 points during the study. In addition, Gamma's high metacognitive skills might also help him identify and overcome obstacles that may impede progress. Specifically, by monitoring his progress in reaching a solution, he adjusted his plan and strategies if needed to successfully solve the problem (Davidson & Sternberg, 1998).

It is worth noting that both Beta and Gamma mentioned that AR crashed once during the study. Beta indicated the glitch happened in the Probe Design room, "Yeah, it crushed, so I had to do it [Probe Design] again," while Gamma got stuck in the Communication Center while writing the recommendation.

Group 2: Finished all 10 steps

Epsilon, Kappa and Mu were all considered as having low metacognition, medium goal orientation, and high solution scores. In addition, all three participants finished all 10 steps during the problem-solving in AR, and had higher scores than learners in Group 1. Since all three participants had similar problem-solving processes, the following paragraphs will describe the detailed steps of one learner.

For example, Epsilon—who had a high SAP score—started by exploring the environment. "I was trying to figure out how the game works first of all." She also wanted to "see what I was… what I had, what I was working with, what I am supposed to be doing." These positive actions were consistent with literatures in that SAP can positively predict learner energy and emotion in class (Brondino et al., 2014; Elliot et al., 2011). Due to these positive energy, Epsilon was not affected by the unfamiliar game control in AR, "I refreshed because I got stuck in the door way. I was like I need to go straighter."

After exploration, she identified the problem, and started step 3 and 4,

Oh, mostly, I use the notebook a lot and then I think I also use the solar system one a lot, cause first I went to the aliens, figured out, like, learned about what they could and couldn't do without dying, and then I went to the… and then I went to the solar systems, and just staying there.

She then came up with the hypothesis where Jakala-Tay could live (i.e., step 5) using the notes she took during step 3 and step 4 as follows: "because I knew if something was, like, had too heavy of a gravity, they would just, like, die"; and she had two hypotheses: "I think I went to Venus and IO; I want to say either that or Titan and IO, because there wasn't as much information about them." Then she also went through step 6,

I went back, because I was, like…, hmm…, let me see if there's anything that I forgot about them, like, I wanted to double-check with the notes I'd taken it…, cross the… uh… I'd like to see it compared to the planet, you know.

After that, she started step 7 and 8,

Yeah, I just, like, I ruled out that everything that wasn't compatible, and then I sent a probe to the ones that, like, didn't right away get scratched off, and I looked at them, and then I just kind of like chose from there, that one I thought would be.

She was also being very cautious in conducting step 9. "I think… yeah… I realize they had a good gravity, and then I went ahead and reviewed it one last time to see if there's anything I missed." Finally, she made the decision in Step 10. Epsilon gained 4 points for the solution she submitted. Different from other learners in Group 1, who made an easy decision at the early stage of problem-solving, Epsilon did not decide until the last step of the problem-solving. Based on the interview, she also demonstrated self-reflection behavior and a deeper understanding for herself—"my notes I took at the beginning weren't very good, I just wanted to [take more notes], so I was taking a more active stuff later." When being asked about the reason for doing so, her response was, "so that way… um… I can make a decision. It sounded a very indecisive person", which was consistent with Ning (2016)'s study on SAP that learners had a higher SAP usually had a deeper understanding for oneself.

Group 3: Solving the problem efficiently

Delta and Theta also had a higher score compared to learners in cluster 1. Theta even earned all 7 points for his answer. Different from learners in cluster 2, Delta and Theta had a higher metacognition score and lower goal orientation scores. In addition, they did not go through all 10 steps, but skipped some steps, which made them solve the problem more efficiently.

For example, Delta almost skipped the step 6 and 7 regarding sending the probe and reading about information in mission control, and he stated that,

I think I was, like, a little bit, like, hesitant to, like, start sending out probes, cause I guess was just… I mean I knew that, like, when I noticed that the budget was, like, kind of insanely high, I was kind of, like, I just don't want to waste too much time because, like… I don't know… it was a weird, like… I did as if there was real kind of things, so I didn't want to, like, you know, having wasted insane amount of money just, like… sending all these probes.

He did not "want to waste too much time", and he also admitted that "I was more for efficiency". He further explained that, *There's maybe game but, like, in terms of, you know, real applications, you won't have that luxury to be able to, like, you know, send in another thing; you kind of feel like there's much you… Especially for something serious as, like space exploration; you can't simply, just… you know… call it back.*

Then he did send three probes after he realized he had the luxury of time and money. In addition, he also demonstrated problem-solving efficiency during other steps. For example, he quickly skipped the redundant information in Mission Database, because he thought this tool was "not really relevant to the task."

As for Theta, he completely skipped step 6 and 7. He mentioned that "[I] tried, and that's it." So, he "didn't send any probes out." By skipping these two steps, he had more time to focus on step 3, 4, and 5 to learn more about Jakala-Tay and planets and develop hypotheses. For example, he could "double-check everything," and he suggested that "it was kind of intriguing saying I think this is because I initially had two options and I ended up going with a different thing than my initial plan." He then described the whole process during the interview:

Originally, when I read about their home world using lots of sulfur, I immediately thought or I checked this whole system. It was Venus and Saturn Titan. So those are what I was going for, but Venus was a little too hot and it didn't have all the correct composition of things, and it was a little too similar in size to earth. Saturn's Titan—the ground seemed muddy, which wouldn't be good composition-wise for building things, and there wasn't a sense of the other necessary requirements for Jakalay-Tay to survive, besides the sulfur and atmosphere, really. Then I came across IO, which had the same things, but it also had the same color scheme that was reminiscent and lots of volcanic activity, which they're used to, which would lead to nitrogen, which allows they need to use the nitrogen for the gases or growing the plants. So it had the two main gases they needed, the rocky *composition, also the temperature highs and lows were similar, and it was smaller than Earth—closer in size to our moon, in fact.*

He admitted that "I was thinking it [Sending probes] could be useful to test my theories, but I never sent any, so I could be totally wrong." However, he finally decided to send Jakala-Tay to IO, and he stated that "I just went through a pretty logical way of going about it—just checking the facts," which turned out to work in this study, and he got the highest score.

In summary, Group 3 participants behaved exactly like what Gourgey (1998) described as "effective learner" who "seek to understand concepts and relationships, monitor their understanding, and choose and evaluate their actions based on whether the actions are leading toward their goals" (p. 89). Particularly, take Delta as an example, he understood the time constrain of the study—"I just don't want to waste too much time", then he also had the concept of space exploration that it is irreversible—"you can't simply, just…you know…call it back.", so he decided not to send the probe at the beginning. After he realized he had the time and money to send the probe, he chose a different action based on the situation, and eventually sent three probes.

Group 4: Outliers

There were four outliers based on the multiple regression analysis result, including Alpha, Chi, Lota, and Omega. Alpha and Chi failed to solve the problem because they neglected important information during step 3 and step 4 regarding learning about Jakala-Tay and the solar system. Lota and Omega got the high scores because they followed all 10 steps during the problem-solving process, and utilized 60 minutes fully during the study.

Alpha and Chi both had a medium metacognition score and considered themselves to have higher TAP and TAV goal orientations compared to other goal orientations. For

Alpha, she only selected the temperature factor for Jakala-Tay and the planets during the step 3 and step 4, because “temperature was one thing I was, like, sure of.” Therefore, she neglected all the other important information, such as atmosphere and gravity, solely focused on the task to locate a planet that had an appropriate temperature, which led her to choose an inappropriate planet for the alien. As for Chi, he solely focused on finding a planet that had an appropriate atmosphere, and he admitted that

So that's the big thing that I guess my mind just focused so much on. If they even breathe a little bit of—I think it was hydrogen—even if they breathe a little bit of, it would die, and so, I guess my mind just kind of focused on, okay, gases.

Both Alpha and Chi are more task goal-orientated. Once they focused on the “wrong” task, it was difficult for them to spare resources towards other aspects during problem-solving.

Different from Alpha and Chi, both Lota and Omega had higher metacognition scores. Lota reported that she had a higher SAV and OAV, while Omega considered that he had a higher TAP, TAV and OAV. Lota and Omega were the perfect problem-solvers in AR, because they followed all 10 steps during problem solving and used up all 60 minutes the study provided (see Table 32).

Table 32. 10 Problem-solving Steps of Lota and Omega

Steps	Lota and Omega
1. Identify the problem correctly	Lota: Well, first I went into the alien room, and I've read their whole intro, like, why they came to us, but then I saw that there were breakdown of the species, and I found those species that I was looking for—Jakala-Tay. Omega: I'm just find the information I need to find [about Jakala-Tay]

Table 32 continued.

Steps	Lota and Omega
2. Explore the 3D environment by visiting all rooms in AR and look over all tools	Lota: I felt like it was pretty straightforward once you walked around and saw what each room was for and had the main room, and then you had the room where you got the information about the aliens, and then you know you could send a probe if you needed to or design a probe, and then go send a probe. I feel like it was very clear what each room was for. Omega: Well, when I first came to this point I was just kind of seeing what's here, so I just started reading through everything. I want to see what they're—what they're trying to do.
3. Discover what Jakala-Tay needs to survive in Alien Database	Lota: Well, first I went into the alien room, and I've read their whole intro like why they came to us, but then I saw that there were breakdown of the species, and I found those species that I was looking for—Jakala-Tay. And I read all of their information. You know, what they needed in the atmosphere. They needed nitrogen to be able to grow their plants that they brought with them; definitely could not have hydrogen, because it's deadly to them, and they live very well in an atmosphere that has sulfur. So then I wrote all that down in my notebook. Omega: [I read] the temperature and then their prior planet, because, I mean, that's the whole; you want to have something that's very similar to their planet they lived on.
4. Search the Solar System Database for possible planets	Lota: And I started looking at all of our solar system, looked all planets first, process of elimination; and then I looked at the moons; one of the planets had sulfur, and one other had sulfur, and then on the moon. Is it IO?

Table 32 continued.

Steps	Lota and Omega
	Omega: So there, you know, I try to find a planet that's similar to their old planet and fits their breathing and, like, their life support. So I'm trying to go over just some moons and planets that I think are good.
5. Develop hypotheses about where Jakala-Tay can live	Lota: I saw that it [IO] had volcanic eruptions, which is OK for them, because they have volcanic eruptions on planet—on their home planet. They experienced that, so it is not like it would be anything very different to them to experience volcano eruptions; and then it didn't have hydrogen, so it already met two of my requirements, you know; and so it just needed a little bit of nitrogen, so they can maybe grow their plant life.
	Omega: I knew it was gonna be IO or Venus, one of those two.
6. Figure out if there is any missing information needed for making a decision	Lota: I already felt like I knew the answer, but I went… I hadn't used the other two rooms yet, you know?
	Omega: I mean, the thing was, I wasn't really limited on money, so I started just buying stuff, so, say, okay, let me just send three dedicated [probes]. So I'm just gonna buy everything that is important.
7. Launch probes to gather the information in the Probe Design room	Lota: So, I designed a lander probe to go to IO and collect the spectroscopy… spectroscopy of it.
	Omega: The first place to send is Venus, because…, so they started saying a lot of stuff about volcanoes and it's hot, a lot of the similarities, especially when they said sulfur.
8. Check the data from the probe in the Mission Control room	Lota: And then when it came back it indeed did have ten percent nitrogen.

Table 32 continued.

Steps	Lota and Omega
	Omega: Well, I was like oh let me go see cause that's the only place left, and let me go see what these—if I can see what the probes [returned]
9. Decide whether the selected planet is a good choice for Jakala-Tay	Lota: So I just felt like that was a… that's the correct answer;… it had all three, yeah, that mentioned in their breakdown, and they could have volcanic eruptions, but it might even feel kind of normal to them to be doing that same kind of living environment. Omega: I thought IO was a good second choice, but the problem with IO was that it didn't have nitrogen for their food, but everything else was fairly decent; the temperatures, if I remember… were a little bit low, there was nothing that was perfect.
10. If so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4.	Lota: After doing some research, I had reason to believe the atmosphere if IO would be suitable for Jakala-Tay. It frequently has volcanic eruptions that shoots out sulfur, putting sulfur in the air. I then sent a land probe to IO and evaluated the percentages in the atmosphere. It has no hydrogen, which is perfect, since hydrogen is deadly to Jakala-Tay. It also has sulfur, which is ideal, and is made up of 10% nitrogen which means, with some alterations, they may be able to grow the plant life from their home planet. The volcanic eruptions and earthquakes make IO scary to us, but Jakala-Tay are used to this from home. I believe that IO is the best option. Omega: Fits most of the specifications of gases temperature and seismic activity.

Despite the similar characteristics and problem-solving processes Lota and Omega had, they had different solutions scores—Lota wrote a comprehensive answer and earned full scores (i.e., 7 points), while Omega wrote a much shorter answer and only earned 4 points.

Chapter 5: Summary and Discussion

Empiricism is the act of making decisions based on what is.

- Ken Schwaber, SCRUM Development Inventor

The goal of this study was to examine the impact of learner characteristics (i.e., metacognition and goal orientation) on learner problem-solving (i.e., problem-solving performances and processes) in a SG environment that adopts PBL pedagogy. Using AR as the SG environment, this study employed a sequential mixed research design, SGA, and multiple data sources to analyze 159 undergraduate learners' metacognition, goal orientations, and problem-solving performances and processes in a laboratory setting. This chapter will summarize the results as presented in Chapter 4. Discussion of findings, study implications, limitations, and future research directions will also be presented.

There are two research questions in the study:

1. To what extent are problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)? There are three sub-questions:

(a) Is there a statistically significant difference in learner problem-solving performances based upon metacognition (high, medium, and low metacognitive levels) and goal orientation (task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goal orientation)?

(b) Can learner metacognition and goal orientation predict problem-solving performances?

(c) What are the reasons for any problem-solving performance differences based on learner characteristics (i.e., metacognition and goal orientation)?

2. To what extent are problem-solving process differences based on learner characteristics (i.e., metacognition and goal orientation)? There are five sub-questions:

(a) What are learner problem-solving process patterns?

(b) Are there any problem-solving process pattern differences among students, based on their metacognition?

(c) Are there any problem-solving process pattern differences among students, based on their goal orientation?

(d) Are there any problem-solving process pattern differences based on the interaction between learner metacognition and goal orientation?

(e) What are the reasons for any problem-solving process pattern differences based on learner characteristics (i.e., metacognition and goal orientation)?

The following sections will present the summary of the results and discussions for each sub-question.

Research Question One

For question (a), cluster analyses in SPSS were used to determine whether there was a statistically significant difference among learner groups based on their metacognition and goal orientations. For question (b), a multiple regression analysis was used to examine whether learner metacognition (MC) and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, and OAV) could predict problem-solving performance differences (i.e., Solution Scores, SS). As for question (c), the grounded theory was used to analyze interview codes and explain the reasons for any possible problem-solving differences among learners.

1.a. Learner performance differences based on learner characteristics

Using cluster analysis, learners were clustered into three groups based on metacognition, including high (*Mean* = 85.17, *N* = 24), medium (*Mean* = 67.46, *N* = 79), and low metacognitive levels (*Mean* = 51.77, N = 56). One-way ANOVA showed that there were statistically significant differences in learner problem-solving performances based on these three groups, $F(2, 156) = 4.848$, $p = .009$, $\eta^2 = .058$; Levene's Test: $F\ (2, 156) = .458$, $p = .633$. Interestingly, learners in the lowest metacognitive level group had the highest solution score (*Mscore* = 3.86), followed by learners in the high metacognitive level group (*Mscore* = 3.08) and medium metacognitive group (*Mscore* = 2.46). The post hoc analysis using Tukey's HSD test indicated that participants in the medium metacognitive group scored significantly lower than those in the low metacognitive group, $p < 0.05$. However, there was no significant difference between the high and low metacognitive level groups, or between high and medium metacognitive groups. In addition, one-way ANCOVA analyses indicated that there were significant effects of learner metacognitive levels on problem-solving performances after controlling for ethnic groups, $F(2, 156) = 4.999$, $p = .008$, $\eta^2 = .061$; Levene's Test: $F\ (2, 156) = .801$, $p = .451$ and gender $F(2, 156) = 4.677$, $p = .011$, $\eta^2 = .057$; Levene's Test: $F\ (2, 156) = .281$, $p = .756$.

