Computers in Human Behavior 102 (2020) 151–165

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

The impact of learner metacognition and goal orientation on problem-solving in a serious game environment

Sa Liu[a,*], Min Liu[b]

[a] Harrisburg University of Science and Technology, USA
[b] University of Texas, Austin, USA

ARTICLE INFO

Keywords:
Metacognition
Goal orientation
Problem-solving
Serious game
Log data
Similarity measure

ABSTRACT

To understand the impact of learner metacognition and goal orientation on problem-solving, this study investigated 159 undergraduate students' metacognition, goal orientations, and their problem-solving performances and processes in a laboratory setting using a Serious Game (SG) environment that adopts problem-based learning (PBL) pedagogy to teach space science. Utilizing cluster analysis, multiple regression, similarity measure and data visualization, this study analyzed multiple data sources, including computer log data, problem-solving performance scores, and survey data. The results show that both learner metacognition and goal orientation affected problem-solving. The findings of this study offer insights of how learner characteristics impact on problem-solving in SG environments with PBL pedagogy. It also contributes to understanding the design of SG environments to benefit learners based on their metacognitive levels.

1. Introduction

Problem-based learning (PBL) is a learner-centered constructivist instructional method, which embeds student learning processes in solving real-life problems (Barrows & Tamblyn, 1980; Hmelo-Silver, 2004; Savery, 2015). 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 gaming and computing technology, researchers argued that Serious Games (SG) 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). By adopting PBL pedagogy in SG environments, instructional designers and researchers hope to increase learner motivation, enhance learning experiences, improve learning performances and help learners develop critical thinking and problem-solving skills (Hou & Li, 2014; Lee & Chen, 2009; Sánchez & Olivares, 2011).

From a constructivist perspective, learner characteristics are important for understanding individual learning. 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 the learning process. Researchers suggested that metacognition is necessary for students to succeed in PBL (Davidson & Sternberg, 1998; Marra, Jonassen, Palmer, & Luft, 2014; Shin, Jonassen, & McGee, 2003). Without adequate metacognition, learners may have difficulty understanding complex topics in learning environments (e.g., SG), as they may fail to plan, set goals, use effective strategies, and monitor and reflect learning processes during learning (Azevedo, Cromley, & Seibert, 2004; Mayer, Griffith, Jurkowitz, & Rothman, 2008). Goal orientation is another important learner characteristic. Previous studies have shown that student goal orientations critically influences their behavior in SG environments (Hsieh, Cho, Liu, & Schallert, 2008; Liu, 2005; Liu, Kang, Lee, Winzeler, & Liu, 2015). Literature also suggested that goal orientation plays an important role at the earliest stage of learner metacognitive regulation, which can guide the entire metacognitive regulatory processes (Moshman, 2017; Schraw & Dennison, 1994; Zimmerman, 2002, 2013).

Although many important studies in the past four decades have revealed learner metacognition affect learning (e.g., Flavell, 1979, 1987; Mihalca, Mengelkamp, & Schnotz, 2017; Schraw & Dennison, 1994; Shin, Jonassen, & McGee, 2003), few studies have 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

* Corresponding author. Interactive Media Program, 326 Market St, Harrisburg, PA, 17101.
E-mail address: saliu@harrisburgu.edu (S. Liu).

https://doi.org/10.1016/j.chb.2019.08.021
Received 12 April 2019; Received in revised form 7 August 2019; Accepted 24 August 2019
Available online 26 August 2019
0747-5632/

problem-solving processes in SG environments is still unclear. In addition, although Gul and Shehzad (2012) have suggested metacognition and goal orientation together could result in learner academic success, i.e., high GPA, 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, to better understand learner during problem-solving, more research is needed on how metacognition and goal orientation would affect learner problem-solving in SG environments, including both problem-solving performances and processes. This study aims to answer this question by studying college students in a SG environment that adopts the PBL pedagogy to teach space science.

2. Related work

In this section, related work is reviewed, including problem-based learning, serious games, and learner characteristics such as learner metacognition and goal orientation.

2.1. Problem-based learning and serious games

PBL has been used in many disciplines throughout K–16 education all over the world. For instance, in Australia, Maitland (1997) 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 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.

Abt (1970) first 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). 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). Mostly recently, Loh, Sheng, and Ifenthaler (2015) defined “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).

To take advantage of PBL, researchers have adopted PBL pedagogy in SG for different subject matters, such as science, mathematics, computer skills, and so on (Hou & Li, 2014; Lee & Chen, 2009; Liu et al., 2014; Sánchez & Olivares, 2011; Spires, Rowe, Mott, & Lester, 2011). For example, Crystal Land was a SG adopting PBL pedagogy (Spires et al., 2011) to teach 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. In university settings, Hou and Li (2014) designed and developed a SG to help university students to learn personal computer assembly through problem-solving process. In this SG, university students were situated in a locked room, and they had to escape the room by finding all computer components in the room and assembling them correctly within 10 min.

2.2. 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. 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).

In addition, a few studies indicated that learner metacognition affected learning in SG (Bogard, Liu, & Chiang, 2013; Liu et al., 2004; Tsai, Huang, Hou, Hsu, & Chiou, 2016). For example, Liu et al. (2004) exploratory study on sixth graders indicated learner tool use patterns in a SG 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. (2013) study took a closer look on 15 advanced learners (i.e., graduate students) cognitive processes while solving the problem using the same SG. They conducted a cross cluster analysis, and the results indicated that leaner metacognition, particularly the self-regulation element, impacted advanced learner problem-solving in the SG. Recently, 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. They found that students demonstrated different metacognitive controls of visual attention based on their comprehension level.

