INSTRUCTION BASED ON COMPUTER SIMULATIONS AND VIRTUAL LABORATORIES

Ton de Jong

INTRODUCTION

In the scientific debate on what is the best approach to teaching and learning, a recurring question concerns who should lead the learning process, the teacher or the learner (Tobias & Duffy, 2009). Positions taken vary from a preference for direct, expository, teacher-led instruction (Kirschner, Sweller, & Clark, 2006) to fully open student-centered approaches that can be called pure discovery methods (e.g., Papert, 1980), with intermediate positions represented by more or less guided discovery methods (e.g., Furtak, Seidel, Iverson, & Briggs, 2012; Mayer, 2004). This discussion is a recurring theme in this chapter.

In addressing the issue of the role of technology and guidance in instruction, the specific technology of computer simulations (or virtual laboratories) occupies a central place (Donnelly, Linn, & Ludvigsen, 2014; Krajcik & Mun, 2014). Computer simulations and virtual labs, through their interactive character, offer a special opportunity for student centered, inquiry-based learning, while at the same time offering options for program or teacher led guidance of the learning process. Thus, a major goal of this chapter is to examine whether people learn better when simulations include guidance (i.e., guided discovery method) or when simulations allow people to learn freely without much guidance (i.e., pure discovery method). Another goal of this chapter is to examine whether individuals learn better with computer simulations than with conventional instructional media or with traditional wet laboratories.

Computer simulations are computer programs that have as their core a computational model of a system or a process. The system or process that is modeled normally has a natural world origin and the model that is created is usually a simplification (i.e., reduction and abstraction) of the real world phenomenon (de Jong & van Joolingen, 2008). When learning with simulations, learners interact with the model through an interface that enables them to change values of input variables and observe the effects of these changes on output variables (see, e.g., de Jong, 2006a); in the case of virtual laboratories the interface of the model resembles a real laboratory (de Jong, Linn, &
Instruction Based on Computer Simulations

Simulations or virtual labs can be used as the basis for training of knowledge or skills (or a combination of both, van Berkum & de Jong, 1991).

In the remainder of this chapter, I provide an historical overview of educational computer simulations, a theoretical framework for learning from computer simulations, and summaries of the current state of research on the effects of incorporating guidance with computer simulations and on the effects of teaching with computer simulations rather than conventional media. The chapter closes with a discussion of practical and theoretical implications and an examination of future directions.

HISTORICAL OVERVIEW

Simulations have long been used in training to avoid risks for the operators (e.g., aviation), subjects (e.g., medicine), or both operators and subjects (e.g., military, business). One of the earliest reports on the use of a simulator for learning dates back to the Ruggles orientator (Jones, 1927)—a device to test pilots on very basic skills. These devices gradually became more sophisticated, with today’s flight simulators offering a very high level of functional and physical fidelity. Besides these full-fledged simulators, computer-based simulators have been very popular, beginning from the landmark introduction of Microsoft Flight Simulator in 1979/1980. These relatively inexpensive computer-based simulators have also shown their value for training pilots (see, for an overview, Koonce & Bramble, 1998). In medicine, the use of simulators and simulations has a century long history, but simulators started to enter medical education at a reasonable level only during the second half of the 20th century. Nowadays a rapid increase in the use of simulators in medical education is underway, ranging from high fidelity simulators to desktop computer-based simulations (Bradley, 2006). Possibly the longest history of simulation use can be found in the military. Wargames were often used, with the earliest known version, the German Kriegspiel, dating back to the early 19th century. Even before that, the military used sand tables with iconic representations. Nowadays, digital games are dominant and they exist in many variations for professional and private usage (Smith, 2010). Business games and simulations have a shorter history that dates back to the 1950s. Still, there are many of these games in existence now using technologies that enable online access. However, there is not much research that examines the outcomes of this type of learning, and the research that exists is not very definitive on the benefits of business games for the acquisition of knowledge (Anderson & Lawton, 2009; Leemkuil & de Jong, 2011).

The traditional application areas for the use of computer simulations for training were focused towards practical skills training (e.g., flying, diagnosing, doing business) and no specific guidance for students was built in. In the late 1970s and early 1980s new systems were developed, often in the form of intelligent tutoring systems which emphasized training of conceptual knowledge along with skills and included forms of student guidance. These simulations also began to involve domains other than the traditional ones sketched above.

