Pedagogy in Educational Simulations and Games

Koen Veermans & Tomi Jaakkola

University of Turku, Turku, Finland

Email: koen.veermans@utu.fi, tomi.jaakkola@utu.fi

Abstract Educational simulations and serious games hold great potential for creating engaging and productive learning environments in science, technology, engineering and mathematics (STEM) domains. In this paper, we present and reflect on some of our research findings from a series of studies on a computer simulation in the domain of electricity. These studies used the same simulation with varying instructional designs and over a range of grades. Interestingly, each design had a unique influence on either student performance or student engagement, or both. We hope our results can provide insight for designers producing simulations (or, serious games) for education and for educators utilizing these designs in practical settings.

1.1 Introduction

From their inception, educational simulations have held the promise of creating engaging and productive learning environments in science, technology, engineering and mathematics (STEM) domains. Some of the advantages that have been put forward in the literature include simulations being learner-centric, scalable, reusable; having affordances related to illustration and visualization; leading to student interest and engagement; and producing desirable learning outcomes, particularly in terms of conceptual knowledge but also with regard to developing understanding about scientific inquiry (Slavin, Lake, Hanley, & Thurston, 2014).
In addition to these advantages, the current learning analytics trend towards obtaining learner data in order to analyze productive learner behavior also adds to renewed and strengthened interest in educational simulations and serious games. However, this trend does not mean that the outcomes of learning with and from educational simulations or games are straightforward or always positive. In this paper, we will present and reflect on some of our research findings from a series of studies with a computer simulation in the domain of electricity (e.g. Jaakkola & Nurmi, 2008; Jaakkola, Nurmi

& Lehtinen, 2010; Jaakkola, Nurmi, & Veermans, 2011; Jaakkola, 2012; Jaakkola & Veermans, 2015; 2018). Through these studies, our aim is to demonstrate that sound pedagogical design can make a simulation (or a game) suitable and effective across a wide age range of pupils, and that different pedagogical decisions can have notable impact on students' learning.

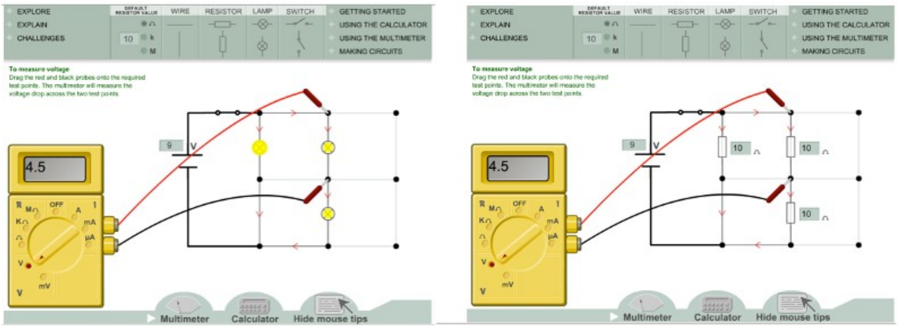

Fig 1.1. The simulation with bulbs (left) and resistors (right) as used in the studies.

1.2 General Settings in the Studies

In the studies that are reviewed in this paper, student participants used the same simulation (see Fig 1.1) to build circuits, observe circuit behavior, and study the properties and underlying principles of electric circuits. The representation level of the simulation was semi-realistic; it displayed circuits schematically but also included light bulbs with dynamically-changing brightness and realistic measuring devices. The simulated model was authentic but for two exceptions: The wires had no resistance and the battery was always ideal (i.e. there was no change in its potential difference with time).
The students' inquiry process with the circuit simulation was supported and guided by instructional worksheets designed to confront and overcome common misconceptions about electric circuits. In general, the worksheets asked students to construct various circuits and conduct electrical measurements with the simulation. The worksheet also provided scaffolding for the students to predict, investigate and infer how the changes and differences in circuit configurations affected circuit behavior. The worksheets began with a very simple and structured task, wherein the students were asked to construct a circuit with one battery, wires, and a bulb. Subsequent tasks were progressively more challenging and open-ended, requiring students to construct more complex circuits that met a specific criterion (e.g. brightness of bulbs has to be A > B = C).

The general procedure was identical across all studies. In the beginning, students took a pre-test designed to assess their baseline knowledge of electric circuits. The pre-test scores were then used to assign students into the different conditions. Matched pairs of students were created based on the pre-test scores, and the students in each pair were allocated randomly to either of the conditions. This procedure ensured relatively small differences in pre-test knowledge between the conditions, which made the assessment of learning gains during the intervention easier between the conditions. After students were allocated into the conditions, random pairs of students were created within conditions. These pairs then had approximately 90 minutes to build and study circuits in the simulation and solve various circuit challenges listed in the worksheets. To assess students' level of engagement during the simulation task, students were asked to indicate their situational interest in the beginning, middle, and end of the intervention in some of the studies. The post-test that was designed to assess changes in students' subject knowledge during the intervention was administered one day after the intervention. Although the students worked in pairs during the intervention, they completed all the tests individually.

