Review of Educational Research

http://rer.aera.net

Scientific Discovery Learning with Computer Simulations of Conceptual Domains

Ton De Jong and Wouter R. Van Joolingen

REVIEW OF EDUCATIONAL RESEARCH 1998 68: 179

DOI: 10.3102/00346543068002179

Review of Educational Research
Summer 1998, Vol. 68, No. 2, pp. 179–201

Scientific Discovery Learning With Computer Simulations of Conceptual Domains

Ton de Jong and **Wouter R. van Joolingen**
University of Twente

Scientific discovery learning is a highly self-directed and constructivistic form of learning. A computer simulation is a type of computer-based environment that is well suited for discovery learning, the main task of the learner being to infer, through experimentation, characteristics of the model underlying the simulation. In this article we give a review of the observed effectiveness and efficiency of discovery learning in simulation environments together with problems that learners may encounter in discovery learning, and we discuss how simulations may be combined with instructional support in order to overcome these problems.

In the field of learning and instruction we now see an impressive influence of the so-called constructivistic approach. In this approach a strong emphasis is placed on the learner as an active agent in the process of knowledge acquisition. As in the objectivistic tradition, where developments were followed and encouraged by computer-based learning environments such as programmed instruction, tutorials, and drill-and-practice programs (Alessi & Trollip, 1985), computer learning environments help to advance developments. Examples are *hypertext environments* (see, e.g., Gall & Hannafin, 1994), *concept mapping environments* (see, e.g., Novak & Wandersee, 1990), *simulations* (De Jong, 1991; Reigeluth & Schwartz, 1989; Towne, 1995), and *modeling environments* (e.g., diSessa & Abelson, 1986; Riley, 1990; Smith, 1986).

In this article we concentrate on the use of *computer simulations* for learning because learning with simulations is closely related to a specific form of constructivistic learning, namely, *scientific discovery learning*. First, we give a short introduction to the two key terms in this article (computer simulation and scientific discovery learning) followed by a short overview of studies that com-

---

Part of the work presented here was carried out within the SAFE/SIMULATE, the SMISLE, and the SERVIVE projects. These projects were partially sponsored by the European Commission in its Telematics programs. We appreciate the contributions of our colleagues from these projects to the work presented here, especially Jules Pieters and Janine Swaak (University of Twente), Melanie Njoo (now at Origin/Instruction Technology), Anja van der Hulst (now at TNO Physics and Electronics Laboratory), and Robert de Hoog (University of Amsterdam). Jules Pieters, Jeroen van Merriënboer (University of Twente), Patricia Alexander (University of Maryland), and Simon King (EDS-Ingévision) provided us with helpful comments on a draft of this article. Please address correspondence concerning this article to the first author.

pared unsupported simulation-based discovery learning to some form of expository teaching. These studies show that the advantages of simulation-based learning are not always evident and suggest that one of the reasons for this is that learners have problems with discovery learning. This conclusion brings us to the main questions discussed in this article: What are the problems that learners have in discovery learning, and how can we design simulation environments that support learners in overcoming these problems?

A *computer simulation* is a program that contains a model of a system (natural or artificial; e.g., equipment) or a process. Computer simulations can broadly be divided into two types: simulations containing *conceptual* models, and those based on *operational* models. Conceptual models hold principles, concepts, and facts related to the (class of) system(s) being simulated. Operational models include sequences of cognitive and noncognitive operations (procedures) that can be applied to the (class of) simulated system(s). Examples of conceptual models can be found in economics (Shute & Glaser, 1990) and in physics (e.g., electrical circuits; White & Frederiksen, 1989, 1990). Operational models can, for example, be found in radar control tasks (Munro, Fehling, & Towne, 1985).

In discovery learning contexts we generally find conceptual rather than operational simulations. (The latter are mainly used for experiential learning.) As a class, conceptual models encompass a wide range of model types—for instance, qualitative versus quantitative models, continuous versus discrete models, and static versus dynamic models (see Van Joolingen & De Jong, 1991a). Models may also differ considerably in complexity. They range from very simple, straightforward models—for example, simple Mendelian genetics (Brant, Hooper, & Sugrue, 1991)—to very complex models—for example, the medical simulation HUMAN (Coleman & Randall, 1986), in which 200 variables and parameters can be changed. Also, specific characteristics, like the place of variables in the model or the distance between theoretical and operational variables, can help to define a conceptual model (Glaser, Schauble, Raghavan, & Zeitz, 1992). In scientific discovery learning the main task of the learner is to infer the characteristics of the model underlying the simulation. The learners' basic actions are changing values of input variables and observing the resulting changes in values of output variables (De Jong, 1991; Reigeluth & Schwartz, 1989). Originally, the means of giving input and receiving output of simulation environments were rather limited, but now increasingly sophisticated interfaces using direct manipulation for input and graphics and animations as outputs are emerging (e.g., Härtel, 1994; Kozma, Russell, Jones, Marx, & Davis, 1996; Teodoro, 1992); virtual reality environments are the latest development (see, e.g., Thurman & Mattoon, 1994).

Discovery learning has its roots in Gestalt psychology and the work of Bruner (1961). The study of discovery learning has, over the last few decades, moved away from concept discovery (as in Bruner's studies) toward what has been called *scientific discovery learning* (Klahr & Dunbar, 1988; Reimann, 1991). Theories of scientific discovery learning are usually based on theories of scientific discovery. Rivers and Vockell (1987), for example, described a cycle that involves planning (designing an experiment), executing (carrying out the experiment and collecting data), and evaluating (analyzing the data and developing a hypothesis). Friedler, Nachmias, and Linn (1990) said that scientific reasoning comprises the following steps: "(a) define a scientific problem; (b) state a hypothesis; (c) design an experiment; (d) observe, collect, analyze, and interpret data; (e) apply the results; and (f) make predictions on the basis of the results" (p. 173). De Jong and Njoo (1992) added the distinction between *transformative processes*—that is, processes that yield knowledge directly (for example, the ones mentioned by Friedler et al. and Rivers and Vockell)—and *regulative processes*—that is, processes that are necessary to manage the discovery process (for example, planning and monitoring). A second group of theories on scientific discovery learning finds its inspiration in the work of Simon (e.g., Kulkarni & Simon, 1988; Qin & Simon, 1990; Simon & Lea, 1974). A major contribution in this field is Klahr and Dunbar's theory of scientific discovery as dual search (SDDS), which takes as central concepts two spaces: *hypothesis space* and *experiment space*. In SDDS theory, hypothesis space is a search space consisting of all rules possibly describing the phenomena that can be observed within a domain. Experiment space consists of experiments that can be performed with the domain and the outcomes of these experiments. Although the first emphasis in SDDS theory is on the structure of the search spaces, Klahr and Dunbar have paid considerable attention to discovery processes.

In an early overview of computer-based education, Bangert-Drowns, Kulik, and Kulik (1985) reported that simulation-based learning does not raise examination scores. Later studies that contrasted (sometimes as part of a larger set of comparisons) learning from "pure" simulation (containing conceptual models) with learning from some form of expository instruction (computer tutorial, classroom) covered a variety of domains, such as biology (Rivers & Vockell, 1987), economics (Grimes & Willey, 1990), Newtonian mechanics (Rieber, Boyce, & Assad, 1990; Rieber & Parmley, 1995), and electrical circuits (Carlsen & Andre, 1992; Chambers et al., 1994). Sometimes a single simulation is compared to expository instruction (Rieber & Parmley, 1995), but quite often a comparison is made between (a) a simulation embedded in a curriculum or expository instruction and (b) the curriculum or expository instruction as such (Carlsen & Andre, 1992; Chambers et al., 1994; Grimes & Willey, 1990; Rieber et al., 1990; Rivers & Vockell, 1987). Also, in some cases the expository instruction to which the simulation is compared is "enhanced"—for example, by "conceptual change features" (Chambers et al., 1994) or by questions (in one condition of Rieber et al., 1990). As an overall picture, favorable results for simulation-based learning were reported by Grimes and Willey (1990), and no difference between simulation-based learning and expository teaching was reported by Carlsen and Andre (1992) and Chambers et al. (1994). A mixture of favorable and no-difference results was found in several substudies by Rivers and Vockell (1987). In Rieber et al. (1990) the group of students receiving a simulation in addition to a tutorial scored higher on a test measuring "application of rules" than the tutorial-only group, but scored at the same level as a tutorial group that received additional questions while learning. In Rieber and Parmley (1995) subjects who received only an unstructured (pure) simulation fell short of the performance of subjects who received a tutorial.