Many of these systems had a focus on diagnosis and trouble shooting. SOPHIE, for example, was an environment for teaching electronic trouble shooting skills, but also aimed at acquisition of conceptual knowledge of electronic laws, circuit causality, and the functional organization of particular devices (Brown, Burton, & de Kleer, 1982). Similarly, QUEST (White & Frederiksen, 1989) focused on electronic circuit troubleshooting, with central attention for knowledge of underlying models from different
perspectives. Another system that combined the learning of operational and conceptual knowledge was STEAMER. This system simulated a complex steam propulsion system for large ships (Hollan, Hutchins, & Weitzman, 1984) and students also had to understand the models underlying this system. Systems such as MACH-III on complex radar devices (Kurland & Tenney, 1988) and IMTS (Towne et al., 1990) also focused on troubleshooting, again with associated conceptual knowledge. A further example of early conceptual simulations for physics education was ARK (Scanlon & Smith, 1988; Smith, 1986). ARK (Alternate Reality Kit) was a set of simulations on different physics topics (e.g., collisions) that provided students with direct manipulation interfaces. In another field, the MYCIN software was an expert system that contained a large rule base for medical diagnosis, which the GUIDON software took up and augmented with teaching knowledge to make this learning software. A student model was created through an overlay approach and students could receive dedicated feedback on their diagnosis process (Clancey, 1986).

Smithtown was one of the first educational computer simulations that targeted an exclusively conceptual domain (i.e., economic laws) and that included a range of support mechanisms for students (Shute & Glaser, 1990). In Smithtown, students could explore simulated markets. They could change such input variables as labor costs and population income and observe the effects on output variables such as prices.

THEORETICAL FRAMEWORK

Why would learning with simulations or virtual laboratories be better than more traditional, expository, explanation-based, real laboratories, or “on the job training” approaches? There are several considerations why simulation-based learning would improve performance.

When simulations are used to help learners acquire conceptual knowledge the rationale is that they foster deeper cognitive processing during learning, which in turn leads to deeper learning outcomes (in comparison to direct instruction). In particular, computer simulations are intended to encourage learners to engage in a process of inquiry and to actively restructure knowledge (e.g., when data are found that are not consistent with a hypothesis, see, for example, Monaghan & Clement, 1999). The main theoretical frameworks, which can be traced back to Piaget’s original ideas (1976), consider learning with simulations as involving an inquiry cycle consisting of processes such as orientation, hypothesis generation, experiment design and data interpretation, drawing conclusions, and reflection (de Jong, 2006a; Friedler, Nachmias, & Linn, 1990; Pedaste et al., 2015). A related approach describes this scientific inquiry as a specific kind of problem solving with associated moves in a problem space—in this case, hypothesis and experiment space (Klahr & Dunbar, 1988; Klahr & Simon, 2001).

A second approach that highlights advantages of simulations for conceptual learning is the work on multiple representations (Ainsworth, 2006; Mayer, 2009; McElhaney, Chang, Chiu, & Linn, 2014). The basic idea behind this approach is that if multiple representations (e.g., graphs, tables, animations) are offered, translations must be made between these representations, leading to deeper and more abstract knowledge. Simulations often offer multiple representations and these are often also dynamically linked (van der Meij & de Jong, 2006).

A third reason to prefer simulations or virtual labs over expository instruction is that in these environments students are in charge of the learning process; they take the decisions on what to manipulate and the route to follow in the material which is expected to lead to a higher motivation and engagement of students (Sengupta-Irving & Enyedy,
2015), and some studies indeed indicate that simulation-based inquiry activities clearly engage students (e.g., Pedersen & Irby, 2014). A necessary condition for staying engaged, however, is that students keep control over the situation and do not get lost. Guidance, as discussed later, plays an important role in this.

Acquiring conceptual knowledge with simulations and virtual labs also shows affordances over learning with real laboratories (de Jong et al., 2013). In virtual labs students can perform many more experiments in a shorter time than in real laboratories and, in addition, in virtual labs students are not distracted by measurement errors. Another affordance of simulations and virtual labs over the real experiment is that in virtual laboratories reality can be augmented by showing otherwise unobservable phenomena which may show specific conceptions (e.g., forces) directly to students. This idea of augmenting reality is very often used in simulations and virtual labs (Olympiou, Zacharia, & de Jong, 2013; Zacharia & de Jong, 2014).

In simulations and virtual labs a range of situations can be offered that may be rare, dangerous, or expensive compared to training in the real world or lab (Alessi & Trollip, 2001). Practicing of skills (e.g., in aviation, medicine) requires variation in tasks and confrontation with critical tasks (van Merriënboer, 1997); simulations, not having to rely on the natural and possibly rare occurrence of differing and critical situations, offer the best opportunity to effect this. Further, simulations can help to speed up or slow down tasks so that practice can take place more deliberately and moments of reflection may be built in.

Critics assert that learning with simulations is not effective because the required processes are too demanding for students, leading to cognitive overload (Kirschner et al., 2006). However, these authors have focused on unsupported pure discovery and do not consider learning environments that offer students support for their inquiry processes. This support is essential for successful inquiry and thus for learning from computer simulations. The next section outlines this issue and discusses the effectiveness of learning with computer simulations.