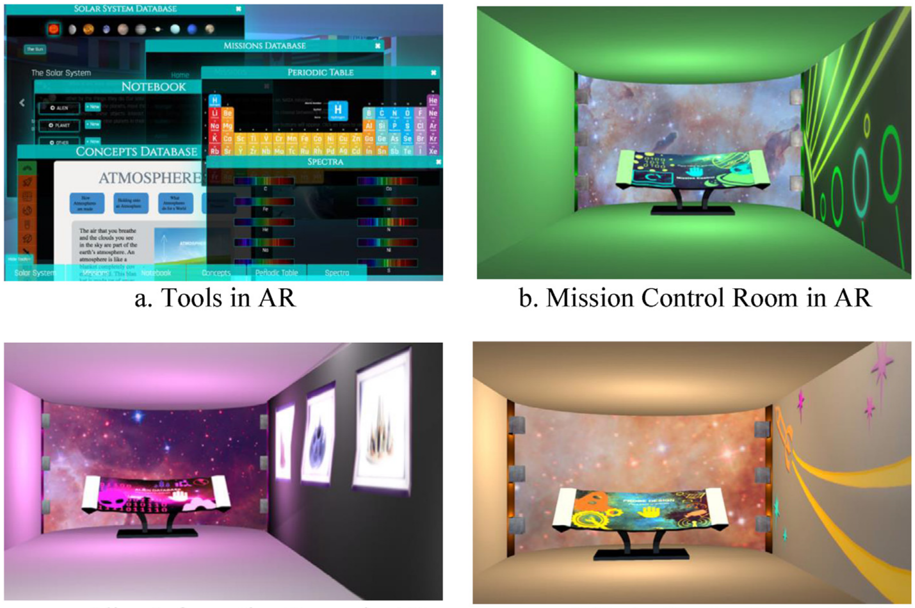
Although it was not clear the causal relationship between metacognition and learning in SG, these preliminary studies suggested that there was a connection between them. Therefore, more research is needed to explore the role of metacognition for learning in SG.

2.3. Goal orientation

Goal orientations state individuals have different reasons or goals for learning (Elliot & Church, 1997; Elliot & McGregor, 2001; Elliot, Murayama, & Pekrun, 2011). Although scholars used different measurements to examine goal orientation, studies indicated that goal orientation affected learn problem-solving. For example, Bereby-Meyer and Kaplan (2005) investigated the effect of children goal orientation on a problem-solving strategy transfer with different task, and suggested, regardless of age and/or perceived ability, participants goal orientation affected the problem-solving strategy transfer.

In addition, goal orientation also affected student achievement and problem-solving behavior within SG (Hsieh et al., 2008; Liu, 2005; Liu et al., 2015; Tran, Smordal, & Conley, 2016). Tran et al. (2016) found that 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.

Furthermore, there are also three studies on a SG called Alien Rescue (AR) and learner goal orientation (Hsieh et al., 2008; Liu, 2005; Liu et al., 2015). Liu (2005) examined 437 sixth-grader goal orientation from a motivational perspective. It adopted *Motivated Strategies for Learning Questionnaire* (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991) to measure student intrinsic and extrinsic goal orientations. The results suggested student science knowledge scores were positively



a. Tools in AR b. Mission Control Room in AR

c. Alien Information Room in AR d. Probe Design Room in AR

Fig. 1. Alien Rescue environment.

related to their intrinsic goal orientation. Hsieh et al. (2008) 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. Results suggested student performance-avoidance goal orientation "moderated the relation between self-efficacy and science achievement" (p. 34).

The most recent study on student goal orientation in AR used *Patterns of Adaptive Learning Scales* (PALS; Midgley et al., 1998) to assess sixth grader mastery, performance-approach, and performance-avoidance goal orientations (Liu et al., 2015). 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 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 and avoidance goal orientation showed an inappropriate behavior pattern, such as exploring more fun tools rather than gathering information to solve the problem.

Besides MSLQ and PALS, most recently, Elliot et al. (2011) constructed a goal orientation model that includes six different goal orientation dimensions, including (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).

2.4. Summary

From a constructivist perspective, learner plays an important role in the learning process; therefore, learner characteristics are important for understanding learning. Literature has shown learner metacognition and goal orientation 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 in a SG environment that adopts PBL pedagogy. In addition, the goal orientation measurement has been further developed into 3×2 measurement (Elliot et al., 2011), which might provide more insights on the impact of goal orientation on problem-solving. Therefore, more research on these topics are needed.

2.5. The current study

To understand learner problem-solving based on metacognition and goal orientation, 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. It employed quantitative research design using multiple data sources to analyze learner problem-solving in a SG environment in a laboratory setting. 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 two sub-questions:
   (a) Is there a statistically significant difference in learner problem-solving performances based upon metacognition and goal orientation?
   (b) Can learner metacognition and goal orientation predict problem-solving performances?
2. To what extent are problem-solving process differences based on learner characteristics (i.e., metacognition and goal orientation)? There are two sub-questions:
   (a) What are learner problem-solving process patterns?
   (b) Are there any problem-solving process pattern differences based on learner metacognition and goal orientation?

3. Material and methods

3.1. Alien Rescue as a serious game environment

This study utilized Alien Rescue (AR, http://alienrescue.edb.utexas.edu; Liu, Lee, Kang, & Liu, 2016) as the SG environment. AR adopts PBL pedagogy to teach middle school students about our solar system and problem-solving. In this environment (see Fig. 1), learners face an ill-structured problem—to save six displaced alien species 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.

There are four rooms in the AR environment including the Main Room, Probe Design Room (Room P), Alien Information Room (Room

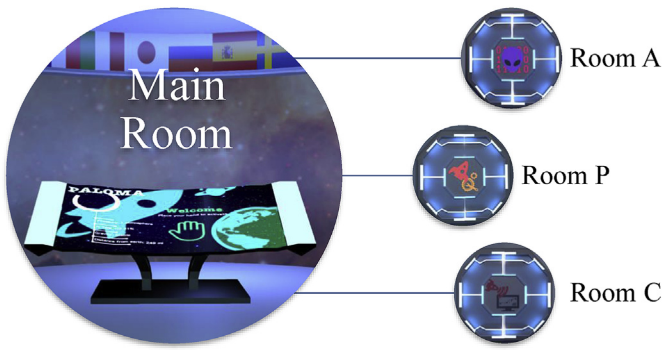


Fig. 2. Alien Rescue room layout.

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 Fig. 2). 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 1 for tool descriptions and locations.

3.2. Participants

The study is approved by Institutional Review Board (IRB) to protect the rights and welfare of participants. The participants were 159 undergraduate students mainly from participant pool at a large public university in the southwestern United States. These participants were majored in various disciplines including education, natural sciences, liberal arts, nursing and fine arts. All participants were new to AR and voluntary to participate 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 $H_1 = 0.15$, $\alpha = 0.20$, $1\text{-}\beta = 0.80$, 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 x 2^k, where k = number of clustering variables (Dolnicar, 2002). This study used multiple types of cluster analysis, and the maximum k is six different goal orientation variables. Therefore, the suggested sample size is 64.