The general conclusion that emerges from these studies is that there is no clear and univocal outcome in favor of simulations. An explanation of why simulation-based learning does not improve learning results can be found in the intrinsic problems that learners may have with discovery learning. Chambers et al. (1994), for example, analyzed videotapes of students working with a simulation and noticed that the students were unable to deal with unexpected results and that the students did not utilize all the experimenting possibilities that were available. Also, studies that have compared the learning behaviors of successful and unsuccessful learners in simulation learning environments (e.g., Schauble, Glaser, Raghavan, & Reiner, 1991) have pointed to specific shortcomings of learners. For this reason, authors of a number of studies have suggested additional instructional measures to help learners overcome the problems that they may have with scientific discovery learning.

In the discussion that follows we provide an overview of potential problems with simulation-based scientific discovery learning and search for guidance in dealing with these problems. In addition, we examine studies that have looked at the effect of combining simulations with various instructional support measures for learners.

The literature that serves as the framework for this discussion comes from several sources. First, we began with documents from two relevant research programs: the Learning Research and Development Center and Carnegie Mellon (e.g., Klahr & Dunbar, 1988; Reimann, 1991; Schauble, Glaser, et al., 1991; Shute & Glaser, 1990). Not only were these documents useful in organizing this review, but they were valuable resources in locating additional studies of scientific discovery learning with computer simulations. Next, we searched on-line retrieval systems (e.g., the ERIC database) using the main descriptor *computer simulation(s).* This yielded (in the June 1997 version of the ERIC database) 2,073 writings. Because limiting this initial search with the additional descriptor *discovery (learning or processes)* gave a set of papers that did not contain some relevant papers we knew of, we examined the ERIC descriptions of all 2,073 papers. We also solicited papers presented at national and international conferences (e.g., American Educational Research Association, European Association for Research on Learning and Instruction, World Conference on Artificial Intelligence in Education, and the International Conference on Intelligent Tutoring Systems) that addressed the topic of computer simulations, and we examined the contents of edited volumes published over the last five years. Furthermore, we engaged in a physical search of selected research journals likely to publish studies dealing with computer simulations. These journals included the *Journal of Research in Science Teaching, Computers and Education,* the *Journal of Computer-Based Instruction, Instructional Science,* and *The Journal of the Learning Sciences.*

On the topic of discovery learning with computer simulations we found four types of papers. First, we found papers that we call engineering studies, in which a learning environment is merely described. The second type is conceptual papers that deal with theoretical issues related to discovery learning and simulations. Third, we found papers in which empirical data were gathered (through, for example, log files or thinking-aloud procedures) on discovery learning processes. In the fourth type of paper, experimental studies are described in which simulation environments are evaluated against expository teaching, or in which different versions of basically the same simulation environment are compared. Our selection process was guided by the following criteria. First, we excluded experimental papers if they did not use carefully controlled experimental designs and/or did not have well defined performance measures. Second, we targeted original studies for this review and excluded subsequent writings that merely recast a previous study or repeated the same argumentation.

Problems That Learners Encounter in Discovery Learning

In the following subsections we identify a number of characteristic problems that learners may encounter in discovery learning, and classify them according to the main discovery learning processes: hypothesis generation, design of experiments, interpretation of data, and regulation of learning.

Hypothesis Generation

Coming up with new hypotheses is generally recognized as a difficult process (Chinn & Brewer, 1993) that clearly distinguishes successful and unsuccessful learners (Schauble, Glaser, et al., 1991). An important problem here is that learners (even university students) simply may not know what a hypothesis should look like. Njoo and De Jong (1993a) assessed the "validity" of the learning processes of 91 students of mechanical engineering who worked on a simulation on control theory. They observed the syntactical correctness of the learning processes that students wrote down on "fill-in forms." For example, for the process of generating a hypothesis they examined whether it consisted of variables and a relation between them, not whether the hypothesis was correct in the domain. Njoo and De Jong found an average rate of correctness of 42% for processes generally, and even lower scores for the specific process of generating hypotheses.

A second problem is that learners may not be able to state or adapt hypotheses on the basis of data gathered. Klahr and Dunbar (1988) found that in 56% of observed cases students failed to draw the right conclusions from disconfirming experiments; that is, hypotheses were retained incorrectly on the basis of a negative experimental result. Other studies also emphasize the resistance of learners to theoretical change. Chinn and Brewer (1993) present seven typical learners' reactions to anomalous data, of which only one is the adaptation of the theory on the basis of the data. Chinn and Brewer also cite a large number of studies in which it was found that learners ignore anomalous data (see also Chambers et al., 1994), reject anomalous data, hold anomalous data in abeyance, reinterpret anomalous data and retain an initial theory, or reinterpret anomalous data and make marginal changes to an initial theory (Chinn & Brewer, 1993, p. 4). Also, Dunbar (1993) found evidence in his studies that subjects have an overall difficulty with dropping an original goal, which leads them to persist in holding an initial hypothesis rather than stating a new one. As an explanation, Dunbar mentions what he calls the "unable-to-think-of-an-alternative-hypothesis" phenomenon, meaning that subjects stick to their current hypothesis (despite conflicting evidence) simply because they have no alternative.

These findings may lead to the general assumption that people have a strong tendency to keep their original ideas. However, Klahr and Dunbar (1988) also found a reverse effect: Learners rejected hypotheses in the absence of disconfirming experimental outcomes. This general problem of translating data into theory is illustrated in a study by Kuhn, Schauble, and Garcia-Mila (1992), who found that subjects (10-year-olds) changed their ideas on the causality of a domain variable many times (10 to 11 times) during an experimentation session. The frequent change of ideas can partly be explained by the fact that subjects in this study employed a large repertoire of what Kuhn et al. call "invalid inferences." For example, subjects made inferences about causality on a single instance or made inferences about a variable that had not been changed in two experiments. One aspect that may well influence the ability to adapt hypotheses on the basis of data is the distance between the theoretical variables and the variables that are manipulated in the simulation (Van Joolingen & De Jong, 1997). Glaser et al. (1992) contrasted (a) the environments Voltaville (on DC circuits) and Refract (on refraction of light), in which a relatively small distance exists between the theoretical variables and the variables that can be manipulated in the simulation, with (b) environments such as Smithtown (on economics), where a larger distance exists between the theoretical variables and the variables that can be manipulated in the simulation. Glaser et al. assert that in the former type of environment it is easier for subjects to see the relation between (a) their manipulations of such variables as lenses, distances, and resistances and (b) the characteristics of the theoretical models.

A third problem in stating hypotheses is that learners can be led by considerations that do not necessarily help them to find the correct (or best) theoretical principles. Van Joolingen & De Jong (1993) describe a phenomenon that they call *fear of rejection*. In an analysis of the use of a so-called hypothesis scratchpad by 31 students, they found that subjects tend to avoid hypotheses that have a high chance of being rejected—for example, a hypothesis in which the relation between variables has a high level of precision. A similar phenomenon was described by Klayman and Ha (1987) and by Klahr, Fay, and Dunbar (1993).

