Evaluating the Effectiveness of a Simulation-based Collaborative Learning Environment and Validating New Ways to Assess Learning Outcomes

Kristina Wieland
Dirk Ifenthaler
Albert-Ludwigs-University Freiburg, Germany

Abstract
Technology-based learning environments offer new ways to support learning in complex domains, but there is still a lack of understanding as to under what conditions learning is most likely to be successful. Moreover, reliable and valid methodologies for assessing the learners’ performance are required. We designed and developed a simulation-based learning environment in the area of school management which includes (1) instructional support to facilitate learning, (2) collaborative features, and (3) a methodology to determine the learning outcome. A study with 66 participants was conducted to evaluate effects of the learning environment on learning outcome and validation of the used instruments. Results showed that students’ performance significantly improved after using the learning environment and they had no problems with the complex domain. Data also revealed validation issues for the applied instruments. We discuss our initial findings and conclude with directions for future research.

Introduction
The rise of computer technology affected the design of learning environments and raised hopes and expectations regarding effective instruction. Features such as the multimedia capability and networking opportunities of computers opened new paths for applying instructional principles in many ways. However, technology alone will not lead to successful learning and it is more the underlying pedagogy than the technology that will facilitate learning which requires thorough design of effective learning environments. And even though technology-based learning environments offer opportunities to support learning in complex domains, there is still a lack of understanding as to under what conditions learning is most likely to be successful (Cannon-Bowers & Bowers, 2008). Moreover, to evaluate effectiveness of a learning environment, methodologies for assessing the learners’ performance are required.

One possible application of technology to enhance learning involves the use of simulations, which play an important role in education. Clariana and Strobel (2008) define a simulation as “an executable (runnable) model, computer software that allows a learner to manipulate variables and processes and observe results” (p. 330). Simulations are particularly useful for learning about complex domains, because they enable learners to form and test hypotheses in a setting as complex as the field of application, but without risking negative consequences when making mistakes. Learners do not per se have the ability to construct meaningful and desired outcomes when working with complex problems (Dörner, 1996; Spector, 2000). Thus, we claim that learning environments which utilize technological enhancements such as simulations are only effective if instructional support is provided.

Moreover, determining the effectiveness of learning environments requires well-defined performance criteria and valid measurement instruments. Christensen, Spector, Sioutine, and Cormack (2000) claim that learning should be assessed through a comparison of learners’ thinking and performance with that of experts.

We developed a simulation-based learning environment in a complex domain which includes (1) instructional support to facilitate learning, (2) collaborative features, and (3) a methodology to determine learning outcome. Our goals were firstly to examine the effect of a simulation-based learning environment on students’ learning outcome and secondly to develop a methodology for assessing students’ performance when working on complex problems.

The next section will focus on rationales for the use of technology-based learning environments including our research questions and hypotheses. Then we will introduce the design and procedure of our research study, followed by the presentation of our results. Finally, we will discuss our findings and present directions for future research.

Rationales for the Use of Technology-based Learning Environments
The mere use of technology alone does not lead to successful learning, but computer-based learning environments have the potential to facilitate authentic learning where students can construct meaningful knowledge. When providing learners with authentic activities, knowledge is not presented in abstract forms, but rather embedded in a meaningful context. Learner-centered activities play a central role where the teacher’s role moves from a directive disposition to a supportive one. This means, that instruction is not broken down into manageable
Leithwood, Lawton, and Cousins (1989) identified seven categories that are indirectly and directly related to school. Schools are systems with many interrelated variables which contribute in different ways to dropout rates. Changing the school itself rather than the student may contribute to preventing at-risk students from dropping out of parents' level of education, etc. (Grob, Jaschinski, & Winkler, 2003). However, some studies have revealed that lead a student to drop out of school may be found in the students themselves, such as socio-economic status, their dropout rate. Student retention is an important issue in schools today (Attewell & Seel, 2003). Many factors which individually as well as collaboratively.

The Learning Environment

The teacher is in the role of a coach who provides information and support when needed. Advocates of situated learning approaches argue that students are not prepared effectively for job performance if learning is separated from practice (Enkenberg, 2001) and claim that decontextualized knowledge also contributes to the problem of inert knowledge which refers to information that a learner may possess, but is not able to recall or apply in order to solve problems (Whitehead, 1929).