3.3. Procedure

Firstly, in a laboratory setting, participants were given an online survey 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 min 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. By using AR in a laboratory setting, this study hoped to control variables that might affect learner problem-solving processes (e.g., teacher guidance and peer influences in the real classroom).

3.4. Data sources

Multiple data sources were included such as student activity logs, problem-solving performance scores, demographic information, metacognition measurements, and goal orientation measurements.

3.4.1. 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 (i.e., open or close), gate access action (i.e., go through), and problem solution texts. Because no direct teaching or guidance was provided during the study, the computer log data indicates individual participant problem-solving processes in AR. To ensure the reliability, the game log files were consistent over time and samples.

3.4.2. Problem solution scores

Learner problem solution scores were used to measure problem-solving performances, which were evaluated by the quality of learner 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 species, which were evaluated using an 8-point (0–7 points) grading rubric. This rubric has been used in multiple previous AR studies (Bogard et al., 2013; Liu et al., 2009, 2015), see Table 2.

In this study, a few learners have submitted multiple home solutions to Jakala-Tay, or even other Aliens. In these cases, similar to Liu et al.

Table 1
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.

Table 2
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

(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 (2014) approach, the solution scores were scored using the rubric by a panel of three trained raters and reached 100% agreement.

3.4.3. 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%).

Most of the participants 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 Business School.

3.4.4. 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 ($a = 0.90$) and significant correlations between these two components in previous studies ($r = 0.54$ and $r = 0.45$ respectively) (Schraw & Dennison, 1994). The measurement also had a high reliability ($a = 0.95$) using the sample data from this study.

3.4.5. Goal orientation measurement

Learner goal orientation was measured using the *3 X 2 Achievement Goal Orientation Inventory* (Elliot et al., 2011). This has six subscales with high reliability, which are task-approach ($a = 0.84$), task-avoidance ($a = 0.80$), self-approach ($a = 0.77$), self-avoidance ($a = 0.83$), other-approach ($a = 0.93$), and other-avoidance goals ($a = 0.91$). Using data from this study, the reliability numbers were task-approach ($a = 0.85$), task-avoidance ($a = 0.90$), self-approach ($a = 0.81$), self-avoidance ($a = 0.81$), other-approach ($a = 0.92$), and other-avoidance goals ($a = 0.92$). For each subscale, there are 3 items for a total of 18 items. Participants can rate these statements on a 7-point scale (1 = not true of me, 7 = extremely true of me) through the online survey.

3.5. Data analysis

For the first research question, there are two sub-questions. For question 1(a), cluster analyses in SPSS were used to see if there is a statistically significant difference among learner groups based on metacognition and goal orientation. Several researchers suggested that k-means cluster analysis is the most popular technique when exploring participant homogeneous groups (Dolnicar, 2002; 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 1(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:

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

For research question two. *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) (Kang, 2017; Liu et al., 2004, 2009). 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, as 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, Li, & Sheng, 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 based on learner characteristics. In addition, two-proportion z-test was used to decide 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 it was used with two-proportion z-test to identify learner problem-solving process differences in AR.

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 et al., 2016; Loh & Sheng, 2014; Van der Loo, 2014). Based on Loh et al. (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 bigram ($q = 2$) 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. 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. 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), see Table 3.

To conduct similarity measure, firstly, learner room visit sequences in AR during the problem-solving processes were converted into strings to facilitate the analysis. Then the data analysis tool *R* and *stringdist*

Table 3
Coefficient formula for different similarity measures.

Similarity Coefficient	Formula
Jaccard	$\frac{\lvert Q(X;q) \cap Q(Y;q)\rvert}{\lvert Q(X;q) \cup Q(Y;q)\rvert}$
Cosine	$\frac{v(X;q) \cdot v(Y;q)}{\lVert v(X;q) \rVert_2 \lVert v(Y;q) \rVert_2}$
Longest Common Substring	$1 - \frac{d_{LCS}(X \cdot Y)}{d_{\max}(X \cdot Y)}$

package were used to calculate the similarity of learner action sequential strings (Van der Loo, 2014). According to Loh et al. (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. Use the above method, this study calculated the similarity coefficients in learner action sequences among all students based on their clustered groups. Particularly, this study compared all learners to learner who had the highest solution score (i.e., 7 points).

4. Results

4.1. Research question 1(a) learner performance differences based on learner characteristics

To examine if there was a statistically significant difference among learners based on metacognition and goal orientations, k-means cluster analyses were used ($k = 3$).

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. 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, 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 ($M = 3.55$, $SD = 2.704$, $N = 84$).

One-way ANOVA analysis indicated that there were statistically significant differences in learner problem-solving performance based on participant ethnic groups, $F(4, 153) = 3.774$, $p = .006$, $\eta^2 = 0.097$; Levene's Test: $F(4, 153) = 0.944$, $p = .440$. 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 Fig. 3).

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 ($N = 61$), 2) Cluster 2: low metacognition and medium multiple goal orientations ($N = 51$), and 3) Cluster 3: medium metacognition and low multiple goal orientations ($N = 46$). According to one-way ANOVA, learner problem-solving performances were statistically significant based on these three clusters, $F(2, 155) = 11.208$, $p = .000$, $\eta^2 = 0.126$; Levene's Test: $F(2, 155) = 0.989$,

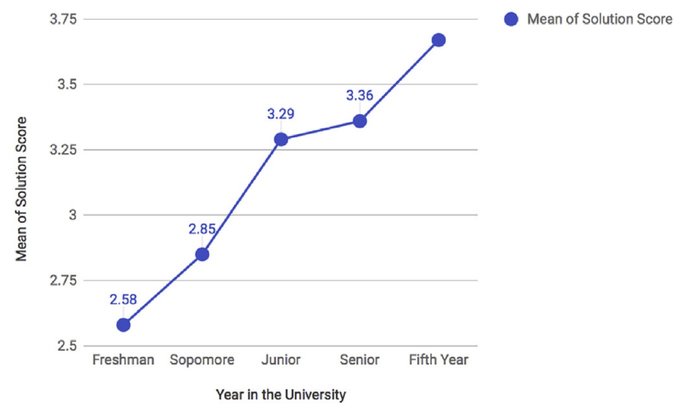


Fig. 3. Learner solution scores based on their year at the university.

$p = .374$. Specifically, learners in Cluster 2 and Cluster 3 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, $p < 0.05$.