As for goal orientation, learners were clustered into five groups, including 1) medium in all six goal orientations (N = 41); 2) low in all six goal orientations (N = 5); 3) high in all six goal orientations (N = 85); 4): high in TAP, TAV, SAP, SAV, but low in OAP and OAV (N = 23); and 5) high in TAP and OAP, Medium in SAP, but low in TAV, SAV and OAV (N = 4). According to one-way ANOVA, there was no significant difference on problem-solving performances based on goal orientation groups among learners, $F(4, 153) = 1.520$, $p = .199$, $\eta^2 = .038$; Levene's Test: $F\ (4, 153) = .884$, $p = .475$.

Considering the impact of both learner metacognition and goal orientation on problem-solving performance, learners were clustered into three groups, including 1) high metacognition and high multiple goal orientations, 2) low metacognition and medium multiple goal orientations, and 3) medium metacognition and low multiple goal orientations. According to one-way ANOVA, learner problem-solving performances were statistically significant based on these three clusters, $F(2, 155) = 11.208$, $p = .000$, $\eta^2 = .126$; Levene's Test: $F\ (2, 155) = .989$, $p = .374$. Specifically, learners in Cluster 2 (i.e., low metacognition and medium multiple goal orientations) and Cluster 3 (i.e., medium metacognition and low multiple goal orientations) had scores two points higher—in an eight-point scale system—compared to learners in Cluster 1. The post hoc analysis using Tukey's HSD test indicated that participants in Cluster 1 scored significantly lower than those in Cluster 2 and 3, $p < 0.05$. In addition, One-way ANCOVA analyses in SPSS indicated that there were significant effects of learner final clusters on problem-solving performances after controlling for ethnic groups, $F(2, 155) = 9.726$, $p = .000$, $\eta^2 = .112$; Levene's Test: $F\ (2, 155) = .711$, $p = .493$ and gender $F(2, 155) = 11.148$, $p = .000$, $\eta^2 = .126$; Levene's Test: $F\ (2, 155) = 1.092$, $p = .338$.

1.b. Can learner characteristics predict problem-solving performance differences?

Multiple regression was conducted in SPSS to identify whether learner metacognition (MC) and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, and OAV) would predict learner problem-solving performances (i.e., SS). Specifically, the regression model is as follows:

$$Y_{ss}=\beta_1 MC+\beta_2 TAP+\beta_3 TAV+ \beta_4 SAP+ \beta_5 SAV+ \beta_6 OAP+ \beta_7 OAV + u$$

The results showed that the model as a whole was significant ($p < .01$), which indicated that learner goal orientation and metacognition were significant predictors for problem-solving performance in this study ($R^2 = .134$, $F(7, 155) = 3.283$, $p < .01$). The regression equation was:

$$Y_{ss}= -.025*MC - 0.34*TAP + 0.252*TAV + 0.006*SAP - 0.229*SAV - 0.003*OAP + 0.065*OAV + u$$

In addition, learner TAP, TAV, and SAV were significant predicators of performance during problem-solving. Specifically, for every point increase in TAP, a 0.34-point decrease in learner problem-solving performance was predicted; for every point increase in TAV, a 0.252-point increase in learner problem-solving performance was predicted; and for every point increase in SAV, a 0.229-point decrease in learner problem-solving performance was predicted. Furthermore, there was a weak relationship between learner metacognition and problem-solving performance (r = -0.19, p = 0.009). SAP and problem-solving performance also showed a weak relationship (r = -0.211, p = 0.004).

The above regression model only had a small R^2 (R^2=.134), which indicated that this model could only predict 13.4% of the data. Therefore, based on the final cluster result, this study proposed a new regression model, which used the three final cluster groups generated from question (a) as variables to predict learner problem-solving performances, as follows:

$$Y_{ss} = \beta_1 Final\_GO + \beta_2 Final\_MC + u$$

The multiple regression results showed that the new model as a whole was significant ($p < .000$), which indicated that high metacognitive level-high goal orientation, low metacognitive level-medium goal orientation and medium metacognition level-low goal orientation were significant predictors for problem-solving performance,

R^2=.445, F(2, 129) = 51.7, $p < .000$. The regression equation was as follows:

$$Y_{ss} = -0.318 * Final\_GO - 0.428 * Final\_MC + u$$

Both Final_GO and Final_MC were significant predicators of performance during problem-solving. Specifically, with every point increase in Final_GO, a 0.928-point decrease in learner problem-solving performance was predicted; and with every point increase in Final_MC, a 1.314-point decrease in learner problem-solving performance was predicted. In addition, there was a strong correlation between Final_GO and problem-solving performance (r = -0.571, p = 0.000). Final_MC and problem-solving performance also showed a weak relationship (r = -0.616, p = 0.000).

1.c. Reasons for differences in learner performances

To understand the reasons for any possible problem-solving differences among learners based on metacognition and goal orientation, the 12 selected participants were divided into four groups based on the final cluster analysis result, including: 1) high metacognition, high multiple goal orientations and low performance (N = 3); 2) low metacognition, medium multiple goal orientations, and high performance (N = 3); 3) medium metacognition, low multiple goal orientations, and high performance (N = 2); and 4) outliers (N = 4).

Interviewees in Group 1 failed to identify the problem correctly at the initial stage of problem-solving, which caused two of them to ultimately be unable to solve the problem, but the student who corrected himself in the early stages of problem-solving did solve the problem and had a high solution score. Group 2 students were considered to have low metacognition and medium goal orientation compared to their peers. Students in this group all demonstrated double-checking behavior during problem-solving, which may have led them to solve the problem and gain a higher score. Interviewees in Group 3 also had high performances due to their double-checking behavior and good time management skills during problem-solving. As for the outliers in Group 4, two learners had zero points, and the other two learners had high solution scores. Participants who failed to solve the problem showed distracted problem-solving behavior, poor time management, over-confidence in their metacognition skills, and inefficient information filtering behavior. As for participants who gained high scores during the study, these individuals demonstrated competent time management skills, double-checking behavior, and care for outcomes.

Discussion of research question one

The data suggested that there were statistically significant differences in learner problem-solving performances based on metacognitive level, which was expected—learner metacognition has been linked to successful problem-solving in PBL (Davidson & Sternberg, 1998; Gourgey, 1998; Marra et al., 2014; Mihalca et al., 2017; Shin et al., 2003). Previous AR studies also indicated learner metacognition affected learning (Bogard et al., 2013; Liu et al., 2004). However, interestingly, learners in the lowest metacognitive level group had the highest solution scores, followed by learners in the high metacognitive level group and medium metacognitive group, which was unexpected—research showed that learner metacognitive skills can help with problem-solving (Davidson & Sternberg, 1998; Marra et al., 2014; Mihalca et al., 2017).

Based on the literature and stimulated recall interviews, there are two possible explanations for this result. The first is the Dunning-Kruger Effect (Kruger & Dunning, 1999), which indicates a cognitive bias in which people of low metacognitive ability usually mistakenly assess their ability as greater than it is. For example, Pajares and Kranzler (1995) tested student confidence of problem-solving ability and problem-solving performances of 329 high school students. They found most (86%) high school students were overconfident on their problem-solving abilities, a few of them (9%) were under-confidence, and only 4% of them successfully predicated their results. For all 18 math problems, students in the overconfidence group erred more often (6.2 problems on average) than did those in the under-confidence group (3.5 problems on average). Moreover, students in the under-confidence group had higher performance scores. Likewise, it is possible that some learners in the high metacognition group overestimated their metacognitive levels during the study. The stimulated interview further verified the explanation—participants Beta and Zeta self-reported higher metacognition levels but did not demonstrate a high metacognition level while solving the problem. Instead, they misunderstood the problem at an early stage of problem-solving and did not self-regulate their behavior appropriately during problem-solving. There may be more learners like Beta and Zeta in high metacognition group that have overestimated their metacognitive levels and led them to have lower problem-solving performances.

The second possible explanation might be relevant to problem complexity in this study. Scholars have pointed out that there are different complexity levels for different problems, and problems in PBL are designed to be ill-structured, complex, open-ended, and relevant to real life (Hmelo-Silver, 2004). In addition, student preconceptions of the problem would affect their problem-solving strategies during problem-solving processes (Phillips, 2001). Particularly, in solving an ill-structured problem, depending on whether they considered the problem as complex or not, one student might insist on a simple answer while another may be open to complex and alternative solutions. As mentioned before, AR is designed for sixth grade students to learn science subjects. Although advanced learners have used it for research purposes, the complexity of the problem might not be challenging or engaging enough for some participants who had a higher metacognition level. Rather, this problem—to help Jakala-Tay find a suitable home within 60 minutes—might be more appropriate for participants who had a lower metacognition level. The interview also further confirmed this hypothesis—participant Omega had a high metacognition level, but wrote a very short answer for the solution score, because Omega might have considered the problem in this study as too simple (he found both correct answers in an early stage of problem-solving), which might have made him unwilling to write a complex rationale for the solution. As a result, he earned a much lower score than those providing more thorough explanations. Likewise, there might be more participants like Omega, who had a high metacognition level, but were not interested in solving this simple problem. Therefore, participants in lower metacognition level had significantly higher problem-solving performances compared to participants that occupy higher metacognition level.

As for goal orientation, based on literature and cluster analysis, this study adopted the multiple goal orientation perspective and clustered learners into five groups, including 1) medium in all six goal orientations; 2) low in all six goal orientations; 3) high in all six goal orientations; 4) high in TAP, TAV, SAP, SAV, but low in OAP and OAV; and 5) high in TAP and OAP, medium in SAP, but low in TAV, SAV and OAV. These groups indicated that learners can adopt multiple goal orientations simultaneously and have different levels of goal orientations during learning (Barron & Harackiewicz, 2001; Bereby-Meyer & Kaplan, 2005; Daniels et al., 2008; Harackiewicz et al., 2008; Jang & Liu, 2012; Meece & Holt, 1993; Pintrich, 2000b; Pintrich et al., 2003; Zusho et al., 2005). In addition, there were no significant differences in problem-solving performances based on goal orientation groups among learners in this study, which indicated learner goal orientation did not directly affect final problem-solving performances in AR. This is consistent with literature stating that different goal orientations may lead to similar learning outcomes (Pintrich, 2000b). This study also verified Pintrich (2000b)'s proposal that learners with different goal orientations might follow different trajectories and have different experiences over time but end up with the same achievement or performance. Admittedly, the stimulated interview showed that learner goal orientation did mediate learner problem-solving processes, which may lead to different learner problem-solving performances. For example, both Omega and Lota, who had high metacognition levels and might have considered the problem too easy, wrote simple answers. Different from Omega, Lota had a high OAV and a higher SAP and SAV score, which might have affected the results, because she told the researcher "I don't wanna do worse [than others]" and "I like to be as good as I think I can be, like, if I studied for four hours, like, I want to see I put in work, and I should have done better." Therefore, Omega wrote a short answer and earned 4 points, while Lota provided detailed evidence in her solution and earned the full scores—7 points.

Few studies have been conducted on the interaction between learner metacognition and goal orientations on problem-solving (Gul & Shehzad, 2012; Ning, 2016); therefore, the results from this study can contribute to the knowledge on this topic. Cluster analysis showed that an interaction between metacognition and goal orientations did affect learner problem-solving in this study. The final cluster of leaners included 1) high metacognition and high multiple goal orientations, 2) low metacognition and medium multiple goal orientations, and 3) medium metacognition and low multiple goal orientations. In addition, learner problem-solving performances were statistically significant based on these three clusters—low metacognition and medium metacognition groups had significantly higher problem-solving performances compared to the high metacognition groups. As mentioned before, this result could also be explained using the Dunning-Kruger effect and appropriate problem complexity—some learners in the high metacognition group might have either overestimated their metacognition levels or not engaged in the problem-solving due to problem complexity.

Furthermore, this study found that learner metacognition and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, and OAV) can predict leaner problem-solving performances (i.e., SS). In addition, learner TAP, TAV, and SAV were significant predicators of performance during problem-solving. Specifically, for every point increase in TAP, a 0.34-point decrease in learner problem-solving performance was predicted. This predication was consistent with the finding of Stoeber et al. (2015) that TAP predicted exam performance. For every point increase in TAV, a 0.252-point increase in learner problem-solving performance is predicted; and for every point increase in SAV, a 0.229-point decrease in learner problem-solving performance was predicted, which is consistent with David's (2014) finding that SAV negatively predicted test performance.

By modifying the multiple regression model based on the final cluster result, this study increased the predication rate from 13.4% to 44.5%. The new model indicated that high metacognitive level-high goal orientation, low metacognitive level-medium goal orientation, and medium metacognition level-low goal orientation were significant predictors for problem-solving performance. Therefore, the results suggested that compared to single goal orientation and metacognition, multiple goal orientations and metacognition together can be used to predict learner problem-solving performances more accurately. This result further supports the multiple goal orientations perspective—1) learners did possess multiple goal orientations during learning and different goal orientations (Barron & Harackiewicz, 2001; Wolters et al., 1996); 2) different patterns in the levels of different goal orientations might lead to the same learning outcomes (Daniels et al., 2008; Pintrich, 2000b).

Research Question Two

Firstly, for question (a), (b), (c), (d), *Tableau* and *R* were used to visualize learner problem-solving processes based on their activity log (i.e., tool use frequency, duration and room visit sequences). Specifically, Chord Diagrams were used to visualize learner tool use frequency and duration actions in AR. Similarity measures were used to analyze room visit sequences during problem-solving based on learner characteristics. Two-proportion z-test were used to identify whether the room visit sequence differences between two groups of learners were statistically significant. For question (e), the grounded theory was used to analyze interview codes and explain the reasons for any possible problem-solving process differences among learners.

2. a. Visualizing learner problem-solving process patterns

This study visualized three types of learner problem-solving process patterns including learner tool use frequency, tool use duration, and room visit sequences. For tool use frequency, the three most frequently used tools were Solar System Database (17,815 times), Probe Design (16,205 times), and Mission Control (12,395 times), while the least frequently used tools were Periodic Table (732 times), Spectra (1,668 times), and Concepts Database (2,361 times). In addition, female students, on average, used all of the tools less frequently compared to male students (101, 121.9 times, respectively). Furthermore, this study found that the highest-scoring students used Solar System Database more frequently compared to the lowest-scoring students, while the lowest-scoring students used Alien Database, Communication Center, Mission Control, and Probe Design more frequently.

As for tool use duration, during the 60-minute study, the top three tools that learners spent time using were Alien Database (M = 19.67), Probe Design (M = 13.93), and Solar System Database (M = 9.06). Learners spent the lowest amount of time using Concepts Database, Missions Database, and Spectra—less than 4 minutes. In addition, male students stayed in the Probe Design tool longer (M = 14.63) than female students (M = 13.15). They also stayed in Mission Control, Notebook, and Spectra longer than female students (M = 8.69, 7.77, 4.56 for male students; M = 6.77, 6.69, 3.17 for female students). Students of both genders stayed in Communication Center and Periodic Table for about the same amount of time (about 7 and 4 minutes). Considering the performance factor, learners who had the highest scores used Mission Control, Concepts Database, Notebook, Probe Design, and Solar Database longer compared to students who had the lowest scores (M = 8.292, 4.583, 8.292, 15.167, 8.917 for the highest-scoring students; M = 6.963, 3.907, 7.204, 13.074, 7.963 for the lowest scoring students). Learners with the lowest scores used Missions Database and Communication Center more than learners with the highest scores

(M = 3.722, 7.278 for the lowest scoring students; M = 3.041, 6.417 for the highest-scoring students). Both types of learners spent similar amounts of time on Alien Database and Periodic Table (around 18 and 4 minutes).

With respect to learner room visit sequences, the data visualization showed that learners who scored zero, one, two, and three points had larger differences on their room visit sequences compared to learners who scored seven points; while learners who scored four, five, and six points had a similar room visit sequence compared to learners who scored seven points. However, the two-proportion z-test results indicated that there were no significant differences in learner room visit sequences between learners who scored seven points and any other groups.