Design of Experiments

A crucial aspect of scientific discovery is the design of experiments that provide information for deciding upon the validity of a hypothesis. Alternatively, if a learner does not yet have a hypothesis, well designed experiments can be used to generate ideas about the model in the simulation. Klahr, Dunbar, and Fay (1991) identified a number of successful heuristics for experimentation in the BigTrak environment, which concerns the operation of a programmable robot. Among their heuristics are the following: Design simple experiments to enable easy monitoring, design experiments that give characteristic results, focus on one dimension of a hypothesis, exploit surprising results, and use the a priori strength of a hypothesis to choose an experimental strategy (Klahr et al., 1991, pp. 388–391). In the literature we find a number of phenomena in which learners use poorly designed experiments.

The first phenomenon, *confirmation bias*, is a learner's tendency to seek information that confirms a hypothesis, instead of trying to disconfirm the hypothesis. In a classic experiment Wason (1960) found confirmation bias for a rule discovery (2-4-6) task in which seeking confirming evidence is not the best strategy to use (Klayman & Ha, 1987). Dunbar (1993) showed, in a simulation environment, that some students have a strong inclination to search for evidence that supports their current hypothesis, and that this inclination may prevent them from stating an alternative hypothesis, even when they are confronted with inconsistent evidence. In an experiment with a simulation of the spread of an influenza epidemic, Quinn and Alessi (1994) found that only in a small number of cases (one out of six in a sample of 179 subjects) did students conduct experiments with the intention of eliminating hypotheses. In this study students were asked before running an experiment to choose the purpose of the experiment from a series of alternatives presented.

The second phenomenon describes learners who design *inconclusive experiments*. One of the best known examples is described in Wason's (1966) card turning experiment. This phenomenon, which is analogous to the phenomenon of confirmation bias, shows that subjects do not always behave as logical thinkers and do not perform the actions that would be most effective for testing a hypothesis. For instance, in the context of discovery learning with simulations, Glaser et al. (1992) point to a frequently observed phenomenon: Learners tend to vary too many variables in one experiment and, as a result, cannot draw any conclusions from their experiments. Reimann (1991) observed in the domain of optics that subjects perform poorly designed experiments that do not allow them to draw univocal conclusions. In two studies, Van Joolingen and De Jong (1991b, 1993) found that learners often manipulated variables that had nothing to do with the hypothesis they were testing. The percentage of effective experiments could be as low as 22%. Shute and Glaser (1990) and Schauble, Glaser, et al. (1991) report that unsuccessful learners do not gather sufficient data before drawing conclusions.

A third phenomenon is that subjects show *inefficient experimentation behavior*. For example, Kuhn et al. (1992) found that subjects did not use the whole range of potential informative experiments that were available, but only a limited set, and moreover designed the same experiment several times.

A fourth phenomenon describes learners that construct *experiments that are not intended to test a hypothesis*. Schauble, Klopfer, and Raghavan (1991) identified what they have called the engineering approach, in which learners attempt to create some desirable outcome instead of trying to understand the model. An engineering approach, as compared to the scientific approach, leads to a much narrower search and to a concentration on those variables where success is expected. As a consequence, this approach may prevent learners from designing experiments that provide well organized data that are sufficient for discovering all relevant domain relations. This engineering approach was also found by Schauble, Glaser, Duschl, Schulze, and John (1995) and Njoo and De Jong (1993a). A comparable phenomenon was found by White (1993), who reported that students created experiments that were fun (students worked with games in White's simulation environment) instead of experiments that provided insight into the model.

Interpretation of Data

Once the correct experiments have been performed, the data that come from these experiments need to be interpreted before the results of the experiments can be translated into hypotheses in the domain. According to Schauble, Glaser, et al. (1991), successful learners are more proficient in finding regularities in the data than unsuccessful learners. Klahr et al. (1993) found that subjects misencoded experimental data; the rate at which subjects made at least one misencoding ranged from a mean of 35% to a high of 63%, depending on the type of actual rule involved. And indeed, as Klahr et al. state, "compared to the binary feedback provided to subjects in the typical psychology experiment, real-world evidence evaluation is not so straightforward" (p. 114). They reported that the misinterpretation of data most likely resulted in a confirmation of the current hypothesis, thus suggesting that the hypothesis that a subject holds may direct the interpretation of data (see also Chinn & Brewer, 1993; Kuhn et al., 1992).

The interpretation of graphs, a frequently needed skill when interacting with simulations, is also clearly a difficult process. Linn, Layman, and Nachmias (1987) compared a group of students who worked in a "microcomputer-based laboratory" (MBL) with students from traditional classes. In the MBL students carried out experiments in the field of heat and temperature. The output of these experiments was given in the form of dynamically generated graphs. Linn et al. found that students' graphing abilities increased as a result of working with the MBL, but that on the more complicated graphing skills (for example, comparing different graphs) difficulties still existed after the MBL course. Mokros and Tinker (1987) placed students in computer labs where they could generate graphs on the basis of experiments and where they were encouraged to make graphical predictions. Mokros and Tinker found that children's initial problems in interpreting graphs disappeared quickly.

Regulation of Discovery Learning

For regulative processes it is frequently reported that successful learners use systematic planning and monitoring, whereas unsuccessful learners work in an unsystematic way (e.g., Lavoie & Good, 1988; Simmons & Lunetta, 1993). Shute and Glaser (1990) claim that successful learners plan their experiments and manipulations to a greater extent and pay more attention to issues of data management. Glaser et al. (1992) reported that successful discoverers followed a plan over experiments, whereas unsuccessful ones used a more random strategy, concentrating at local decisions, which also gave them the problem of monitoring what they had been doing (see also Schauble, Glaser, et al., 1991). Though Glaser et al. mentioned persistence in following a goal as a characteristic of good learners, these successful subjects were also ready to leave a route when it apparently would not lead to success. Goal setting is also reported as a problem (for subjects with low prior knowledge) by Charney, Reder, and Kusbit (1990). In a more general way Veenman and Elshout (1995) found that, over a number of studies, individuals with a high intellectual ability showed a better working method than individuals with a low intellectual ability, but also that working method made its own contribution to learning outcome on top of intellectual ability. For the process of monitoring, differences between successful and unsuccessful learners are reported by Lavoie and Good, who found that good learners make more notes during learning, and by Schauble, Glaser, et al., who found that successful learners recorded data more systematically.

Combining Simulations and Instructional Support

The previous section presented a number of characteristic problems encountered in scientific discovery learning. A number of researchers and designers have recognized these problems and, in line with the developments in concept discovery learning (see, e.g., Mayer, 1987), have provided learners with support in addition to simulations. In the current section we summarize a number of methods of supporting learners in the discovery process. The first means of support we describe is to provide the learner with direct access to domain information. Subsequently, we present methods that aim to support the learner in specific discovery processes.

Direct Access to Domain Knowledge

A frequently uttered claim about learning with simulations is that learners should know something beforehand if discovery learning is to be fruitful. Insufficient prior knowledge might be the reason that learners do not know which hypothesis to state, cannot make a good interpretation of data, and engage in unsystematic experimentation (Glaser et al., 1992; Schauble, Glaser, et al., 1991). Several authors have introduced access to extra information as a support measure in a simulation environment, quite often in the form of a (more or less sophisticated) hypertext/hypermedia system (Glaser, Ragahvan, & Schauble, 1988; Lajoie, 1993; Shute, 1993; Thomas & Neilson, 1995). Shute (1993) described an intelligent tutoring system (ITS) on basic principles of electricity in which learners could ask for a definition of a concept (e.g., ammeter, ampere, charge, circuit, current) by selecting a term from a menu and follow hypertext links. Shute reported positive effects of the use of this on-line hypertext dictionary on a composite posttest measuring declarative and conceptual knowledge, problem solving, and transfer of knowledge and skills.

