Simulation & Gaming
Volume 40 Number 1
February 2009 110-133

10.1177/1046878107308078
http://sg.sagepub.com
hosted at
http://online.sagepub.com

Putting Theory-Oriented Evaluation Into Practice

A Logic Model Approach for Evaluating SIMGAME

Jan Hense
Ludwig Maximilians University, Munich, Germany
Willy Christian Kriz
University of Applied Sciences Vorarlberg, Austria
Joseph Wolfe
Experiential Adventures LLC, Seattle, USA

Evaluations of gaming simulations and business games as teaching devices are typically end-state driven. This emphasis fails to detect how the simulation being evaluated does or does not bring about its desired consequences. This paper advances the use of a logic model approach, which possesses a holistic perspective that aims at including all elements associated with the situation created by a game. The use of the logic model approach is illustrated as applied to SIMGAME, a board game created for secondary school level business education in six European Union countries.

Keywords: *business games; evaluation methods; logic model; secondary education*

One can look back at a very long history of gaming simulations and their use for educational purposes. The history of efforts to study and prove their effectiveness and efficiency as a learning tool for business education purposes is also long (Wolfe & Crookall, 1998) and has been accompanied by changes in what has been accepted as valid evaluation criteria and methodologies (Feinstein & Cannon, 2002). Accordingly, a frequent critique in the gaming simulation literature concerns the conduct and results of its evaluation activities. It has been argued that past evaluations have been based on anecdotal evidence (Greenblat, 1989), lacked methodological rigor (Dorn, 1989; Remus, 1981), failed to use valid outcome criteria (Anderson & Lawton, 1992; Feinstein & Cannon, 2002; Wolfe, 1990), or used psychometrically unproven measurement techniques (Gosen & Washbush, 2004; Spector, 2000; Spector, Christensen, Sioutine, & McCormack, 2001).

Even when the picture is confined to research considered methodologically sound, traditional evaluation shortcomings become apparent. These failings have resulted in widely varying appraisals of the value of games and simulations for teaching business subjects. This can be exemplified by reviews conducted by those such as Faria (2001), Gosen and Washbush (2004), Keys and Wolfe (1990), Wolfe (1997), and Greenlaw and Wyman (1973). Although the overall picture suggests gaming simulation is an effective learning tool, the results of the relevant studies are far from ambiguous and are often contradictory.

Such heterogeneous results indicate that mediating factors may exist, and that these factors exert an important influence on what makes for an effective experiential learning environment. For practitioners, such as simulation designers and trainers who use them, it would be of value to learn more about these influences. However, because most studies concentrate on showing whether simulations are effective as compared with an alternative teaching method (Wolfe, 1997), they fall short of divulging the specific conditions and factors that must be met to make simulation an effective learning tool. Thus, the problem of knowing *whether* an educational method such as simulation gaming works is relevant only if we also know *how* it works. Despite the considerable efforts that have been spent on evaluating the simulation gaming method, too little attention has been paid to generating knowledge about the mechanisms that underlie effective learning in such situations.

This problem has a striking parallel in the history of evaluation theory in general (Rossi, Lipsey, & Freeman, 2004; Shadish, Cook, & Leviton, 1991). We can observe a growing awareness over the last three decades that mere outcome-focused studies are only of limited value for everyday practice. As Cronbach and his colleagues state: "Everybody seems to think that evaluation is not rendering the service it should" (Cronbach et al., 1980, p. 44). One of the most important developments in the field of general program evaluation, which took place in reaction to these concerns, was the emergence of theory-oriented evaluation approaches. The commonality associated with these approaches is their emphasis on the use of logic models or theories in studying and explaining how a program or a learning method produces both its intentional and unintentional outcomes.

This paper's following sections introduce the theory-oriented approach to evaluation. In this presentation, a discussion of its potential merits in regard to evaluating gaming simulations and how the method tries to ameliorate the problems of the more traditional evaluation approaches that have dominated the field's past work will be given. Subsequently, the implementation of a theory-oriented evaluation will be demonstrated and applied to the case of SIMGAME*,* a business education simulation game.