2. b. Problem-solving process patterns based on metacogniton

The Chord Diagram showed that learners with low metacognition levels had higher tool use frequency, followed by learners with medium and high metacognition levels. It is worth noting that learners in the high metacognition group used nine tools (out of ten tools) less often than learners who had lower metacognition, and the Missions Database was the only tool that learners in the high metacognition group used more than those in the low metacognition group.

Considering tool use duration, data visualization showed that learners in the low metacognition group used all of the tools for longer periods compared to learners in the medium, and high metacognition groups. In addition, for individual tools, learners in low metacognition group stayed in Spectra, Probe Design, Notebook, and Mission Control longer than learners in the high metacognition group. However, they stayed in Alien Database for less time compared to the other two groups. They also stayed in Solar System Database for slightly less time compared to the high metacognition group, but longer than students in the medium metacognition group. Furthermore, these three groups all stayed in Communication Center for about the same amount of time.

The similarity measure was also conducted for learner room visit sequences based on metacognition groups. Based on this visualization, the distances among all learners were similar, which indicated that learners in all three metacognition level groups had a similar room visit sequences.

2. c. Problem-solving process patterns based on goal orientation

The problem-solving process patterns based on five goal orientation groups were also examined. The five groups were 1) medium in all six goal orientations (N = 41); 2) low in all six goal orientations (N = 5); 3) high in all six goal orientations (N = 85); 4) high in TAP, TAV, SAP, SAV, but low in OAP and OAV (N = 23); and 5) high in TAP and OAP, Medium in SAP, but low in TAV, SAV and OAV (N = 4).

The data showed that learners in Group 5 had the highest tool use frequency, followed by Group 4, 3, and 1. Learners in Group 2 had the lowest tool use frequency. In particular, learners in Group 5 visited Solar System Database, Probe Design, Notebook, and Alien Database most frequently, while learners in Group 1 visited Periodic Table the most, learners in Group 2 visited Missions Database tool and Concepts Database the most, and learners in Group 4 visited Spectra and Mission Control the most.

As for learner tool use duration, the data showed that learners in Group 5 stayed in all the tools for longer periods (95.25 minutes) compared to learners in the other four groups, and learners in Group 2 stayed in all the tools for the shortest length of time (64.40 minutes). In addition, learners in Group 5 stayed in Spectra, Notebook, Mission Control, Concepts Database, Alien Database, and Communication Center longer compared to the other four groups. Group 4 learners stayed in Probe Design, Solar System Database, and

Missions Database the longest compared to other learners. Group 2 stayed in Periodic Table the longest compared to their peers.

Based on the visualization, the room visit sequences were slightly different between Group 1 and Group 3, but there were bigger differences between Group 1 and Group 2. The Cosine measure also indicated that there were bigger differences between Group 1 and Group 4, as well as between Group 1 and Group 5. In addition, two-proportion z-test indicated there were no significantly differences in learner room visit sequences between learners in Group 1 and Group 2, 3, 4. Although the Cosine method did show a significant difference in room visit sequences between Group 5 and group 1, the LCS and Jaccard methods did not suggest a significant difference in this area.

2. d. Problem-solving process patterns based on the interaction between metacogniton and goal orientaion

Learner problem-solving process patterns were also examined based on the three cluster groups. According to the visualization, the Chord Diagram showed that learners in Cluster 2 (i.e., low metacognition and medium multiple goal orientations) had the highest tool use frequency (544.57 times), followed by Cluster 3 (i.e., medium metacognition and low multiple goal orientations, 487.41 times) and Cluster 1 (high metacognition and high multiple goal orientations, 477.66 times). In addition, learners in Cluster 2 used Probe Design and Mission Control tools most frequently; learners in Cluster 3 used Periodic Table, Notebook, Concepts Database, and Alien Database tools most frequently; and learners in Cluster 1 used Missions Database and Communication Center most frequently.

According to the data visualization of learner tool use duration based on the final clusters, learners in Cluster 2 had the longest tool use duration (84.06 minutes), while learners in Cluster 1 had the shortest tool use duration (77.72 minutes). In addition, learners in Cluster 2 stayed in Spectra, Solar System Database, Probe Design, Notebook, and

Mission Control longer compared to the other four groups. Cluster 3 learners stayed in Periodic Table, Mission Control, Concepts Database, and Communication Center the longest among the three groups. Cluster 1 stayed in Missions Database and Alien Database longer compared to their peers.

The similarity measures of room visit sequences suggested that Group 1 and Group 3 were very similar, but there were bigger differences between Group 2 and Group 3. However, two-proportion z-test indicated that there were no significant differences in learner room visit sequences among learners in the three clusters.

2. e. The reasons for the learner problem-solving process differences

To understand the reason for learner problem-solving process differences, 12 learners who participated in stimulated interviews were grouped into 4 groups, including 1) multiple goal orientations and low performance group (N = 3); 2) low metacognition, medium multiple goal orientations, and high performance group (N = 3); 3) medium metacognition, low multiple goal orientations, and high performance group (N = 2); and 4) outliers (N = 4). In addition, this study summarized 10 steps that a successful problem-solver would usually complete during this study based on 12 stimulated recall interviews and the literature, including: 1) identify the problem correctly; 2) explore the 3D environment by visiting all rooms in AR and look over all tools; 3) discover what Jakala-Tay needs to survive in Alien Database; 4) search the Solar System Database for possible planets; 5) develop hypotheses about where Jakala-Tay can live; 6) figure out if there is any missing information needed for making a decision; 7) launch probes to gather the information in the Probe Design room; 8) check the data from the probe in the Mission Control room; 9) decide whether the selected planet is a good choice for Jakala-Tay; 10) if so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4.

Learners in Group 1 had a higher metacognition and goal orientations compared to learners in the other two clusters, but they all failed at the very first step (i.e., Identify the problem correctly) during problem-solving. It is worth noting that only one learner eventually solved the problem due to his high self-approach goal orientation and self-corrective behavior, while the other two learners could not solve the problem because they might have overestimated their metacognitive levels, demonstrated poor time management, and were not able to self-correct during the problem-solving process.

Learners in Group 2 were all considered to have low metacognition, medium goal orientation, and high solution scores. They each finished all 10 steps during the problem-solving in AR. Learners in Group 3 had higher metacognition and lower goal orientation scores compared to Group 2, but they also had high solution scores like the learners in Group 2. In addition, they did not go through all the 10 steps as the learners in Group 2 did, thus they solved the problem more efficiently.

As for the outliers in Group 4, they either failed to solve the problem, because they neglected important information during steps 3 and 4 regarding learning about the alien and solar system, or they solved the problem, because they followed all 10 steps during the problem-solving, and fully utilized all the 60 minutes allotted by the study.

Discussion of research question two

This study used tools such as *Tableau* and *R* to visualize learner problem-solving processes based on computer log data (i.e., tool use frequency, duration and room visit sequences), which helped the researcher to identify learner problem-solving process patterns. Specifically, Chord Diagrams (Flajolet & Noy, 2000) and the *R circlize* package

(Gu et al., 2014) were used to visualize learner tool use frequency and duration action in AR. Just like Wei et al. (2016) described, these diagrams did provide a compact way of representing information in this study. In addition, using the similarity measures and two proportion *z* test, this study further examined whether there were statistically significant differences in learner room visit sequences among different groups of learners. Therefore, combining data visualization, similarity measure and two proportion *z* test, this study presented the data in a comprehensive way to help understand learner problem-solving processes.

Particularly, this study visualized three types of learner problem-solving process patterns including learner tool use frequency, tool use duration, and room visit sequences. Previous AR studies suggested that there were 10 cognitive tools that can be categorized into four types, including (a) sharing cognitive load (Alien Database, Solar System Database, Missions Database, Concepts Database, Spectra, and Periodic Table), (b) supporting cognitive processes (Notebook), (c) supporting otherwise out-of-reach activities (Probe Design Center), and (d) supporting hypothesis testing (Mission Control Center and Communication Center) (Liu & Bera, 2005; Liu et al., 2014, 2016). In addition, Alien Database, Solar System Database, Notebook, Probe Design and Mission Control were considered as critical tools for solving the problem and were used more by learners, while Missions Database, Concepts Database, Spectra, Periodic Table, and Communication Center tools were less critical for problem-solving and were used less by learners in previous AR studies (Liu et al., 2015). Furthermore, Alien Database, Probe Design, and Mission Control tools were also considered as more fun tools and were used more by learners in previous AR studies (Liu et al., 2015; Liu et al., 2016).

In this study, the three most frequently used tools were Solar System Database, Probe Design, and Mission Control, while the least frequently used tools were Periodic

Table, Spectra, and Concepts Database. This result is consistent with previous AR findings (Liu et al., 2015). Specifically, previous AR studies suggested that Probe Design and Mission Control were the most fun tools for learners, and they also allow learners to conduct otherwise out-of-reach activities and supporting hypothesis testing—equip a probe with scientific instruments and receive data from a launched probe (Liu et al., 2015; Liu et al., 2016). Solar System Database was "needed to understand what each planet in our solar system can offer" (Liu et al., 2015, p. 192). Therefore, learners in this study used the critical and fun tools most frequently, and the tools for supporting otherwise out-of-reach activities and supporting hypothesis testing most frequently, while used three of the sharing cognitive load tool the least frequently. As for tool use duration, during the 60-minute study, the top three tools that learners spent time using were for sharing cognitive load—Alien Database and Solar System Database, and supporting otherwise out-of-reach activities—Probe Design. Learners spent the least time on Concepts Database, Missions Database, and Spectra tools. These results are consistent with previous AR studies, because Alien Database, Solar System Database, Probe Design, and Mission Control are the most important tools for solving the problem; consequently, learners usually spent more time with these tools and used them more frequently (Liu et al., 2015).

With respect to the effect of learner metacognition, the Chord Diagram showed that learners with a low metacognition level had higher tool use frequency, followed by learners with medium and high metacognition. Since tool use frequency was positively related to engagement (Liu et al., 2004, Liu et al., 2013), learners in the low metacognition level group might engage with problem-solving in this study more, which may have led them to have higher problem-solving performances. It is worth noting that Missions Database was the only tool that learners in the high metacognition group used more frequently than those who were in the low metacognition group. This tool is designed using seven levels of accordion menu—learners must click each tab to view the information, which is different from all the other tools. Each time the learner clicks the tab, it is logged as using Mission Database once. Therefore, learners in the high metacognition level group probably checked this less useful tool more frequently compared to other learners, which could not effectively help them with problem-solving in AR. For tool use duration, data showed that learners in the low metacognition group stayed in all the tools longer than learners in higher metacognition groups. In addition, regarding individual tools, learners in the low metacognition group stayed in Spectra, Probe Design, Notebook, Mission Control—tools in all the four categories— longer than learners in the high metacognition group. Therefore, it is possible that learners in the low metacognition level group considered the study more engaging and used the tools more appropriately; this sense of engagement may eventually have resulted in higher performance for these learners.

As for room visit sequences, the data suggested that learners in all three metacognition level groups had similar room visit sequences based on the visualization of the similarity measures. It is interesting because the literature suggested that learner navigational sequences can be used to differentiate learners in SG (Loh et al., 2016). The explanation might either be due to the characteristics of the room visit sequences or the characteristics of the SG. Firstly, the room visit sequences in AR only consisted of four units due to the limited rooms (i.e., M for Main Room, A for Alien Information Room, P for Probe Design Room, and C for Mission Control Center), while in Loh et al. (2016)'s study, there were 25 units in the sequences. Mathematically, there were far less variations in room visit sequences. Secondly, AR was a problem-based learning SG, in which learners demonstrated more free exploration behavior during problem-solving, while Loh et al. (2016)'s study used a maze puzzle SG, in which learners demonstrated more behavior based on logical decision. Therefore, if a future study could find a way to exclude the free exploration behavior, the room visit sequences might show larger differences based on learner metacognition levels.

As for goal orientation, the Chord Diagram showed learners in different goal orientation groups had different tool use frequency—learners in Group 5 (i.e., high in TAP and OAP, Medium in SAP, but low in TAV, SAV and OAV) had the highest tool use frequency (738.25 times) and longest duration (95.25 minutes), while learners in Group 2 (i.e., low in all six goal orientations) had the lowest tool use frequency (444.60 times) and stayed in all the tools for the shortest duration (64.40 minutes). Literature suggested that learner TAP was significantly and positively correlated to engagement in learning (Elliot et al., 2011; Gillet et al., 2015). Since learners in Group 5 had high TAP and learners in Group 2 had low TAP, it is reasonable that Group 2 learners had the least tool use frequency and duration. Both TAP and OAP were also positively related to the motive for success in learning (Diseth, 2015); therefore, Group 5 learners might have higher motivation to use tools more during their problem-solving. For individual tools, learners in Group 5 visited Solar System Database, Probe Design, Notebook, and Alien Database most frequently, and stayed in Spectra, Notebook, Mission Control, Concepts Database, Alien Database, and Communication Center longer compared to the other four groups. Learners in Group 2 visited Missions Database and Concepts Database most frequently and stayed in Periodic Table the longest compared to their peers. As discussed in the previous sections, Solar System Database, Probe Design, Notebook, Alien Database, and Mission Control are all critical tools for problem-solving in AR. These tools were also considered as more useful compared to Missions Database and Concepts Database. Therefore, learners in Group 5 had more positive tool use, which is consistent with previous studies in that learners TAP positively predicted task value and strategic learning strategies (Diseth, 2015)—Group 5 learners used these useful tools most frequently and spent the most time with them.

Furthermore, the room visit sequences were also different among learner goal orientation groups. However, the two-proportion z-test indicated that there were no significant differences in learner room visit sequences based on their goal orientation groups. It was similar to learner room visit sequences based on metacognition—no significant differences, which might have been caused either by the characteristics of the room visit sequences or the characteristics of AR. Although there were significant differences in the room visit sequences based on learner goal orientations, interview results suggested that goal orientations affected learner problem-solving processes in AR. For example, Beta demonstrated help-seeking behavior by asking the researcher some clarification questions during problem-solving, which was consistent with literature about learners who have a higher TAP score (Ning, 2016). The positive tone in Epsilon's explanation was consistent with literature on learners who have a high SAP. These learners usually feel positive activity-related emotions (Brondino et al., 2014). Mu had a higher OAV score compared to his other five goal orientations, which was consistent with literature indicating that learners who have high OAV have a higher level of academic achievement (Diesth, 2015).

The results showed that the interaction between learner metacognition and goal orientation also affected learner problem-solving processes in AR. Particularly, the Chord Diagram showed that learners in low metacognition-medium multiple goal orientations had the highest tool use frequency (544.57 times) and the longest tool use duration (84.06 minutes), followed by medium metacognition-low multiple goal orientations learners (487.41 times, 78.61 minutes) high metacognition-high multiple goal orientations learners (477.66 times, 77.72 minutes). This result is consistent with previous section about learner metacognition—learners who had low metacognition had the highest tool use frequency and duration, while learners who had high metacognition had the lowest tool use frequency and duration.

It is worth noting that the five goal orientation groups evolved into three groups after interacting with goal orientation, which indicated that learner metacognition affected learner goal orientation grouping. Particularly, influenced by metacognition, medium multiple goal orientation learners had the highest tool use frequency and duration as opposed to high multiple goal orientation learners. This might be due to the fact that low metacognition learners were more engaged with the SG, so they could achieve the highest tool use frequency and duration, even though they only had medium multiple goal orientation scores. Likewise, high multiple goal orientations learners were affected by their high metacognition, which caused them to have the lowest tool use frequency and duration due to lower engagement with the SG. In addition, learners who had low metacognition-medium multiple goal orientations used Probe Design and Mission Control tools the most. They also stayed at these two tools longer than other learners. Learners who had medium metacognition-low multiple goal orientations used Periodic Table, Notebook, Concepts Database, and Alien Database most frequently and stayed in Periodic Table, Mission Control, Concepts Database, and Communication Center the longest among all the three groups. It appeared that when learners had lower metacognition level and lower multiple goal orientations, they were engaging in more appropriate tool use. In addition, these two clusters of learners had higher problem-solving performances. This is consistent with previous results indicating that learners who had more appropriate tool use patterns would have better problem-solving performances.

