Novel Simulation-Based Learning Modules for Teaching Database Concepts

Dr. Sabahattin Gokhan Ozden, Penn State Abington

Dr. Ozden is an assistant professor of Information Sciences and Technology at Penn State Abington. He has a PhD and MS from Auburn University in Industrial and Systems Engineering and a B.Sc. in Software Engineering and double major in Industrial Systems Engineering from Izmir University of Economics. His research interests are warehousing, optimization, and information systems. Dr. Ozden has published in various journals/conferences including Computers and Operations Research. His website address is https://www.gokhanozden.com.

Dr. Omar M. Ashour, Penn State Erie, The Behrend College

Dr. Omar Ashour is an Associate Professor of Industrial Engineering at Pennsylvania State University, The Behrend College. Dr. Ashour received the B.S. degree in Industrial Engineering/Manufacturing Engineering and the M.S. degree in Industrial Engineering from Jordan University of Science and Technology (JUST) in 2005 and 2007, respectively. He received his M.Eng. degree in Industrial Engineering/Human Factors and Ergonomics and the Ph.D. degree in Industrial Engineering and Operations Research from Pennsylvania State University (PSU) in 2010 and 2012, respectively. Dr. Ashour was the inaugural recipient of William and Wendy Korb Early Career Professorship in Industrial Engineering in 2016. Dr. Ashour's research areas include applied decision making, modeling and simulation, virtual reality, and process improvement. He contributed to research directed to improve engineering education.

Dr. Ashkan Negahban, The Pennsylvania State University, School of Graduate Professional Studies

Dr. Ashkan Negahban is an Assistant Professor of Engineering Management at The Pennsylvania State University, School of Graduate Professional Studies. He has a PhD and MS from Auburn University and a BS from University of Tehran (all in Industrial Engineering). His research interest is in the application of agent-based, discrete event, and Monte Carlo simulation in design and operation of complex systems. His web address is http://ashkannegahban.com.

Novel Simulation-Based Learning Modules for Teaching Database Concepts

Abstract

This paper presents novel immersive simulation-based learning (I-SBL) modules as an alternative active-learning method for teaching and learning fundamental concepts related to database design. I-SBL involves a 3-dimensional simulated environment that resembles a real-world system. Students can navigate through the simulated environment (in low- and high-immersion modes), observe and understand the underlying dynamics, evaluate the effect of the changes on the system, and learn by doing. The use of such modules is especially important when access to the real system is limited or impossible due to geographical barriers and/or regulations and safety considerations. We assess the impact of the proposed approach by implementing a sample I-SBL module in an undergraduate database class. The study involves two groups of students: control and test groups. Students in the control group complete a traditional problem-based learning (PBL) assignment, while the test group uses the I-SBL version of the same assignment. The assessment data collected include demographics, prior preparation, motivation, usability tests, and pre/post quizzes to measure knowledge gain. Statistical analysis of the results suggests that I-SBL performs at least as well as PBL. The results also provide important insights into the effective design and implementation of I-SBL.

Introduction and Background

Today, databases are ubiquitous and can be found in virtually any information system. Appropriate database design is critical for businesses as it enables informed decision-making, better monitoring of operations, and enhanced service to customers. The increased importance of databases in today's business world has made database design one of the core topics in computer science, information science, and information systems programs [1]. Teaching and learning database design is challenging because designing implies a creative thinking process without exact formulas. Both knowledge and skills are identified as key points in database design. Database design knowledge is technical as well as practical. Many skills are required including problem-solving, critical thinking, creativity, communication, team working, and time management. Traditionally, expositional lectures or closed and hands-on laboratories are used to teach database design. Exams are usually used to evaluate knowledge and skills required in the database design process. In the rest of this article, we refer to such methods as *traditional methods*. Unfortunately, despite their wide use, traditional methods are found to be ineffective for teaching and learning the abstract and complex domain of database design [2, 3]. While existing teaching methods for design learning provide some clues, there is no single, well-known solution. This causes a frustration among students [3] since they look for a definitive answer [4]. Students normally understand reasoned design solutions explained by an instructor, but do not feel confident when doing similar exercises by themselves [4]. They need a thorough explanation of the problem by the instructor sometimes or even an initial

solution that may help them start. Moreover, students struggle with interpreting user requirements [3] and convert them into a database design. Especially, identifying entity types is a challenging task for students since the lack of understanding the problem domain [5].

Problem-based learning (PBL) has the potential to offer an effective solution. PBL is an instructional (and curricular) learner-centered approach that empowers learners to conduct research, integrate theory and practice, and apply knowledge and skills to develop a viable solution to a problem. Successful implementation of PBL relies on appropriate definition of the problem (often ill-structured and interdisciplinary) and an instructor that guides the learning process and conducts a thorough debriefing at the conclusion of the learning experience [6]. PBL is based on theoretical educational and psychological foundations [7, 8, 9] with a cohesive body of research that shows its effectiveness for all learner groups, including K-12 [10], undergraduate [11], graduate and professional [12], and online learners [13]. Existing studies have also investigated PBL's effectiveness in a wide range of fields such as medical and health [14], chemistry [15], mathematics [16], and different engineering disciplines, namely electrical [17], mechanical [18], civil [19], and software [20] engineering. In particular, PBL is shown to be an effective active-learning approach in teaching and learning database design compared to traditional methods [4].

In this paper, we propose and assess a novel method called *immersive simulation-based learning* (I-SBL) for teaching and learning database design. We start by a formal definition of I-SBL and its components, and summarize the supporting pedagogical and psychological theories. We then describe a sample I-SBL module specifically designed for teaching and learning fundamental concepts related to database design. The experimental design and assessment instruments are then discussed. We then compare the effectiveness of I-SBL with PBL in terms of *knowledge gain* and *motivation*. Finally, we conclude the paper by providing a discussion on the main findings, lessons learned, potential improvement opportunities, and future research extensions.