Theory-Oriented Evaluation Approaches

When addressing theory-oriented approaches to evaluation, it is important to make two clarifications. The first clarification is terminological, as a variety of labels for such approaches is in use. Among such labels are "theory-driven evaluation" (Chen, 1990; Chen & Rossi, 1983), "theory-based evaluation" (Weiss, 1997), "program theory evaluation" (Rogers, Petrosino, Huebner, & Hacsi, 2000), and "theory-oriented evaluation" (Stame, 2004). Within this article's context, the phrase "theory-oriented evaluation" will be employed, even though the labels stated above have been used interchangeably in literature.

A second clarification concerns the term "theory." Here it does not refer to a theory of evaluation. Instead, it refers to the theoretical assumptions that underlie the program evaluated and/or its evaluation design. Terms frequently used interchangeably with "program theory" are "theory of change," "program model," or "logic model." Because the term "theory" is easily misunderstood, and tends to be too strong in practical contexts, our preferred terms are "program model" or "logic model."

The central thesis in the theory-oriented approach is that evaluations should be based on a logic model that explains how the interaction of a program, its participants, and its environment is expected to elicit the program's desired outcomes. In the case of gaming simulations as learning tools, their logic models must be based on theoretical assumptions and empirical evidence derived from learning psychology, education, and simulation gaming theory. The logic model should also describe how the simulation's features, learner characteristics, and context conditions interact to generate its intended learning outcomes.

Logic models typically consist of several variables that are relevant to the context of the evaluated program, as well as stating their mutual relationships. Usually, at least three kinds of variables are contained in a logic model. These are antecedent variables ("input"), variables related to program activities ("process"), and variables related to program effects ("outcome"). Often, logic models are represented graphically with boxes and arrows depicting variables and their mutual causal relationships, as will be exemplified by the logic model of SIMGAME, which we will present in a later section of this paper (cf. Figure 3).

Chen (1990, 2005) differentiates two basic aspects of program theory. *Normative theories* (or *action models*) guide practitioners during the development and implementation of a program. Because they are normally taken for granted from previous experience, normative theories are often only implicitly stated and thus have to be reconstructed indirectly from interviewing the project's staff and examining the materials they used. An action model typically contains a sketch of important steps that need to be carried out to put the program into practice. Chen (2005) defined them as "a systematic plan for arranging staff, resources, settings, and support organizations in order to reach a target population and deliver intervention services" (p. 23). Accordingly, action models often contain explicitly outlined chains of events that summarize how the status quo is intended to be transformed in light of the program's goals. Evaluating a program's action model is accomplished by assessing if it is put into practice within the program structure. Thus the evaluation delivers information for immediately improving the program's implementation.

Causative theories (or *change models*), on the other hand, are purely descriptive in nature. Their function is to explain *why* a given program brings forth its effects under certain conditions. Although they are not needed in everyday practice, they are essential in identifying weaknesses of the concept of a program for future improvements. Given the complexity of most social programs and their context mechanisms, change models are often less explicit than action models in representing the events which lead to the program's desired changes. To explain why certain outcomes result from a program, many factors have to be considered at the same time. So change models often represent a general summary of the context, input, and process factors that are expected to exert an influence on the outcomes of a program. Evaluating causative theories requires analyzing the causal interdependencies between program elements.

Conducting Theory-Oriented Evaluations of Gaming Simulations

What are the basic steps if one wants to conduct a theory-oriented evaluation? Reynolds (1998) developed a comprehensive approach, which he calls confirmatory program evaluation. Its basic steps will be outlined below. Further helpful guides and examples for constructing and using logic models in evaluation can be found in Chen (1990), Donaldson and Gooler (2003), W. K. Kellogg Foundation (2004a, 2004b), and Yampolskaya, Nesman, Hernandez, and Koch (2004).