A number of authors have pointed to the critical aspect of timing of the availability of information. Berry and Broadbent (1987) found that providing information at exactly the moment it is needed by the learner is much more effective than providing all necessary information before interaction with a simulation begins. Leutner (1993) used a simulation of a fairly complex agricultural system; the students' assignment was to optimize agricultural production. Some students were given information (consisting of domain concepts, facts, rules, and principles) *before* interacting with the simulation, whereas other students were given information (background information on system variables) *while* interacting with the simulation. Leutner found that permanently available information helped learners to acquire domain knowledge but that information provided before the simulation was not effective. For acquiring functional knowledge (the ability to optimize the outcome of the simulation) the same pattern was found, but here the results were less direct, because providing the information before or during the interaction with the simulation was combined with more or less elaborate experimentation advice. Also, Elshout and Veenman (1992) reported that subjects who received domain information before working in a simulation environment (on heat theory) did not profit from this information.

Information must not only be provided by the learning environment but also be invoked from learners' memory. Support measures can stimulate learners to confront their prior knowledge with experimental outcomes. In order to achieve this, Lewis, Stern, and Linn (1993) provided learners with an electronic notation form for noting "everyday life examples" of phenomena they observed in a simulation environment (on thermodynamics).

Support for Hypothesis Generation

Hypothesis generation is a central process in discovery learning. Several studies have created support to overcome the problems that learners have with this process. The Smithtown environment (Shute & Glaser, 1990) offers the learner support for hypothesis generation by means of a hypothesis menu. This menu consists of four windows which present parts of a hypothesis—for example, variables, verbs to indicate change, and connectors. A similar means of support is the hypothesis scratchpad (Van Joolingen & De Jong, 1991b, 1993). Here learners are offered different windows for selecting variables, relations, and conditions. These two approaches offer learners elements of hypotheses that they have to assemble themselves.

A more directive support for creating hypotheses can be found in CIRCSIM-TUTOR (Kim, Evens, Michael, & Rovick, 1989), an ITS in the domain of medicine—specifically, on treating problems associated with blood pressure. In this simulation students are asked to state qualitatively what will happen to seven components of the cardiovascular system. As a means of support, they are given a predefined spreadsheet in which to provide their ideas. One step further is to offer learners complete hypotheses. In Pathophysiology Tutor (PPT; Michael, Haque, Rovick, & Evens, 1989) learners can select from lists of predefined hypotheses, arranged in nested menus. Njoo and De Jong (1993a, 1993b) have used similar techniques. They conclude that offering predefined hypothesis to learners positively influences the learning process and the performance of learners. Quinn and Alessi (1994) forced students to write down, before experimenting, a single most plausible hypothesis or a list of multiple plausible hypotheses. The idea is that having more hypotheses available will lead to a strategy of elimination, which could be better than focusing on one hypothesis at a time. Their data showed that the multiple hypotheses strategy did indeed lead to more effective performance (reaching a required state of the simulation), but only if the complexity of the simulation was low. At higher levels of complexity, no advantage of the multiple hypotheses strategy over the single hypothesis strategy could be found. The higher effectiveness of the multiple hypotheses strategy could have been enhanced by the fact that one of the variables included had a counterintuitive result.

Support for the Design of Experiments

To support a learner in designing experiments the learning environment can provide *experimentation hints*. Rivers and Vockell (1987) gave some examples of such hints, like "It is wise to vary only one variable at a time." They provided learners with general experimentation hints of this sort before the learners worked with computer simulations. This did not affect the learning outcome, but it had an effect on the students' experimentation abilities. Hints can also be generated dynamically on the basis of the actual experimentation behavior of learners. Hints are then presented if a learner displays nonoptimal learning behavior. An example of a system containing this type of hint is Smithtown (Shute & Glaser, 1990). Leutner (1993) studied the effect of providing learners with adaptive advice of this kind. He found that if the advice had a limited character, it helped to increase the learner's domain knowledge, but hindered the acquisition of functional knowledge. If the advice contained greater detail, then it also helped to increase the learner's functional knowledge, though the effect was less clear since it was combined with giving extra domain information.

Support for Making Predictions

Whereas a hypothesis is a statement about the relations between variables in a theoretical model, a prediction is a statement about the value(s) of one or more dependent variables under the influence of one or more independent variables, as they can actually be observed in the simulation. One specific way to help learners express predictions is to give them a graphing tool in which they can draw a curve that depicts the prediction. Lewis et al. (1993) provided learners with such a tool. Feedback was given to learners by drawing the correct curve in the same diagram in which the learner's prediction was drawn. Tait (1994) described a similar mechanism, but in his case feedback also included explanations of the differences between the system's curve and the learner's curve. Reimann (1991), in an environment on the refraction of light, provided learners with the opportunity to give predictions at three levels of precision: as numerical data, as a drawn graph, and as an area in which the graph would be located.

Support for Regulative Learning Processes

Regulative processes are the processes that manage the learning process. Regulative aspects such as planfulness and systematicity are regarded as central characteristics of successful discovery learning (Glaser et al., 1992; Schauble et al., 1995). The two most central regulative processes are planning and monitoring (De Jong & Njoo, 1992), both of which can be supported by introducing model progression into the simulation environment. In addition to model progression, we found specific measures for supporting planning or monitoring. Finally, regulative processes can be supported by structuring the discovery process.

Model progression. The basic idea behind model progression is that presenting the learner with the full complexity of the simulation at once may be too overwhelming. In model progression the model is introduced gradually, step by step. White and Frederiksen's (1989, 1990) work on QUEST is one of the best known applications of *model progression*. In QUEST, electrical systems and models of electrical circuits differ in *order* (qualitative or quantitative models), *degree of elaboration* (number of variables and relations between variables), and *perspective*. While learning with QUEST, learners are confronted with models that advance from a qualitative to a quantitative nature, that are more elaborated, and that transform from a functional to a physical perspective. In this respect the instructional sequence follows the (assumed) transition from a novice knowledge state to an expert one. As far as we know, no controlled evaluation of QUEST has been undertaken. Model progression in which the model increases in complexity for the learner was studied by Swaak, Van Joolingen, and De Jong (1998). SETCOM is a simulation on harmonic oscillation where the model develops from free oscillation, through damped oscillation, to oscillation with an external force. Swaak et al. found that model progression was successful in enlarging the students' intuitive knowledge (but not their definitional knowledge) as compared to an environment without model progression. In a study in a different domain, but with the same type of environment, De Jong et al. (in press) found no effect of providing learners with model progression on top of giving them assignments.

Quinn and Alessi (1994) performed a study in which students had access to a simulation (on the spread of a disease within a population) with four input variables. One group started off with access to all four input variables; one group practiced with three variables before proceeding to the full simulation; and the last group started with access to two variables, proceeded to three, and ended with all four. In all cases students had to minimize the value of one of the output variables. The data revealed that model progression had no overall positive effect on performance. Moreover, model progression proved to be less efficient than providing the students with full complexity at the outset. It should be noted that the domain used by Quinn and Alessi was quite simple; the variables in the model did not interact. In another study on a more complex simulation of a multimeter, Alessi (1995) found that gradually increasing the level of complexity of the interface was beneficial for initial learning and for transfer. Also, Rieber and Parmley (1995) found, in the area of Newtonian motion, that subjects learning with a simulation that provided increasing control over variables scored significantly higher on a test measuring application of rules than did subjects who could exercise control in its full complexity from the start.

Planning support. Planning support may, as Charney et al. (1990) have postulated, be especially helpful for subjects who have low prior knowledge. Planning support takes away decisions from learners and in this way helps them in managing the learning process. Support for planning can be given in different ways. Even in the early days of the use of simulations for scientific discovery learning, Showalter (1970) recommended using *questions* as a way to guide the learner through the discovery process. His questions—for example, "Do rats ever reach a point at which they don't learn more?" (p. 49)—focused the learner's attention on specific aspects of the simulation. Zietsman and Hewson (1986) used similar types of questions in conjunction with a simulation on velocity, and Tabak, Smith, Sandoval, and Reiser (1996) used such questions with the aim of setting goals in a biological simulation.