Computer-based learning environments can be effective tools when it comes to implementing situated learning activities. One possible application is the use of simulations of complex systems as a means for training and research. Educational technologists have been employing computer-based simulations and games in a variety of educational contexts to utilize technology in supporting effective instruction. Simulations have become a prevalent tool in learning and instruction since they provide a broad applicability in different instructional settings (Ifenthaler, 2009). The following elements are essential to an effective instructional simulation: complexity, learning objectives, interaction, real-world processes, and embedded, and non-random outcomes (Shelton, 2007).

Additionally, the meaning making process in situated learning environments often includes collaboration between learners, because learning is a social, dialogical process based on the negotiation between learners (Jonassen, Davidson, Collins, Campbell, & Haag, 1995). The influence of social interaction and individual meaning making is reciprocal, because personal understanding and social knowledge building are interdependent and cannot exist separately (Stahl, 2006). A different kind of interaction can be found in a continuing acquisition of knowledge and skills when novices learn from experts in the context of practical activities. Experts can either act as role models or coaches or both. Lave and Wenger (1991) propose to let the participant move from being an observer to a fully functioning agent within the culture of the group. This will enable learners to learn how to speak about and within a community of practice.

Instructional designers need to be developing learning environments which enable learners to construct meaningful knowledge. This can be achieved by providing learners with collaborative activities that are embedded in a context meaningful to them with room for reflection with others about their learning progress (Jonassen et al., 1995). An example for an instructional model that applies principles derived from the situated learning approach such as learning activities and collaboration is problem-based learning (Jonassen, 1997; Savery & Duffy, 1995a, 1995b). Problem-based learning describes an instructional method where learning occurs when learners are working on a solution for a given, authentic problem. Learning is considered the process of creating content knowledge and developing problem-solving as well as self-directing skills (Hung, Jonassen, & Liu, 2008). Other models implementing principles from the situated learning approach are cognitive apprenticeship (Collins, Brown, & Newman, 1989), project-based science (PBS), and communities of practice (Lave, 1991). Cognitive apprenticeship refers to a way of learning, where novice learners first observe an expert and are then guided by a more experienced person to utilize their own cognitive and metacognitive skills. The experienced person may demonstrate the skills and levels, provide examples, and support the novice through feedback and assistance (Dennen & Burner, 2008). Project-based science is characterized by a driving question, investigations, artifact development, collaborations and use of tools to support inquiry (Singer, Marx, Krajcik, & Chambers, 2000). A community of practice (CoP) is a sustained social network of people which can be bound in a formal or informal way. They are engaged in a common practice and share a common set of values and knowledge (Dennen & Burner, 2008). All approaches have in common that learners actively work on a complex task in interaction with peers and/or a more experienced person. The teacher is in the role of a coach who provides information and support when needed.

The Learning Environment

Our modular learning environment is concerned with the area of school management. It was derived from principles of the situated learning approach and includes questionnaires to assess dependent variables, instructional materials, communication tools, and a simulation which is embedded in an authentic scenario. Students work individually as well as collaboratively.

The learner’s mission is to improve a school’s efficiency, which is represented by means of the school’s dropout rate. Student retention is an important issue in schools today (Attewell & Seel, 2003). Many factors which lead a student to drop out of school may be found in the students themselves, such as socio-economic status, their parents’ level of education, etc. (Grob, Jaschinski, & Winkler, 2003). However, some studies have revealed that changing the school itself rather than the student may contribute to preventing at-risk students from dropping out of school. Schools are systems with many interrelated variables which contribute in different ways to dropout rates. Leithwood, Lawton, and Cousins (1989) identified seven categories that are indirectly and directly related to
dropout, such as teachers, programs & instruction, and school-community relations. Their empirical model was used as the underlying causal model of the simulation (see Ifenthaler, 2009). Complex systems are characterized by interconnected variables which influence each other (Funke, 1991; Seel, Ifenthaler, & Pirnay-Dummer, 2009). In the simulation, all seven categories of the underlying model are represented as investment fields, i.e. learners can allocate money from a given budget to different areas with each standing for a particular attribute of one variable of the causal model.

The main learning objective is to explore the underlying causal model of the simulation. Students can enter data into the simulation and receive an output, but calculations as well as causal relationships between the variables are not known by the user. Thus, students need to infer the causal model from the output. To facilitate the exploration, they will be given information about the variables which allows for conclusions regarding their interrelationships. Moreover, students will learn about the underlying model through feedback from the system (simulation output such as the school’s dropout rate) and collaboration with peers. The collaborative part enables them to exchange ideas, reflect on past results, and further refine their mental model. Collaboration may facilitate learning in complex domains, but evidence about its effectiveness has not yet been demonstrated clearly. Nevertheless collaboration is effective insofar as it enhances active learner engagement, which is generally crucial to improved learning outcomes (see Christensen et al., 2000).