Reynolds (1998) proposed seven key steps for conducting a theory-based evaluation. The following lists these steps as applied to evaluating simulation games:

1. Use theoretical and empirical knowledge from gaming simulation literature, learning theory, and education to specify a logic model that shows how features of the gaming simulation, learner characteristics, and context conditions interact to generate the desired learning outcomes, as well as possible side effects. If you are not involved in the game's development, consider the original developers' notions on the games' mechanisms relevant for learning.
2. Measure the effects of participation in the gaming simulation in regard to the logic model's outcomes, for example primarily learning results.
3. Collect data on mediating factors and key background factors, for example learners' previous knowledge, trainer characteristics, simulation features, and gaming process.
4. Estimate main effects of participation for the total group and any relevant subgroups.
5. If main effects are detected, test the logic model's causal mechanisms to explain them. If effects are not detected, conduct a causal analysis to understand the lack of effects.
6. Interpret the pattern of findings to facilitate generalization of the evaluation results to other contexts.
7. Identify formative uses of findings for program improvement.

The most fundamental step, which needs some further specification, is the development of the logic model. This first step is heavily dependent on previous conditions of the gaming simulation that is to be evaluated. While in some cases there already may exist a well-designed theoretical basis that served as the basis of the simulation game's development, others are more experimental and intuitively conceptualized. In any case, determining a logic model not only requires reviewing empirical and theoretical knowledge on gaming simulations, but also most often involves interaction with the game's developers and/or users. Although useful techniques for supporting this process and for resolving possible problems are described in Chen (1990), Leeuw (2003), Weiss (2000), or Yampolskaya et al. (2004), the second half of this article will demonstrate how these were applied to an actual evaluation.

Potential Merits of Theory-Oriented Approaches

Why and for what purposes should logic models in the evaluation of gaming simulations be used? The following list points out the benefits that have been cited in the literature (Petrosino, 2000; Rogers et al., 2000). This list is presented according to the evaluation process's key stages.

Choice of focus. Logic models can fulfill a guiding function during the evaluation process's early stages. A well-designed logic model, which depicts how learning is expected to happen, can help to identify the relevant questions and variables when designing the evaluation study.

Vague and distant outcomes. Programs often aim at outcomes that are difficult to measure and that are only expected as long-term effects. Gaming simulations often seek to foster general skills such as creativity or systems thinking. Because the logic model shows antecedent conditions for such long-term outcomes, it can help find theoretically sound indicators for such learning outcomes. Thus, even if it is difficult to measure these types of outcomes, it is possible to draw well-substantiated conclusions regarding their existence by looking at their antecedents.

Stakeholder participation. Evaluators are normally dependent to a degree on the stakeholders' participation in developing a program's logic model. In the case of gaming simulations this includes game designers, trainers, participants, and possibly administrative staff. One positive side effect can be that practitioners, such as simulation designers or trainers, need to make their assumptions about the workings and effects of the gaming simulation explicit (Huebner, 2000). This can help improve the simulation even before the evaluation itself takes place because implicit expectations have to be made explicit, and naive and simplistic expectations are put into question.

Implementation monitoring. Because a gaming simulation's logic model contains all the factors that are considered relevant to make it work, the model itself allows the monitoring of these factors when they are implemented during its execution. For example, if the logic model postulates that the induction of cognitive conflicts is crucial for a player's learning outcomes, it directs the evaluators to measure if these conflicts are really present in the simulation.

Interpretation of results. If the evaluation proves the simulation game to be effective, the logic model helps to locate the causal explanations for those effects (Mark, Hoffman, & Reichardt, 1992). Although it is never possible to "prove" a logic model (Weiss, 1997), it can be considered valid if two conditions are met. At least some of the most important relationships postulated by the logic model exist, and alternative models that could explain how effects are produced differently have been ruled out by theoretical considerations (Davidson, 2000; Weiss, 1997).