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, $F(2, 155) = 9.726$, $p = .000$, $\eta^2 = 0.112$; Levene's Test: $F(2, 155) = 0.711$, $p = .493$ and gender $F(2, 155) = 11.148$, $p = .000$, $\eta^2 = 0.126$; Levene's Test: $F(2, 155) = 1.092$, $p = .338$.

4.2. Research question 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). After eliminating three outliers, assumptions of linearity, reliability of measurement, homoscedasticity, multicollinearity and normality for multiple regression were met using the data (Osborne & Waters, 2002). The Cronbach's Alpha for this study were task-approach ($a = 0.85$), task-avoidance ($a = 0.90$), self-approach ($a = 0.81$), self-avoidance ($a = 0.81$), other-approach ($a = 0.92$), other-avoidance goals ($a = 0.92$), and metacognition ($a = 0.95$).

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, $R^2 = 0.134$, $F(7, 155) = 3.283$, $p < .01$. 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 4 for the regression results.

Table 4
Summary of multiple regression results.

Predictor	Solution Score (N = 156)							
	Pearson Correlation	B	R^2	β	Tolerance	VIF	t	Sig.
Model			.134**				5.253	.000
Metacognition	−0.19**	−0.025		−0.122	0.902	1.108	−1.517	.000
TAP	−0.277***	−0.34		−0.299**	0.457	2.186	−2.644	.009
TAV	−0.073	0.252		0.275*	0.366	2.731	2.180	.031
SAP	−0.211**	0.006		0.006	0.346	2.893	.048	.961
SAV	−0.207**	−0.229		−0.259	0.252	3.967	−1.701	.091
OAP	−0.112	−0.003		−0.004	0.48	2.083	.969	.969
OAV	−0.045	0.065		0.097	0.406	2.465	.418	.418

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Although the regression model was considered significant, it only had a small (R^2 = 0.134), which indicated this model could only predict 13.4% of the data. Therefore, the researchers modified the regression model based on previous cluster analysis. Specifically, using the three final cluster groups as variables to predict learner problem-solving performances. 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 = 0.445, F(2, 129) = 51.7, $p < .000$. The regression equation was as follows:

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

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 5 for the regression results.

4.3. Research question 2(a)problem-solving process differences: visualizing learner problem-solving process patterns

In addition to examine the problem-solving performance, this study visualized three types of learner problem-solving process patterns including learner tool use frequency (Liu & Bera, 2005; Liu et al., 2004), duration (Liu et al., 2009, 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. 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 (2361 times), see Fig. 4.

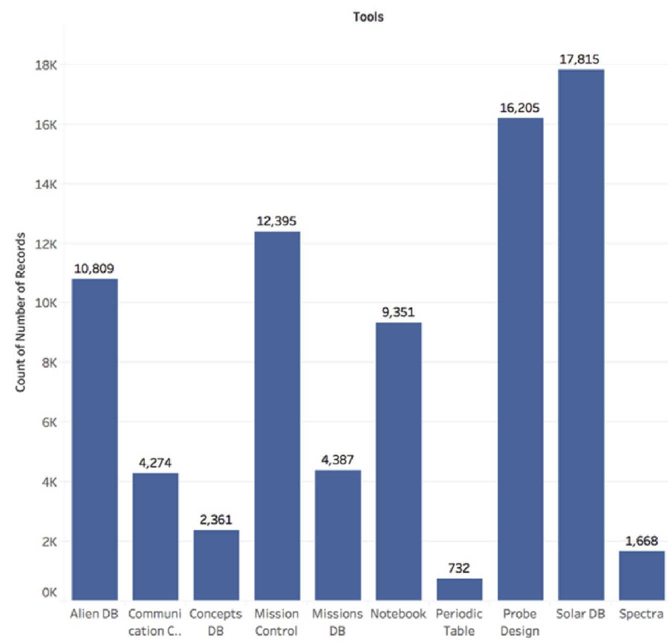


Fig. 4. Tool use frequency among all learners.

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-min of problem-solving in AR. Female student average also used Probe Design 28.4 times fewer than male student (see Fig. 5).

To further understand tool use frequency pattern, Tableau was used to visualize learner average tool use frequency based on solution score. The visualization shows that the highest-scoring student used Solar System Database more frequently than the lowest scoring students, while the lowest scoring students used Alien Database, Communication

Table 5
Summary of multiple regression results using a new model.

Predictor	Solution Score (N = 132)							
	Pearson Correlation	B	R^2	β	Tolerance	VIF	t	Sig.
Model			.445***				16.510	.000
Final_MC	−0.571***	-.928		−0.318***	0.648	1.542	−3.897	.000
Final_GO	−0.616***	−1.314		0.428***	0.648	1.542	−5.251	.000

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

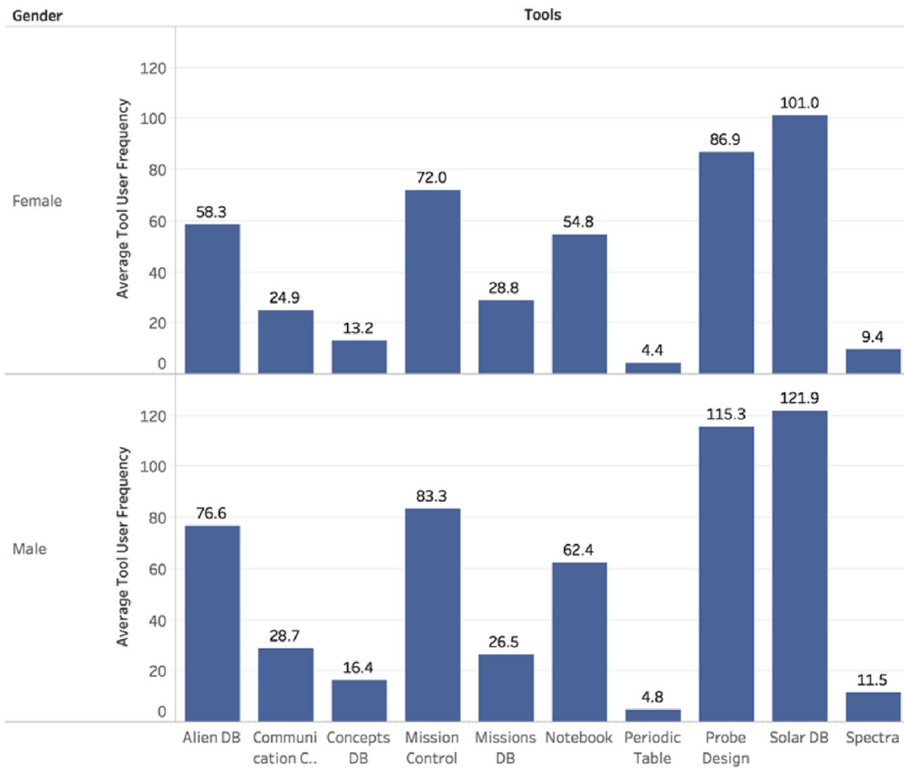


Fig. 5. Tool use average frequency based on gender.