Research Questions and Hypotheses

According to Johnson-Laird (1989), it is important to understand the learning-dependent progression from a novice’s to an expert’s understanding of a complex domain (Ifenthaler & Seel, 2005). Therefore, we will compare the learner’s understanding of the underlying model of the simulation with that of the expert to assess the final learning outcome. Accordingly, we will address the following research questions in our empirical investigation:

1. Is the simulation-based learning environment an effective instructional method for improving the understanding of complex domains (systems)?

2. Is the comparison of novice models and expert models a valid instrument for assessing learning-dependent progression in complex domains?

We claim that our simulation-based learning environment is an appropriate tool for enhancing students’ complex problem solving skills. Interacting with a simulation that models the behavior of a complex dynamic system enables learners to create and test hypotheses in order to refine their mental model about the problem. Collaborating with peers provides learners with new perspectives, facilitates reflection, and thus helps them to develop a solution to the given problem. We hypothesize that students will perform better on the posttests than on the pretests.

We argue that learning can be viewed as moving along a continuum towards expertise. Therefore, the similarity between students’ models and the expert model (the causal model of the simulation) should increase, thus indicating improved learning. To measure these outcomes, we have two different instruments. The simulation output, which calculates the new dropout rate of the school, provides a direct indication about the quality of a learner’s decision. Success is reflected by decreased dropout rates and failure by increased dropout rates. Secondly, a questionnaire will be given to the learners in which they will be asked to rate the impact of the variables on each other and on dropout rates. The data from the questionnaire will be transformed into a causal model which will then be compared to the expert model, i.e. the underlying model of the simulation. The higher the similarity score, which ranges from 0-1, the more similar are the models.

Method

Participants

The sample consisted of 66 students (54 female and 12 male) from a German Gymnasium (high school). Their average age was 17.94 years (SD = 1.28). It included three classes from grades 11-13. The Gymnasium participating in the study is specialized in pedagogy, which means that the area of school management is part of the students’ curriculum, might be related to future job opportunities, and is likely of personal interest to them.

Design and Procedure

The study was conducted during two regular 90 minute class sessions and was divided in four main phases (see Table 1). The study started with a short introduction in phase 1 and then students logged into the online learning environment and completed a demographic survey and pretests. Then they were given more information about the study and scenario along with instructions in class. The students’ first learning task was to work through the provided instructional materials and learn about the categories involved in school management in phase 3. Students were randomly assigned to discussion groups of five for the collaborative part in phase 3. During the second 90 minute class session, students were given 60 minutes to discuss their results and collaboratively explore the causal model of the simulation in order to achieve an optimal result in the last simulation run. After the collaborative activities, the study closed with the posttests and an evaluation survey (see Table 1).
In the evaluation survey, students were asked to evaluate different aspects of the learning environment on a five-point Likert scale (1 = I disagree, 5 = I fully agree). It served to assess usability of the learning environment, the benefit of instructional support, and success of the collaboration. Both pretest and posttest consisted of two scores that will be further explained in the following paragraph on outcome variables.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Activities</th>
<th>Obtained Data</th>
<th>Point in time</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pretest and demographic survey</td>
<td>dropout rate</td>
<td>Day 1</td>
<td>45 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>causal model similarity score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Learning activities I: individual tasks</td>
<td>dropout rate</td>
<td>Day 1</td>
<td>45 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>causal model similarity score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Learning activities II: collaborative task</td>
<td>dropout rate</td>
<td>Day 2</td>
<td>60 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>causal model similarity score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>students’ evaluations of the learning environment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Posttest and evaluation survey</td>
<td>dropout rate</td>
<td>Day 2</td>
<td>30 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>causal model similarity score</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome Variables

Two outcome variables were used for this study, namely (1) dropout rate and (2) causal model similarity score. The simulation returned a new calculated dropout rate after each run based on the learner’s input which provides performance information. Good decisions from the learners and effective placement of investments will lead to a lower dropout rate. Non-effective investments result in an increased dropout rate. This means that the lower the dropout rate, the better the performance. Participants were asked to rate the impact of variable that are part of the causal model underlying the simulation on each other with a score between 1 and 10. Each score was transformed into a directed relation. All obtained relations combined resulted in a causal network. This was then compared to the causal model underlying the simulation using the SMD (Surface, Matching, Deep Structure) software developed by Ifenthaler (2008b). The software calculates a score ranging from 0 (no similarity) to 1 (identical) indicating similarity of a participant’s causal model with the underlying model of the simulation. The closer a participant’s causal model is to the expert model, i.e. the underlying model of the simulation, the better the person performs.