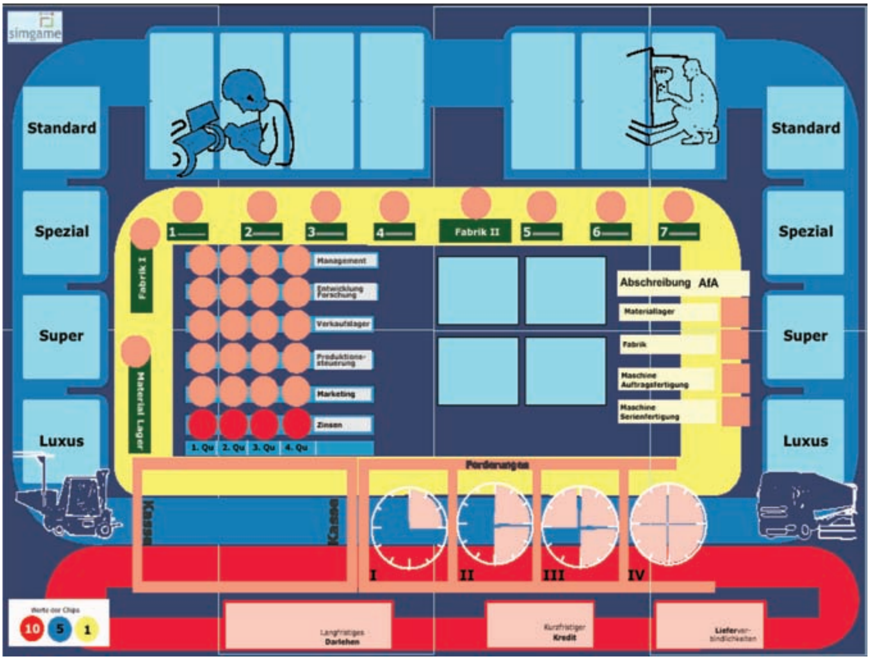
Understanding failure and success. More than just successful program effects can be analyzed via logic models. It is also possible to find explanations for cases when gaming simulations do *not* succeed. Sometimes a program has considerable merit, but fails to achieve its declared goals. Logic models can help discover these otherwise hidden effects and supply information for making future improvements in the simulation.

Some caveats, nonetheless, have to be mentioned. First, developing a logic model demands considerable content knowledge in the evaluated object's domain (Chen, 1990). Although more traditional method-centered approaches to evaluation often refrain from reflecting on causal relationships, developing a logic model involves intense theoretical considerations. Within the domain of gaming simulations, most evaluations are carried out by simulation experts, so there should be no lack of knowledge concerning simulation theory. To explain learning with simulations, additional expertise in the domains of educational and learning theory can be seen as an essential prerequisite for developing logic models.

Another caveat concerns unintended outcomes. Developing a logic model purely from the program's viewpoint can make the evaluation unable to detect negative as well as accidentally positive outcomes and side effects. This caveat is especially appropriate in regards to gaming simulations. Because of the open nature of their learning environments, their learning processes are often less predictable than in conventional settings. Scriven's (1973) so-called goal-free approach to evaluation indicates it is advisable for evaluators to keep an open eye for such unpredicted outcomes and not confine their evaluations to consider only official aims. As a consequence, we must simultaneously consider the logic that influenced the program's development and its logic from theoretical and empirical sources when constructing a logic model (Davidson, 2000).

The following section illustrates how a theory-oriented approach to evaluation is put into practice. Its central aspects will be applied by presenting the logic model based evaluation of the business game SIMGAME.

Figure 1
SIMGAME Playing Board (German-Language Version)



Evaluation of the Business Game SIMGAME

The business simulation game SIMGAME was a result of a project sponsored in 2003 and 2004 by the European Union's Leonardo da Vinci program. The project's overall goals were to (a) improve the quality of business education, particularly in the domain of entrepreneurial thinking, and (b) enhance students' employability and competitiveness in the European labor market. SIMGAME was developed by gaming simulation experts and subsequently tested in 30 schools in the five European countries of Austria, Czech Republic, Germany, Italy, and Slovakia.