White (1984) helped learners to set goals in a simulation of Newtonian mechanics by introducing *games.* Games, as White uses them, ask learners to reach a specific state in the simulation—for example, to get a spaceship in the simulation around a corner without crashing into any walls (p. 78). White found that learners who learned with a simulation that contained games outperformed learners who worked with the pure simulation on a test of qualitative problems (containing questions of the form "What would happen if . . . ?" or "How could one achieve . . . ?" [p. 81]). The ThinkerTools environment (White, 1993) employs games in a similar way.

De Jong et al. (1994) describe different types of *assignments* that can be used in combination with simulations. Types of assignments include (a) investigation assignments, which prompt students to find the relation between two or more variables, (b) specification assignments, which ask students to predict a value of a certain variable, and (c) "explicitation" assignments, which ask the student to explain a certain phenomenon in the simulation environment. De Jong et al. (in press), using a simulation on collisions, Swaak et al. (1998), using a simulation on harmonic oscillation, and De Jong, Härtel, Swaak, and Van Joolingen (1996), using a simulation on the physics topic of transmission lines, found that students (who were free to choose) used assignments very frequently and that using assignments had a positive effect on learners' gaining what these researchers call intuitive knowledge.

Monitoring support. Support for monitoring one's own discovery process can be given by overviews of what has been done in the simulation environment. Reimann (1991) provided learners in Refract with a notebook facility for storing numerical and nominal data from experiments. Data in the notebook could be manipulated, so that (a) experiments could be sorted on values for a specific variable, (b) experiments could be selected in which a specific variable had a specified value, and (c) an equation could be calculated over experiments. Also, the student could replay experiments from the notebook. Similar notebook facilities are present in Smithtown (Shute & Glaser, 1990) and Voltaville (Glaser et al., 1988). In SHERLOCK learners can receive upon request an overview of all the actions they have taken so far (Lesgold, Lajoie, Bunzo, & Eggan, 1992). Schauble, Raghavan, and Glaser (1993) presented monitoring support that not only provided an overview of the student's actions, but also offered the opportunity to group actions under goals and to ask for an "expert view" that gave the relevance of the student's actions in the context of a specific goal (e.g., to find the relation between two variables). (These methods in fact supported both monitoring and planning.) In all the examples presented here, learners have to select previous experiments for comparison from the complete set of experiments themselves. Reimann and Beller (1993) propose a system (CABAT) that selects a previous experiment on the basis of similarity and proposes this experiment to the learner for comparison.

Structuring the discovery process. Regulative processes can also be supported by leading the learner through different stages of the process. Several studies have compared the effects of structured environments (where structuring is quite often combined with several other measures) with so-called unstructured environments. Linn and Songer (1991) found that providing students with a sequence of experimentation steps ("Before doing the experiment . . . ," "Now do the experiment," "After doing the experiment . . .") and with more detailed directions in each of these steps was effective. They report that up to two and four times as many students were able to distinguish between central concepts from the domain (heat and temperature), as compared to students who used an unstructured version of the environment. Njoo and De Jong (1993b) had learners (students of mechanical engineering) work with a simulation (on control theory) in conjunction with forms that had separate cells for writing down the following: variables and parameters, hypotheses, experiment, prediction, data interpretation, and conclusion. On a test that measured "qualitative insight" the structured group outperformed a group who worked with the simulation environment alone.

Gruber, Graf, Mandl, Renkl, and Stark (1995) gave half of their subjects (60 students of a vocational economics school) instruction in making predictions, comparing predictions to outcomes, and drawing inferences. The other half received no guidance. The simulation used was in the field of economics and involved a jeans factory for which profit should be maximized. On a knowledge test in which students had to make predictions in new situations, the guidance group outperformed the no-guidance group. White (1993), in her ThinkerTools environment, forced subjects to follow a four-phase sequence of activities— "asking questions, doing experiments, formulating laws, and investigating generalizations" (p. 53)—and provided detailed scaffolding in each phase. White found a clear advantage for a simulation-based curriculum compared to a traditional curriculum on a test that measured qualitative predictions in real-world situations. In a number of experiments, Veenman and Elshout compared the learning behavior and learning result of learners working with structured and unstructured simulation environments. In the unstructured simulation subjects did not receive any instructional guidance. In the structured (or "meta-cognitive mediation") condition, subjects received "task assignments" and were prompted to "paraphrase the question, to generate a hypothesis, to think out a detailed action plan, and to make notes of it." Also, after they had performed a series of actions, they were "requested to evaluate their experimental outcomes," to "draw a conclusion elaborating on the subject matter, and to make notes" (e.g., Veenman, Elshout, & Busato, 1994, p. 97). The domains involved were simple electrical circuits, heat theory, and statistics. In an overall analysis of the data of four of their studies, Veenman and Elshout (1995) found no overall effect of structuring the environment. At a more detailed level, they found evidence that low-intelligence subjects with poor working methods profited from structured environments, whereas this was not true for low-intelligence subjects with good working methods or for high-intelligence subjects regardless of their working methods. In this overall analysis, several performance measures (including tests of factual knowledge and problem solving tasks) were combined into a single performance score.

We found two studies in which a comparison was made between a structured simulation environment and traditional, expository instruction. Lewis et al. (1993) required learners using a structured environment to make predictions before doing an experiment and to write down "graph comparisons" and "conclusions" after the experiment. Additionally, these learners were encouraged to write down "everyday examples," "important points," "confusion about" notes, and "example of concept" notes (p. 48). This was done in electronic form using a "Post-it note" metaphor. Lewis et al. found that, in responding to items that required a fundamental understanding of the difference between heat and temperature, students who had used the structured environment outperformed students who had followed the traditional curriculum in the preceding year. In Smithtown (Shute & Glaser, 1990) learners are taken by the hand and led through a fixed sequence of actions that is a little less strict than, for example, the sequence from Lewis et al. (1993). In Smithtown, learners are asked only if they want to make a prediction before experimentation, and they are not forced to do this. Smithtown includes not only structuring, but also a wealth of other supportive measures. An evaluation of Smithtown, using a test that required recall of concepts, failed to show an advantage of Smithtown over a traditional lesson (though learning with Smithtown was far more efficient).

Conclusion and Discussion

In this article we have given an overview of studies in scientific discovery learning with computer simulations of conceptual domains. From studies that empirically examined the discovery learning process we can conclude that a number of specific skills are needed for a successful discovery. Generally, one can say that successful discovery learning is related to reasoning from hypotheses, to applying a systematic and planned discovery process (like systematic variation of variable values), and to the use of high-quality heuristics for experimentation. These skills may have a general character, but can also be more closely related to a given domain (Glaser et al., 1992). Several problems characteristically encountered in the discovery process were identified. For the process of hypothesis generation, a learner's potential weaknesses include choosing hypotheses that seem "safe" and unsuccessfully transforming data into a hypothesis, both when the data are confirming and when they are disconfirming. For designing experiments, we found reports of learners who design inconclusive experiments, who show inefficient experimentation behavior, who follow a confirmation bias, and who apply an engineering approach instead of a scientific one. Furthermore, learners quite often have trouble with the interpretation of data as such. A final reported problem is that students are not very capable of regulating the learning process; this is evident in unstructured behavior driven by local decisions rather than an overall plan and in insufficient monitoring of the learning process.

We also examined instructional measures that are used in conjunction with simulations. Quite a few of the studies in which instructional measures were introduced were still in the engineering phase and did not evaluate the effect of the instructional measure in a controlled manner. Other studies in which the effects of adding instructional measures were evaluated used combinations of instructional measures, so that the effect of a specific measure could not be traced. On the basis of the remaining studies, three individual instructional measures can be seen as holding the promise of positively influencing learning outcomes. First, providing direct access to domain information seems effective as long as the information is presented concurrently with the simulation, so that the information is available at the appropriate moment. Secondly, providing learners with assignments (or questions, exercises, or games) seems to have a clear effect on the learning outcome. Thirdly, learners who use an environment that includes model progression perform better than learners who use the same environment without model progression, though it seems that the model needs to be sufficiently complex if this effect is to be evident. For other individual measures (e.g., hypothesis support, experimentation hints, monitoring tools, prediction support) the evidence is not substantial enough to warrant general conclusions. Finally, a number of studies on structuring the environment show that this may lead to more effective learning than using an unstructured environment, though it should be noted that structuring the environment in all these studies not only involved dividing up the learning process into distinct steps, but also included other instructional measures.