Although learner room visit sequences were not statistically significant based on the interaction of learner metacognition and goal orientations, it is interesting that high metacognition-high multiple goal orientations learners and medium metacognition-low multiple goal orientations learners were very similar on their room visit sequences, while they had significantly different problem-solving performances. In addition, there were bigger differences between low metacognition-medium multiple goal orientations learners and medium metacognition-low multiple goal orientations learners, while they both had high problem-solving performances. This indicated that two similar room visit sequences could lead to completely different problem-solving performances, while two different room visit sequences could lead to similar problem-solving performance—high problem-solving performance in this study.

To better understand problem-solving differences based on learner characteristics, this study summarized 10 steps that a successful problem-solver would usually perform during this study, in addition to analyzing the problem-solving process patterns. Jonassen (2000) suggested that to solve a problem, there must be "someone [who] believes that it is worth finding the unknown" (p. 65), which indicated that the proposed problem "must have some social, cultural, or intellectual value" (Jonassen, 2000, p. 65). In this study, all participants joined the study voluntarily to "find the unknown" home for one of the alien. Based on the four conceptual stages of problem-solving in AR (Liu et al., 2004) and stimulated recall interviews, this study further suggested 10 operational steps for solving the problem successfully for one alien. Different from the four conceptual stages, these 10 steps were practical—several successful problem-solvers demonstrated these actual behaviors during the 60-minute study. Therefore, these steps could be used to guide learner problem-solving while taking into consideration metacognition levels and goal orientation groups. Admittedly, it is possible that learners could solve the problem without going through each of the 10 steps; however, if they skip these steps, they will have to simply guess and depend on luck to solve the problem—just like Theta, who earned the top score in the SG but did not send any probes to verify his hypothesis. He admitted that "I was thinking it [sending probes] could be useful to test my theories. But I never sent any, so I could be totally wrong." So he was not completely sure whether he found the correct planet for the alien, rather he "guessed" the right answer based on the evidence he gathered. Therefore, he might not be able to find the appropriate planet for a different alien. In addition, it is also possible for learners to solve the problem without following the exact order of the steps; however, this would consume more time during problem-solving, because they might need to go through these steps back and forth multiple times.

Other Factors Affecting Learner Problem-solving

Besides learner metacognition and goal orientations, this study found that there might be at least other three factors affecting learner problem-solving during the study, including AR design glitch and game controls, learner previous knowledge, and the incentives for participating in the study. Some interviewees indicated there was one game design glitch in the environment—they had to refresh the browser to exit the consoles, which might have affected learner problem-solving. In addition, this glitch might have affected log data regarding the tool use duration and frequency for the four consoles (i.e., Probe Design, Mission Control, Communication Center and Alien Database). It might have changed the room visit sequence data slightly as well. However, since this glitch only happened once or twice during the 60-minute study, and it did not happen to everyone, the study did not exclude the data collected from participants who encountered glitches. A few participants also mentioned they had trouble in using the game controls, although they indicated it did not affect them during the study. Still, this might have caused some frustration during the problem-solving.

The second factor that might have affected learner problem-solving in this study was learner previous knowledge about the space science topic. On the one hand, some participants who considered themselves to have insufficient knowledge, made more effort to take notes, and utilized the information about the solar system in AR to solve the problem. On the other hand, some participants who considered themselves to have known about the solar system were seeking solutions solely based on their preexisting knowledge, which caused them to find the incorrect planets. Therefore, learner previous knowledge on the subject did not necessary help solve the problem; rather, it could hinder problem-solving, especially for some novice problem-solvers, because they might have simply leapt into problem-solving action rather than taking the time to understand the problem (Dominowski, 1998).

Lastly, for all 159 participants in this study, they were told that they had a chance to get a $20 Amazon gift card. This financial incentive might have affected the sample composition in this study, because some participants might have joined the study because of the potential gift card. In addition, 116 of them from the participant pool and 21 of them from one class in the College of Education were told they could get an extra two credits for one of their classes. This academic incentive might have also affected the sample composition—participants might have joined the study because of the extra credits. Therefore, learners who did not care for gift cards or extra credits might not have joined the study. However, to attract more participants, the researcher had to provide these incentives. In addition, since this study focused on the effects of learner metacognition and goal orientation on learner problem-solving, these incentives did not directly affect learner problem-solving in the laboratory setting.

CONCLUSION

Using AR as the SG environment, this study investigated 159 undergraduate learners' metacognition, goal orientations, problem-solving performances and processes in a laboratory setting using a sequential mixed research design. The conclusions of this study are presented in the following sections.

Learner metacognition affects problem-solving

Firstly, with respect to learner metacognition, the results showed that there were statistically significant differences in learner problem-solving performances based on metacognition. In addition, learners in the lowest metacognitive level group had the highest problem-solving performances, followed by learners in the high metacognitive level group and the medium metacognitive group. Based on stimulated interviews and literature, learners in the lowest metacognitive level might find the problem in this study more engaging, while some learners in the highest metacognitive level might find it too easy to provide a complex solution, or some learners in the highest metacognitive level might have overestimated their metacognitive level in the self-reported survey (i.e., Dunning-Kruger effect).

As for learner problem-solving processes, the Chord Diagram showed that learners with a low metacognitive level had higher tool use frequency and duration, followed by learners with medium and high metacognitive levels, which again indicated that learners with a low metacognitive level engaged in the problem-solving more compared to learners in the high metacognitive level. In addition, for individual tool use duration, learners in the low metacognition group stayed in the critical tools for problem-solving (i.e., Probe Design, Notebook, and Mission Control) longer than learners in the high metacognition group. Furthermore, learner metacognitive level did not significantly affect learner room visit sequences—learners in all three metacognitive levels had similar room visit sequences. This was either caused by 1) the characteristics of the SG—a PBL environment, which required free exploration during the problem-solving or 2) the characteristics of the sequences—limited units in the sequences, which offered less variations for calculation.

Learner goal orientations affect problem-solving

Considering learner goal orientations, this study adopted the multiple goal-orientations perspective, which indicated that learners can possess multiple goal orientations simultaneously and have different levels of goal orientations to benefit their learning. In this study, learner goal orientation did not directly affect final problem-solving performances in AR. This is consistent with literature that different goal orientations might lead to the same learning outcomes (Pintrich, 2000b). As for learner problem-solving processes, there were no significant differences in problem-solving processes based on learner room visit sequences, which verified Pintrich (2000b)'s proposal that learners with different goal orientations might follow different trajectories and have different experiences over time but end up with the same achievement or performance.

In addition, the data showed that learners in different goal orientation groups had different tool use frequency and duration—learners who had higher TAP and OAP had the highest tool use frequency and duration, while learners who had the lowest goal orientations had the lowest tool use frequency and duration. This is mainly because learner TAP and OAP were positively related to the motive for success in learning (Diseth, 2015), which might have stimulated learner tool use frequency and duration during the study. Furthermore, this study showed that learners who had higher TAP also used critical tools (i.e., Solar System Database, Probe Design, Notebook, Alien Database and Mission Control) most frequently and spent the most time on them, which is consistent with previous studies that learner TAP positively predicted task value and strategic learning strategies (Diseth, 2015).

The interaction between metacognition and goal orientations affect learner problem-solving

Using cluster analysis, the interaction between metacognition and goal orientations was shown to affect learner problem-solving in this study. The final cluster of learners based on the interaction of learner metacognition and goal orientations included 1) high metacognition and high multiple goal orientations, 2) low metacognition and medium multiple goal orientations, and 3) medium metacognition and low multiple goal orientations. In addition, learner problem-solving performances were statistically significant based on these three clusters. Particularly, learners in the low metacognition-medium multiple goal orientations cluster and medium metacognition-low multiple goal orientations cluster had significantly higher problem-solving performances compared to learners in the high metacognition-high multiple goal orientations group.

Furthermore, learner metacognition (MC) and goal orientation (i.e., TAP, TAV, SAP, SAV, OAP, and OAV) can predict leaner problem-solving performances (i.e., SS). Specifically, learner TAP, TAV, and SAV were significant predictors of performance during problem-solving. By modifying the multiple regression model based on the final cluster result, this study increased the prediction rate from 13.4% to 44.5%, which indicated that high metacognitive level-high goal orientation, low metacognitive level-medium goal orientation and medium metacognition level-low goal orientation were significant predictors for learner problem-solving performance in AR.

The interaction of learner metacognition and goal orientation not only significantly affected problem-solving performance but also affected learner problem-solving processes in AR, especially tool use duration and frequency. Particularly, low metacognition-medium multiple goal orientations learners had the highest tool use frequency and duration because they were more engaged with problem-solving, followed by medium metacognition-low multiple goal orientations learners and high metacognition-high multiple goal orientations learners. In addition, low metacognition-medium multiple goal orientations learners and medium metacognition-low multiple goal orientations learners used the critical tools (i.e., Probe Design, Mission Control, Notebook, and Alien Database) most frequently and stayed in these tools longer, while high metacognition-high multiple goal orientations learners used less useful tools (i.e., Missions Database) most frequently and stayed longer.

As for room visit sequences, learner metacognition and goal orientations did not significantly affect learner room visit sequences in AR. Therefore, learners who had different characteristics could either have different room visit sequences or similar room visit sequences. In addition, learners could all have high problem-solving performances despite that having different room visit sequences in AR—all roads lead to Rome. This study also found a successful learner usually would go through 10 steps in AR to help one alien species to locate a suitable home, as follows: 1) identify the problem correctly; 2) explore the 3D environment by visiting all rooms in AR and look over all tools; 3) discover what one alien species needs to survive in Alien Database; 4) search the Solar System Database for possible planets; 5) develop hypotheses about where this alien species can live; 6) figure out if there is any missing information needed for making a decision; 7) launch probes to gather information in the Probe Design room; 8) check the data from the probe in the Mission Control room; 9) decide whether the selected planet is a good choice for the selected alien species; 10) if so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4. Admittedly, the learner could also simply guess and depend on the luck to solve the problem or solve the problem without following the exactly order of the 10 steps exactly. However, these 10 operational steps could be used to guide learner problem-solving and help researchers further understand learner problem-solving in AR.

LIMITATIONS

Although the researcher designed the study based on previous studies, collected reliable and valid data, and conducted rigorous data analysis, there are at least three limitations for this study.

First, although the sample size in this study met the basic requirement for conducing multiple regression and cluster analysis, the participants were all undergraduate students from one university, which indicates that the results from this study might not be generalized into other institutions or learners in different age groups.

Secondly, the study used problem solution scores as the only measurement for learner problem-solving performances, which might not reflect learner authentic problem-solving abilities. In addition, although the rubric for grading these solution scores has been used in many previous AR studies, it could not capture all aspects of learner problem-solving processes.

Finally, this study asked participants to fill out the survey before playing the game, so the researcher would have time to select 12 participants for conducting a stimulated recall interview based on their survey results. However, this ordering might have increased the Dunning-Kruger effect, because participants were less likely to objectively self-evaluate their metacognitive level before encountering the problem. Therefore, reversing the order of survey completion and game play might reduce the effect and increase the validity of the metacognition survey—because the participants might have a better understanding of their own metacognitive levels after playing the problem-solving game.

IMPLICATIONS AND FUTURE RESEARCH

This study offered implications for future SG design and development, particularly in designing PBL based SG environments for undergraduate students. The study indicated that it is important to consider the problem complexity factor in designing a PBL environment. Otherwise, learners might not find the problem worth solving. In addition, the study results can help AR designer, developer and researcher to improve the current environment, including (a) fixing the current game glitches, such as the navigation problem and program crush issue, (b) developing in-game learner problem-solving performance rubric to better evaluate learner problem-solving performance, and (c) providing appropriate scaffolding mechanism during the 10 steps problem-solving processes.

Based on the results and limitations of the current study, the researcher proposes the following four recommendations for future research. First, this study suggested that problem-solving in AR might be too easy for learners who had higher metacognition levels. To verify or generalize this finding, a larger sample size of participants in other settings is needed. In future studies, researchers could either design a more challenging problem-solving SG for this group of learners (i.e., undergraduate students), or use the same learning environment in a laboratory setting but recruit a different group of learners who are in an appropriate age group; i.e., middle school students. Furthermore, more qualitative data are needed to understand learner problem-solving processes differences based on their metacognition and goal orientation. For example, future researchers can interview more participants who have similar metacognition and goal orientation level, but have different problem-solving performances to better understand their problem-solving.

Secondly, this study only used solution score to represent learner problem-solving performance, which might not truly reflect learner problem-solving performance. In future studies, on the one hand, researchers could develop a more comprehensive way to measure learner problem-solving performances, such as creating a rubric considering learner problem-solving processes. On the other hand, designers and developers in AR could also explore alternative ways to measure learner solution scores in the environment. For example, the solution form could ask learners to explain their problem-solving processes rather than just submitting a final answer with justifications, so the learner might be able to provide more information regarding their problem-solving processes, then the researchers could grade learner problem-solving performances based on the problem-solving processes rather than just using learner final answers alone.

Third, researchers could reverse the data collection order of gameplay and self-reported survey to potentially minimize the Dunning-Kruger Effect during the study. However, researchers might need to randomly select stimulated recall interview participants or asked the participants to stay for an extended time so that the researcher could purposely select suitable participants.

Finally, this study found learner metacognition and goal orientation both affected learner problem-solving in SG at a macro-level. This study also explored the possible reasons for these effects based on stimulated recall interviews and literature. However, how learner metacognition and goal orientation affected learner problem-solving performance and processes in detail are still unknown. Future studies need to further investigate how learner metacognition and goal orientation affect learner problem-solving in SG environments at a micro-level.

Appendices

APPENDIX A: METACOGNITIVE AWARENESS INVENTORY (MAI)

Instructions: Think of you are engaging in the learning activities. Read each statement carefully. Consider if the statement is true or false as it generally applies to you when you are in the role of a learner, with 0 being "totally untrue of me" and 100 being "totally true of me." All your responses will be kept anonymous and confidential. There are no right or wrong responses, so <u>please be open and honest</u>.

1. I ask myself periodically if I am meeting my goals.
2. I consider several alternatives to a problem before I answer.
3. I try to use strategies that have worked in the past.
4. I pace myself while learning in order to have enough time.
5. I understand my intellectual strengths and weaknesses.
6. I think about what I really need to learn before I begin a task
7. I know how well I did once I finish a test.
8. I set specific goals before I begin a task.
9. I slow down when I encounter important information.
10. I know what kind of information is most important to learn.
11. I ask myself if I have considered all options when solving a problem.
12. I am good at organizing information.
13. I consciously focus my attention on important information.
14. I have a specific purpose for each strategy I use.
15. I learn best when I know something about the topic.
16. I know what the teacher expects me to learn.
17. I am good at remembering information.
18. I use different learning strategies depending on the situation.
19. I ask myself if there was an easier way to do things after I finish a task.
20. I have control over how well I learn.
21. I periodically review to help me understand important relationships.
22. I ask myself questions about the material before I begin.
23. I think of several ways to solve a problem and choose the best one.
24. I summarize what I've learned after I finish.
25. I ask others for help when I don't understand something.
26. I can motivate myself to learn when I need to
27. I am aware of what strategies I use when I study.
28. I find myself analyzing the usefulness of strategies while I study.
29. I use my intellectual strengths to compensate for my weaknesses.
30. I focus on the meaning and significance of new information.
31. I create my own examples to make information more meaningful.
32. I am a good judge of how well I understand something.