The Proposed Immersive Simulation-Based Learning (I-SBL)

I-SBL involves PBL in an immersive simulation environment. An I-SBL module is a learning environment specified by:

a. A **virtual setting** (simulation model) that resembles a real system/environment and enables contextually enriched, technology-enhanced active-learning. The simulation models in our proposed I-SBL modules can be explored on a 2D display or via a VR headset (if available);

b. **Processes** in the virtual (simulation) environment that include multiple stations, and comprise technical as well as organizational aspects;

c. A set of **products** or **entities** that flow throughout the virtual environment and are processed (e.g., manufactured, assembled, stored, transported);

d. A **didactical concept** that comprises formal and informal learning, enabled by own actions of the learners during and after virtual site visits/field trips based on *remote active-learning* instead of *on-site learning*. These represent problems/projects inspired by real-world situations that the learner may face at future workplace.

Since I-SBL essentially involves PBL enhanced by an immersive simulated environment, the ped-

agogical and psychological theories supporting PBL also apply to I-SBL. Here, we summarize a sample of the supporting theories for PBL that apply to and/or are enhanced by I-SBL. The interested reader is referred to [7, 8, 9] for more details about these theories.

- I-SBL drives long-lasting development of critical thinking and problem-solving skills by: (a) activating relevant prior knowledge; (b) providing a *context*, via immersive simulations, that resembles future professional context; and, (c) stimulating learners to elaborate on their knowledge. These are the three principles of *Information Processing Approach to Learning* theory [21].

- The *Constructivism Theory* [22] suggests learners construct their own interpretations of the physical world through cognitive, interpretive activities that construct mental models by accommodating new ideas and phenomena with prior beliefs and knowledge. Knowledge is constructed via interactions with the environment and indexed by relevant contexts. The immersive simulations in I-SBL provide both the *context* and an *environment* to interact with, which are often missing in STEM education.

- Based on the *Self-determination Theory* [23], the proposed I-SBL method supports improved student motivation. I-SBL promotes *autonomous motivators* by taking into account students' views, giving opportunities for student choice, and supporting students in taking greater responsibility for their own learning. This is in contrast to traditional methods that promote *controlled motivators* represented by rewards/punishments (such as passing a test), which can cause a sense of pressure and anxiety, and lead to superficial learning.

- Learning via I-SBL is self-directed and problem-centered, draws on previous experiences, and integrates into professional learners' everyday life as the problems/projects used are inspired by and resemble real-world situations. These are the pillars of the *Adult Learning Theory* and shown to enhance professional and continuing education [24].

Learning outcomes in I-SBL include competency development in *cognitive* (recognition of knowledge) and *affective* (interest or attitudes of the learner) performance areas – two of the domains in fundamental taxonomies of learning goals [25, 26, 27]. This includes development in all competency classes, namely *professional and methodological competencies*, *socio-communicative competencies*, *personal competencies*, and *activity and action competencies* [28]. The experiments discussed in the following section are designed to support some of these expected outcomes and represent a first step to collect evidence on the effectiveness of the proposed I-SBL method.

In this work, the Simio® simulation package [29] is used to develop the immersive simulations. The advantages of the technology used in these I-SBL modules can be summarized as follows:

- Learners can navigate through the simulation model in 3D using a typical personal computer or laptop (*low-immersion* mode) but also via a virtual reality (VR) headset (if available) for enhanced immersive experience (*high-immersion* mode). Therefore, access to special equipment (e.g., VR headsets and powerful graphic cards) or special learning spaces and laboratories for immersive and active-learning experiences is not a requirement.

- Simio LLC provides free licenses for installing the software on college/university computers. The company also offers a *free* evaluation license for installation on students' personal
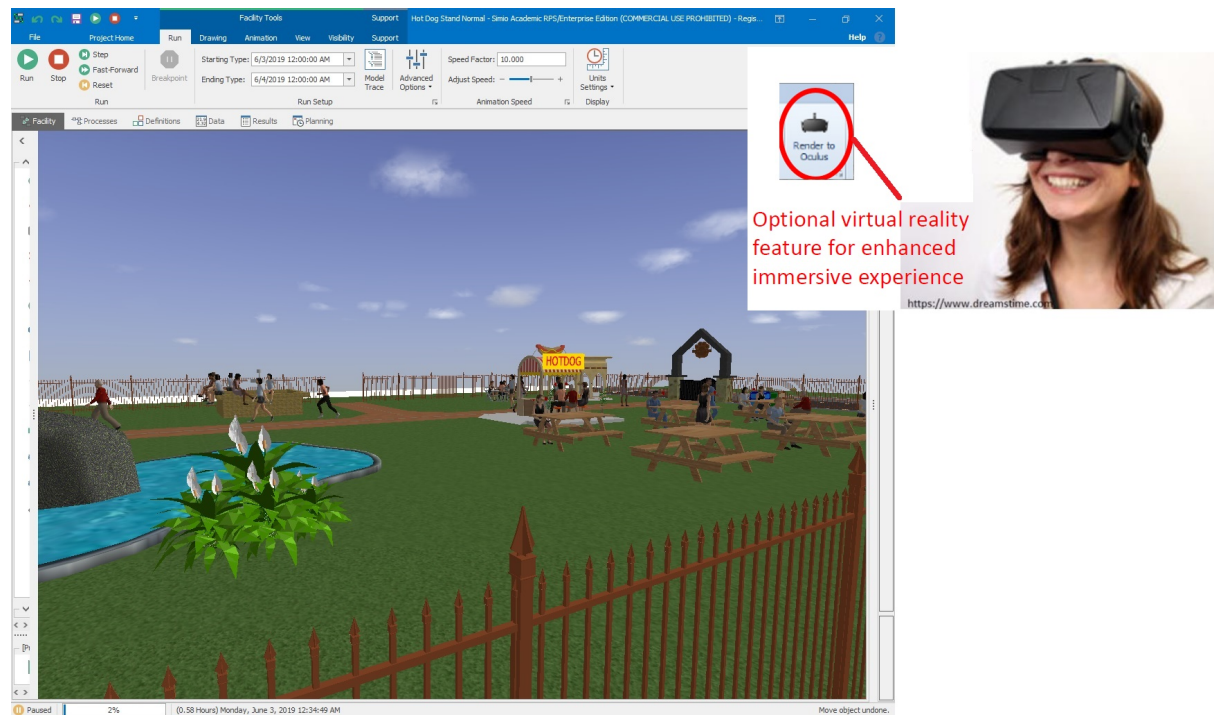
Figure 1: A snapshot of the immersive simulation model used in the sample I-SBL module

computer. As a result, the technology used here enables free access and large-scale implementation in both formal and informal online and in-person education settings.