Interestingly, though the overall impact was predominantly positive, each design had a unique influence on student performance and/or engagement. We hope our results can provide some new insights for designers when designing simulations (or, serious games) for education and for educators utilizing these designs in practical settings.

1.3 Learning Outcomes, Interest and Learning Time

The general goal of our studies has always been to study learning outcomes across different settings, but gradually, due to reports and literature indicating that students' motivation and interest towards science start declining from their early years at school (European Commission, 2011; Osborne & Dillon, 2008; Vedder-Weiss & Fortus, 2011, 2012), interest in science (more specifically, interest during science tasks) became an integral part of our investigations (Tapola, Veermans, & Niemivirta, 2013; Tapola, Jaakkola, & Niemivirta, 2014). In other words, the goal should be to design learning environments that are both productive (good learning outcomes) and engaging (motivating from students' perspectives).

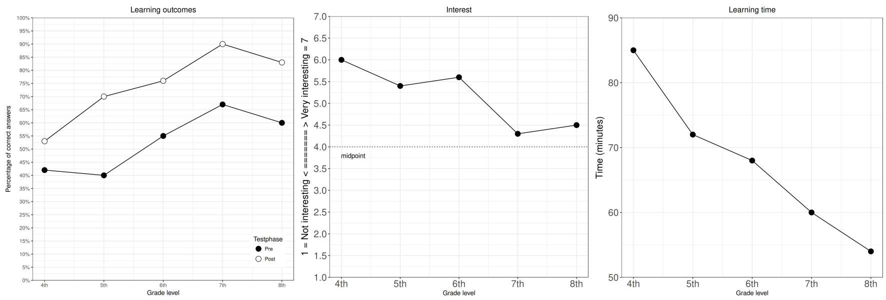

Fig 1.2. Trends in learning outcomes (pre- and post-), interest and learning time across grades 4-8 (9-15 years)

As a starting point, we present the global findings on learning outcomes and interest that were obtained in studies across a range of grades (Fig 1.2) As can be seen from the first graph, regardless of their initial knowledge, students gained knowledge while interacting with the simulation-based learning environment in all five grades, with the smallest gain occurring in grade 4, the largest in grade 5, and similar gains in grades 6 and 8 (the overall outcome shows a significant linear interaction between grade level and post-test scores; students scored higher as a function of grade level, $p<.001$). Interest was above midpoint for all grades, but the second graph shows the presence of a tendency for decreasing interest with increasing grade.

Based on the overall outcomes, it can be argued that the learning environments used in these studies fulfilled the aims of the learning environments both in terms of learning outcomes and interest. However, the results also showed that different outcomes regarding learning and interest were obtained over the different grades. This was the first indication that simulation design was not the only factor that affected participants' learning outcomes and experiences. More indications will be pointed out in the following sections in which we investigate the impacts that different design elements and decisions had on the outcomes in greater detail.

1.4 Instructional Support

In the previous section, we argued that it was not the simulation alone that mattered, since the results showed clear differences across grades. Learner support in the instruction was more structured in the beginning but became gradually less structured and more open-ended towards the end. The results of the 7th and 8th graders in terms of learning process, outcomes and interest levels supported the idea that this shift from structured to open-ended lay at the core of the successful use of the same simulation learning environment over a range of grades. The 7th and 8th graders were quick (about 20% less time needed to complete the learning tasks than elementary school students), but not highly interested during the early phases of the learning task with relatively structured instruction. Their perceptions changed when the tasks became more open-ended, resulting in higher perceived interest towards the end (interest at time 1=4.1 while interest at time 3=4.5) and good learning outcomes ($p<.001$, see Fig 1.2). This showed that an environment with instructional support that built on the idea of fading from structured to open-ended was flexible towards individual differences among learners and could thus address students with different proficiency levels within a grade, but also provide learners across a wider age range with engaging and productive learning experiences.