33. I find myself using helpful learning strategies automatically.
34. I find myself pausing regularly to check my comprehension.
35. I know when each strategy I use will be most effective.
36. I ask myself how well I accomplish my goals once I'm finished.
37. I draw pictures or diagrams to help me understand while learning.
38. I ask myself if I have considered all options after I solve a problem.
39. I try to translate new information into my own words.
40. I change strategies when I fail to understand.
41. I use the organizational structure of the text to help me learn.
42. I read instructions carefully before I begin a task.
43. I ask myself if what I'm reading is related to what I already know.
44. I reevaluate my assumptions when I get confused.
45. I organize my time to best accomplish my goals.
46. I learn more when I am interested in the topic.
47. I try to break studying down into smaller steps.
48. I focus on overall meaning rather than specifics.
49. I ask myself questions about how well I am doing while I am learning something new.
50. I ask myself if I learned as much as I could have once I finish a task.
51. I stop and go back over new information that is not clear.
52. I stop and reread when I get confused.

APPENDIX B: GOAL ORIENTATION QUESTIONNAIRE

Instructions: The following statements represent types of goals that you may or may not have for this class. Circle a number to indicate how true each statement is of you (1 – "Not true of me", 7 – "Extremely true of me"). All your responses will be kept anonymous and confidential. There are no right or wrong responses, so please be open and honest.

Task-approach goal items

To get a lot of questions right on the exams in this class.
To know the right answers to the questions on the exams in this class.
To answer a lot of questions correctly on the exams in this class.

Task-avoidance goal items

To avoid incorrect answers on the exams in this class.
To avoid getting a lot of questions wrong on the exams in this class.
To avoid missing a lot of questions on the exams in this class.

Self-approach goal items

To perform better on the exams in this class than I have done in the past.
To do well on the exams in this class relative to how well I have done in the past.
To do better on the exams in this class than I typically do in this type of situation.

Self-avoidance goal items

To avoid doing worse on the exams in this class than I normally do on these types of exams.
To avoid performing poorly on the exams in this class compared to my typical level of performance.
To avoid doing worse on the exams in this class than I have done on prior exams of this type.

Other-approach goal items

To outperform other students on the exams in this class.
To do well compared to others in the class on the exams.
To do better than my classmates on the exams in this class.

Other-avoidance goal items

To avoid doing worse than other students on the exams in this class.
To avoid doing poorly in comparison to others on the exams in this class.
To avoid performing poorly relative to my fellow students on the exams in this class.

APPENDIX C: CODEBOOK

Codes	Definitions	Examples
Perceptions on AR	Participant overall perceptions about the AR game and environment.	
AR game control	Participant comments on game control including keyboard and mouse.	Epsilon: then, I got stuck, then, I was, like, I need to go straighter. Gamma: I actually not quite sure how to use control. I used to use the space (key for triggering the console).
AR game glitch	The glitch that participants encountered during the game play.	Beta: yeah it crushed, so I had to do it again. Lota: No, not really [getting stuck some point at somewhere].
Positive aspects	Participant positive comments on AR.	Lota: I felt like it was pretty straightforward once you walked around and saw what each room was for, and had the main room, and then you had the room where you got the information about the aliens, and then you know you could send a probe if you needed to, or design a probe and then go send a probe. I feel like it was very clear what each room was for. Mu: I felt like more were just kind, like, like, I had more fun than I thought I would for a game it's just pretty much reading.
Negative aspects	Participant negative comments on AR.	Kappa: Just twice [got stucked in the environment]. It did [brother me], because I was… I mean… I just had to go back, but it wasn't that bad. Omega: Eh, poor graphics [when I first enter this environment].

Obstacles in AR	The obstacles participants encountered in AR	Lota: I did keep forgetting where my tools were, so I thought that I had to go to the main one, to like, look at my solar system and look at my tools, that kind of stuff, but I figured it out eventually that I didn't have to do that. Chi: I was pretty confused in the beginning. It's been quite a bit of time and I guess that's mainly my fault because I don't think…. I don't think I read many of the instructions well as I should have.
Redundant Information	The information that in the environment, but is not helpful for solving the problem in AR	Alpha: Yeah, no [not helpful for solving the problem]… but, like, I want to make sure that I was thorough in trying to find what I could. Delta: I think there was, like, I think one of the tools it was, like, missions or something that I clicked on, it was, like, not really relevant to the task.
Problem-solving process	Learner problem-solving process during the study. It approximately lasted 60 minutes, and included the following 10 steps.	
Step 1. Identify the problem correctly	Whether the participant identify the problem correctly–helping Jakala-Tay to find a suitable home.	Beta: Yay, and then I was gonna, like, I did, when I did, is gonna show you, because I started doing it on another alien, after you said, then I was, like, coming back. Delta: Yeah. I was, yeah, this is where, like, where I pretty much started in find information [about Jakala-Tay], or, like, finding my answers first.
Step 2. Explore the 3D environment by visiting all rooms in AR and look over all tools	Participants explore the 3D environment by visiting all the room in AR, and looking over all the tools.	Alpha: I was just trying to get a sense of, like, what everything looked, like, how everything worked. Beta: Yeah, I wanted to look around first, to see where kinda I was, to see what I can do, or where I would go.
Step 3. Discover	Participants visit Alien	Delta: I think it's been a good, like, a good amount of my time was spent in that room. That

what Jakala-Tay needs to survive in Alien Database	Database to find out what Jakala-Tay needs to survive in new environments.	[Alien Database] was the most. It was just most…, I think the thought process was just finding information about, like, this species and then…, about, like, just both the species themselves and their home planet. Gamma: I just read about the elements that they need, temperature, and the gravity, so I just read them all.
Step 4. Search the Solar System Database for possible planets	Participants search the Solar System Database for possible planets	Chi: I was looking and then to see which planets…, sort of certain kinds of gases that existed. Kappa: So then I went through each one and I was, like, reading, oh that's what I felt like that's what took the most time cause I read one now, okay, does this one have sulfur? does it have hydrogen? does it have nitrogen?
Step 5. Develop hypotheses about where Jakala-Tay can live	Participants come up with hypotheses about where Jakala-Tay can live.	Lota: I saw that it [IO] had volcanic eruptions which is OK for them because they have volcanic eruptions on planet, on their home planet. They experienced that so it is not like it would be anything very different to them to experience volcano eruptions, and then it didn't have hydrogen so it already met two of my requirements. You know? And so it just needed a little bit of nitrogen, so they can maybe grow their plant life. Omega: I knew it was gonna be IO or Venus, one of those two.
Step 6. Figure out if there is any missing information needed for making a decision	Participants figure out if there is any missing information for making a final decision.	Delta: it was, was comparison, and then after that I…umm, that's why I use other tools to, like, further, like, validate and see if, like, it was (correct). Lota: I already felt like I knew the answer but I went… I hadn't used the other two rooms yet, you know, so I designed a lander probe to go to IO and collect the spectroscopy… spectroscopy of it.
Step 7. Launch	Participants launch a	Epsilon: yeah, I just like, I ruled out that everything that wasn't compatible and then I sent

probes to gather the information in the Probe Design room	probe to gather the information in Probe Design Room.	a probe to the ones that, like, didn't right away get scratched off. Zeta: I eventually send the probes to Mars, like, I saw on the, like, the atmosphere levels that I had a lot of nitrogen.
Step 8. Check the data from the probe in the Mission Control room	Participants check the data from the probe in the Mission Control Room.	Alpha: So I was looking at the missions (control), I was, like, I don't think this is really gonna help me, so I left. Beta: I went back to… like, the alien, no…this one, Mission Control room…, and I saw if it is hydrogen, I was, like, okay.
Step 9. Decide whether the selected planet is a good choice for Jakala-Tay	Participants decide if selected planet is a good choice for Jakala-Tay	Omega: I thought IO was a good second choice, but the problem with IO was that it didn't have nitrogen for their food. Theta: I liked the fact that you had to really cross-check all the facts make sure something was correct before you were spending. Double-check everything. It was kind of intriguing saying I think this is because I initially had two options and I ended up going with a different thing than my initial plan.
Step 10. If so, write a recommendation message with the justification in the Communication Center—if not, go back to step 4.	If so, participants write a recommendation message with the justification in the Communication Center. If not, go back to step 4.	Mu: I would have chosen to send a recommendation and I kind of did both IO and Venus, but I wasn't sure if the game would actually allow me to actually send multiple recommendations. So I just wrote kinda what I thought. Theta: I kind of had it figured out earlier then I wanted to double-check and I'm glad I did, because that's when I switched to IO and I thought I had that one figured out, but I kept checking for another 10 minutes probably.
Metacognition	Participant knowledge and cognition about cognitive phenomena, including both metacognitive knowledge and metacognitive regulation. There are four types of metacognitive knowledge, including: tasks, self, strategies and interactions. Metacognitive regulation ran through the entire problem-solving process in three phases, including	

	forethought, performance, and self-reflection. In the forethought phase, there are two major categories: task analysis and self-motivation belief. During task analysis, problem-solvers would engage in goal setting and strategic planning.	
Metacognitive knowledge—tasks	Participant knowledge about how the nature of the task influences the task performance	Epsilon: Uh-emm, technically, but not really, like, I don't know anything about the planets moons, but I know, like, that Jupiter has…it is a gas planet. You know? A very basic information. Lota: Especially the… the probes and stuff, and satellites, like, I'm already kind of knew, what each one was for, what it found, but I'm not like an expert, I'm taking [astronomy] as an elective, I'm not like an expert or anything.
Metacognitive knowledge—self	Participant knowledge about one's own skills, strengths, and weaknesses	Alpha: I did, just because, like, I knew I was gonna be really weak as far as my natural planet knowledge. So I was like, maybe if I have, like, really, really extensive notes, I could finally match things up. Lota: I'm a senior now, so when I first started, you know, all the professors tell you, oh you, like, write everything down everyone brings our laptops and types their notes, and types everything that you know, everything on slides everything and I don't do that anymore, I listen. because that's how I absorb everything best… I've learned a lot about myself in the way that I learned.
Metacognitive knowledge—strategies	Participant knowledge regarding the alternative strategies for performing the task	Epsilon: I went back because I was, like, hmm, let me see if there's anything that I forgot about them, like I wanted to double-check with the notes I'd taken. It crosses the… uh… I'd like to see it compared to the planet, you know? Lota: Yeah, it was just reading and kind of process of elimination, what they couldn't have and what they needed to have.
Metacognitive knowledge—	Participant knowledge about the preceding	Lota: I think it made me be able to go, like, to skim more, I was able to skim a little more because, like, I saw down here, concepts, like, I didn't need that I didn't need anything in

interactions	types of knowledge interact with one another to influence the outcome of cognitive performance	there, because I know gravity and the fact that it has and I know what the spectra is for and stuff like that. Zata: Uh.. maybe, but I, like, I feel like Mars is a good fix, cause I feel like just saying, like, the colors and stuff, like, what it liked was kind of just like an obvious fit. and like the other ones just, like, previous knowledge I did not know they would really work out, I'm not really know too much about, like, the other planets, they are not as commonly talked about, so I was just, like, I'm just gonna talk about Mars, because I feel, like, that's like what I am familiar with, like, I feel, like, he'd be fine with Mars.
Goal Orientation	Participant general purposes toward learning either in AR or their class, which includes task-approach (TAP), task-avoidance (TAV), self-approach (SAP), self-avoidance (SAV), other-approach (OAP), and other-avoidance (OAV) goal orientations.	
TAP	Participant goal regarding finishing the task.	Alpha: So… um… like, for one thing I'm a pretty competitive person even just with myself. So what's one of those things were, like, especially in video games, I just, I really want to do it well, one of those things I'm looking through it. I'm, like… uh… I should have paid more attention in science class. Then I will just have the natural talent. Beta: I didn't look at her screen just because, like, I was kind of, like, focused on trying to make sure that I could figure it out because I know that, that was a big thing for it. Also, I did want, I wanted to make sure that where I was, like,… where I was going with it. I did not want to be influenced, that anyone be, like, "oh, she did this," you know, changing yourself, I just don't.
TAV	Participant goal regarding avoiding incorrect answers.	Chi: So after I figured that out, I visited another room I just wanted to visit each room before I… [submit a solution]. Delta: Yeah, it was a lot in there. I think I was, like, a little bit, like, hesitant to, like, start sending out probes cause I guess was just… I mean, I knew that, like, when I noticed that

		the budget was, like, kind of insanely high, I was kind of, like, I just don't want to waste too much time because, like… I don't know… it was a weird, like… I did as if there was real kind of things, so I didn't want to, like, you know, having wasted insane amount of money just, like… sending all these probes.
SAP	Participants goal regarding wanting to perform better compare to self.	Epsilon: Not really, cause, like, when I saw people were leaving, I was, like, I still have, like, 17 minutes, you know, so and then I think after this, I went and did a probe. Kappa: Actually, no, I was looking at the this guy's (another student) screen. Because he… he was writing a lot of notes and then so I was, like, thinking in my head, okay, maybe, I should write notes, but then when I started trying to figure it out for myself then I realized okay, I don't need to, I don't need to write that many notes but I just remember he was writing a lot of notes, and so I was, like, okay maybe I'm doing something wrong.
SAV	Participant goal regarding avoiding performing worse compare to self.	Kappa: If I were gonna do a flyby, there are certain things that, there are certain instruments that only work for flybys, so I was, like, I'm not gonna waste my money putting that if I'm gonna land on there, so I tried to, like… you know… make the best use of my money even though I had a lot of money. Lota: I always, like, if I'm working on an assignment, I always double-check, like, the key requirements and when it's due, and I look at the grading rubric, and, like, the checklist, kinda.
OAP	Participant goal regarding wanting to outperform peers.	Theta: It was challenging and kind of fun because you had lots of options right away. You could even submit an answer immediately which puts a little pressure on time knowing that the faster you are. Well, I assume that's better problem solving wise. I don't know what your average time is. Does people usually finish this in less time?
OAV	Participants goal regarding avoiding	Mu: yeah just double-checking, a lot of my time. I was like, I should… yeah, so I guess if I have extra time, I would usually go back and check things.

	doing worse than peers.	Omega: Just checking some more, like, it's just… it's just, like, you know, the double, triple checking, making sure everything is right.

References

Abt, C. C. (1970). *Serious games*. New York, NY: The Viking Press.

Akin, A., Abaci, R., & Çetin, B. (2007). The validity and reliability of the Turkish version of the metacognitive awareness inventory. *Educational Sciences: Theory & Practice*, *7*(2), 671-678.

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, *68*(1), 52-81.

Alyaz, Y., Spaniel-Weise, D., & Gursoy, E. (2017). A Study on Using Serious Games in Teaching German as a Foreign Language. *Journal of Education and Learning*, *6*(3), 250-264. doi:10.5539/jel.v6n3p250

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, *84*(3), 261-271. doi:10.1037/0022-0663.84.3.261

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of educational psychology*, *80*(3), 260-267. doi:10.1037/0022-0663.80.3.260

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M. T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, *53*(1), 74-85. doi:10.1016/j.compedu.2008.12.020

Archer, J. (1994). Achievement goals as a measure of motivation in university students. *Contemporary educational psychology*, *19*(4), 430-446. doi:10.1006/ceps.1994.1031

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, *29*(3), 344–370. doi: 10.1016/j.cedpsych.2003.09.002

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, *40*(1), 193-211. doi:10.1007/s11251-011-9167-4

Barendregt, W., & Bekker, T. M. (2011). The influence of the level of free-choice learning activities on the use of an educational computer game. *Computers & Education*, *56*(1), 80-90. doi:10.1016/j.compedu.2010.08.018

Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, *28*(1), 51-75. doi: 10.1177/0265532210376379

Barron, K. E., & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: testing multiple goal models. *Journal of personality and social psychology*, *80*(5), 706-722. doi:10.1037/0022-3514.80.5.706

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, *1996*(68), 3-12.

Barrows, H. S. (2000). *Problem-based learning applied to medical education*. Chicago, IL: Southern Illinois University School of Medicine.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*. New York, NY: Springer Publishing Company.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press.

Bendixen, L. D., & Hartley, K. (2003). Successful learning with hypermedia: The role of epistemological beliefs and metacognitive awareness. *Journal of Educational Computing Research, 28*(1), 15-30.