THE EFFECTIVENESS OF SIMULATION-BASED LEARNING

One of the central research questions addressed in this chapter is whether simulation-based instruction fosters an effective form of learning. When considering this question it is important to realize that it is not the technology per se that causes learning but rather the instructional method (Clark, 1994). The question of effectiveness also depends on what is being measured (e.g., different types of knowledge, inquiry skills, attitudes) as an outcome of the learning process, and may vary with the students’ characteristics. A short overview of research is presented in the following sections. The first section presents studies comparing different versions of basically the same simulation learning environment, what Mayer (2014) calls value-added research. Here, for example, different types of guidance are compared. The second section follows what Mayer (2014) call media comparison research and presents studies in which completely designed simulation environments are compared to alternative didactic approaches (i.e., expository instruction, real laboratories or on the job training).

The Effects of Providing Students with Cognitive Guidance

Work claiming that direct instruction is superior to inquiry learning generally involves unguided inquiry (e.g., Klahr & Nigam, 2004). What is very clear from the literature is that unguided simulation- (inquiry-) based learning is not effective (Mayer, 2004)
but that adding guidance to inquiry makes this a productive instructional approach (e.g., Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Furtak et al., 2012). Although guidance was already present to some degree in the early computer simulation systems cited here, most specifically Smithtown, recent research has shown that learning with simulations is most effective when the student is sufficiently supported (de Jong & van Joolingen, 1998; Lazonder, 2014; Mayer, 2004; Zacharia et al., 2015). Guidance can be integrated with the software or can be provided by the teacher; it can aim at a specific inquiry process (e.g., hypothesis generation) or at the structuring of the entire process. Contemporary inquiry learning as well as simulation-based learning environments contain all kinds of guidance (Hmelo-Silver, Duncan, & Chinn, 2007). Several overview studies have indicated which types of guidance are available and how they make learning from simulations effective (e.g., Chang, Chen, Lin, & Sung, 2008; de Jong, 2005, 2006a, 2006b, 2010; Fund, 2007; Quintana et al., 2004; Zacharia et al., 2015). Recently, de Jong and Lazonder (2014) distinguished the following types of guidance: process constraints (that limit the range of activities students can do), performance dashboards (that provide students with an overview of their activities and/or learning products), prompts (direction for students to perform certain actions), heuristics (suggestions on how to perform certain actions), scaffolds (electronic tools that take over parts of the learning process), and, finally, direct presentation of information (presenting parts of the to-be-learned information directly).

Orientation

When working with computer simulations, students may need to be supported in identifying the main variables in the domain and find the topic that will be their focus before they continue in the inquiry process (Sharples et al., 2014). According to Zacharia et al. (2015) orientation is one of the least supported inquiry phases. Still, it is an important phase; what happens in this phase may influence the rest of the inquiry process. Activation of prior knowledge was used in a number of studies. Reid, Zhang, and Chen (2003), for example, studied students who learned with a simulation on buoyancy and provided them with interpretative support, which consisted of multiple-choice questions to activate prior knowledge and to make students think about the variables that played a role in the simulation, together with access to a database of background knowledge. This support had a positive effect on students’ intuitive understanding (students’ ability to predict effects of changes), which effect was confirmed in a follow-up study (Zhang, Chen, Sun, & Reid, 2004). Holzinger, Kickmeier-Rust, Wassertheurer, and Hessinger (2009) compared three versions of a course on blood flow for medical students: a text version, a computer simulation, and a computer simulation preceded by a short video explaining the main parameters. Students who received the video plus simulation treatment outperformed the simulation only and text groups (that scored at the same level) on a short multiple-choice test of conceptual knowledge. Lazonder, Hagemans, and de Jong (2010) compared students (without prior knowledge of the domain) who received no information on domain variables involved with students who received this information before interaction with an inquiry task or before and during this task. Offering this information resulted in more hypothesis-driven behavior and higher conceptual knowledge scores. Typical forms of guidance to
support students in the orientation phase are having them create of complete a concept map, having them fill in a quiz, or directly presenting the relevant information to the students.

**Hypothesis and Question Design**

A specific scaffold for hypothesis generation is the hypothesis scratchpad (Shute & Glaser, 1990; van Joolingen & de Jong, 1991). In such a scratchpad, learners can compose hypotheses by filling in if-then statements and by selecting variables and relations to fill in the slots. For each hypothesis, they can indicate whether it was tested or not, and whether the data confirmed the hypothesis. In an empirical study, van Joolingen and de Jong (2003) found that working with such a scratchpad is not very easy for students, but recent versions using a more drag and drop and flexible interface seem more promising. As a variant, when students do not have hypotheses, they like to test a similar tool, the question scratchpad, which helps students to formulate more open, not directed, questions (see, for both, de Jong, Sotiriou, & Gillet, 2014). Following work by Njoo and de Jong (1993), Gijlers and de Jong (2009) provided students with complete, predefined hypotheses. They created three experimental groups of collaborating dyads of learners. One group (control) did not receive specific scaffolds, one group had a shared hypothesis scratchpad combined with a chat, and the final group received a large set of propositions about the domain. As an overall result, students in the proposition group outperformed students in the other two groups on a measure of intuitive knowledge.