For younger learners, it was not only about instruction being more or less structured (the 4th graders' learning outcomes might indicate that the more open-ended tasks at the end were too difficult for them) but also about how that structure was presented, as was illustrated by the results from a study that varied the type of instruction in terms of the level of scaffolding (Jaakkola, Nurmi & Veermans, 2011). One condition received implicit instruction–procedural guidance (e.g. what kind of circuit to construct and what to measure)–while the other condition received more explicit instruction in which the implicit instruction was accompanied with the rationale behind the explorations (e.g. students received guidance on where to focus their attention). The results showed clear differences between the two conditions both in terms of learning outcomes (higher in the explicit condition: post=11.29, pre-post $p=.013$, $d=.78$, vs. post=9.75, pre-post $p=.31$, $d=.22$) and in terms of learning time (longer in the explicit condition: 79 min. vs. 66 min.). This result also highlighted that it was not just about the simulation, but also about the pedagogy and what they together triggered in the learner. For learning to occur, a learner should not just be interacting with a learning environment but also be cognitively engaged in the activity. Both the shorter learning time and the lower learning gains suggested that this was not happening in the condition with implicit instruction.

1.5 Perceptual Concreteness of Simulation Elements

An important but often overlooked design consideration for the development of educational simulations is the determination of how concrete the simulation (elements) should be, since the perceptual and conceptual concreteness of any representation can greatly affect what the students learn and how they can utilize that knowledge in other situations. A simulation with concrete elements is easier to relate to and understand, but extensive amount of detail can also hinder extraction of relevant information and result in overly contextualized understanding. More abstract elements in a simulation can highlight relevant information and support generalization, but at the cost of becoming more difficult to understand. As a result, concreteness fading–starting with concrete elements to ensure proper contextualization, then switching to abstract elements to ensure that understanding is less bound to specific contextual details–has been proposed as an optimal solution (Fyfe, McNeil, Son & Goldstone, 2014).

In the studies for which the outcomes were summarized above, learners either used a simulation environment that employed this idea of concreteness fading or used a simulation with solely concrete elements. In the concrete condition, learners used concrete elements (bulbs) throughout the learning phase while in the fading condition, they started with concrete elements (bulbs) but switched to abstract elements (resistors) after some time. Though bulbs and resistors could theoretically be argued to be equally concrete or abstract (equivalent in the physics sense–bulbs are a special case of resistor), our empirical evidence shows that from a learner's point of view (i.e. perceptually and conceptually) they are not. In the pre-test, questions related to bulbs were answered correctly significantly more often than questions on resistors (both 5^{th} and 6^{th} and 7^{th} and 8^{th}, $p<.01$). Moreover, the first study with this design comparison of the two conditions in terms of learning outcomes showed that learners (5^{th} and 6^{th} graders) in the concrete condition outperformed learners in the fading condition on the post-test ($p<.05$, consistent over both grade levels; Jaakkola & Veermans, 2015). As it turned out, a considerable number of the 5^{th} graders in the fading condition failed to complete the learning tasks within the given time frame, indicating a clear difficulty compared to the concrete condition where practically all participants were able to complete the tasks within the given time frame. In the fading condition, 11% of the 5^{th} graders and 75% of the 6^{th} graders completed all tasks while in the concrete condition, 75% of the 5^{th} graders and 100% of the 6^{th} graders completed all tasks (χ^2 for all $p<.001$). This result indicated that students' interaction with resistors rather than bulbs caused the learning process rate and learning outcomes to deteriorate. During this study, perceived interest was also assessed at several points in time, with an in-depth analysis of interest (Tapola, Veermans, & Niemivirta, 2013) showing an interaction between the level of interest for students and condition ($p=.002$, $\eta2=.13$; interest in the concrete condition increased while interest in the fading condition decreased during the conditions' interactions with the learning environment), revealing that concrete elements were to be preferred for these students from an interest perspective at this age, in addition to being better from a learning perspective as found in Jaakkola & Veermans (2015).

In a follow-up study, the fading from concrete to abstract elements was slightly delayed (i.e. students used bulbs a bit longer before moving to resistors), with the underlying hypothesis that the delay might help students align and link the representations better (Jaakkola & Veermans, 2018). The results showed that delayed fading was able to improve the learning outcomes of 6th graders but not 5th graders (among 5th graders, the concrete condition continued to outperform the fading condition).

One design consideration that could be derived from these results is therefore that concrete representations are especially important for younger students and that their benefits may outweigh the potential benefits of generalization through abstract representations even when fading is employed. Generalization benefits might occur only at a later age–there were some indications of this effect in the study on 7th and 8th graders.