- The technology used here also has a key advantage over 360° videos that are becoming popular in higher education and are widely available via public sites such as YouTube. While 360° videos allow users to look in different directions as the video is playing, users are unable to *move* in other directions. In the proposed immersive simulations, the learner has total control over how they explore the virtual environment, where they spend more time, and the perspective from which they make observations.

Other I-SBL modules developed as part of this ongoing project are publicly available through the project's website at https://sites.psu.edu/immersivesimulationpbl.

A Sample I-SBL Module for Database Design

The I-SBL module presented here strives to mimic a situation where the student is hired as an intern to help design a database for a hypothetical company that owns food stands across a city. The student is tasked to visit one of the company's food stands located in a small city park to learn more about the system and then come up with a database design accordingly. The immersive simulation model is treated as the real-world system, mimics the operation of a food stand, and includes enhanced 3D animation features to resemble the actual scenario. Instead of physically visiting the real system, students will explore the simulation model and make observations as the simulation is running. Figure 1 provides a screenshot of the simulation model.

After exploring the simulation and becoming familiar with the system under study, the student is asked to create an *entity-relationship (ER) diagram* and *relational schema* by identifying relevant entity types, relationships, and attributes. As for the learning objectives, the students will be able to do the following after successful completion of this I-SBL module:

1. Identify relevant entity types and their relationships (including cardinalilty ratios).

2. Identify and select relevant attributes for each entity type and their primary key.

3. Develop an appropriate ER diagram based on the identified entity types and attributes by following the normalization rules (no multi-valued attribute, no composite attribute).

4. Convert many-to-many relationships to composite entity types.

5. Develop relational schema for an ER diagram using *Stable* and *Mapped* translation methods.

The general logic of the simulation is shown in Figure 2 and can be summarized as follows. The park has three entrances and the food stand is located in the center of the park. Some of the people entering the park stop by the food stand to purchase snacks while others will walk by the food stand and exit from a different entrance. The inter-arrival times and order processing times are modeled using random variables. After receiving their food, customers either continue walking toward an exit or or sit down at the tables in front of the food stand. Customers that decide to sit at a table, will choose the table with the minimum number of people already sitting on it (ties are broken arbitrarily). After finishing their food, customers walk toward the trash cans to dispose of any trash item and then exit the park or go back in the line to order more food.

Research and Experiment Design

The study compared two educational settings using a *control* and a *test* group. The two groups were assigned an identical assignment with the same learning objectives and material. The only difference between the groups was the use of the simulation model along with the assignment for the test group. In this experiment, all students in the test group opted to use the simulation in low-immersion mode (by watching the simulation run on a typical display). Both groups were taught by the same instructor in the same section and students were randomly assigned to each group. It is hypothesized that any difference between the groups will be attributed to the use of simulation in the I-SBL module. Figure 3 shows the experiment procedure that was carried out. IRB approval was obtained prior to the experiment and data collection.

We used the following instruments to measure the effect of the simulation:

1. Demographics survey: The survey collects data about the subject's age, gender, race, grade point average (GPA), grade in a prerequisite course, major, semester standing, work experience, and experience with computer simulation and video games in general. The survey was collected before the assignment at the start of the semester.

2. Reduced Instructional Materials Motivation Scale (RIMMS): This instrument is a 12-item questionnaire to assess the level of student motivation. Motivation is measured by four factors: attention (A), relevance (R), confidence (C), and satisfaction (S). Each factor has 3 items in the questionnaire [30]. This questionnaire was collected immediately after finishing
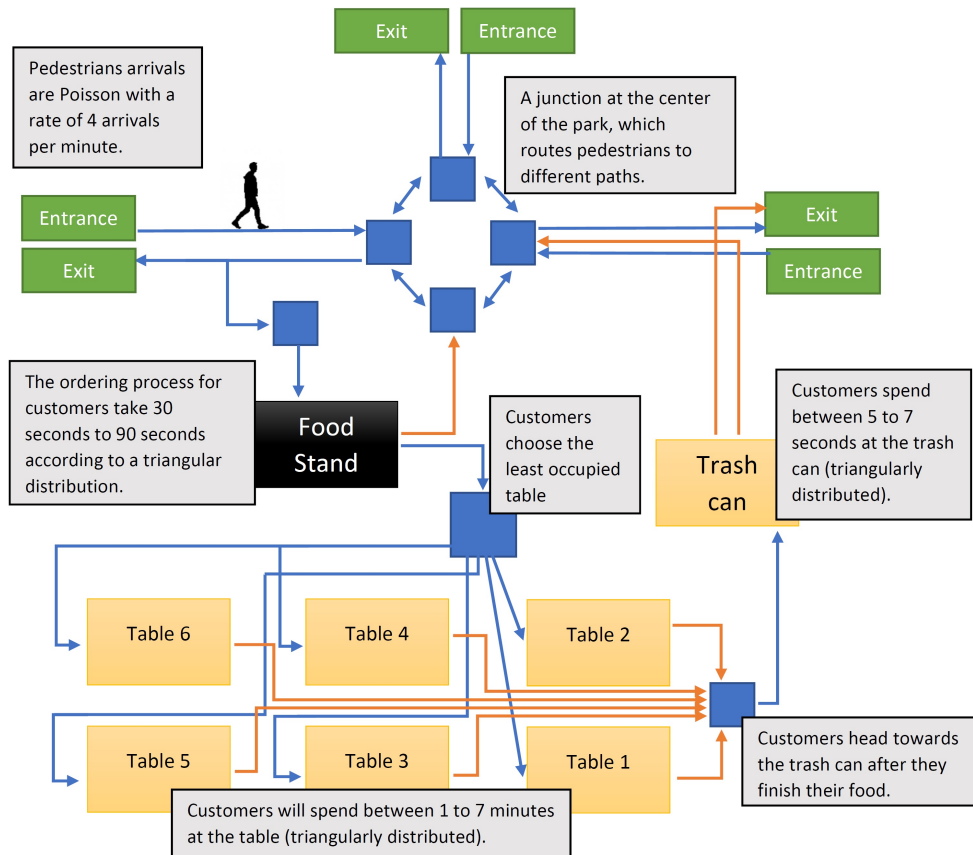
Figure 2: A logic of the simulation model of the park with food stand