In a follow-up study, Gijlers and de Jong (2013) found that when the use of propositions was combined with the construction of a shared concept map, learning processes and results of students further improved. A broad conclusion might be that for beginning students, support in the form of complete hypotheses is more beneficial than having them compose hypotheses themselves, but actively creating hypotheses with the help of an interactive scaffold may be of benefit for more advanced students.

**Investigation**

There is a range of heuristics that can be used to design “good experiments” (Baker & Dunbar, 2000) of which the control of variables strategy (CVS, Chen & Klahr, 1999) is the best known (Wilhelm & Beishuizen, 2003). This strategy implies that only one variable at the time is varied in experiments and the others are kept constant, so that effects can be attributed to that variable that was varied. Klahr, Zimmerman, and Jirot (2011) assert that CVS can best be directly instructed to students but research shows that there are many subtle ways to teach students CVS and to see an effect of this on resulting conceptual knowledge. Kuhn and Dean (2005), for example, showed that even a simple prompt to focus on one variable at a time may help to increase students’ performance level. Keselman (2003) had students work with a simulation-based learning environment on a complex domain with multivariable causality (i.e., risks of earthquakes). One group of students received extensive practice on manipulating variables and making predictions, whereas a second group also observed a teacher modeling the design of good experiments. Both groups improved compared to a control group that received no support, but the modeling group improved most on knowledge gained and on skill in designing good experiments. Künsting, Kempf, and Wirth (2013) gave students metacognitive support in the form of a 20-minute training preceding the interaction with a virtual lab on buoyancy and support during the learning process and
found effects of this training on both the use of CVS during the lab and of conceptual knowledge acquired.

There are also indications that this kind of support is specifically beneficial for students with poor prior knowledge. Reid et al. (2003) gave students explanations on good experimental procedures and helped students structure their experiments. Overall they found no effect of this support on a range of measures, including intuitive knowledge (on which only a marginal effect was found), transfer, and knowledge integration (how far students had related their new knowledge to existing knowledge and how far they acquired deep principles). A follow-up study Zhang et al. (2004) added the nuance of taking learners’ reasoning abilities into account. Zhang et al. found that support that aimed at the experimentation behavior of students was most effective for low reasoning ability students, indifferent for high ability students, and detrimental for the mid-range ability students.

Veermans, van Joolingen, and de Jong (2006) compared two simulation-based learning environments (containing implicit or explicit heuristics for designing experiments) on the physics topic of collision. They found no overall difference between the two conditions on knowledge gained and strategies acquired, but found indications that the explicit condition favored weak students. Recently, van Riesen, Gijlers, Anjewierden, and de Jong (submitted) designed a specific scaffold that helped students to perform unconfounded experiments. In this online scaffold students can choose dependent and independent variables and the design of the scaffold is such that only unconfounded experiments must be designed. Van Riesen et al. (submitted) found that no overall effect of the scaffold compared to conditions that did only contain guiding questions, but found that, compared to the other conditions, the scaffold increased the conceptual knowledge of students with relative poor prior knowledge.

**Drawing Conclusions**

In the conclusion phase, students combine hypotheses or questions with data and observation in order to answer the question or evaluate the hypothesis. This is a difficult process, and people have, for example, many reasons not to refute a hypothesis for which disconfirming evidence was found (Chinn & Brewer, 2001). Woolf et al. (2002) presented guidance in the form of a tool that presents students an overview of all hypotheses and observations that they can link together and use to write a report. In Go-Lab, a similar tool is present that automatically presents all hypotheses and questions generated by a student together with data and observations. Students can link these together, complement this with a written conclusion, and, in case of a hypothesis, indicate how much the conclusion has influenced their original confidence in the hypothesis (de Jong et al., 2014). The latter leaves room for students to indicate if they were insecure of the evidence or found the evidence incomplete.

McNeill, Lizotte, Krajcik, and Marx (2006) provided students with a structured template that prompted students to write down a claim (hypothesis), two pieces of evidence, and a reasoning of the evidence related to the claim. They compared a continuous and a fading condition (in which the prompts became gradually less specific) and found that all students increased their reasoning capacities considerably, but that students in the fading condition were stronger in reasoning at a post-test. This same triple scheme was used by Delen and Krajcik (2015), who found that students’ reasoning when using self-generated data was superior to students using “second hand” data. A similar tool was used by Kyza, Constantinou, and Spanoudis (2011) in the
inquire software STOCHASMOS; students could attach interpretations to data and write which hypotheses were confirmed or disconfirmed by these data. Pallant and Lee (2015) used a four-step scheme that included the claim, an explanation (rationale), an uncertainty indicator (like in Go-Lab above), and a rationale (explanation) of the uncertainty. They found that students’ uncertainty rating was based on a mixture of a judgment of their personal knowledge and the perceived quality of the evidence.