Bereby-Meyer, Y., & Kaplan, A. (2005). Motivational influences on transfer of problem-solving strategies. *Contemporary Educational Psychology*, *30*(1), 1-22. doi:10.1016/j.cedpsych.2004.06.003

Berkson, L. (1993). Problem-based learning: Have the expectations been met? *Academic Medicine*, *68*(10), S79-88.

Bogard, T., Liu, M., & Chiang, Y. H. V. (2013). Thresholds of knowledge development in complex problem solving: A multiple-case study of advanced learners' cognitive processes. *Educational Technology Research and Development*, *61*(3), 465-503. doi:10.1007/s11423-013-9295-4

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., ... & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, *94*, 178-192. doi:10.1016/j.compedu.2015.11.003

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C: National Academy Press.

Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver: A guide for improving thinking, learning, and creativity*. New York, NY: Freeman.

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior*, *21*(3), 487-508. doi:10.1016/j.chb.2004.10.005

Brondino M., Raccanello D., & Pasini M. (2014) Achievement Goals as Antecedents of Achievement Emotions: The 3 X 2 Achievement Goal Model as a Framework for Learning Environments Design. In T. D. Mascio, R. Gennari, P. Vitorini, R. Vicari, F. de la Prieta (Eds.), *Methodologies and Intelligent Systems for Technology Enhanced Learning* (pp. 53-60). Cham, Switzerland: Springer. doi:10.1007/978-3-319-07698-0_7

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1, pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher, 18*(1), 32-42.

Brush, T., & Saye, J. (2001). The use of embedded scaffolds with hypermedia-supported student-centered learning. *Journal of Educational Multimedia & Hypermedia,* 10, 333–356.

Bulu, S. T., & Pedersen, S. (2012). Supporting problem-solving performance in a hypermedia learning environment: The role of students' prior knowledge and metacognitive skills. *Computers in Human Behavior*, *28*(4), 1162-1169. doi:10.1016/j.chb.2012.01.026

Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational behavior and human decision processes*, *67*(1), 26-48. doi:10.1006/obhd.1996.0063

Calderhead, J. (1981). Stimulated recall: A method for research on teaching. *British Journal of Educational Psychology*, *51*(2), 211-217. doi:10.1111/j.2044-8279.1981.tb02474.x

Cheng, M. T., Su, T., Huang, W. Y., & Chen, J. H. (2014). An educational game for learning human immunology: What do students learn and how do they perceive? *British Journal of Educational Technology*, *45*(5), 820-833. doi:10.1111/bjet.12098

Chowdhry, S. (2016). Student's perception of effectiveness of a technology enhanced problem based learning environment in a Mechanical Engineering module. *Journal on Today's Ideas - Tomorrow's Technologies, 4*(1), 15-32. doi:10.15415/jotitt.2016.41002

Ciani, K. D., & Sheldon, K. M. (2010). Evaluating the mastery-avoidance goal construct: A study of elite college baseball players. *Psychology of Sport and Exercise*, *11*(2), 127-132. doi:10.1016/j.psychsport.2009.04.005

Clark, R., Kirschner, P. A., & Sweller, J. (2012). Putting students on the path to learning: The case for fully guided instruction. *American Educator*, Spring 2012. 6-11.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Colliver, J. A. (2000). Effectiveness of problem-based learning curricula: Research and theory. *Academic Medicine*, *75*(3), 259-266. doi: 10.1097/00001888-200003000-00017

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661-686. doi:10.1016/j.compedu.2012.03.004

Corliss, S. B. (2005). *The effects of reflective prompts and collaborative learning in hypermedia problem-based learning environments on problem solving and metacognitive skills* (Doctoral dissertation). Retrieved from ProQuest. (Access No. 3187842)

Cunningham, D., & Duffy, T. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 170-198). New York, NY: Simon and Schuster.

Dahlgren, M. A., & Dahlgren, L. O. (2002). Portraits of PBL: Students' experiences of the characteristics of problem-based learning in physiotherapy, computer engineering and psychology. *Instructional Science*, *30*(2), 111-127. doi:10.1023/A:1014819418051

Daniels, L. M., Haynes, T. L., Stupnisky, R. H., Perry, R. P., Newall, N. E., & Pekrun, R. (2008). Individual differences in achievement goals: A longitudinal study of cognitive, emotional, and achievement outcomes. *Contemporary Educational Psychology*, *33*(4), 584-608. doi: 10.1016/j.cedpsych.2007.08.002

David, A. P. (2012). Structural validation of the 3 x 2 achievement goal model. *Educational Measurement and Evaluation Review*, *3*, 50-59.

David, A. P. (2014). Analysis of the separation of task-based and self-based achievement goals in a Philippine sample. *Psychological Studies*, *59*(4), 365-373. doi:10.1007/s12646-014-0266-6

Davidson, J. E., & Sternberg, R. J. (1998). Smart problem solving: How metacognition helps. In D. J. Hacker, J. Dunlosky. A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 47-68). Abingdon, UK: Routledge

deChambeau, A. L., & Ramlo, S. E. (2017). STEM high school teachers' views of implementing PBL: An investigation using anecdote circles. *Interdisciplinary Journal of Problem-Based Learning*, *11*(1), 1-13. doi:10.7771/1541-5015.1566

Dengfeng, L., & Chuntian, C. (2002). New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. *Pattern Recognition Letters, 23*(1), 221-225. doi: S0167-8655(01)00110-6

Diseth, Å. (2015). The advantages of task-based and other-based achievement goals as standards of competence. *International Journal of Educational Research*, *72*, 59-69. doi:10.1016/j.ijer.2015.04.011

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and instruction*, *13*(5), 533-568. doi:10.1016/S0959-4752(02)00025-7

Dolnicar, S, A Review of Unquestioned Standards in Using Cluster Analysis for Data-driven Market Segmentation, *CD Conference Proceedings of the Australian and New Zealand Marketing Academy Conference 2002 (ANZMAC 2002),* Deakin University, Melbourne, 2-4 December 2002.

Dominowski, R. L. (1998). Metacognition in educational theory and practice, In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds.), *Verbalization and problem solving* (pp.25-45). Abingdon, UK: Routledge.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040-1048. doi:10.1037/0003-066X.41.10.1040

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*(2), 256-273. doi:10.1037/0033-295X.95.2.256

Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *The power of problem-based learning: A practical" how to" for teaching undergraduate courses in any discipline*. Sterling, VA: Stylus Publishing.

Duffy, T. M., & Jonassen, D. H. (2013). Constructivism: New implications for instructional technology. In T.M. Duffy & D. H. Jonassen (Eds.). *Constructivism and the technology of instruction: A conversation* (pp. 1-16). Abingdon, UK: Routledge.

Elliot, A. J. (1994). *Approach and avoidance achievement goals: An intrinsic motivation analysis.* Unpublished doctoral dissertation. University of Wisconsin-Madison, Madison, WI.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational psychologist*, *34*(3), 169-189. doi:10.1207/s15326985ep3403_3

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. doi:10.1037/0022-3514.72.1.218

Elliot, A. J., & Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. *Journal of personality and social psychology*, *66*(5), 968-980. doi:10.1037/0022-3514.66.5.968

Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 achievement goal framework. *Journal of personality and social psychology*, *80*(3), 501-519. doi:10.1037//0022-3514.80.3.501

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of educational psychology*, *91*(3), 549-563. doi:10.1037/0022-0663.91.3.549

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613-628. doi:10.1037/0022-0663.100.3.613

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3× 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632-648. doi: 10.1037/a0023952

Elliot, A., Murayama, K., Kobeisy, A., & Lichtenfeld, S. (2015). Potential-based achievement goals. *British Journal of Educational Psychology*, *85*(2), 192-206. doi:10.1111/bjep.12051

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of personality and social psychology*, *54*(1), 5-12. doi:10.1037/0022-3514.54.1.5

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, *67*, 156-167. doi:10.1016/j.compedu.2013.02.019

Ertmer, P. A., & Glazewski, K. D. (2015). Essentials for PBL implementation: Fostering collaboration, transforming roles, and scaffolding learning. In A. E. Walker, H. Leary, C. E. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential readings in problem-based learning Exploring and Extending the Legacy of Howard S. Barrows* (pp.89-106). West Lafayette, IN: Purdue University Press.

Ertmer, P. A., & Newby, T. J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional science*, *24*(1), 1-24. doi:10.1007/BF00156001

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43-71. doi:10.1002/piq

Ertmer, P. A., & Simons, K. D. (2006). Jumping the PBL implementation hurdle: Supporting the efforts of K–12 teachers. *Interdisciplinary Journal of Problem-based learning*, *1*(1), 40-54. doi:10.7771/1541-5015.1005

Evensen, D. H., Salisbury-Glennon, J. D., & Glenn, J. (2001). A qualitative study of six medical students in a problem-based curriculum: Toward a situated model of self-regulation. *Journal of Educational Psychology*, *93*(4), 659-676. doi:10.1037/0022-0663.93.4.659

Flajolet, P., & Noy, M. (2000). Analytic combinatorics of chord diagrams. In D. Krob, A.A. Mikhalev, A.V. Mikhalev (Eds.), *Formal Power Series and Algebraic Combinatorics* (pp. 191-201). Springer, Berlin: Heidelberg. doi: 10.1007/978-3-662-04166-6_17

Flanagan, M. J., Putwain, D. W., & Caltabiano, M. L. (2015). The Relationship Between Goal Setting and Students' Experience of Academic Test Anxiety. *International Journal of School & Educational Psychology*, *3*(3), 189-201. doi:10.1080/21683603.2015.1060910

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist, 34*(10), 906-911. doi:10.1037/0003-066X.34.10.906

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21-29). Hillsdale, NJ: Lawrence Erlbaum Associates.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fortunato, I., Hecht, D. Title, C. K. & Alvarez, L. (1991). Metacognition and problem Solving. *Arithmetic teacher, 39*(4), 226-229.

Foster, A. N., & Mishra, P. (2009). Games, claims, genres, and learning. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 33-50). Hershey, PA: Information Science Reference.

Freer, P. (2017). Problem-Based Learning and Structural Redesign in a Choral Methods Course. *Contributions to Music Education*, *42*, 53-72.

Gagne, R. (1985). *The Conditions of Learning and Theory of Instruction*. New York, NY: Holt, Rinehart and Winston.

Gauld, A., & Stephenson, G. M. (1967). Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology*, *58*(1-2), 39-49.

Ge, X., Law, V., & Huang, K. (2016). Detangling the Interrelationships Between Self-Regulation and Ill-Structured Problem Solving in Problem-Based Learning. *Interdisciplinary Journal of Problem-Based Learning, 10*(2). 1-14. doi:10.7771/1541-5015.1622

Gee, J. P. (2007). *Good video games+ good learning: Collected essays on video games, learning, and literacy*. New York, NY: Peter Lang Publishing.

Gillet, N., Lafrenière, M. A. K., Huyghebaert, T., & Fouquereau, E. (2015). Autonomous and controlled reasons underlying achievement goals: Implications for the 3× 2 achievement goal model in educational and work settings. *Motivation and Emotion*, *39*(6), 858-875. doi:10.1007/s11031-015-9505-y

Glesne, C. (2015). *Becoming qualitative researchers: An introduction.* Boston, MA: Pearson.

Gourgey, A. F. (1998). Metacognition in basic skills instruction. *Instructional science*, *26*(1), 81-96. doi:10.1023/A:1003092414893

Graafland, M., Schraagen, J. M., & Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. *British journal of surgery*, *99*(10), 1322-1330. doi:10.1002/bjs.8819

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). circlize implements and enhances circular visualization in R. *Bioinformatics*, *30*(19), 2811-2812.

Gul, F., & Shehzad, S. (2012). Relationship between metacognition, goal orientation and academic achievement. *Procedia-Social and Behavioral Sciences*, *47*, 1864-1868. doi:j.sbspro.2012.06.914

Hacker, D. J. (1998) Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.). *Metacognition in educational theory and practice* (pp. 1-23). Hillsdale, NJ: Lawrence Erlbaum Associates.

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), 1284-1295. doi:10.1037/0022-3514.73.6.1284

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. doi:10.1207/s15326985ep3301_1

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of Achievement Goal Theory: Necessary and Illuminating. *Journal of Educational Psychology*, *94*(3), 638-645. doi:10.1037//0022-0663.94.3.638

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*(2), 316-330. doi:10.1037/0022-0663.92.2.316

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of educational psychology*, *100*(1), 105-122. doi:0.1037/0022-0663.100.1.105

Harr, R., Buch, T., & Hanghøj, T. (2008). Exploring the discrepancy between educational goals and educational game design. In T. Connolly & M.Stanfield (Eds.), *Proceedings of the 2nd European Conference on Games-based Learning* (pp. 16-17). Sonning Common, England: Academic Publishing International.

Hmelo, C. E. (1998). Problem-based learning: Effects on the early acquisition of cognitive skill in medicine. *The Journal of the Learning Sciences*, *7*(2), 173-208. doi:10.1207/s15327809jls0702_2

Hmelo-Silver, C. E. (2002). Collaborative ways of knowing: Issues in facilitation. In G. Stahl (Ed,), *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community* (pp. 199-208). Seattle, WA: International Society of the Learning Sciences.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235-266. doi:10.1023/B:EDPR.0000034022.16470.f3

Hmelo-Silver, C. E., & Barrows, H. S. (2015). Problem-based learning: Goals for learning and strategies for facilitating. In A. E. Walker, H. Leary, C. E. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential readings in problem-based learning Exploring and Extending the Legacy of Howard S. Barrows* (pp. 69-84). West Lafayette, IN: Purdue University Press.

Hsieh, P. P. H., Cho, Y., Liu, M., & Schallert, D. L. (2008). Examining the interplay between middle school students' achievement goals and self-efficacy in a technology-enhanced learning environment. *American Secondary Education*, *36*(3), 33-50.

Horton. L. (2014). *The effects of problem-based learning scaffolds on cognitive load, problem-solving, and student performance within a multimedia- enhanced* *learning environment* (Doctoral dissertation). Retrieved from UT Electronic Theses and Dissertations. (http://hdl.handle.net/2152/25003)

Hou, H. T., & Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior*, *30*, 29-38. doi:10.1016/j.chb.2013.07.052

Howard, B. C., McGee, S., Shia, R., & Hong, N. S. (2000, April). *Metacognitive self-regulation and problem-solving: Expanding the theory base through factor analysis*. Paper presented at the annual meeting of the American educational research association, New Orleans, LA.

Hu, J., & Gao, X. A. (2017). Using think-aloud protocol in self-regulated reading research. *Educational Research Review*, *22*, 181-193. doi: 10.1016/j.edurev.2017.09.004

Huang, W. H., Huang, W. Y., & Tschopp, J. (2010). Sustaining iterative game playing processes in DGBL: The relationship between motivational processing and outcome processing. *Computers & Education*, *55*(2), 789-797. doi:10.1016/j.compedu.2010.03.011

Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. T. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning*, *25*(4), 332-344. doi: 10.1111/j.1365-2729.2009.00316.x

Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage.

IBM. (2017a). *Smarter Games for Military.* Retrieved from https://www-935.ibm.com/services/us/gbs/gaming/military/

IBM. (2017b). *Smarter Games for Government*. Retrieved from https://www-935.ibm.com/services/us/gbs/gaming/government/

Jagacinski, C. M., & Duda, J. L. (2001). A comparative analysis of contemporary achievement goal orientation measures. *Educational and Psychological Measurement*, *61*(6), 1013-1039. doi:10.1177/00131640121971626

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, *31*(8), 651-666.

Jang, L. Y., & Liu, W. C. (2012). 2× 2 Achievement goals and achievement emotions: a cluster analysis of students' motivation. *European Journal of Psychology of Education*, *27*(1), 59-76. doi: 10.1007/s10212-011-0066-5

Jenkins, H., Klopfer, E., Squire, K., & Tan, P. (2003). Entering the education arcade. *Computers in Entertainment (CIE)*, *1*(1), 1-11. doi: 10.1145/950566.950591

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college what evidence is there that it works?. *Change: the magazine of higher learning, 30*(4), 26-35.