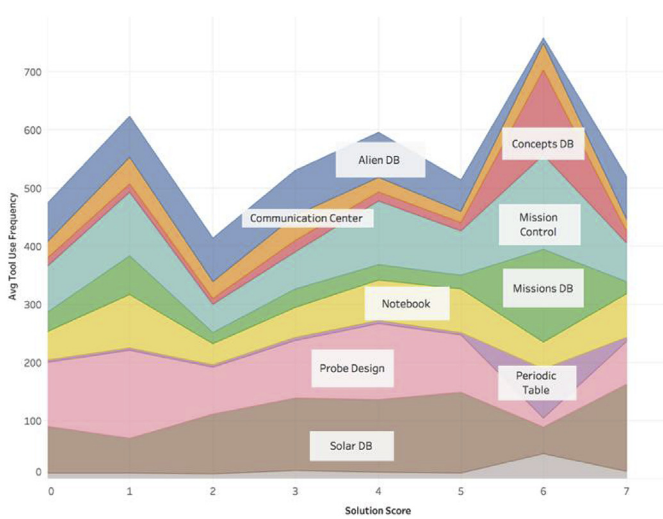


Fig. 6. Tool use frequency based on solution scores.

Center, Mission Control, and Probe Design more frequently than the highest-scoring students, see Fig. 6.

Tool use duration: Besides tool use frequency among all learners, Tableau was also used to visualize tool use duration for all learners. During 60-min 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 min (see Fig. 7). 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.

Considering the gender factor, data visualization showed male

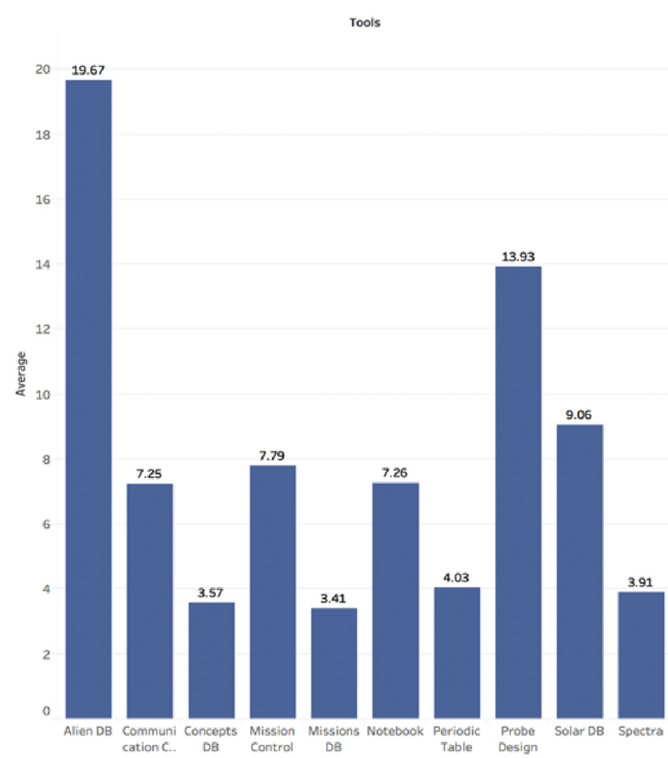


Fig. 7. Tool use duration averages for all learners.

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; $M = 6.77$, 6.69, 3.17 for female). Both genders stayed in Communication Center and Periodic Table about the same amount of time—about 7 and 4 min (see Fig. 8).

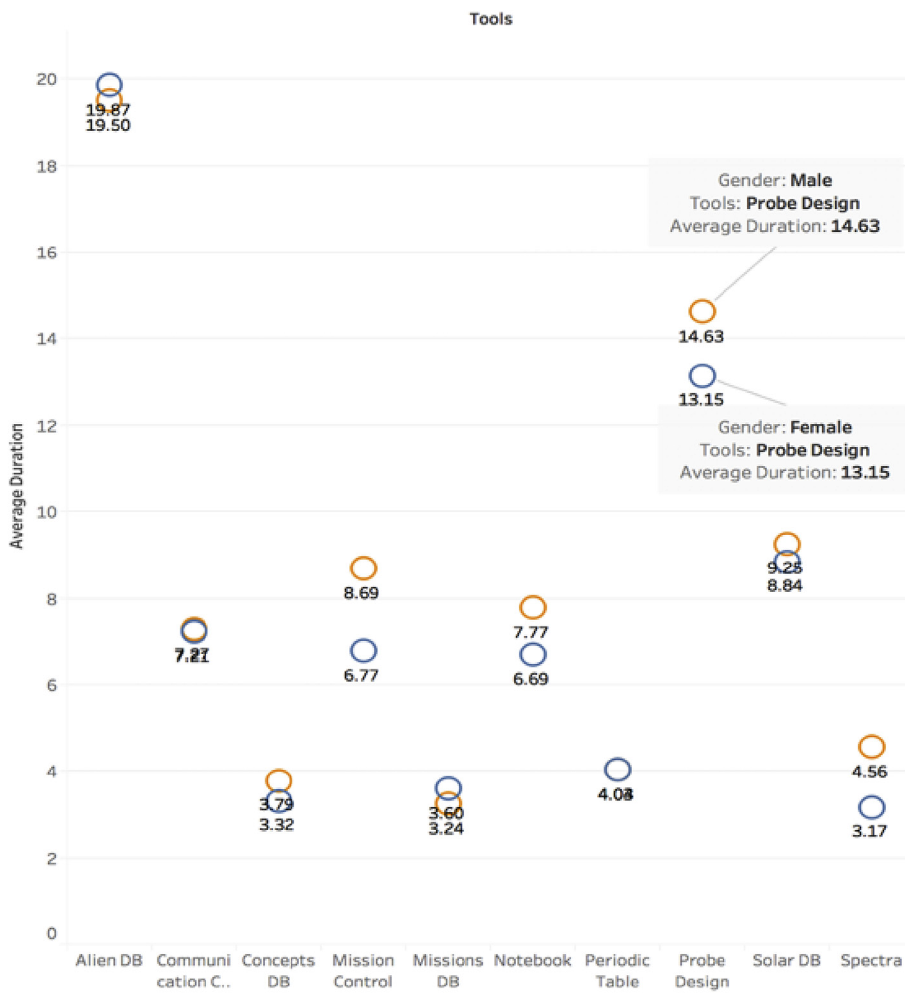


Fig. 8. Tool use duration averages based on gender.