Reflection

A number of studies have systematically examined the effect of reflection guidance, as in the work by Lin and Lehman (1999). Davis (2000) examined the effects of activity and self-monitoring prompts on project progress and knowledge integration, in the context of the KIE (Knowledge Integration Environment) inquiry environment. Activity prompts encouraged students to reflect on the content of their activities. An activity prompt may, for example, “ask students to justify their decision or write a scientific explanation of a decision” (Davis, 2000, p. 822). Self-monitoring prompts activated students to express their own planning and monitoring by giving them sentence openers to complete. A sample prompt would be “Pieces of evidence or claims in the article we didn’t understand very well included . . .” (Davis, 2000, p. 824). Three studies were conducted in the domain of heat and temperature. Two studies compared conditions with different types of reflection prompts, while the third study looked deeper into students’ reactions to prompts. Overall, self-monitoring prompts helped more with creating a well-connected conceptual understanding (knowledge integration) than activity prompts did, although Davis also concluded that similar prompts led to quite different reactions from different learners.

Zhang et al. (2004) performed a study in which they gave learners reflective support, i.e., support that “increases learners’ self-awareness of the learning processes and prompts their reflective abstraction and integration of their discoveries” (p. 270). The learning environment centered around a simulation on floating and sinking of objects. The treatment consisted of: a) showing the students their inquiry processes (goals of experiments, predictions, and conclusions); b) reflection notes that students had to fill in, asking them to reflect on the experiment; and c) a fill-in form after the experiment that asked students to think over the process they had gone through and the discoveries they had made.

Students who received this type of evaluation support outperformed students who did not receive this support on a number of performance measures. Wichmann and Leutner (2009) studied the effects of reflective prompts in the context of a simulation on photosynthesis. These prompts were related to stating hypotheses, interpreting results, and thinking of new research questions. The students who received these prompts outperformed students from two control groups who did not receive these prompts or who only received explanation prompts. These results emerged for both a knowledge test and a scientific reasoning test. Eckhardt, Urhahne, Conrad, and Harms (2013) found that asking students to reflect on their learning process through an online questionnaire increased their knowledge gain.

Regulation

One way to support students in regulating their inquiry process is by providing them with an overall structure (Njoo & de Jong, 1993). In Sci-Wise (Shimoda, White, & Frederiksen, 2002) and the follow-up Thinkertools/Inquiry Island (White et al., 2002),
the inquiry process was divided into question, hypothesize, investigate, analyze, model, and evaluate. Learners had differently structured tab sheets for each of these tasks and dedicated advisors they could call upon to receive domain independent advice on how to perform a specific inquiry task. Manlove, Lazonder, and de Jong (2009b) report a series of studies on a process coordinator, a tool that structures the inquiry process for students and that helps and prompts student to plan, monitor, and evaluate. Their work shows that these tools stimulate students to plan, but that students are not very much inclined to monitor their work. The studies also indicated a negative relation between use of regulative facilities and students’ model scores. The authors make clear that this could be caused by the general character of the tool and that a stronger embedding of the tool in the domain would be necessary.

In a similar vein, Chang et al. (2008) found that providing students with step-by-step guidance in a simulation environment on optics was less effective than giving them detailed experimentation hints and hypothesis support. Providing students with sets of assignments or exercises that give them ideas for questions to ask and variables to manipulate and that give structure in the inquiry process is a very powerful form of support. This conclusion is substantiated by many studies with both an experimental character (Swaak & de Jong, 2001) and more practice-oriented work (Adams et al., 2008).

As an alternative to restricting or prompting students on steps in the inquiry process, students can also be shown (as in a dashboard) what they have been doing so that they can decide on the learning route themselves. Hagemans, van der Meij, and de Jong (2013) gave students a concept map overview of the domain topics in a simulation and showed the students which parts had been (successfully) completed. This influenced the knowledge that students acquired in a positive way.

Simulation-Based Learning Compared to Other Instructional Approaches

The above studies in majority show that if we compare guided with unguided simulations or virtual labs, students in the guided versions outperform the students in the unguided ones. A next step then is to compare guided simulations or labs to learning from direct instruction or from learning from real laboratories. There is now an emerging set of studies that present (large-scale) comparative evaluations that make these comparisons. In these cases the simulation is often embedded in a larger instructional arrangement, and it includes a composite of different types of guidance, which means that no specific data on specific types of guidance are available for these large-scale evaluations.

Simulations vs. Traditional Teaching

One of the first examples of such a large-scale evaluations concerns Smithtown, a supportive simulation environment in the area of economics, which underwent a large-scale evaluation with a total of 530 students. Results showed that after 5 hours of working with Smithtown, students reached a degree of micro-economics understanding that would require approximately 11 hours of traditional teaching (Shute & Glaser, 1990).