A crucial aspect of scientific discovery learning is the instructional goal for which it is used. Following the earliest ideas on discovery learning, it is frequently claimed that scientific discovery learning leads to knowledge that is more intuitive and deeply rooted in a learner's knowledge base (Berry & Broadbent, 1984; Laurillard, 1992; Lindström, Marton, Ottosson, & Laurillard, 1993; Swaak & De Jong, 1996) and that has a more qualitative character (White, 1993). It is also claimed that the results of simulation-based learning are properly measured only by "tests of application and transfer" (Thomas & Hooper, 1991, p. 500). Support for this claim is found in studies by Berry and Broadbent (1984), who showed that while simulations can be effective in teaching the ability to acquire a certain state in the simulation, this does not necessarily mean that the associated conceptual knowledge is learned as well. This lack of a relation between "explicable" knowledge and "functional" knowledge has also been found for simulations on business (Anderson & Lawton, 1992), Newtonian motion (with children; Flick, 1990), kinematics (McDermott, 1990), collisions (De Jong et al., in press; Whitelock et al., 1993), agriculture (Leutner, 1993), a subdomain of economics (Mandl, Gruber, & Renkl, 1994), acceleration and velocity (Rieber, 1996; Rieber et al., 1996), and harmonic oscillations (Swaak et al., 1998).

In the studies that we cited in this overview we find support for the importance of "intuitive" or "deep" knowledge for discovery learning. In studies that compared simulation with expository teaching, Grimes and Willey (1990), for example, used a test containing items that asked for "recognition and understanding," "explicit application," or "implicit application." In this study the simulation group, which showed an overall advantage over the control group, was specifically successful in items measuring implicit application. In Carlsen and Andre (1992), simulation groups performed no better on a posttest than did a no-simulation group; however, when the items were analyzed (by looking at the alternatives chosen) on the mental models that students had acquired, students from the simulation groups showed more advanced models. Rieber et al. (1990) used a test to measure the ability to apply rules from the domain. The simulation group took significantly less time in answering the posttest questions than did a group who had received a tutorial enhanced with questions. According to Rieber et al., this points to more deeply processed knowledge.

In studies where different versions of simulation environments were compared, we again see an effect of the type of knowledge test used. De Jong et al. (in press) and Swaak et al. (1998) used a test of definitional knowledge and also a test measuring "intuitive" knowledge. In the latter test, subjects had to predict what would happen after a change was introduced in a situation, and they had to make this prediction as quickly as possible (see also Swaak & De Jong, 1996). Though learners improved in definitional knowledge when learning with the simulation environments (which also contained expository information), the gain in intuitive knowledge was larger. Also, differential effects of simulation environments came out only on the test of intuitive knowledge.

Finally, the type of knowledge test used also seems to play a role in the studies that compared structured simulation environments with unstructured ones or with the normal curriculum. In Linn and Songer (1991) and Lewis et al. (1993) a test was used that measured qualitative distinctions between central concepts. Njoo and De Jong (1993a, 1993b) used items that measured qualitative insight, and Gruber et al. (1995) and White (1993) used tests in which predictions had to be given (as in De Jong et al., in press, and Swaak et al., 1998). All these studies showed an advantage for the structured simulation environments. In Veenman and Elshout (1995), where learners were tested on a combination of qualitative and definitional knowledge, no overall effect of structuring the environment was found, with an exception for specific group of learners. Finally, in an evaluation of Smithtown (Shute & Glaser, 1990), no difference in effectiveness could be found between a structured simulation environment and a traditional lesson, but here a test measuring recall of concepts was applied. Advantages of simulations seem clear when the instructional goal is the mastery of discovery skills. In Rivers and Vockell (1987), not only was domain knowledge assessed, but also discovery abilities were measured by a number of general tests (e.g., the Watson-Glaser

Critical Thinking Appraisal) and by analyzing the trend in scores on a domain pretest. Rivers and Vockell concluded that students from the simulation curricula outperformed the control subjects, especially if the simulations contained guidance in the form of hints that pointed to good discovery behavior (see also Faryniarz & Lockwood, 1992; Woodward, Carnine, & Gersten, 1988).

The development of environments that invite learners to engage in self-directed (discovery) learning and that provide support tools for the learning process continues (see, e.g., Suthers, Weiner, Connelly, & Paolucci, 1995). In our view, therefore, a further and deeper analysis of problems that learners encounter in discovery learning and a further evaluation of specific ways to support learners should be the principal items on the current research agenda in this area. Studies should aim to find out when and how to provide learners with means to overcome their deficiencies in discovery learning—in other words, when and how to provide scaffolding for the discovery learning process.

For these evaluation studies there are three additional points of interest. The first one is that introducing additional support tools is not only meant to enable the learner to perform certain actions, but can also be used to prevent cognitive overload (Glaser et al., 1988, p. 63). However, some instructional measures may also raise cognitive load, by introducing more complexity into the environment. Gruber et al. (1995), for example, suggest that an increase in cognitive load occurs when multiple perspectives are introduced into a simulation environment. Further research on support measures should take into consideration the effects of additional support measures on cognitive load (see, e.g., De Jong et al., in press; Swaak et al., 1998). A second aspect of support tools is that in learning environments these tools can be used as unobtrusive measures, as was recognized by Glaser et al. (1988) in the design of Voltaville. For example, in SHERLOCK (Lesgold et al., 1992) the student goes through the diagnostic problem solving process by choosing from menus of actions. On the one hand, this helps the student in the planning process; on the other hand, this helps the researcher (the system) to assess the student's intentions. In the SHERLOCK environment, information from this measure, which otherwise serves as a planning tool for the learner, is utilized for generating adequate hints. Van Joolingen (1995) describes some principles of how information gathered through a hypothesis scratchpad can be used for assessing the learner's actual state of knowledge. The third point of interest is that the place of simulations in the curriculum should be investigated. Lavoie and Good (1988) suggest that a "Piagetian" approach be used, which implies that simulations are introduced in a first phase of learning, where exploration is allowed, and that concepts are formally introduced later, followed finally by concept application (see also Brant et al., 1991; White, 1993). This suggests a potential use of computer simulation that differs from the classical, hypothesis-driven approach.

Only after sufficient research along the lines sketched in this section has been conducted might an appropriate design theory for instructional simulations arise. Current attempts at such a theory, though interesting, are necessarily fragmentary and incomplete (see, e.g., Thurman, 1993). But when such a theory does indeed arise, discovery learning with simulations can take its place in learning and instruction as a new line of learning environments, based on technology, in which more emphasis is given to the learner's own initiative.

References

Alessi, S. M. (1995, April). *Dynamic vs. static fidelity in a procedural simulation*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Alessi, S. M., & Trollip, S. R. (1985). *Computer based instruction, methods and development*. Englewood Cliffs, NY: Prentice-Hall.

Anderson, P. H., & Lawton, L. (1992). The relationship between financial performance and other measures of learning on a simulation exercise. *Simulation & Gaming, 23*, 326–340.

Bangert-Drowns, R., Kulik, J., & Kulik, C. (1985). Effectiveness of computer-based education in secondary schools. *Journal of Computer-Based Instruction, 12*, 59–68.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology, 36A*, 209–231.

Berry, D. C., & Broadbent, D. E. (1987). Explanation and verbalization in a computer-assisted search task. *The Quarterly Journal of Experimental Psychology, 39A*, 585–609.