### Results

The collected data were quantitatively analyzed to first explore the effectiveness of the simulation-based learning environment on students’ learning outcome. It was expected that students performed better on the posttests than on the pretests. Second, a correlation analysis between the two outcome variables, dropout rate and similarity score, was conducted to cross-validate the two measurement instruments.

### Effectiveness of simulation-based learning environment

A prior version of the simulation (DIVOSA; see Ifenthaler, 2002, 2009) had been used before and showed promising results. The instructional materials along with the collaborative part should have enhanced students’ understanding of the complex problem, namely school management. We thus expected indicators for learning from pretest to posttest reflected by a higher similarity between the learner’s model and the expert model on the one hand and a lower dropout rate on the other hand. Paired t-tests were conducted to assess learning gains (posttest - pretest comparison). The descriptive data included in Table 2 show that the mean 6.86 (SD = 2.23) of the dropout score in the pretest is higher than the mean 5.71 (SD = 1.81) of the score in the posttest.

<table>
<thead>
<tr>
<th>Table 2: Paired samples statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropout rate</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>pretest score</td>
</tr>
<tr>
<td>posttest score</td>
</tr>
</tbody>
</table>

Note. A lower score indicates better performance
A paired-samples t-test was employed to test the hypothesis that there are differences between the dropout rate obtained in the posttest compared to that in the pretest. The t-test revealed significant differences between the dropout rates obtained in the pretest and those from the posttest with $M$ difference $= 1.15$ ($SD = 2.54$), $t(65) = 3.69$, $p < .01$, $d = .61$ (medium effect). Therefore, the data support the stated hypothesis which means that students performed better on the posttest than the pretest.

Not all students completed the questionnaire for the similarity score and therefore, six datasets had to be excluded from the analysis due to the missing values. The descriptive data of the remaining sample ($N_{rem} = 60$) in Table 3 show that the mean $M = .248$ ($SD = .179$) of the posttest causal model similarity score is higher than the mean $M = .230$ ($SD = .178$) of the pretest causal model similarity score.

<table>
<thead>
<tr>
<th>Similarity score</th>
<th>N</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest score</td>
<td>60</td>
<td>.23</td>
<td>.18</td>
</tr>
<tr>
<td>posttest score</td>
<td>60</td>
<td>.25</td>
<td>.18</td>
</tr>
</tbody>
</table>

A paired-samples t-test was employed to test the hypothesis that there are differences between the similarity score obtained in the posttest compared to that in the pretest. The t-test revealed no significant differences between the dropout rates obtained in the pretest and those from the posttest with $M$ difference $= .02$ ($SD = .22$), $t(59) = .59$, $p = .56$. This means that the data do not support the hypothesis and the similarity score did not improve during the study.

Table 4 shows the results of the survey where students were asked to evaluate the learning environment with 1 representing a low score and 5 a high score. The survey covered three different areas, namely usability, instructional support, and collaboration. Each of them was represented by several items as presented in Table 4. The data show that students positively rated the learning environment since all three factors received high average ratings with acceptable standard deviations.

<table>
<thead>
<tr>
<th>Factor</th>
<th>No. of Items</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>usability</td>
<td>21</td>
<td>3.80</td>
<td>.48</td>
</tr>
<tr>
<td>instructional support</td>
<td>14</td>
<td>3.63</td>
<td>.45</td>
</tr>
<tr>
<td>collaboration</td>
<td>13</td>
<td>3.40</td>
<td>.56</td>
</tr>
</tbody>
</table>

Comparison of expert and novice models for assessing learning gains

It was argued that learning is regarded as a progression from novice to expert models. We assumed that a success in the simulation game (low dropout rate) will be associated with a high similarity score between the learner’s model and the expert model (simulation’s causal model). Thus, we expected a high validity of our research instruments using a cross-validation approach between the two instruments (simulation output, and causal model assessment) respectively. We expected a negative correlation between the two measurement instruments. Correlations between the two scores were calculated and revealed a positive correlation of $r = .26$ between the two posttest scores and a negative correlation $r = -.03$ between the pretest scores. None of the two correlations were significant which indicates that there is no correlation between the measurement instruments. The results raise concerns regarding validity of the two applied measurement instruments that requires further examination which will be taken on and discussed in the discussion section.