the assignment. The assignment was completed in eleven days. Both groups complete the RIMMS questionnaire.

3. Knowledge test (quiz): This test was created by the instructor to assess students' understanding of the following concepts: databases (definition and usage), entities, relationships, relational schemas, and stable and mapped translation methods (ER diagram to relational schema conversion). The tests were given twice; pre- and post- the assignment. The pre-quiz was conducted while the students were working on the assignment (day 9). The post-quiz was conducted four days after completing the assignment. Both groups took the tests.

4. Open-ended questions: Three questions were asked: (1) *"What changes would you recommend in this assignment to enhance/improve your learning experience?"* (2) *"Do you think you can use or apply what you have learned in this assignment in your professional life? Please explain how."* and (3) *"What helped you learn the concepts in this assignment? Please explain. (Examples: Instructor, textbook, lectures, simulation model, Google, etc.)"* The goal was to capture additional insights not captured by the other instruments. This questionnaire was conducted for both groups immediately after completing the assignment.

We implemented the sample I-SBL module in an introductory 200-level undergraduate database course entitled *Organization of Data*, offered as part an Information Sciences and Technology
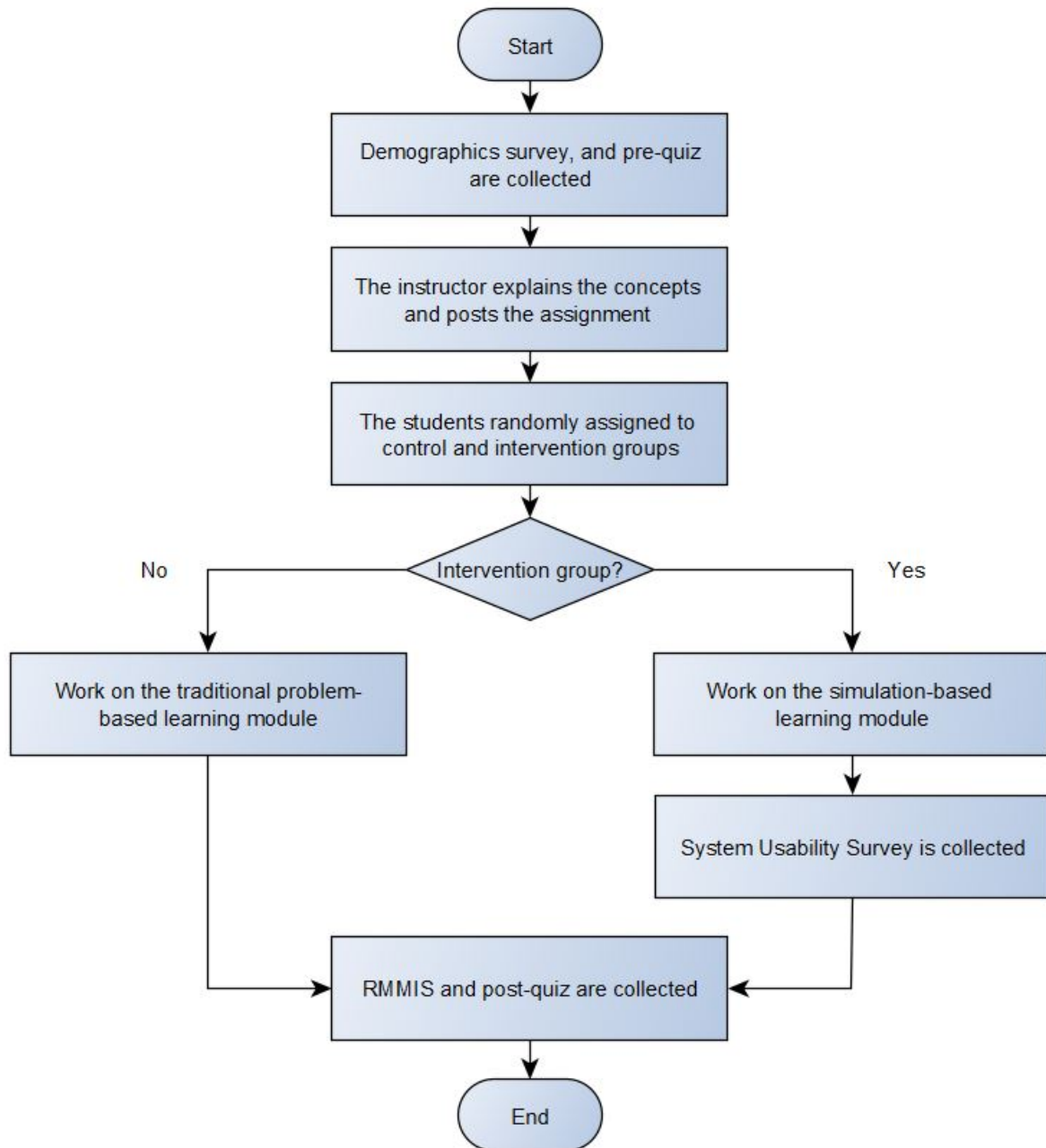
Figure 3: General design of the assessment experiments

(IST) program at The Pennsylvania State University. This is a 3-credit hour course with two 75-minute lectures per week, where one-third of the lectures are devoted to computer lab exercises. By the end of the course, students are expected to be able to develop an effective database satisfying a given set of data requirements. This course encompasses requirements gathering and specification, conceptual database design using ER diagrams, logical database design using relational schema, normalization, and physical database design. In addition, the course discusses the use of *Oracle Application Express* for database oriented web application development.