Brant, G., Hooper, E., & Sugrue, B. (1991). Which comes first, the simulation or the lecture? *Journal of Educational Computing Research, 7*, 469–481.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*, 21–32.

Carlsen, D. D., & Andre, T. (1992). Use of a microcomputer simulation and conceptual change text to overcome students' preconceptions about electric circuits. *Journal of Computer-Based Instruction, 19*, 105–109.

Chambers, S. K., Haselhuhn, C., Andre, T., Mayberry, C., Wellington, S., Krafka, A., Volmer, J., & Berger, J. (1994, April). *The acquisition of a scientific understanding of electricity: Hands-on versus computer simulation experience; conceptual change versus didactic text.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Charney, D., Reder, L., & Kusbit, G. W. (1990). Goal setting and procedure selection in acquiring computer skills: A comparison of tutorials, problem solving, and learner exploration. *Cognition and Instruction, 7*, 323–342.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1–51.

Coleman, T. G., & Randall, J. E. (1986). *HUMAN-PC: A comprehensive physiological model* [Computer software]. Jackson: University of Mississippi Medical Center.

diSessa, A., & Abelson, H. (1986). Boxer: A reconstructible computational medium. *Communications of the ACM, 29*, 859–868.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397–434.

Elshout, J. J., & Veenman, M. V. J. (1992). Relation between intellectual ability and working method as predictors of learning. *Journal of Educational Research, 85*, 134–143.

Faryniarz, J. V., & Lockwood, L. G. (1992). Effectiveness of microcomputer simulations in stimulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29*, 453–470.

Flick, L. B. (1990). Interaction of intuitive physics with computer simulated physics. *Journal of Research in Science Teaching, 27*, 219–231.

Friedler, Y., Nachmias, R., & Linn, M. C. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal of Research in Science Teaching, 27*, 173–191.

Gall, J. E., & Hannafin, M. J. (1994). A framework for the study of hypertext.

Instructional Science, *22*, 207–232.
Glaser, R., Raghavan, K., & Schauble, L. (1988). Voltaville, a discovery environment to explore the laws of DC circuits. In *ITS-88* [Proceedings of the 1988 conference on intelligent tutoring systems] (pp. 61–66). Montreal, Quebec, Canada: Université de Montréal, Département d'informatique et de recherche opérationnelle.
Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). Scientific reasoning across different domains. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 345–373). Berlin, Germany: Springer-Verlag.
Grimes, P. W., & Willey, T. E. (1990). The effectiveness of microcomputer simulations in the principles of economics course. *Computers & Education*, *14*, 81–86.
Gruber, H., Graf, M., Mandl, H., Renkl, A., & Stark, R. (1995, August). *Fostering applicable knowledge by multiple perspectives and guided problem solving.* Paper presented at the conference of the European Association for Research on Learning and Instruction, Nijmegen, The Netherlands.
Härtel, H. (1994). COLOS: Conceptual learning of science. In T. de Jong & L. Sarti (Eds.), *Design and production of multimedia and simulation based learning material* (pp. 189–219). Dordrecht, The Netherlands: Kluwer Academic Publishers.
de Jong, T. (1991). Learning and instruction with computer simulations. *Education & Computing*, *6*, 217–229.
de Jong, T., Härtel, H., Swaak, J., & van Joolingen, W. (1996). Support for simulation-based learning: The effects of assignments in learning about transmission lines. In A. Díaz de Ilarazza Sánchez & I. Fernández de Castro (Eds.), *Computer aided learning and instruction in science and engineering* (pp. 9–27). Berlin, Germany: Springer-Verlag.
de Jong, T., Martin, E., Zamarro, J.-M., Esquembre, F., Swaak, J., & van Joolingen, W. R. (in press). The integration of computer simulation and learning support: An example from the physics domain of collisions. *Journal of Research in Science Teaching.*
de Jong, T., & Njoo, M. (1992). Learning and instruction with computer simulations: Learning processes involved. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 411–429). Berlin, Germany: Springer-Verlag.
de Jong, T., van Joolingen, W., Scott, D., de Hoog, R., Lapied, L., & Valent, R. (1994). SMISLE: System for Multimedia Integrated Simulation Learning Environments. In T. de Jong & L. Sarti (Eds.), *Design and production of multimedia and simulation based learning material* (pp. 133–167). Dordrecht, The Netherlands: Kluwer Academic Publishers.
van Joolingen, W. R. (1995). QMaPS: Qualitative reasoning for intelligent simulation learning environments. *Journal of Artificial Intelligence in Education*, *6*, 67–89.
van Joolingen, W. R., & de Jong, T. (1991a). Characteristics of simulations for instructional settings. *Education & Computing*, *6*, 241–262.
van Joolingen, W. R., & de Jong, T. (1991b). Supporting hypothesis generation by learners exploring an interactive computer simulation. *Instructional Science*, *20*, 389–404.
van Joolingen, W. R., & de Jong, T. (1993). Exploring a domain through a computer simulation: Traversing variable and relation space with the help of a hypothesis scratchpad. In D. Towne, T. de Jong, & H. Spada (Eds.), *Simulation-based experiential learning* (pp. 191–206). Berlin, Germany: Springer-Verlag.
van Joolingen, W. R., & de Jong, T. (1997). An extended dual search space model of learning with computer simulations. *Instructional Science*, *25*, 307–346.
Kim, N., Evens, M., Michael, J. A., & Rovick, A. A. (1989). CIRCSIM-TUTOR: An

intelligent tutoring system for circulatory physiology. In H. Maurer (Ed.), *Computer assisted learning: Proceedings of the 2nd International Conference ICCAL* (pp. 254–267). Berlin, Germany: Springer-Verlag.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1–48.

Klahr, D., Dunbar, K., & Fay, A. L. (1991). Designing experiments to test 'bad' hypotheses. In J. Shrager & P. Langley (Eds.), *Computational models of discovery and theory formation* (pp. 355–401). San Mateo, CA: Morgan-Kaufman.

Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*, 111–146.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211–228.

Kozma, R. B., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology supported learning environments* (pp. 41–61). Hillsdale, NJ: Erlbaum.

Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction, 9*, 285–327.

Kulkarni, D., & Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science, 12*, 139–175.

Lajoie, S. P. (1993). Cognitive tools for enhancing learning. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 261–289). Hillsdale, NJ: Erlbaum.

Laurillard, D. (1992). Learning through collaborative computer simulations. *British Journal of Educational Technology, 23*, 164–171.

Lavoie, D. R., & Good, R. (1988). The nature and use of predictions skills in a biological computer simulation. *Journal of Research in Science Teaching, 25*, 335–360.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. H. Larkin & R. W. Chabay (Eds.), *Computer-assisted instruction and intelligent tutoring systems: Shared goals and complementary approaches* (pp. 201–239). Hillsdale, NJ: Erlbaum.

Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction, 3*, 113–132.

Lewis, E. L., Stern, J. L., & Linn, M. C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*, 45–58.

Lindström, B., Marton, F., Ottosson, T., & Laurillard, D. (1993). Computer simulations as a tool for developing intuitive and conceptual understanding in mechanics. *Computers in Human Behavior, 9*, 263–281.

Linn, M. C., Layman, J., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Journal of Contemporary Educational Psychology, 12*, 244–253.

Linn, M. C., & Songer, N. B. (1991). Teaching thermodynamics to middle school students: What are appropriate cognitive demands? *Journal of Research in Science Teaching, 28*, 885–918.

Mandl, H., Gruber, H., & Renkl, A. (1994). Problems of knowledge utilization in the development of expertise. In W. J. Nijhof & J. N. Streumer (Eds.), *Flexibility in training and vocational education* (pp. 291–305). Utrecht, The Netherlands: Lemma BV.

Mayer, R. E. (1987). *Educational psychology: A cognitive approach.* Boston: Little, Brown and Co.

McDermott, L. C. (1990). Research and computer based instruction: Opportunity for interaction. *American Journal of Physics, 58*, 407–415.