1.6 Virtual alone or together with real?

Traditionally, virtual (computer-based) and real (physical equipment) labs and learning environments have been positioned as competitors, with proponents of real environments emphasizing the importance of tangible experiences for the development of learning and understanding while proponents of virtual environments emphasize that manipulation rather than physicality is at the core of this learning and development and that physicality aside, virtual environments have many advantages over real environments. The main focus of theorizing and empirical research on both sides has been to show that one is better and thus to be preferred. In one of our studies (Jaakkola, Nurmi & Veermans, 2011), we explored the possibilities of combining the two approaches in a 2x2 design comparing the two types of instruction mentioned in the previous section (though the previous section discussed only the simulation conditions of this 2011 study) and comparing virtual to a virtual-real combination where students did everything with both the simulation and real equipment (twice the amount of the same work). The instruction dimension extended an earlier study in 2008 that investigated real, virtual and virtual-real environments, but the real condition was excluded due to its showing the least favorable results (Jaakkola & Nurmi, 2008).

The results of the 2011 study showed that learning outcomes were better in the combination condition regardless of the instructional support (explicit or implicit), but also that there was an interaction between instruction and learning environments. In the simulation-only condition, explicit instruction prolonged learning time and enhanced learning outcomes (79 min, post=11.29, pre-post: $p=.013$, $d=.78$, vs. 66 min, post=9.75, pre-post: $p=.31$, $d=.22$), while the same instruction in the combination condition prolonged learning time (similar in magnitude) without enhancing learning outcomes (90 min, post=12.33, pre-post: $p<.001$, $d=1.24$ vs. 73 min, post=12.67, pre-post: $p<.001$, $d=1.51$). In fact, the implicit combination outperformed the explicit simulation condition, and did so within a shorter learning time. This result highlighted once more that it was about the whole learning environment and what it triggered in the learner. Based on explicit or implicit instruction alone with simulation or based on a comparison between explicit simulation and explicit combination, one might have concluded that the students in the lower outcome condition might have learned more if they had spent more time. However, the comparison between implicit combination and explicit simulation showed that it was not that simple.

Implicit instruction alone was not sufficient in the virtual setting to trigger cognitive involvement, and the addition of explicit instruction improved students' performance. Adding real circuits into the combination was even more effective and seemed to make explicit instruction unnecessary. In contrast, adding explicit instruction seemed to slow down the learning process with no apparent learning outcome benefits.

While these results stem from real and virtual environments, it has been argued that the distinction between real and virtual is construed rather than definite (Burbules, 2004). More and more of our present day realities are in fact virtual, and trying to maintain a strict dichotomy is therefore no longer helpful. A more flexible framework of thinking that allows virtual experiences to be real experiences allows us to focus more on the nature of the experiences rather than whether they are real or virtual. In the case of the real-virtual setting, experiencing two distinct representations might be of bigger importance for the experience than one being real and the other virtual. If that is the case, the real and virtual experiences from the real-virtual setting described above might very well be incorporated in a similar way in a fully virtual environment, as there is no limitation that prevents combining virtually-real and simulation the same way that real and simulations were combined in this study.

1.7 How to Combine Multiple Representations?

There are principally two different ways to combine laboratories and simulations (or any mix of multiple representations). In a sequential combination, laboratories and simulations are used in different phases while in a parallel combination, both laboratories and simulations are available at all times. In other words, the main difference is that in a sequential combination, representations are not co-present (most of the time), whereas in a parallel combination, both representations are always co-present. In the studies described here, real and virtual circuits were combined in parallel (Jaakkola & Nurmi, 2008; Jaakkola, Nurmi, & Lehtinen, 2010; Jaakkola, Nurmi & Veermans, 2011) while concrete and abstract simulation elements were combined sequentially (bulbs-then-resistors; Jaakkola & Veermans, 2015; 2018). While we have not explicitly compared the two types of combination in a single study, our results across studies (see also Gentner, Loewenstein, & Thompson, 2003) suggest that parallel combination may be the safer option. This may be because parallel combinations are less sensitive to order, while sequential combinations must consider order in the study design (e.g. laboratory or simulation first; concrete or abstract first). Order is far more critical and sensitive in the sequential combination as the representations are used in isolation. Ideally, this entails that each representation should be used when the benefits it can offer learners is higher than the benefits offered by the other representation(s). However, it is not an easy task to determine in practice which representation can offer the highest benefit in a particular context or at a particular moment, because the true benefit is determined by both the informational and computational properties of the representations[1], which depend on the individual learner's characteristics to some degree.