Table 1: Gender composition per group

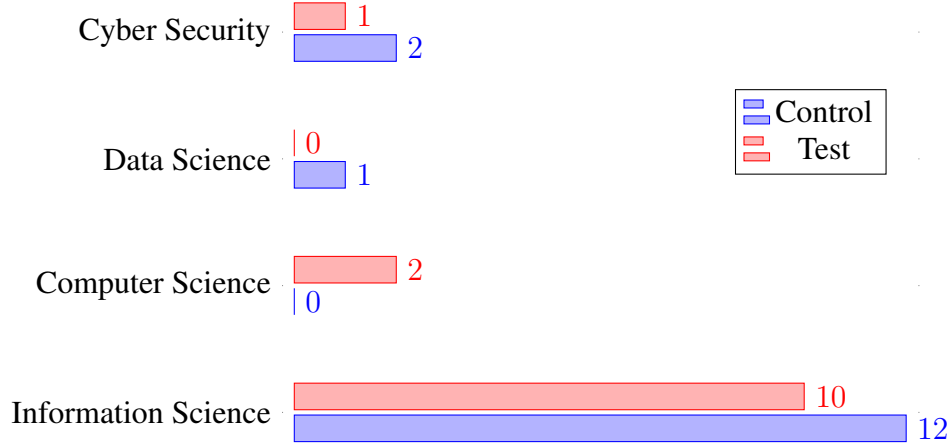|	Male	Female	Other
Control group	86.67%	13.33%	0%
Test group	15.38%	84.62%	0%

Figure 4: Number of students by major

Student Population

Table 1 shows the gender composition of the student participants. As shown in Figure 4, the majority of the students in both groups are in IST with a few students being from other majors. The majority of the students in both groups have no previous work experience (Table 2). As shown in Figure 5, most of the students are sophomores. We also asked about the level of experience with computer simulation (Table 3) and computer games (Table 4). Figure 6 compares the GPA of students in the two groups. There was no significant statistical difference between two group in terms their GPA ($H_0 : \mu_{GPA}^{test} - \mu_{GPA}^{control} = 0$, the p-value is 0.249 using a Mann-Whitney test).

Research Hypotheses

We use our experiments to investigate the following hypotheses:

1. I-SBL is at least as effective as PBL with respect to student motivation as measured by the RIMMS instrument.

2. The students in the test group would perform at least as well as the control group in terms of understanding the concepts related to entities & relationships, and definition & usage of databases.

3. The students in the test group would perform at least as well as the control group in developing ER diagrams and relational schema for the assignment.
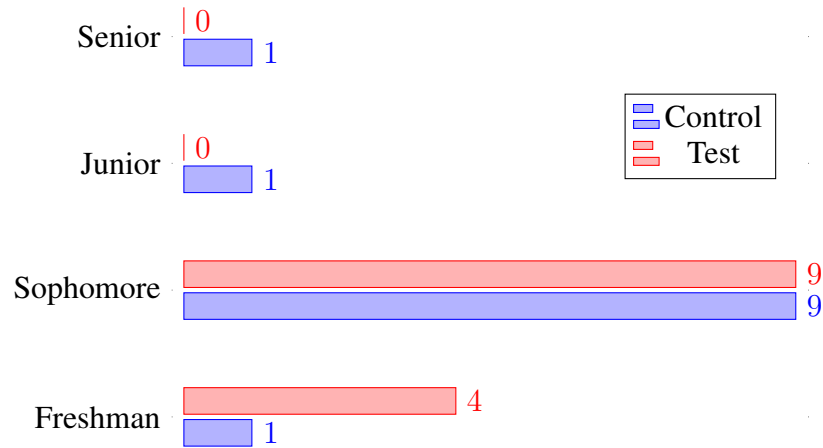
Figure 5: Number of students by their year of school

Table 2: Percentage of students broken down by previous work experience

Group	Has Work Experience	No Work Experience
Control	13.33%	86.67%
Test	15.38%	84.62%

We also investigated the usability of the immersive simulation model. The scores from the usability survey and the feedback from the open-ended questions will be used to improve the design of future I-SBL modules. For the sake of conciseness, we do not provide the usability survey results in this paper and focus on motivation and knowledge gain.

Statistical Comparisons and Results

Table 5 shows the mean, median, and standard deviation of the four dimensions related to motivation measured by the RIMMS instrument for the control and test groups. Statistical tests were performed for every dimension to test our null hypothesis ($H_0 : \mu_{RIMMS}^{test} - \mu_{RIMMS}^{control} = 0$). No significant statistical difference was detected between the control and test group overall. Therefore, based on our sample data, I-SBL is statistically as effective as PBL in terms of motivation, answering the first research hypothesis posed in the previous section.

Table 6 shows the mean and standard deviation of score for each pre-/post-quiz question. Each quiz question is worth 20 points. Each question was graded based on the following rubric: 0 points for inaccurate description (none of the important keywords found in student's answer), 10 points for partially correct answer (some of the important keywords found in student's answer), 20 points for a complete description (almost all of the important keywords found in student's answer). The test group performed significantly better than control group in their pre-quiz performance (p-value is 0.2364 using a two-sample t-test). There was no statistically significant difference between the control group and the test group in their post-quiz performance (p-value is 0.6838 using a Mann-Whitney U test).