This study also examined the tool use duration differences based on learner problem-solving performances (i.e., solution scores), see Fig. 9. The horizontal axis indicated learner solution score groups, which contained 8 groups (0–7 points). The vertical axis indicated the average minutes of each tool use duration. According to Fig. 9, 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 min). 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).

Room visit sequences: Besides analyzing tool use duration and frequency, this study adopted similarity measure. Particularly, software R was used 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 6 for the average similarity measure coefficients based on learner groups.

Using a simple line graph in R, this study visualized the similarity among learners (see Fig. 10). 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.

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.

4.4. Research question 2(b) problem-solving process patterns based on learner characteristics

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

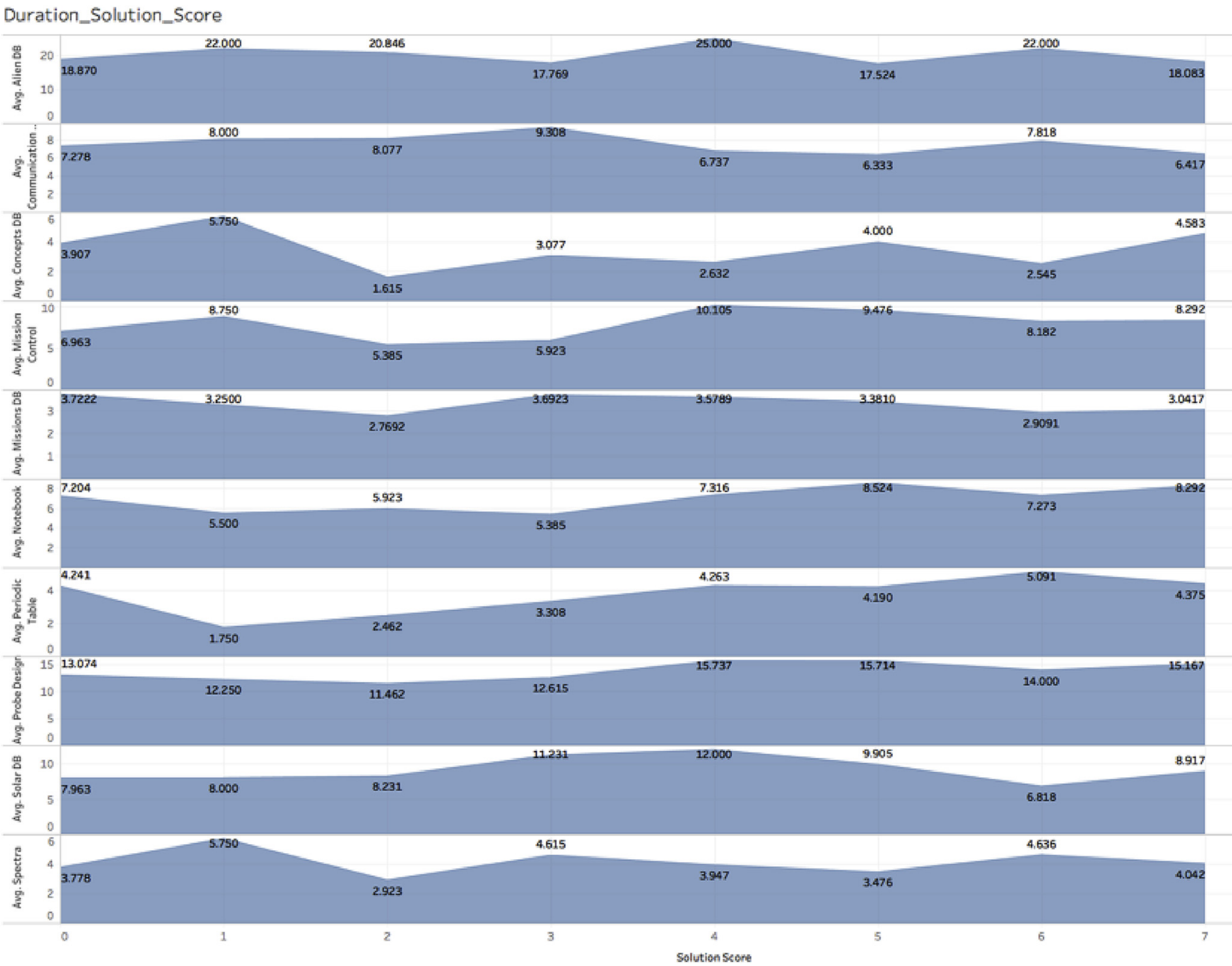


Fig. 9. Tool use duration averages based on solution scores.

Table 6
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

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 7 for detailed learner average tool use frequency in each group.

To help better understand the data, Chord Diagram was used to visualize the data in Table 7. 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 Fig. 11). 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.

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 8). Just as previously analyzed, the data was based on the average tool use duration minutes of each cluster group.