White and Frederiksen (1998) describe the ThinkerTools Inquiry Curriculum, a simulation-based learning environment on the physics topic of force and motion. In the ThinkerTools software, students are guided through a number of inquiry stages that
include experimenting with the simulation, constructing physics laws, critiquing each other’s laws, and reflecting on the inquiry process. ThinkerTools was implemented in 12 classes with approximately 30 students each. Students worked daily with ThinkerTools over a period of a little more than 10 weeks. A comparison of the ThinkerTools students with students in a traditional curriculum showed that the ThinkerTools students performed significantly better on a conceptual test (68% vs. 50% correct).

With regard to large-scale comparisons in the domain of science, Hickey, Kindfield, Horwitz, and Christie (2003) assessed the effects of the introduction of a simulation-based inquiry environment on the biology topic of genetics (GenScope). In GenScope, students can manipulate genetic information at different levels: DNA, chromosomes, cells, organisms, pedigrees, and populations. A large-scale evaluation was conducted involving 31 classes (23 experimental, 8 comparison) taught by 13 teachers, and a few hundred students in total. Overall, the evaluation results showed better performance by the GenScope classes compared to the traditional classes on tests measuring genetic reasoning. A follow-up study with two experimental classes and one comparison class also showed significantly higher gains for the two experimental classes on a reasoning test, with a higher gain for students from the experimental group in which more investigation exercises were offered.

Linn, Lee, Tinker, Husic, and Chiu (2006) evaluated modules created in the TELS (Technology-Enhanced Learning in Science) center. These modules are inquiry-based and contain simulations (e.g., on the functioning of airbags). Over a sample of 4328 students and 6 different TELS modules, an overall effect size of 0.32 was observed in favor of the TELS subjects over students following a traditional course for items measuring how well students’ knowledge was integrated.

In a study on the effects of introducing WISE units (e.g., combining simulations, discussion boards, classroom experiments) with 27 teachers over a set of different science topics, Lee, Linn, Varma, and Liu (2010) found that when the WISE units were introduced, scores of students on knowledge integration tests raised significantly as compared to the traditional form of teaching. Deslauriers and Wieman (2011) compared 124 students who either followed an 11-week traditional curriculum in quantum mechanics based on lectures (from a highly rated lecturer) or a course based on active learning that included a larger set of simulation. It was found that the traditional group scored 19% lower on a conceptual knowledge test. Plass et al. (2012) performed a study including 718 students from 25 classrooms. Students worked in a simulation-based curriculum or a text-based curriculum on a set of chemistry topics. Plass et al. (2012) found that the simulation-based group outperformed the textual group on both conceptual knowledge and transfer.

The domain of mathematics has also been an area for large-scale comparisons. Eysink et al. (2009) compared four technology-based learning environments on the same topic of probability theory. These environments were based on hypermedia learning, observational learning, explanation-based learning, and simulation-based learning. The study involved a total of 624 participants who all received the same knowledge tests with items on situational, intuitive, procedural, and conceptual knowledge. Overall results show that the explanation-based learning environment led to the highest performance, followed by the simulation-based learning environment and then the observational and hypermedia learning environments. However, explanation-based learning was the least efficient of the four approaches, hypermedia learning and observational learning the most efficient, and simulation-based learning held an intermediate position.
Tatar et al. (2008) compared the performance of a few hundred students and 25 teachers using a SIMCALC-based curriculum vs. a standard curriculum. The SIMCALC curriculum was primarily based on mathematic simulations. Results showed large advantages for the SIMCALC students on mathematical knowledge; the teachers’ knowledge in the SIMCALC classes also improved significantly over that of the teachers in the standard curriculum groups. de Jong, Hendrikse, and van der Meij (2010) compared a simulation-based mathematics course to traditional instruction. Over 12 weeks of lessons, students followed either their traditional regular course or an experimental course in which simulation-based exercises were used in conjunction with the book. The data from a total of 418 students from 20 classes could be analyzed. Results indicated that students from both groups scored similarly well overall on a post-test, but students from the traditional lectures scored especially high on procedural knowledge, while the simulation-based students tended to get better scores on conceptual items. In the related area of statistics (correlations), Liu, Lin, and Kinshuk (2010) found that students using simulation-based training clearly outperformed students who followed a lecture-based approach in repairing misconceptions and on tests measuring understanding.

The studies above involved a large number of students, covered a longer teaching period, or both. There is also a larger series of smaller experiments that compared inquiry (simulation-based or similar) with more traditional teaching that are summarized in a set of meta-analyses and overview studies that conform that (guided) inquiry outperforms traditional instruction (Alfieri et al., 2011; McElhaney et al., 2014; Minner, Levy, & Century, 2010; Rutten, van Joolingen, & van der Veen, 2012; Scalise et al., 2011; Slavin, Lake, Hanley, & Thurston, 2014; Smetana & Bell, 2012).