Table 3: Percentage of students broken down by experience with computer simulation

Group	No Experience	Some Experience	Expert
Control	33.33%	66.67%	0.00%
Test	53.85%	38.46%	7.69%

Table 4: Percentage of students broken down by experience with computer games

Group	No Experience	Some Experience	Expert
Control	13.33%	46.67%	40.00%
Test	7.69%	30.77%	61.54%

Table 7 shows the knowledge gain (i.e., post-quiz score - pre-quiz score) within control and test groups. Observations for individual questions and overall score do not follow a Normal distribution, hence we performed a Wilcoxon Signed-Rank test to investigate our null hypothesis ($H_0 : \mu_{quiz}^{post} - \mu_{quiz}^{pre} = 0$) for each group. The overall knowledge gain (total score) was statistically significant in both control and test groups.

We performed a Mann-Whitney test to test the null hypothesis ($H_0 : \mu_{knowledgeGain}^{test} - \mu_{knowledgeGain}^{control} = 0$). There was no statistically significant difference between the control and test groups in their knowledge gain (the difference between post-quiz and pre-quiz performance). This confirms the second research hypothesis posed in the previous section.

Figure 7 compares the control and test groups' scores on their PBL and I-SBL assignment, respectively. We test the data for normality using the Shapiro-Wilk test. Both samples passed the normality test, hence we performed a two-sample t-test and found no statistical significant difference between the control and test groups in terms of their assignment scores (the p-value is 0.69). Therefore, based on our sample data, I-SBL is statistically as effective as PBL in terms of assignment score. This confirms the third research hypothesis posed in the previous section.

Qualitative Assessment of Open-Ended Questions

Student responses to the open-ended questions helped us derive additional insights, especially in understanding why the test group did not perform significantly better than the control group. Here are some of the responses from control (C) and test (T) groups for the following question: *"What changes would you recommend in this assignment to enhance/improve your learning experience?"*

  C: *"To improve my learning experience with this assignment, I would remove any unnecessary information that may lead students to be confused."*

  C: *"I felt that the directions and the way the assignment was conveyed was confusing. I was unsure what parts of the assignment were supposed to be used and what weren't."*

  T: *"I felt like even though it was supposed to be a realistic project, there was a lot of extra information which could make it a little confusing for someone who does not understand what we're learning fully."*
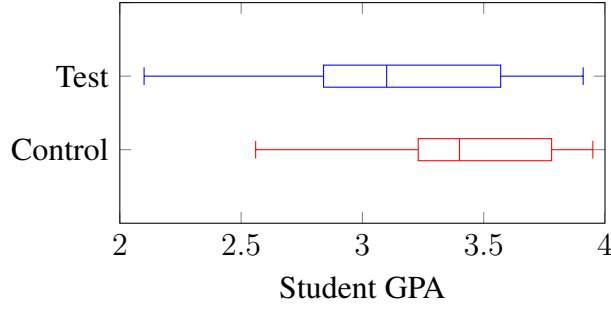
Figure 6: Box plot for student GPAs

Table 5: RIMMS analysis ($H_0 : \mu_{RIMMS}^{test} - \mu_{RIMMS}^{control} = 0$)

	Control			Test			Difference (between groups)	
RIMMS	Mean	Median	SD	Mean	Median	SD	p-value	Stat. Test.
Attention	11.20	13.00	3.21	11.08	11.00	2.43	0.9110	Two-sample t-test
Relevance	12.40	12.00	1.84	11.31	12.00	1.89	0.1341	Two-sample t-test
Confidence	12.07	12.00	2.94	10.38	11.00	2.81	0.0496	Mann-Whitney
Satisfaction	11.47	11.00	3.52	9.69	10.00	2.46	0.1671	Two-sample t-test
Overall	47.13	51.00	10.80	42.46	46.00	8.35	0.2170	Two-sample t-test

T: *"I feel that this assignment is very well constructed and it provides the challenge of reading and comprehending what was being asked and transforming it into the diagram that was required. I feel that some information could be removed as it is information that is wasted. I feel as if the flowchart that was provided in the word document didn't add any new information that wasn't already obtained."*

Some of the comments above can be attributed to the fact that the problem statement in PBL involves certain planned imperfections to mimic real-world projects, which are often ill-structured and vague. Here are some of the responses from control (C) and test (T) groups for the following question: *"Do you think you can use or apply what you have learned in this assignment in your professional life? Please explain how."*

C: *"I believe I can use this to set up my business and help my parents to set up their restaurant structure."*

C: *"I think I can use or apply what I have learned in this assignment in my professional life, because I can come across a situation where someone has a question on queries to help them set up a database, and I could help them accordingly. And if someone brings up as a topic conversation on queries, I can contribute to that topic, because I know what they are talking about, because of this class."*

T: *"Yes it would allow me to see how systems work in the world and compare them to ER diagram tables or their inner workings and data."*

T: *"I would find this incredibly helpful for general database management in the workplace like keeping track of equipment and recording the equipment information and the process of how*

Table 6: Pre-quiz and post-quiz results

| | Pre-quiz | | | | Post-quiz | | | |
| | Control | | Test | | Control | | Test | |
Question	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Q1	9.33	8.84	13.08	9.47	18.00	4.14	18.46	5.55
Q2	18.00	5.61	18.46	5.55	17.33	7.04	20.00	0.00
Q3	11.33	7.43	11.54	6.89	19.33	2.58	17.69	4.39
Q4	16.00	7.37	20.00	0.00	18.00	4.14	16.15	5.06
Q5	9.33	9.61	10.00	9.13	14.00	8.28	16.15	7.68
Total Score	64.00	26.13	73.08	19.74	86.67	11.13	88.46	12.14

Q1: What is an entity type?

Q2: What is a relationship?

Q3: What is a database?

Q4: Why do we need databases?

Q5: Explain one-to-one, one-to-many, and many-to-many relationships in your own words.