Johnson, M. L., & Kestler, J. L. (2013). Achievement goals of traditional and nontraditional aged college students: using the 3× 2 achievement goal framework. *International Journal of Educational Research*, *61*, 48-59. doi:10.1016/j.ijer.2013.03.010

Johnson, W. L., Vilhjalmsson, H., & Marsella, S. (2005). Serious games for language learning: How much game, how much AI? In C.-K. Looi, G. McCalla, B. Bredeweg (Eds.), *Artificial Intelligence in Education: Supporting Learning Through Intelligent and Socially Informed Technology* (pp. 306-313), Amsterdam, Netherlands: IOS Press.

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm?. *Educational technology research and development, 39*(3), 5-14.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational technology research and development*, *48*(4), 63-85. doi:10.1007/BF02300500

Kang, J. (2017). *Examining scientific thinking processes in open-ended serious games through gameplay data* (Doctoral dissertation). Retrieved from UT Electronic Theses and Dissertations. (http://hdl.handle.net/2152/61557)

Kang, J., Liu, S., & Liu, M. (2017). Tracking students' activities in serious games. In F. Lai, J. D. Lehman (Eds.), *Learning and knowledge analytics in open education* (pp. 125-137). New York, NY: Springer International Publishing.

Kaplan, A., & Flum, H. (2010). Achievement goal orientations and identity formation styles. *Educational Research Review*, *5*(1), 50-67. doi:10.1016/j.edurev.2009.06.004

Kaplan, A., Gheen, M., & Midgley, C. (2002). Classroom goal structure and student disruptive behaviour. *British journal of educational psychology*, *72*(2), 191-211. doi:10.1348/000709902158847

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational psychology review*, *19*(2), 141-184. doi:10.1007/s10648-006-9012-5

Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary educational psychology*, *28*(1), 37-58. doi:10.1016/S0361-476X(02)00012-7

Kasim, R. M. (1999). What can studies of problem-based learning tell us? Synthesizing and modeling PBL effects on national board of medical examination performance: Hierarchical linear modeling meta-analytic approach. *Advances in Health Sciences Education*, *4*(3), 209-221.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education*, *51*(4), 1609-1620. doi: 10.1016/j.compedu.2008.03.003

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, *16*(1), 99-111. doi:10.1007/s10956-006-9038-y

Kickmeier-Rust M.D., Hillemann E., Albert D. (2011) Tracking the UFO's Paths: Using Eye-Tracking for the Evaluation of Serious Games. In Shumaker R. (Ed.), *Virtual and Mixed Reality - New Trends* (pp. 315-224). Berlin, Heidelberg: Springer

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: In search of the meaning of cognitive tools. *Instructional Science*, *35*(3), 207-256.

doi:10.1007/s11251-006-9005-2

Kirschner, P. A., & Drachsler, H. (2012). Learner Characteristics. In N. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1743-1745). New York, NY: Springer US. doi:10.1007/978-1-4419-1428-6_347

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75-86. doi:10.1207/s15326985ep4102_1

Klopfer, E., & Haas, J. (2012). *The more we know: NBC News, educational innovation, and learning from failure*. Cambridge, MA: MIT Press.

Klopfer, E., & Yoon, S. (2005). Developing games and simulations for today and tomorrow's tech savvy youth. *TechTrends*, *49*(3), 33-41. doi: 10.1007/BF02763645

Krebs, P., Burkhalter, J. E., Snow, B., Fiske, J., Ostroff, J. S., & Eysenbach, G. (2013). Development and Alpha Testing of QuitIT: An Interactive Video Game to Enhance Skills for Coping With Smoking Urges. *Journal of Medical Internet Research Research Protocols*, *2*(2), e35. doi:10.2196/resprot.2416

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, *77*(6), 1121-1134.

Land, S. M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, *48*(3), 61-78.

doi:10.1007/BF02319858

Latham, G. P., & Brown, T. C. (2006). The effect of learning vs. outcome goals on self-Efficacy, satisfaction and performance in an MBA program. *Applied Psychology*, *55*(4), 606-623. doi:10.1111/j.1464-0597.2006.00246.x

Latham, G. P., & Locke, E. A. (1979). Goal setting—A motivational technique that works. *Organizational Dynamics, 8*(2), 68-80.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational behavior and human decision processes, 50*(2), 212-247.

Lebow, D. (1993). Constructivist values for instructional systems design: Five principles toward a new mindset. *Educational technology research and development*, *41*(3), 4-16.

Lee, J. (2015). *Effects of fantasy and fantasy proneness on learning and engagement in a 3D educational game* (Doctoral dissertation). Retrieved from UT Electronic Theses and Dissertations. http://hdl.handle.net/2152/30927

Lee, C. Y., & Chen, M. P. (2009). A computer game as a context for non-routine mathematical problem solving: The effects of type of question prompt and level of prior knowledge. *Computers & Education*, *52*(3), 530-542. doi:10.1016/j.compedu.2008.10.008

Lee, H. W., Lim, K. Y., & Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research and Development, 58*(6), 629-648. doi:10.1007/s11423-010-9153-6

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., ... & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem-solving: PISA 2003 Results From the US Perspective.* (NCES 2005-003). Washington, DC: US Department of Education, National Center for Education Statistics.

Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B. W., & Lobene, E. V. (2014). Designing game-based learning environments for elementary science education: A narrative-centered learning perspective. *Information Sciences*, *264*, 4-18. doi:10.1016/j.ins.2013.09.005

Lin, L. F. (2017). Impacts of the Problem-based Learning Pedagogy on English Learners' Reading Comprehension, Strategy Use, and Active Learning Attitudes. *Journal of Education and Training Studies*, *5*(6), 109-125. doi:10.11114/jets.v5i6.2320

Linnenbrink, E. A., Pintrich, P. R. (2001). Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses. In V. Simone, J. Sanna, (Eds.). *Motivation in learning contexts: Theoretical advances and methodological implications* (pp. 251-269). Elmsford, NY: Pergamon Press.

Liu, M. (2004). Examining the performance and attitudes of sixth graders during their use of a problem-based hypermedia learning environment. *Computers in Human Behavior, 20*(3), 357-379.

Liu, M. (2005). The effect of a hypermedia learning environment on middle school students' motivation, attitude, and science knowledge. *Computers in the Schools*, *22*(3-4), 159-171.

Liu, M., & Bera, S. (2005). An analysis of cognitive tool use patterns in a hypermedia learning environment. *Educational Technology Research and Development*, *53*(1), 5-21.

Liu, M., Bera, S., Corliss, S. B., Svinicki, M. D., & Beth, A. D. (2004). Understanding the connection between cognitive tool use and cognitive processes as used by sixth graders in a problem-based hypermedia learning environment. *Journal of Educational Computing Research*, *31*(3), 309-334. doi:10.2190/LK2G-8K25-RB8U-PGE9

Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, *55*(2), 630-643. doi:10.1016/j.compedu.2010.02.023

Liu, M., Hsieh, P., Cho, Y., & Schallert, D. L. (2006). Middle school students' self-efficacy, attitudes, and achievement in a computer-enhanced problem-based learning environment. *Journal of Interactive Learning Research*, *17*(3), 225-242.

Liu, M., Horton, L. R., Corliss, S. B., Svinicki, M. D., Bogard, T., Kim, J., & Chang, M. (2009). Students' problem solving as mediated by their cognitive tool use: A study of tool use patterns. *Journal of Educational Computing Research*, *40*(1), 111-139.

Liu, M., Horton, L., Lee, J., Kang, J., Rosenblum, J., O'Hair, M., & Lu, C. W. (2014). Creating a multimedia enhanced problem-based learning environment for middle school science: Voices from the developers. *Interdisciplinary Journal of Problem-Based Learning*, *8*(1), 1-14. doi:10.7771/1541-5015.1422

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development*, *59*(2), 249-265. doi:10.1007/s11423-011-9192-7

Liu, M., Hsieh, P., Cho, Y., & Schallert, D. L. (2006). Middle school students' self-efficacy, attitudes, and achievement in a computer-enhanced problem-based learning environment. *Journal of Interactive Learning Research*, *17*(3), 225-242.

Liu, M., Kang, J., Lee, J., Winzeler, E., & Liu, S. (2015). Examining through visualization what tools learners access as they play a serious game for middle school science. In C. S. Loh, Y. Sheng, & Ifenthaler (Eds.), *Serious Games Analytics* (pp. 181-208). New York, NY: Springer International Publishing.

Liu, M., Lee, J., Kang, J., & Liu, S. (2016). What we can learn from the data: A multiple-case study examining behavior patterns by students with different characteristics in using a serious game. *Technology, Knowledge and Learning*, *21*(1), 33-57. doi: 10.1007/s10758-015-9263-7

Liu, M., Toprac, P., & Yuen, T. T. (2009). What factors make a multimedia learning environment engaging: A case study. In R. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 173-192). Hershey, PA: IGI Global. doi: 10.4018/978-1-60566-158-2.ch010

Lo, M. T., Chen, S. K., & Lin, S. S. (2017). Groups holding multiple achievement goals in the math classroom: Profile stability and cognitive and affective outcomes. *Learning and Individual Differences*, *57*, 65-76. doi: 10.1016/j.lindif.2017.06.001

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Eaglewood 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*(9), 705-717.

Locke, E. A., & Latham, G. P. (2015). Breaking the rules: a historical overview of goal-setting theory. *Advances in Motivation Science, 2*(15), 99-126.

Loh, C. S., Li, I. H., & Sheng, Y. (2016). Comparison of similarity measures to differentiate players' actions and decision-making profiles in serious games analytics. *Computers in Human Behavior*, *64*, 562-574. doi:10.1016/j.chb.2016.07.024

Loh, C. S., & Sheng, Y. (2014). Maximum Similarity Index (MSI): A metric to differentiate the performance of novices vs. multiple-experts in serious games. *Computers in Human Behavior*, *39*, 322-330.

Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. In C. S. Loh, Y. Sheng, & Ifenthaler (Eds.), *Serious games analytics* (pp. 3-29). New York, NY: Springer International Publishing.

Lou, S. J., Shih, R. C., Diez, C. R., & Tseng, K. H. (2011). The impact of problem-based learning strategies on STEM knowledge integration and attitudes: An exploratory study among female Taiwanese senior high school students. *International Journal* *of Technology and Design Education*, *21*(2), 195-215. doi:10.1007/s10798-010-9114-8

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, *29*(6), 861-878. doi: 10.1080/0141192032000137349

Maitland, B. (1997). Problem-based learning for architecture and construction management. In D. Boud, G. Feletti (Eds.), *The challenge of problem-based learning* (pp. 211-217). Abingdon, UK: Routledge.

Maltais, C., Duchesne, S., Ratelle, C. F., & Feng, B. (2015). Attachment to the mother and achievement goal orientations at the beginning of middle school: The mediating role of academic competence and anxiety. *Learning and Individual Differences*, *39*, 39-48. doi:10.1016/j.lindif.2015.03.006

Marra, R. M., Jonassen, D. H., Palmer, B., & Luft, S. (2014). Why problem-based learning works: Theoretical foundations. *Journal on Excellence in College Teaching*, *25*(3-4), 221-238.

Masek, A., & Yamin, S. (2011). The effect of problem based learning on critical thinking ability: A theoretical and empirical review. *International Review of Social Sciences and Humanities*, *2*(1), 215-221.

Maurer, H., & Neuhold, C. (2014). Problem-Based Learning in European Studies. In S. Baroncelli., R.Farneti., I. Horga., S. Vanhoonacker (Eds.), *Teaching and learning the European union* (pp. 199-215). Netherlands: Springer.

Maxwell, N. L., Bellisimo, Y., & Mergendoller, J. (2001). Problem-based learning: Modifying the medical school model for teaching high school economics. *The Social Studies*, *92*(2), 73-78. doi:10.1080/00377990109603981

Mayer, R. E. (2013). Problem-solving. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 769-779). Oxford, UK: Oxford University Press.

Mayer, R. E., Griffith, E., Jurkowitz, I. T., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, *14*(4), 329-339. doi: 10.1037/a0013835

Mayo, M. J. (2007). Games for science and engineering education. *Communications of the ACM*, *50*(7), 30-35. doi:10.1145/1272516.1272536

Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology*, *3*(2), 51-59. doi:10.20982/tqmp.03.2.p051

Mavridis, A., Katmada, A., & Tsiatsos, T. (2017). Impact of online flexible games on students' attitude towards mathematics. *Educational Technology Research and Development*, 1-20. doi:10.1007/s11423-017-9522-5

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement-relevant processes prior to task engagement. *Journal of Educational Psychology*, *94*(2), 381-395. doi:10.1037/0022-0663.94.2.381

Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of educational psychology*, *80*(4), 514-523. doi:10.1037/0022-0663.80.4.514

Meece, J. L., & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of educational psychology*, *85*(4), 582-590. doi:10.1037/0022-0663.85.4.582

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Abingdon, UK: Routldege.

Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The Creative Experiences Questionnaire (CEQ): a brief self-report measure of fantasy proneness. *Personality and Individual Differences*, *31*(6), 987–995. doi: 10.1016/S0191-8869(00)00201-4

Michael, D. R., & Chen, S. L. (2006). *Serious games: Games that educate, train, and inform*. Boston, MA: Thomson Course Technology.

Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of educational psychology*, *89*(4), 710-718. doi:10.1037/0022-0663.89.4.710

Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., ... & Roeser, R. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary educational psychology*, *23*(2), 113-131. doi:10.1006/ceps.1998.0965

Mihalca, L., Mengelkamp, C., & Schnotz, W. (2017). Accuracy of metacognitive judgments as a moderator of learner control effectiveness in problem-solving tasks. *Metacognition and Learning*, 1-23.

Miller, R. B., Behrens, J. T., Greene, B. A., & Newman, D. (1993). Goals and perceived ability: Impact on student valuing, self-regulation, and persistence. *Contemporary Educational Psychology*, *18*(1), 2-14. doi:10.1006/ceps.1993.1002

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054. doi: 10.1111/j.1467-9620.2006.00684.x

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of educational psychology, 94*(2), 249-259. doi:10.1037/0022-0663.94.2.249

Morganti, L., Pallavicini, F., Cadel, E., Candelieri, A., Archetti, F., & Mantovani, F. (2017). Gaming for Earth: Serious games and gamification to engage consumers in pro-environmental behaviours for energy efficiency. *Energy Research & Social Science*, *29*, 95-102. doi:10.1016/j.erss.2017.05.001

Moshman, D. (2017). Metacognitive Theories Revisited. *Educational Psychology Review*, 1-8. doi: 10.1007/s10648-017-9413-7

Nazry, N. N. M., & Romano, D. M. (2017). Mood and learning in navigation-based serious games. *Computers in Human Behavior*, *73*, 596-604. doi:10.1016/j.chb.2017.03.040

NESTA Futurelab. (2004). *Literature review in games and learning: A NESTA Futurelab Research report - report 8*. Bristol, UK: Kirriemuir, J., & McFarlane, A.

Neufeld, V. R., & Barrows, H. S. (1974). The" McMaster Philosophy": an approach to medical education. *Academic Medicine*, *49*(11), 1040-50.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-hall.

Newman, M. (2003). *A pilot systematic review and meta-analysis on the effectiveness of problem based learning*. Newcastle, UK: University of Newcastle, Learning and Teaching Support Network.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological review*, *91*(3), 328-346. doi:10.1037/0033-295X.91.3.328

Nicholls, J. G., Patashnick, M., & Nolen, S. B. (1985). Adolescents' theories of education. *Journal of Educational Psychology*, *77*(6), 683-692. doi:10.1037/0022-0663.77.6.683

Ning, H. K. (2016). Psychometric Properties of the 3×2 Achievement Goal Questionnaire in a Hong Kong Sample. *Journal of Psychoeducational Assessment*, 1-12. doi:10.1177/0734282916677658

Norman, G. R., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic medicine*, *67*(9), 557-565.

Oliveira, A. M. C. A., dos Santos, S. C., & Garcia, V. C. (2013, October). PBL in teaching computing: An overview of the last 15 years. *Frontiers in Education Conference, 2013 IEEE,* 267-272. doi:10.1109/FIE.2013.6684830

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, *13*(1), 48-63.