**Virtual Laboratories Compared to Real Laboratories**

There is now a growing number of studies comparing learning from virtual laboratories to learning in real laboratories that, as an overall outcome, show that, when referring to conceptual knowledge, in virtual laboratories students acquire the same or more conceptual knowledge than in real laboratories (Brinson, 2015; de Jong et al., 2013; Ma & Nickerson, 2006). In a similar vein, Slavin et al. (2014) report that studies that involved science kits for children didn’t report a positive effect on learning outcomes, whereas technology did contribute to science achievements.

An example study comparing real and virtual labs can be found in Chang, Chen, Lin, and Sung (2008), who compared learning of the physics topic of optics with three simulation-based environments (that included support in the form of experimentation prompts and hypothesis support) with learning in a laboratory, and concluded that students in all three simulation environments scored better on a test of conceptual knowledge than the laboratory students. Huppert et al. (2002) compared a group of students who followed a combination of traditional lecture and laboratory-based instruction on microbiology with a group who learned with a computer simulation integrated in the laboratory. They found better scores on a conceptual test for the simulation group and some advantages for the simulation group on acquiring science process skills.

A number of other studies found no differences in outcomes between simulated environments and real laboratories, but in these cases simulation-based training was more efficient. Klahr, Triona, and Williams (2007) compared students’ performance in a simulated and a real environment in which students had to design a car. They
found no difference between the two conditions in resulting knowledge about factors contributing to the car’s performance. Triona and Klahr (2003) compared students’ mastery of the control of variables strategy in physics domains after learning with a simulation or with physical material and found no differences. In a somewhat different but comparable setting, Zacharia and Constantinou (2008) compared two groups of students working in either a physical or a virtual laboratory on heat and temperature; both groups learned equally well. Zacharia and Olympiou (2011) compared a condition where the instruction centered around lectures and textbooks with four experimental conditions in which over 200 students learned about the physics topic of heat and temperature using either a physical laboratory, a simulation, or combinations of a physical laboratory and simulation. All courses were inquiry based, followed the same instructional principles, and lasted 15 weeks. No differences among the four experimental conditions were found on a test measuring conceptual knowledge, but a clear difference in favor of these conditions over the more traditional approach was evident.

Conclusions on the Effectiveness of Simulation and Virtual Lab-Based Learning

In summary, and as a very general conclusion, large-scale evaluations of carefully designed simulation-based learning environments show advantages of simulations and virtual labs over traditional forms of expository learning and over laboratory classes (see also Plass & Schwartz, 2014). In the studies reported, these results mainly referred to conceptual (intuitive) knowledge (Honey & Hilton, 2011). There are, however, other relevant forms of domain knowledge that may be stimulated by learning with simulations, such as inquiry skills, nature of science, and attitudes toward science. These areas, although having already received attention in work cited above, need more research, because there are indications that learning with simulations may have impact on such aspects as systems thinking (Evagorou, Korfiatis, Nicolaou, & Constantinou, 2009).

A second nuance to the overall conclusion that simulations and virtual labs show favorable results is that this seems mainly valid for students with low prior knowledge as reported a few times above. Based on a qualitative analysis of students working with a computer simulation on electrochemistry, for example, Liu, Andre, and Greenbowe (2008) found that students with higher prior knowledge provided more verbal explanations during their work than students with lower prior knowledge. The authors speculate that a highly structured environment is more suited for lower prior knowledge students, whereas those with higher prior knowledge profit more from an open environment. There are indications that prior knowledge is especially important when the simulations are highly interactive (i.e., they have many variables to manipulate, Park, Lee, & Kim, 2009). Other individual differences need more research. For example, gender is scarcely considered in the reported studies, although there is work suggesting that simulations are more favorable for boys than for girls (de Jong et al., 2010; Holzinger et al., 2009). Also spatial ability seems to influence learning with simulations or the effects of guidance. An example of the influence of spatial ability is seen in work by Urhahne, Nick, and Schanze (2009), who found that spatial ability had a strong influence on the level of conceptual knowledge that students gained from two- and three-dimensional simulations in chemistry. Finally, students also need sufficient regulative skills to be able to work in simulation-based learning environments (Hsu & Thomas, 2002).
PRACTICAL IMPLICATIONS

Overall, results encourage the use of computer simulations and virtual labs in the classroom. Upscaling and implementing in actual teaching practices is, however, a challenge. First, schools need software that is stable enough and well enough tested to survive in the classroom climate. Compared to a few years ago, we now see a number of online repositories that do offer such software; examples are PhET (e.g., Wieman, Adams, & Perkins, 2008), the Concord Consortium Collection, CREATE labs (Schwartz, Milne, Homer, & Plass, 2013), and ChemCollective (Yaron, Karabinos, Lange, Greeno, & Leinhardt, 2010). These collections provide teachers with extensive and stable software to work with. One remaining issue with these collections is that they often only present the labs and not the guidance that is needed as was described in this chapter. The newly developed Go-Lab system (de Jong et al., 2014; Gillet, de Jong, Sotiriou, & Salzmann, 2013) now offers a large set of online labs together with an easy-to-use authoring system that allows teachers to combine a lab with guidance, including scaffolds that can be adapted to the teachers’ needs. In this way, teachers can create full-fledged learning environments. One aspect that still needs attention in this respect is that teachers, with their skills and opinions, play a central role in implementing simulations and labs in the classroom (Hmelo-Silver, Liu, Gray, & Jordan, 2015). Training teachers not only on technical skills but also on didactic techniques such as inquiry learning is therefore necessary. Within the Go-Lab project, an extensive tutoring system is under development (Cao, Govaerts, Dikke, Faltin, & Gillet, 2014).