Table 7: Knowledge gain within groups ($H_0 : \mu_{quiz}^{post} - \mu_{quiz}^{pre} = 0$)

| | Control | | | | Test | | | |
Question	Mean	SD	p-value	Stat. Test	Mean	SD	p-value	Stat. Test
Q1	8.67	8.34	0.009*	Wilcoxon	5.38	8.77	0.100	Wilcoxon
Q2	-0.67	2.58	1.000	Wilcoxon	1.54	5.55	1.000	Wilcoxon
Q3	8.00	7.75	0.009*	Wilcoxon	6.15	8.70	0.050	Wilcoxon
Q4	2.00	9.41	0.447	Wilcoxon	-3.85	5.06	0.059	Wilcoxon
Q5	4.67	13.56	0.266	Wilcoxon	6.15	7.68	0.036*	Wilcoxon
Total Score	22.67	26.85	0.008*	Wilcoxon	15.38	18.98	0.018*	Wilcoxon

it is checked out and managed."

In general, the authors observed that students find both PBL and the I-SBL assignment useful in future professional settings. Below are some of the responses from control (C) and test (T) groups for the following question: *"What helped you learn the concepts in this assignment? Please explain. (Examples: Instructor, textbook, lectures, simulation model, Google, etc.)"*

C: *"The instructions and the explanation by the professor was outstanding."*

C: *"The instructor and the Power Points that the instructor made and model diagram."*

T: *"The simulation model helped a lot along with the material that the instructor taught in class about diagrams, tables, etc."*

T: *"The lectures and the simulation made the assignment much more clear."*

In general, the authors observed that students in the test group found the simulation model to be helpful in visualizing and contextualizing the problem.

Table 8: Knowledge gain between groups ($H_0 : \mu^{test}_{knowledgeGain} - \mu^{control}_{knowledgeGain} = 0$).

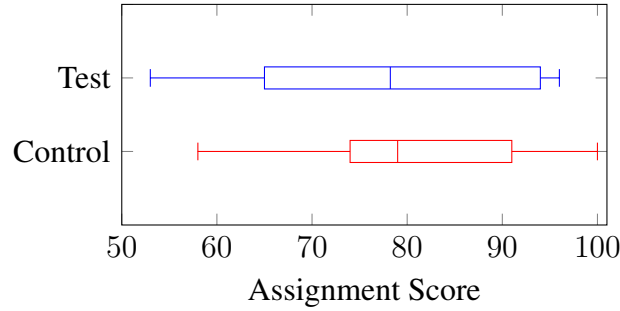| Question	Mean	p-value	Stat. Test
Q1	-3.28	0.253	Mann-Whitney
Q2	2.21	0.180	Mann-Whitney
Q3	-1.85	0.622	Mann-Whitney
Q4	-5.85	0.083	Mann-Whitney
Q5	1.49	1.000	Mann-Whitney
Total Score	-7.28	0.607	Mann-Whitney

Figure 7: Box Plot for Student Assignment Scores

Discussion and Conclusions

In this paper, an immersive simulation-based learning (I-SBL) approach was proposed and assessed. The approach presents a novel method for teaching and learning fundamental concepts related to database design. The approach was implemented an I-SBL module in an undergraduate database design class, and through a set of experiments with control and test groups, it was shown that I-SBL performs (statistically) as well as PBL with respect to knowledge gain and motivation. This is an encouraging first step for I-SBL given that PBL is a well-established active-learning approach that is proven to be significantly more effective than traditional methods (e.g., lectures). The results provide a few important insights on effective design and implementation of I-SBL:

1. There has to be a clear purpose for the simulation in order for learners to truly appreciate its value. In the sample I-SBL module used in this paper, the simulation's role was primarily to help students *visualize* the context, i.e., students would watch the simulation run without having to perform any extensive interaction with the virtual environment. The connection between the simulation and the assignment can be made stronger by adding a follow-up part to the I-SBL module where students implement their database and populate it with sample data collected from the simulation. The authors hypothesize that such interactions with the simulation would significantly enhance the effectiveness of I-SBL. The authors are currently developing I-SBL modules where students will need to collect data from the simulation model to parameterize the analytical models that they need to develop to answer the questions being asked. In the next phase of this research, it is planned to use these I-SBL modules to test the above hypothesis. All I-SBL modules developed as part of this ongoing project are publicly available through the project's website at https://sites.psu.edu/immersivesimulationpbl.

2. Relevance to potential future career paths is another critical consideration when designing an I-SBL module. For the food stand module, the authors realized that students whose family members are in the restaurant business were by far more excited about the assignment and found more value in it as evident in their response to the open-ended questions and based on their conversations with the instructor outside of class. The authors hypothesize that having multiple equivalent I-SBL modules that students can choose from would significantly increase the effectiveness of the activity. It is planned to test this hypothesis by developing several I-SBL modules that are similar to the food stand example in all aspects except the application area (say, an ATM or an electric vehicle charging station). All of these models represent a single-server queueing system and there is no fundamental difference between the ER diagrams and relational schema, but some students may find these applications much more attractive and relevant to their future job.

3. It is also important to note that the I-SBL module used in this study was intended for an *introductory* level undergraduate database course and used as an assignment after students were introduced to the relevant concepts for the first time. Therefore, it was necessary to: (a) use a familiar context such as a food stand with which most students have some experience with; and, (b) keep the complexity level of the system relatively low to focus on learning database-related concepts rather than on understanding the underlying dynamics of a complex system/simulation. For advanced database courses, it is recommended to use contexts with more complex entities and relations that may not have simple visual representations (e.g., database design for human resource management or medical records). This would further enhance how the I-SBL module reflects a future professional context.

This paper presents a first attempt to develop, implement, and assess I-SBL, and like any other first step, has several limitations. Perhaps the most important limitation of the work presented here is a relatively small sample size. This experiment will be replicated in future offerings of the course and see how/if the findings would change based on a larger sample size. Especially, the control group in this experiment had a higher average GPA than the test group that used I-SBL. A larger sample size should help eliminate such potential effects that might exist here. Moreover, the students in the same section of the class were divided into two groups. While this ensures that many other factors were common among the two groups (e.g., both groups received the exact same lectures and material), the students in the test group would clearly see that some of their classmates are able to complete the assignment without the simulation model. This might have created some negative attitude toward the I-SBL version of the assignment as these students had to do additional steps, namely installing the simulation software and observing the simulated environment.