Organization for Economic Cooperation and Development (OECD). (2014). *PISA 2012 results: creative problem-solving: Students' skills in tackling real-life problems* (volume V). Retrieved from http://www.oecd.org/pisa/keyfindings/PISA-2012-results-volume-V.pdf

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, *8*(2), 1-9.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary educational psychology*, *20*(4), 426-443. doi:10.1006/ceps.1995.1029

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education*, *52*(1), 1-12. doi: 10.1016/j.compedu.2008.06.004

Pedersen, S., Liu, M., & Williams, D. (2002). Alien Rescue: Designing for student-centered learning. *Educational Technology*, *42*(5), 11-14.

Pedersen, S. J. (2000). *Cognitive modeling during problem-based learning: The effects of a hypermedia expert tool* (Doctoral dissertation). Retrieved from http://ezproxy.lib.utexas.edu/login?url=http://search.proquest.com/docview/304622266?accountid=7118.

Perkins, A. C. (2016). *Earthquake: Game-Based Learning for 21st Century STEM Education* (Doctoral dissertation). Retrieved from Texas A & M University library. http://oaktrust.library.tamu.edu/handle/1969.1/157955

Perkins, D. (1999). The many faces of constructivism. *Educational leadership, 57*(3), 6-11.

Perini, S., Margoudi, M., Oliveira, M. F., & Taisch, M. (2017). Increasing middle school students' awareness and interest in manufacturing through digital game-based learning (DGBL). *Computer Applications in Engineering Education*. 2017, 1-15. doi:10.1002/cae.21836

Perrotta, C., Featherstone, G., Aston, H., & Houghton, E. (2013). *Game-based learning: Latest evidence and future directions* (NFER Research Programme: Innovation in Education). Slough, UK: NFER.

Perry, J., & Klopfer, E. (2014). UbiqBio: adoptions and outcomes of mobile biology games in the ecology of school. *Computers in the Schools*, *31*(1-2), 43-64. doi:10.1080/07380569.2014.879771

Phillips, F. (2001). A research note on accounting students' epistemological beliefs, study strategies, and unstructured problem-solving performance. *Issues in Accounting Education*, *16*(1), 21-39. doi:10.2308/iace.2001.16.1.21

Piaget, J. (1953). *To understand is to invent.* New York, NY: Grossman.

Pintrich, P. R. (2000a). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation: Theory, research and applications* (pp. 451-502). San Diego, CA: Academic. doi:10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2000b). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of educational psychology*, *92*(3), 544-555. doi:10.I037//0022-O663.92.3.544

Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire* (Technical Report 91-B-004). Ann Arbor, MI: The Regents of the University of Michigan.

Poitras, E., Lajoie, S., & Hong, Y. J. (2012). The design of technology-rich learning environments as metacognitive tools in history education. *Instructional science, 40*(6), 1033-1061. doi:10.1007/s11251-011-9194-1

Porath, C. L., & Bateman, T. S. (2006). Self-regulation: from goal orientation to job performance. *Journal of Applied Psychology*, *91*(1), 185-192. doi:10.1037/0021-9010.91.1.185

Porter, R. L., & Latham, G. P. (2013). The effect of employee learning goals and goal commitment on departmental performance. *Journal of Leadership & Organizational Studies*, *20*(1), 62-68.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Prensky, M. (2001). Fun, play and games: What makes games engaging. *Digital game-based learning*, *5*, 1-31.

Reese, D. D., & Tabachnick, B. G. (2010). *The moment of learning: Quantitative analysis of exemplar gameplay supports CyGaMEs approach to embedded assessment.* Washington, D.C: Society for Research on Educational Effectiveness, 2010. Retrived from http://files.eric.ed.gov/fulltext/ED513110.pdf

Reese, D. D., Tabachnick, B. G., & Kosko, R. E. (2015). Video game learning dynamics: Actionable measures of multidimensional learning trajectories. *British Journal of Educational Technology*, *46*(1), 98-122. doi: 10.1111/bjet.12128

Resnick, L. B. (1987). The 1987 presidential address learning in school and out. *Educational researcher, 16*(9), 13-54.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational technology research and development*, *44*(2), 43-58. doi:10.1007/BF02300540

Rogers, S. A. (2017). *A massively multiplayer online role-playing game with language learning strategic activities to improve English grammar, listening, reading, and vocabulary* (Doctoral dissertation, University of South Alabama). Retrieved from ProQuest.

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments.

International Journal of Artificial Intelligence in Education, *21*(1-2), 115-133. doi:10.3233/JAI-2011-019

Ryan, T. A. (1970). *Intentional behavior: An approach to human motivation*. New York, NY: Ronald Press.

Sánchez, J., & Olivares, R. (2011). Problem solving and collaboration using mobile serious games. *Computers & Education*, *57*(3), 1943-1952. doi:10.1016/j.compedu.2011.04.012

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. In A. E. Walker, H. Leary, C. E. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential readings in problem-based learning Exploring and Extending the Legacy of Howard S. Barrows* (pp. 5-15). West Lafayette, IN: Purdue University Press.

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, *35*(5), 31-38.

Schmidt, S. J. (2008). Memory and remembrance: a constructivist approach. In A. Eril, & A. Nünning (Eds.), *Media and Cultural Memory* (pp.191-201). Berlin, Germany: Walter de Gruyter.

Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology*, *56*(2), 405-429. doi:10.1111/j.1744-6570.2003.tb00156.x

Schmidt, H. G., Rotgans, J. I., & Yew, E. H. (2011). The process of problem-based learning: what works and why. *Medical Education*, *45*(8), 792-806. doi:

10.1111/j.1365-2923.2011.04035.x

Schommer-Aikins, M., Duell, O. K., & Hutter, R. (2005). Epistemological beliefs, mathematical problem-solving beliefs, and academic performance of middle school students. *The Elementary School Journal*, *105*(3), 289-304. doi:10.1086/428745

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary educational psychology, 19*(4), 460-475. doi:10.1006/ceps.1994.1033

Segrelles, J. D., Martinez, A., Castilla, N., & Moltó, G. (2017). Virtualized Computational Environments on the cloud to foster group skills through PBL: A case study in architecture. *Computers & Education*, *108*, 131-144. doi:10.1016/j.compedu.2017.02.001

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, *46*(1), 26-47. doi:10.1080/00461520.2011.538646

Senko, C., & Miles, K. M. (2008). Pursuing their own learning agenda: How mastery-oriented students jeopardize their class performance. *Contemporary Educational Psychology*, *33*(4), 561-583. doi: 10.1016/j.cedpsych.2007.12.001

Seijts, G. H., & Latham, G. P. (2005). Learning versus performance goals: When should each be used?. *The Academy of Management Executive*, *19*(1), 124-131. doi: 10.5465/AME.2005.15841964

Seijts, G. H., & Latham, G. P. (2011). The effect of commitment to a learning goal, self-efficacy, and the interaction between learning goal difficulty and commitment on performance in a business simulation. *Human Performance*, *24*(3), 189-204. doi:10.1080/08959285.2011.580807

Seijts, G. H., Latham, G. P., Tasa, K., & Latham, B. W. (2004). Goal setting and goal orientation: An integration of two different yet related literatures. *Academy of management journal*, *47*(2), 227-239. doi: 10.2307/20159574

Seijts, G.H., Latham, G. P., & Woodwark, M. (2013). Learning Goals. In E. A. Locke, G. P. Latham (Eds.), *New developments in goal setting and task performance* (pp. 195-212). New York, NY: Taylor & Francis.

Shim, S., & Ryan, A. (2005). Changes in self-efficacy, challenge avoidance, and intrinsic value in response to grades: The role of achievement goals. *The Journal of Experimental Education*, *73*(4), 333-349. doi:10.3200/JEXE.73.4.333-349

Shin, N., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of research in science teaching*, *40*(1), 6-33. doi:10.1002/tea.10058

Sideridis, G. D. (2008). The regulation of affect, anxiety, and stressful arousal from adopting mastery-avoidance goal orientations. *Stress and Health*, *24*(1), 55-69. doi:10.1002/smi.1160

Smith, M., Duda, J., Allen, J., & Hall, H. (2002). Contemporary measures of approach and avoidance goal orientations: Similarities and differences. *British Journal of Educational Psychology*, *72*(2), 155-190. doi:10.1348/000709902158838

Sosik, J. J., Chun, J. U., & Koul, R. (2017). Relationships Between Psychological Well Being of Thai College Students, Goal Orientations, and Gender. *Psychology in the Schools*. *54*(7), 703-717. doi:10.1002/pits.22024

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and self-regulated learning constructs. *Educational Research and Evaluation, 10*(2), 117-139.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary educational psychology*, *27*(1), 51-79. doi:10.1006/ceps.2001.1091

Spires, H. A., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, *44*(4), 453-472. doi:10.2190/EC.44.4.e

STATA. (2018). Cluster—Introduction to cluster-analysis commands. Retrieved from https://www.stata.com/manuals13/mvcluster.pdf

Stoeber, J., Haskew, A. E., & Scott, C. (2015). Perfectionism and exam performance: The mediating effect of task-approach goals. *Personality and individual differences*, *74*, 171-176. doi:10.1016/j.paid.2014.10.016

Strobel, J., & Van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-based Learning*, *3*(1), 44-58. doi:10.7771/1541-5015.1046

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Thomas, D. I., & Vlacic, L. B. (2012). The business of collaborating: Designing and implementing a group decision-making scenario using the TeamMATE collaborative computer game. In M. M. Cruz-Cunha (Eds.), *Handbook of Research on Serious Games as Educational, Business and Research Tools* (pp. 446-477). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-0149-9.ch023

Toprac, P. K. (2008). *The effects of a problem-based learning digital game on continuing motivation to learn science*. (Doctoral Dissertation) Retrieved from ProQuest.

Tran, C., Smordal, O., & Conley, A. (2016). The Interaction between Design Features and Learners' Goals: A Case Study of a Science Museum Game. *Interaction Design and Architecture* (29), 24-51.

Trespalacios, J. H., & Chamberlin, B. (2012). 21st century learning: The role of serious games. In M. M. Cruz-Cunha (Eds.), *Handbook of Research on Serious Games as Educational, Business and Research Tools* (pp. 782-799). Hershey, PA: IGI Global.

Tsai, M. J., Huang, L. J., Hou, H. T., Hsu, C. Y., & Chiou, G. L. (2016). Visual behavior, flow and achievement in game-based learning. *Computers & Education*, *98*, 115-129. doi: 10.1016/j.compedu.2016.03.011

Utman, C. H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review*, *1*(2), 170-182.

Van Yperen, N. W. (2006). A novel approach to assessing achievement goals in the context of the 2× 2 framework: Identifying distinct profiles of individuals with different dominant achievement goals. *Personality and social psychology bulletin*, *32*(11), 1432-1445. doi:10.1177/0146167206292093

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, *57*(6), 995-1015.

Vedder-Weiss, D. (2017). Teaching Higher and Lower in Mastery Goal Structure: The Perspective of Students. *The Elementary School Journal*, *117*(4), 566-592. doi:10.1086/691584

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning. Studies in Mathematics Education Series: 6.* Bristol, PA: The Falmer Press.

Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wei, H., Wu, S., Zhao, Y., Deng, Z., Ersotelos, N., Parvinzamir, F., ... & Dong, F. (2016). Data Mining, Management and Visualization in Large Scientific Corpuses. In A. El Rhalibi, F. Tian, Z. Pan, B. Liu (Eds.) *International Conference on Technologies for E-Learning and Digital Entertainment* (pp. 371-379). Cham, Switzerland: Springer.

Weinstein C.E. & Mayer R.E. (1986). The teaching of learning strategies. In M Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327), New York, NY: Macmillan.

Williams, D., Lai, G., Ma, Y. & Prejean, L. (2008). Using an Educational Computer Game to Teach History in a Pedagogical Laboratory. In K. McFerrin, R. Weber,

R. Carlsen & D. Willis (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2008* (pp. 1847-1852). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Willingham, D. T. (2007). *Cognition: The thinking animal*. Englewood Cliffs, NJ: Pearson/Prentice Hall.

Winburg, K., Chamberlain, B., Valdez, A., Trujillo, K., & Stanford, T. B. (2016). Impact of math snacks games on students' conceptual understanding. *Journal of Computers in Mathematics and Science Teaching*, *35*(2), 173-193.

Winters, D., & Latham, G. P. (1996). The effect of learning versus outcome goals on a simple versus a complex task. *Group & Organization Management*, *21*(2), 236-250.

Wolters, C. A. (1998). Self-regulated learning and college students' regulation of motivation. *Journal of educational psychology*, *90*(2), 224-235. doi:10.1037/0022-0663.90.2.224

Wolters, C. A., Yu, S. L., & Pintrich, P. R. (1996). The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences, 8*(3), 211-238. doi:10.1016/S1041-6080(96)90015-1

Won, S., Wolters, C. A., & Mueller, S. A. (2017). Sense of Belonging and Self-Regulated Learning: Testing Achievement Goals as Mediators. *The Journal of Experimental Education*, 1-17. doi:10.1080/00220973.2016.1277337

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. doi: 10.1037/a0031311

Yadav, A., Subedi, D., Lundeberg, M. A., & Bunting, C. F. (2011). Problem-based Learning: Influence on students' learning in an electrical engineering course. *Journal of Engineering Education*, *100*(2), 253-280. doi: 10.1002/j.2168-9830.2011.tb00013.x

Yang, Y. T. C. (2012). Building virtual cities, inspiring intelligent citizens: Digital games for developing students' problem solving and learning motivation. *Computers & Education*, *59*(2), 365-377. doi: 10.1016/j.compedu.2012.01.012

Yang, Y., & Cao, L. (2013). Differential influences of achievement approach goals and intrinsic/extrinsic motivation on help-seeking in e-learning. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, *5*(2), 153-169.

Yang, J. C., Quadir, B., & Chen, N. S. (2016). Effects of the Badge Mechanism on Self-Efficacy and Learning Performance in a Game-Based English Learning Environment. *Journal of Educational Computing Research*, *54*(3), 371-394. doi: 10.1177/0735633115620433

Yang, Y., Taylor, J., & Cao, L. (2016). The 3 x 2 Achievement Goal Model in Predicting Online Student Test Anxiety and Help-Seeking. *International Journal of E-Learning & Distance Education*, *31*(1). 1-16.

Yew, E. H., & Schmidt, H. G. (2012). What students learn in problem-based learning: A process analysis. *Instructional Science*, *40*(2), 371-395. doi: 10.1007/s11251-011-

9181-6

Young, A., & Fry, J. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning, 8*(2), 1-10.

Young, J., & Upitis, R. (1999). The microworld of Phoenix Quest: social and cognitive considerations. *Education and Information Technologies*, *4*(4), 391-408. doi:10.1023/A:1009600528811

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, *25*(1), 3-17. doi:10.1207/s15326985ep2501_2

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70. doi:10.1207/s15430421tip4102_2

Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 233-262). New York, NY: Cambridge University Press.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational psychologist, 48*(3), 135-147. doi:10.1080/00461520.2013.794676

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing students' use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614-628. doi:10.2307/1163093

Zimmerman, B. J., & Martinez-Pons, M. (1992). Perceptions of efficacy and strategy use in the self-regulation of learning. In D. H. Schunk & J. L. Meece (Eds.), *Student* *perceptions in the classroom* (pp.185-207). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zimmerman, B. J., & Paulsen, A. S. (1995). Self-monitoring during collegiate studying: An invaluable tool for academic self-regulation. *New directions for teaching and learning, 1995*(63), 13-27. doi:10.1002/tl.37219956305

Zusho, A., Pintrich, P. R., & Cortina, K. S. (2005). Motives, goals, and adaptive patterns of performance in Asian American and Anglo American students. *Learning and Individual differences*, *15*(2), 141-158. doi:10.1016/j.lindif.2004.11.003

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, *38*(9), 25-32. doi:10.1109/MC.2005.297