Discussion

Purpose of this study was to examine effects of a simulation-based collaborative learning environment as well as methodologies to assess learning outcome. The learning environment included instructional materials, communication tools, and a simulation which was embedded in a complex, authentic scenario. The content evolved around the topic area of school management and learners were asked to reduce the dropout rate of a school and thus to improve its efficiency. 66 students from three classes from a German gymnasium specializing in pedagogy participated in the study. Two outcome variables were used to assess students’ performance. The first one was the dropout rate returned by the simulation with a low score indicating success. The second one was a score representing similarity between students’ causal models and the simulation’s underlying model with a high similarity indicating success.
The first research question concerned the effectiveness of the learning environment and was addressed by a pretest-posttest comparison for each of the two outcome variables. Additionally, an evaluation form was given to students to rate usability of the learning environment. Results varied between the first outcome variable, the dropout rate and the second outcome variable, the similarity score. On the one hand, paired samples t-tests on the first output variable showed that students were able to significantly reduce the dropout rate of the simulation from pre- to posttest thus indicating learning gain. On the other hand, paired samples t-test on the second output variable, showed that the similarity between students’ causal models and the simulation’s underlying model did not increase. This means that students’ performance did not improve after students began working with the simulation-based learning environment given the assumption that that learning is a process whereby novices move towards expertise.

At first, it seems odd that the two outcome variables revealed contrary results regarding the effectiveness of the learning environment. The contrary results also affect the second research question concerning the applicability of the used instruments for assessing learning-dependent progression in complex domains. Results revealed that a cross-validation between the two instruments failed since no correlation between the scores was found. A possible explanation might be validity problems with either one of the outcome variables or both. The following five reasons suggest that the similarity score has validity issues: first, software errors occurred during that data collection process which resulted in missing datasets. Second, students asked many questions when filling out the online questionnaire which indicates that the interface needs improvement in its user friendliness and usability. Third, students reported that the similarity questionnaire made not much sense to them which might be a result of missing prerequisites to work with that instrument. They were not particularly trained regarding system thinking, modeling, and other prerequisites that might be necessary to understand a school and the involved areas such as teachers and curriculum as a complex system that consists of interrelated components. Fourth, the diagnosis of internal knowledge structures is one of the most challenging tasks and remains unsolved in many respects (Al-Diban, 2008; Ifenthaler, 2008a). Measuring an internal encoding requires the process of externalization which can be seen as a re-representation of the person’s understanding about the world. Each assessment technique may exacerbate the externalization process and thus lead to invalid results (Ifenthaler, 2008a, in press; Ifenthaler, Masduki, & Seel, 2009). A fifth reason for the validity problems with the similarity score might be, that students’ mental models might be organized in a way that enables them to operate and understand the simulation, but differs from the simulation’s underlying model. This would explain why students were able to reduce the dropout rate without increasing the similarity of their causal network to the model underlying the simulation. There might be a knowledge structure or mental model that enables students to successfully manage the school not being analogous to a causal network like the one underlying the simulation.

Based on the assumption that the similarity score is not a valid measure, but the dropout rate is a valid measure for students’ performance in the present study, we claim that our simulation-based learning environment is an effective learning tool. The significant reduction of the dropout rate from pretest to posttest suggests learning gains. Students’ positive attitude towards the learning environment reflected by the evaluation questionnaire indicates an adequate challenge level, which means that students did benefit from this particular implementation of technology in education. Students were coping well in the new territory and seemed to enjoy the complex, authentic activities. These findings support the benefit of authentic activities and are in alignment with researchers who advocate inquiry-based learning and recommend to have students work on complex problem-solving rather than providing them with well-defined tasks (Collins et al., 1989; Harley, 1996; Jonassen, 1997; Karagiorgi & Symeou, 2005; Savery & Duffy, 1995b).

Future research will include a further development of our learning environment. Based on our empirical findings, we will continue developing adequate assessment techniques and investigate their effectiveness in future research studies.

References