CONCLUSIONS AND FUTURE DIRECTIONS

This chapter gave an overview of advantages and disadvantages of simulation-based learning. The overall conclusion is that simulation-based inquiry learning can be effective if the students are provided with the appropriate guidance. In those cases where adequate support is given, simulation and virtual lab-based learning may lead to better results than direct instruction or laboratory-based exercises.

The discussion on the relative effectiveness of direct instruction vs. simulation-based learning (or inquiry learning in general) will remain a lively one (Kirschner et al., 2006; Klahr & Nigam, 2004; Kuhn & Dean, 2005). This discussion, however, will have no final “winner.” There are surely domains and/or learners for which a more direct approach is favorable (Kirschner et al., 2006) and even within domains and learners, a variation in instructional methods covering both inquiry approaches and direct instruction might be necessary. The better question is for which goals and for which learners in which circumstances is which instructional approach the best approach (Schwartz & Bransford, 1998) and also which combination of approaches is the most fruitful one. Combining simulations and more expository instructional texts and videos is, for example, possible in the Go-Lab system (de Jonget al., 2014). Myneni, Narayanan, Rebello, Rouinfar, and Puntambekar (2013) describe the combination of a simulation with an online tutor, and several studies point to the virtues of combining virtual and real laboratories (Chini, Madsen, Gire, Rebello, & Puntambekar, 2012; Jaakkola, Nurmi, & Veermans, 2011; Zacharia & de Jong, 2014).

The relative effectiveness of instructional approaches may also depend on the time frame taken into account. Dean and Kuhn (2007), for example, found that when introducing a more extended time frame (10 weeks) and several moments of delayed testing, direct instruction loses its initial advantages and inquiry learning takes over results,
which contradicts the results presented by Klahr and Nigam (2004). Students normally lack extensive experience with inquiry learning and some inquiry skills need longer exposure to be developed (Kuhn, Black, Keselman, & Kaplan, 2000). There are indications from more prolonged work (e.g., Fund, 2007, which lasted over six months) that the effects of scaffolds increase over time. Related work on inquiry learning that is not simulation based (Sadagh & Zion, 2009) shows that effects may only appear after an extended experience with this form of learning. Studies that use a single-shot evaluation may therefore not do justice to inquiry environments; more work that looks at students over a longer period of time is necessary.

A promising road for research and development is collaborative learning with simulations (see also Chapter 18, in this volume, on cooperative learning). Studies have shown the potential advantages of doing inquiry with simulations in a collaborative way (Bell, Urhahne, Schanze, & Ploetzner, 2010; Gijlers & de Jong, 2009; Gijlers, Saab, van Joolingen, de Jong, & van Hout Wolters, 2009; Karlsson, Ivarsson, & Lindström, 2013; Kollöffel, de Jong, & Eysink, 2011; Manlove, Lazonder, & de Jong, 2009a) and more research is needed on identifying conditions that optimize confrontations of opinions between students, on how to design shared representations, and on tools and scaffolds specifically geared towards collaborative inquiry with simulations.

Another route is to add gamification elements to labs or simulations. This may concern the embedding of the lab exercise in a larger narrative context (Bonde et al., 2014), badges, or competition. Kollöffel and de Jong (in press), for example, informed students who worked with a simulation-based math environment on an intermediate score and presented these scores either as a norm-related score or as a score in comparison to other students. After having received their score information, students continued to work in the simulation. The students who received norm-related information scored lower on a post-test than students who received the comparison information. These features open new possibilities to enhance the motivating character of labs and simulations.

Simulations and virtual labs play a specific role in education because they allow student actions (fast interactions with complicated models) that often are hard or impossible to be realized in another way. This chapter showed that guided simulations and virtual labs may form the basis of effective forms of teaching. Now that more stable collections of simulations and virtual labs are available and teachers can embed this technology in specific guidance, there are ample opportunities for introducing these technologies into the actual classroom at a large scale.

ACKNOWLEDGMENT

The preparation of this chapter was partially funded by the European Union in the context of the Go-Lab project (Grant Agreement no. 317601) under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D (FP7). This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.

REFERENCES


Triona, I. M., & Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students’ ability to design experiments. *Cognition and Instruction, 21*, 149–173. doi:10.1207/S1532690XCIC2102_02