The authors hope that this paper and its extensions would encourage the use of immersive simulations in conjunction with PBL in engineering education.

References

[1] Russell Shackelford, Andrew McGettrick, Robert Sloan, Heikki Topi, Gordon Davies, Reza Kamali, James Cross, John Impagliazzo, Richard LeBlanc, and Barry Lunt. Computing curricula 2005: The overview report. *ACM SIGCSE Bulletin*, 38(1):456–457, 2006.

[2] M Mercedes Martínez-González and Gérald Duffing. Teaching databases in compliance with the european dimension of higher education: Best practices for better competences. *Education and Information Technologies*, 12(4):211–228, 2007.

[3] M Pamela Neely. Mastery level learning and the art of database design. In *Proceedings of the Thirteenth Americas Conference on Information Systems*, 2007. URL `https://scholarworks.rit.edu/other/638`.

[4] César Domínguez and Arturo Jaime. Database design learning: A project-based approach organized through a course management system. *Computers & Education*, 55(3):1312–1320, 2010.

[5] Richard T Watson. The essential skills of data modeling. *Journal of Information Systems Education*, 17(1), 2006.

[6] John R Savery. Overview of problem-based learning: definition and distinctions, the interdisciplinary. In *Journal of Problem-based learning*. Citeseer, 2006.

[7] Rose M. Marra, David H. Jonassen, Betsy Palmer, and Steve Luft. Why problem-based learning works: Theoretical foundations. *Journal on Excellence in College Teaching*, 25(3&4):221–238, 2014.

[8] Clare Onyon. Problem-based learning: a review of the educational and psychological theory. *The Clinical Teacher*, 9(1):22–26, 2012.

[9] Cindy E. Hmelo-Silver. Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3):235–266, 2004.

[10] Clarice Wirkala and Deanna Kuhn. Problem-based learning in k–12 education: Is it effective and how does it achieve its effects? *American Educational Research Journal*, 48(5):1157–1186, 2011.

[11] Gery C. Karantzas, Michelle R. Avery, Susie Macfarlane, Alexander Mussap, Gregory Tooley, Zoe Hazelwood, and Julie Fitness. Enhancing critical analysis and problem-solving skills in undergraduate psychology: An evaluation of a collaborative learning and problem-based learning approach. *Australian Journal of Psychology*, 65(1):38–45, 2013.

[12] Tara J. Fenwick. Problem-based learning, group process and the mid-career professional: Implications for graduate education. *Higher Education Research & Development*, 21(1):5–21, 2002.

[13] Meltem Huri Baturay and Omer Faruk Bay. The effects of problem-based learning on the classroom community perceptions and achievement of web-based education students. *Computers & Education*, 55(1):43 – 52, 2010.

[14] Elaine H.J. Yew and Karen Goh. Problem-based learning: An overview of its process and impact on learning. *Health Professions Education*, 2(2):75 – 79, 2016.

[15] Dylan P. Williams, Jonathan R. Woodward, Sarah L. Symons, and David L. Davies. A tiny adventure: the introduction of problem based learning in an undergraduate chemistry course. *Chemistry Education Research and Practice*, 11:33–42, 2010.

[16] Fengfeng Ke. An implementation of design-based learning through creating educational computer games: A case study on mathematics learning during design and computing. *Computers & Education*, 73:26 – 39, 2014.

[17] Aman Yadav, Dipendra Subedi, Mary A. Lundeberg, and Charles F. Bunting. Problem-based learning: Influence on students' learning in an electrical engineering course. *Journal of Engineering Education*, 100(2):253–280, 2011.

[18] Thenkurussi Kesavadas. V-Learn-Fact: A new approach for teaching manufacturing and design to Mechanical Engineering students. In *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, 2013. doi:10.1115/IMECE2013-63817.

[19] A. A. Ahern. A case study: Problem-based learning for civil engineering students in transportation courses. *European Journal of Engineering Education*, 35(1):109–116, 2010.

[20] Lyn Brodie, Hong Zhou, and Anthony Gibbons. Steps in developing an advanced software engineering course using problem based learning. *Engineering Education*, 3(1):2–12, 2008.

[21] R C Anderson. The notion of schemata and the educational enterprise: general discussion of the conference. In R C Anderson, R J Spiro, and W E Montague, editors, *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Erlbaum, 1977.

[22] David H. Jonassen. Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39(3):5–14, 1991.

[23] Mark Albanese. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Medical Education*, 34(9):729–738, 2000.

[24] Sharan B. Merriam. Andragogy and self-directed learning: Pillars of adult learning theory. *New Directions for Adult and Continuing Education*, 2001(89):3–14, 2001.

[25] L. W. Anderson, D. R. Krathwohl, and P. W. Airasian. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman, 2001.

[26] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. *Taxonomy of educational objectives: The classification of educational goals (Book I: Cognitive domain)*. New York: McKay, 1956.

[27] D. R. Krathwohl, B. S. Bloom, and B. Masia. *Taxonomy of educational objectives*. London: Longman, 1964.

[28] Michael Tisch, Eberhard Abele, and Joachim Metternich. *Learning Factories : Concepts, Guidelines, Best-Practice Examples*. Springer International Publishing, 1st edition, 2019.

[29] Jeffrey S. Smith, D. T. Sturrock, and W. D. Kelton. *Simio And Simulation: Modeling, Analysis, Applications*. Simio LLC, Pittsburgh, PA, 5th edition, 2018.

[30] Nicole Loorbach, Oscar Peters, Joyce Karreman, and Michaël Steehouder. Validation of the instructional materials motivation survey (imms) in a self-directed instructional setting aimed at working with technology. *British journal of educational technology*, 46(1):204–218, 2015.