Michael, J. A., Haque, M. M., Rovick, A. A., & Evens, M. (1989). The pathophysiology tutor: A first step towards a smart tutor. In H. Maurer (Ed.), *Computer assisted learning. Proceedings of the 2nd International Conference ICCAL* (pp. 390–400). Berlin, Germany: Springer-Verlag.

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*, 369–383.

Munro, A., Fehling, M. R., & Towne, D. M. (1985). Instruction intrusiveness in dynamic simulation training. *Journal of Computer-Based Instruction*, *2*, 50–53.

Njoo, M., & de Jong, T. (1993a). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching*, *30*, 821–844.

Njoo, M., & de Jong, T. (1993b). Supporting exploratory learning by offering structured overviews of hypotheses. In D. Towne, T. de Jong, & H. Spada (Eds.), *Simulation-based experiential learning* (pp. 207–225). Berlin, Germany: Springer-Verlag.

Novak, J. D., & Wandersee, J. H. (Eds.). (1990). Perspectives on concept mapping [Special issue]. *Journal of Research in Science Teaching*, *27*(10).

Qin, Y., & Simon, H. A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science*, *14*, 281–312.

Quinn, J., & Alessi, S. (1994). The effects of simulation complexity and hypothesis generation strategy on learning. *Journal of Research on Computing in Education*, *27*, 75–91.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction*, *16*, 1–10.

Reimann, P. (1991). Detecting functional relations in a computerized discovery environment. *Learning and Instruction*, *1*, 45–65.

Reimann, P., & Beller, S. (1993). Computer-based support for analogical problem solving and learning. In D. M. Towne, T. de Jong, & H. Spada (Eds.), *Simulation-based experiential learning* (pp. 91–105). Berlin, Germany: Springer-Verlag.

Rieber, L. P. (1996). Animation as feedback in a computer-based simulation: Representations matter. *Educational Technology Research & Development*, *44*, 5–23.

Rieber, L. P., Boyce, M., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction*, *17*, 46–52.

Rieber, L. P., & Parmley, M. W. (1995). To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. *Journal of Educational Computing Research*, *14*, 359–374.

Rieber, L. P., Smith, M., Al-Ghafry, S., Strickland, B., Chu, G., & Spahi, F. (1996). The role of meaning in interpreting graphical and textual feedback during a computer-based simulation. *Computers & Education*, *27*, 45–58.

Riley, D. (1990). Learning about systems by making models. *Computers & Education*, *15*, 255–263.

Rivers, R. H., & Vockell, E. (1987). Computer simulations to stimulate scientific problem solving. *Journal of Research in Science Teaching*, *24*, 403–415.

Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *The Journal of the Learning Sciences*, *4*, 131–166.

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences*, *1*, 201–239.

Schauble, L., Klopfer, L., & Raghavan, K. (1991). Students' transitions from an engineering to a science model of experimentation. *Journal of Research in Science*

Teaching, *28*, 859–882.
Schauble, L., Raghavan, K., & Glaser, R. (1993). The discovery and reflection notation: A graphical trace for supporting self regulation in computer-based laboratories. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 319–341). Hillsdale, NJ: Erlbaum.
Showalter, V. M. (1970). Conducting science investigations using computer simulated experiments. *The Science Teacher*, *37*, 46–50.
Shute, V. J. (1993). A comparison of learning environments: All that glitters... In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 47–75). Hillsdale, NJ: Erlbaum.
Shute, V. J., & Glaser, R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments*, *1*, 51–77.
Simmons, P. E., & Lunetta, V. N. (1993). Problem-solving behaviors during a genetics computer simulation: Beyond the expert/novice dichotomy. *Journal of Research in Science Teaching*, *30*, 153–173.
Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L. W. Gregg (Ed.), *Knowledge and cognition* (pp. 105–128). Hillsdale, NJ: Erlbaum.
Smith, R. B. (1986). The Alternate Reality Kit: An animated environment for creating interactive simulations. In *Proceedings of IEEE Computer Society Workshop on Visual Programming* (pp. 99–106). Silver Spring, MD: IEEE Computer Society Press.
Suthers, D., Weiner, A., Connelly, J., & Paolucci, M. (1995). Belvedere: Engaging students in critical discussion of science and public policy issues. In J. Greer (Ed.), *Proceedings of the AI-Ed 95, the 7th World Conference on Artificial Intelligence in Education* (pp. 266–273). Charlottesville, VA: AACE.
Swaak, J., & de Jong, T. (1996). Measuring intuitive knowledge in science: The what-if test. *Studies in Educational Evaluation*, *22*, 341–362.
Swaak, J., van Joolingen, W. R., & de Jong, T. (1998). Supporting simulation-based learning: The effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction, 8,* 235–253.
Tabak, I., Smith, B. K., Sandoval, W. A., & Reiser, B. J. (1996). Combining general and domain-specific strategic support for biological inquiry. In C. Frasson, G. Gauthier, & A. Lesgold (Eds.), *Intelligent tutoring systems* (pp. 288–297). Berlin, Germany: Springer-Verlag.
Tait, K. (1994). DISCOURSE: The design and production of simulation-based learning environments. In T. de Jong & L. Sarti (Eds.), *Design and production of multimedia and simulation-based learning material* (pp. 111–133). Dordrecht, The Netherlands: Kluwer Academic Publishers.
Teodoro, V. D. (1992). Direct manipulation of physical concepts in a computerized exploratory laboratory. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 445-465). Berlin, Germany: Springer-Verlag.
Thomas, R., & Hooper, E. (1991). Simulations: An opportunity we are missing. *Journal of Research on Computing in Education*, *23*, 497–513.
Thomas, R., & Neilson, I. (1995). Harnessing simulations in the service of education: The Interact simulation environment. *Computers & Education*, *25*, 21–29.
Thurman, R. A. (1993). Instructional simulation from a cognitive psychology viewpoint. *Educational Technology Research & Development*, *41*, 75–89.
Thurman, R. A., & Mattoon, J. S. (1994). Virtual reality: Towards fundamental improvements in simulation-based training. *Educational Technology*, *34*, 56–64.
Towne, D. M. (1995). *Learning and instruction in simulation environments*. Englewood

Cliffs, NJ: Educational Technology Publications.
Veenman, M. V. J., & Elshout, J. J. (1995). Differential effects of instructional support on learning in simulation environments. *Instructional Science*, *22*, 363–383.
Veenman, M. V. J., Elshout, J. J., & Busato, V. V. (1994). Metacognitive mediation in learning with computer-based simulations. *Computers in Human Behavior*, *10*, 93–106.
Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.
Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, UK: Penguin.
White, B. Y. (1984). Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction*, *1*, 69–108.
White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, *10*, 1–100.
White, B. Y., & Frederiksen, J. R. (1989). Causal models as intelligent learning environments for science and engineering education. *Applied Artificial Intelligence*, *3*(2–3), 83–106.
White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*, 99–157.
Whitelock, D., Taylor, J., O'Shea, T., Scanlon, E., Sellman, R., Clark, P., & O'Malley, C. (1993). Challenging models of elastic collisions with a computer simulation. *Computers & Education*, *20*, 1–9.
Woodward, J., Carnine, D., & Gersten, R. (1988). Teaching problem solving through computer simulations. *American Educational Research Journal*, *25*, 72–86.
Zietsman, A. I., & Hewson, P. W. (1986). Effect of instruction using microcomputers simulations and conceptual change strategies on science learning. *Journal of Research in Science Teaching*, *23*, 27–39.

Authors

TON DE JONG is Professor, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands; jong@edte.utwente.nl. He specializes in discovery learning and problem solving.
WOUTER R. VAN JOOLINGEN is Assistant Professor, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands; joolingen@edte.utwente.nl. He specializes in discovery learning and artificial intelligence and education.

Received September 16, 1996
Revision received October 20, 1997
Accepted January 29, 1998