To help better understand the data, Chord Diagram was used to visualize tool use duration based on final cluster (see Fig. 12). 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 min), while learners in Cluster 1 (i.e., high metacognition and high multiple goal orientations) had the shortest tool use duration (i.e., 77.72 min). 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

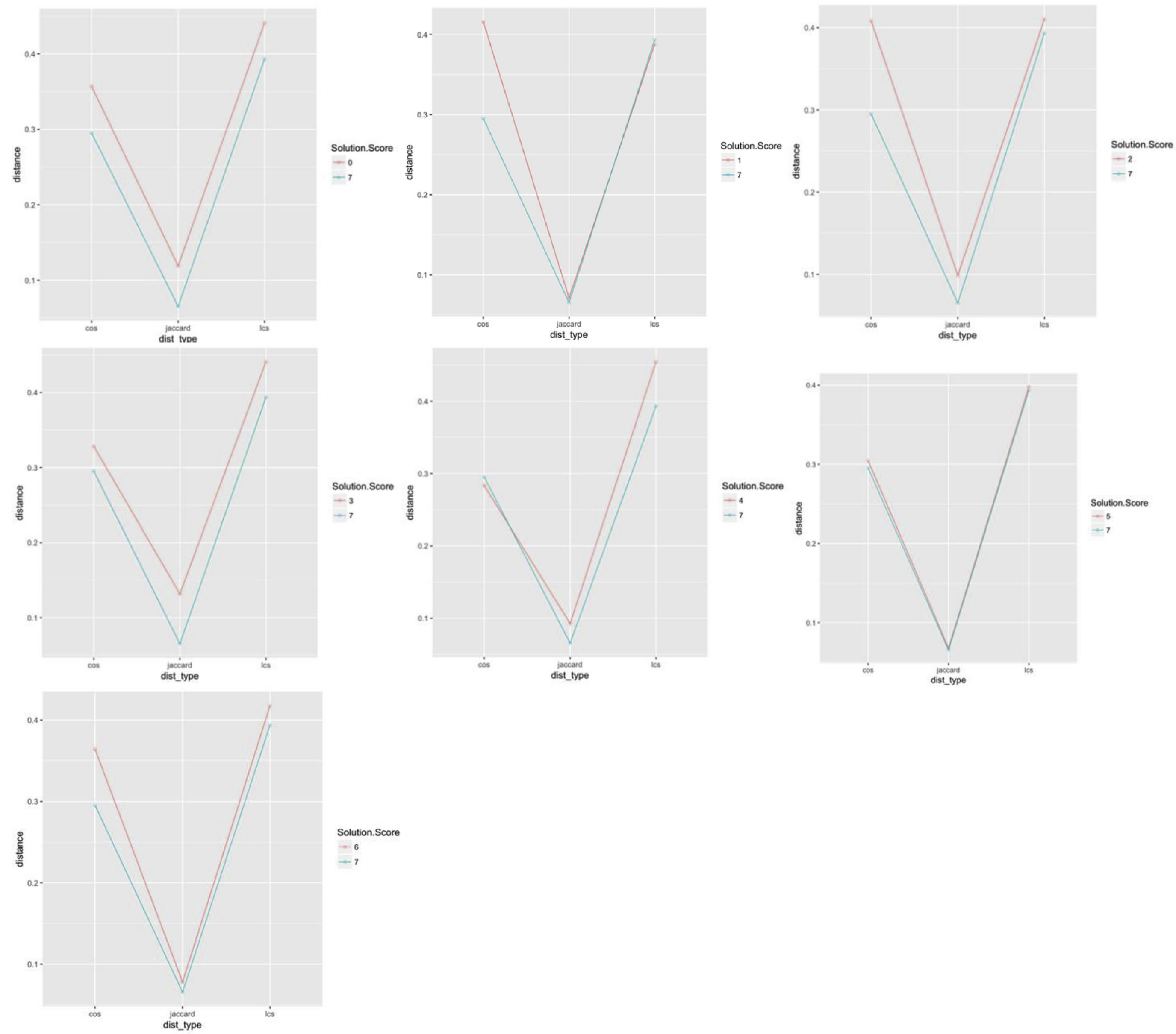


Fig. 10. Visualization for similarity measures based on solution scores.

Table 7
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

the longest compared to other groups. Cluster 1 learners stayed in Missions Database and Alien Database the longest relative to their peers.

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 9 for the average similarity measure coefficients using three methods (i.e., LCS, Cosine, and Jaccard).

Using simple line graphs in R, this study visualized the similarity among learners based on similarity coefficient values (see Fig. 13). 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.

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.

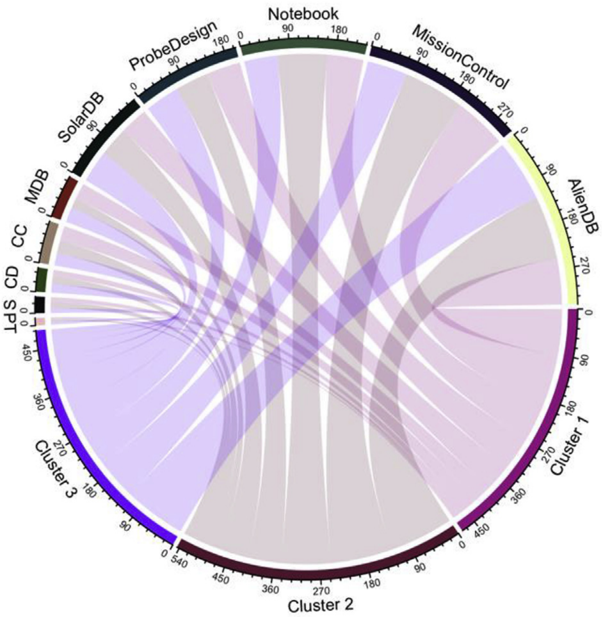


Fig. 11. Tool use frequency based on final cluster.

Table 8
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

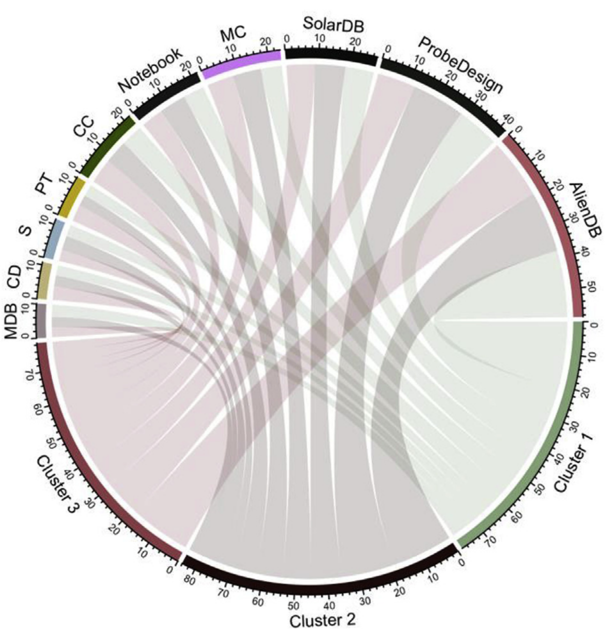


Fig. 12. Tool use duration based on final cluster.