1.8 Conclusion

The previous sections presented several key themes from the outcomes of research on a simulation-based learning environment in the domain of electricity. This conclusion will draw on the findings to connect the design of learning environments and the design of educational games. The positive news from this research is that a single learning environment that progresses from more structured in the beginning to more open-ended later on can provide learners across a wide age range with engaging and productive learning experiences. While this matches well with the idea of levels in game design, it is important to focus explicitly on the function of these levels in terms of their learning outcomes (apart from their function for maintaining interest) since the findings across grades from the studies show that there does not seem to be a straightforward relation between interest and learning outcomes. As can be seen in Fig 1.2, 4th graders combined the highest interest with the lowest learning outcomes while in the higher grades where interest was lower, learning outcomes were still considerable. This suggests that the relation between interest and learning outcomes may be a threshold rather than a continuous function in these kinds of learning environments. Game design may not necessarily need to strive to maximize learners' interest, but instead follow a 'good enough' principle. This would leave room for other design decisions that may be more important for obtaining good learning outcomes. One area that may be important to consider in this respect is if there are ways to include elements that trigger reflection, as the studies examined suggest that reflection (either instigated by instruction or by parallel representations) is more important than higher

1 According to Larkin and Simon (1987), the informational effectiveness of a representation is determined by how much information it contains, whereas the computational effectiveness of a representation is determined by how easily relevant information can be extracted and applied from it.

levels of interest. Given the nature of games, instruction might not be the easiest option but parallel (or sequential) representations can be designed into games. This could provide games with elements that trigger reflection without greatly affecting the game's nature.

An area where the design of simulation learning environments could benefit from practices common in game design is in their attention to use and user data for adjusting initial designs. Analyzing use in reference to the intended outcomes makes it more natural to look for aspects in the environment that may be varied and lead to better learner experiences and outcomes. For instance, we found that even a slight delay in the fading from bulbs to resistors led to better learning outcomes and a more efficient learning process with younger learners (5^{th} and 6^{th} graders) in an experimental study. Once an environment has a substantial number of users, the timing of change in sequential designs (e.g. going from concrete to abstract representations) can be randomized in order to find out whether this is an important factor that influences learning outcomes. Adopting this practice in the building of simulations and serious games could steer designers away from development over a single cycle or a few cycles towards more flexible and continuous iterative development, which may better fit the complicated reality of the interaction between learning environment and learner.

Acknowledgements

This work was supported by grants no 116393 and 252580 from the Academy of Finland to the first author and grant no 266189 from the Academy of Finland to the second author.

References

Burbules, N.C. (2004). Rethinking the Virtual. E-Learning, 1(2), 162-183European Commission. (2011). Science Education in Europe: National Policies, Practices and Research.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness Fading in Mathematics and Science Instruction: A Systematic Review. Educational Psychology Review, 26, (1), p9-25.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. Journal of Educational Psychology, 95(2), 393–408.

Jaakkola, T. (2012). Thinking Outside the Box: Enhancing Science Teaching by Combining (Instead of Contrasting) Laboratory and Simulation Activities. Annales Universitatis Turkuensis B 352. Turku: Painosalama. Academic Doctoral Dissertation.

Jaakkola, T. & Nurmi, S. (2008). Fostering Elementary School Students' Understanding of Simple Electricity by Combining Simulation and Laboratory Activities. Journal of Computer Assisted Learning, 24(4), 271-283.

Jaakkola, T., Nurmi, S. & Veermans, K. (2011). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. Journal of Research In Science Teaching., 48, (1) 71-93.

Jaakkola, T. & Veermans, K. (2015). Effects of abstract and concrete simulation elements on science learning. Journal of Computer-Assisted Learning, 31(4), 300-313.

Jaakkola, T., & Veermans, K. (2018). Exploring the effects of concreteness fading across grades in elementary school science education. Instructional Science, 46(2), 185-207.

Larkin, J.H. & Simon, H.A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. Cognitive Science, 11(1), 65-100.

Osborne, J., & Dillon, J. (2008). Science education in Europe: Critical reflections (Vol. 13). London: The Nuffield Foundation.

Slavin, R. E., Lake, C., Hanley, P., & Thurston, A. (2014). Experimental evaluations of elementary science programs: A best-evidence synthesis. Journal of Research in Science Teaching, 51(7), 870–901. doi: 10.1002/tea.21139

Tapola, A., Jaakkola, T. & Niemivirta, M. (2014). The influence of achievement goal orientations and task concreteness on situational interest. The Journal of Experimental Education, 82(4), 455-479.

Tapola, A., Veermans, M., & Niemivirta, M. (2013). Predictors and outcomes of situational interest during a science learning task. Instructional Science, 41, 1047-1064.

Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' Declining Motivation to Learn Science: Inevitable or Not? Journal of Research in Science Teaching, 48 (2), 199-216.

Vedder-Weiss, D., & Fortus, D. (2012). Adolescents' declining motivation to learn science: A follow-up study. Journal of Research in Science Teaching, 49 (9), 1057-1095.