Table 9
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

5. Discussion and conclusion

This study investigated the impact of learner metacognition and goal orientation on problem-solving performances and processes in a SG environment. For research question one, 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.

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. It is possible that some learners in the high metacognition group overestimated their metacognitive levels during the study. 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 min—might be more appropriate for participants who had a lower metacognition level. Therefore, participants in lower metacognition level had significantly higher problem-solving performances compared to participants that occupy higher metacognition level.

The results also indicated 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.

For the second research question, 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-

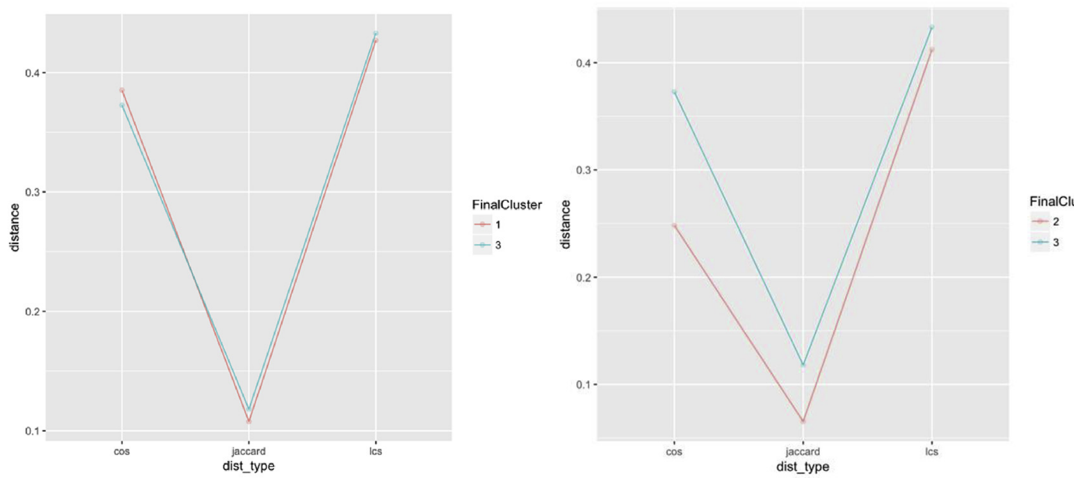


Fig. 13. Visualization for similarity measures based on final cluster.

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 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, 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, 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 hypothesis testing most frequently, while used three of the sharing cognitive load tools the least frequently. As for tool use duration, during the 60-min 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).

The results showed that the learner characteristics 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 min), followed by medium metacognition-low multiple goal orientations learners (487.41 times, 78.61 min) high metacognition-high multiple goal orientations learners (477.66 times, 77.72 min). 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.

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 different 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.

6. Limitations

Although the researcher designed the study based on previous studies, collected reliable and valid data, and conducted rigorous data analysis, there are a few limitations for this study. First, although the sample size in this study met the requirements 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 and capture all aspects of learner problem-solving processes. Furthermore, the self-reported survey that might also affect the external validity of the study.

Declarations of interest

None.

Acknowledgement

We thank the Alien Rescue Research Team at Learning Technologies program in University of Texas at Austin for developing the serious game.

References

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

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. https://doi.org/10.1016/j.cedpsych.2003.09.002.

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

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

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 & Development, 61*(3), 465–503. https://doi.org/10.1007/s11423-013-9295-4.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., et al. (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. https://doi.org/10.1016/j.compedu.2015.11.003.

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. https://doi.org/10.15415/jotitt.2016.41002.

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. https://doi.org/10.1016/j.compedu.2012.03.004.

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.

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.

Dolnicar, S. (December 2002). *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).* Melbourne: Deakin University2–4.

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. https://doi.org/10.1037/0022-3514.72.1.218.

Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519. https://doi.org/10.1037//0022-3514.80.3.501.

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

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). Berlin: Heidelberg: Springer. https://doi.org/10.1007/978-3-662-04166-6_17.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/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.

Freer, P. (2017). Problem-based learning and structural redesign in a choral methods course. *Contributions to Music Education, 42*, 53–72.

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

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.

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

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. https://doi.org/10.1016/j.chb.2013.07.052.

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.

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

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.

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.

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. https://doi.org/10.1016/j.compedu.2008.10.008.

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. https://doi.org/10.11114/jets.v5i6.2320.

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 & 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. https://doi.org/10.2190/LK2G-8K25-RB8U-PGE9.

Liu, M., Horton, L. R., Corliss, S. B., Svinicki, M. D., Bogard, T., Kim, J., et al. (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., et al. (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. https://doi.org/10.7771/1541-5015.1422.

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. https://doi.org/10.1007/s10758-015-9263-7.

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.

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. https://doi.org/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.

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.

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.
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. https://doi.org/10.1080/00377990109603981.
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. https://doi.org/10.1037/a0013835.
Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology, 3*(2), 51–59. https://doi.org/10.20982/tqmp.03.2.p051.
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. https://doi.org/10.1037/0022-0663.89.4.710.
Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., et al. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology, 23*(2), 113–131. https://doi.org/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 Learning1–23.
Moshman, D. (2017). Metacognitive theories revisited. *Educational Psychology Review,* 1–8. https://doi.org/10.1007/s10648-017-9413-7.
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. https://doi.org/10.1109/FIE.2013.6684830.
Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research and Evaluation, 8*(2), 1–9.
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. https://doi.org/10.2308/iace.2001.16.1.21.
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.
Prensky, M. (2001). Fun, play and games: What makes games engaging. *Digital game-based learning, 5*, 1–31.
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. https://doi.org/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: Purdue University Press.
Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033.
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. https://doi.org/10.1002/tea.10058.
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. https://doi.org/10.2190/EC.44.4.e.
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. https://doi.org/10.7771/1541-5015.1046.
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.
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. https://doi.org/10.1016/j.compedu.2016.03.011.
Van der Loo, M. P. (2014). The stringdist package for approximate string matching. *The R Journal, 6*(1), 111–122.
Wei, H., Wu, S., Zhao, Y., Deng, Z., Ersotelos, N., Parvinzamir, F., et al. (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.
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. https://doi.org/10.1080/00220973.2016.1277337.
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. https://doi.org/10.1002/j.2168-9830.2011.tb00013.x.
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2.
Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676.
Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer, 38*(9), 25–32. https://doi.org/10.1109/MC.2005.297.