SIMGAME is a board-based game in which participants are placed on teams to simulate a company's production and marketing decision cycles (Figure 1). Its target group is students in economic branches in upper secondary education and trainees in business professions. Two versions of the game were developed—A "static version" where team-members interact with each other within their own team and a "dynamic version" where teams compete against each other in a common market. This paper is concerned with those parts of the SIMGAME evaluation covering its static version. The implementation of the static version was carried out in schools in the 2003 to 2004 school year. Before implementing the game in classes, a train-the-trainer workshop was organized to prepare the project teachers adequately. Eighty percent of the teachers included in the following analyses attended the workshop.

Game play lasts from 4 to 12 hours, depending on the game version, previous experience of trainers and participants, and, of course, time available for playing the game. Teachers were encouraged to finish playing SIMGAME with a debriefing phase. During debriefing, learners review what happened during the game and reflect on their learning experience (Peters & Vissers, 2004).

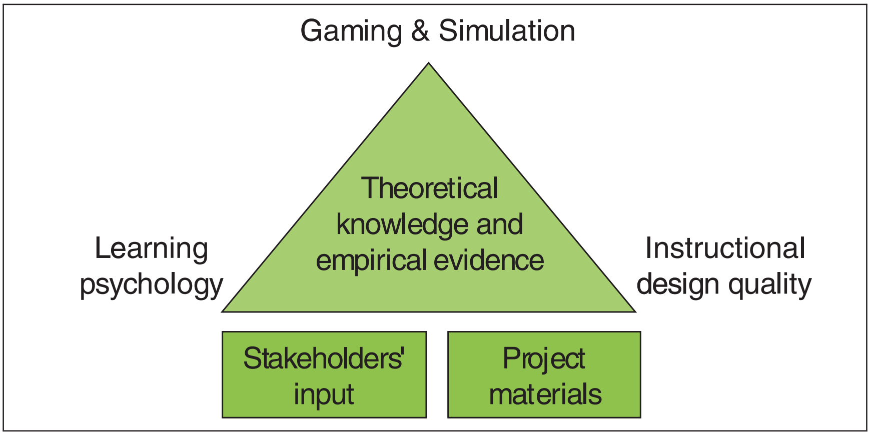
Official and formal goals for the SIMGAME project were created on three hierarchical levels. On the learners' level, SIMGAME was designed to create a problem-oriented and experiential learning environment that enabled students to gain knowledge and skills related to business and management. At the classroom level, SIMGAME was meant to contribute to a change from the traditional learning culture of passive reproduction of facts to a culture that fosters the active production of applicable knowledge. Therefore, on a third level, SIMGAME also was intended to promote the acquisition by teachers of the innovative didactic skills that are needed to realize this new learning and instruction paradigm.

Evaluation and quality assurance was implemented based on the program's three levels. The evaluation of SIMGAME's educational aspects, which also included the study reported here, was implemented at an advanced stage of program development, when a prototype of the game was already in existence. In the context of the overall evaluation activities, the study was aimed primarily at showing if SIMGAME succeeded in bringing forth its desired learning outcomes, as well as identifying game improvement aspects after the first rounds of testing. Accordingly, a theory-based approach was chosen not only to study the learning effects of SIMGAME, but also to explain those mechanisms that make participation in the game a successful learning experience. However, given the advanced state of program progress, it was too late to conjointly work out a program theory with program staff that would steer program development and evaluation at the same time. The following section will discuss how the theory for evaluating SIMGAME was developed for this situation.

Developing a Logic Model for SIMGAME

According to the procedure proposed by Reynolds (1998), the first step in a theory-based evaluation consists of developing a logic model. The evaluation study reported here was directed primarily at improving the game, not on supervising its implementation. Consequently, in the context of this study, the causative theory or *change model* of SIMGAME was of major interest.

Figure 2
Sources of SIMGAME Program Theory



Chen (1990) and other scholars propose a mixed procedure, which simultaneously takes into account domain-specific social science theories and stakeholders' views (Huebner, 2000). Because there was no mature program theory of SIMGAME already available, the social science side was used first to develop an initial version that could be used subsequently for discussion with stakeholders (cf. Figure 2). Three sources were used as starting points for deriving the logic model's theoretical aspects and empirical evidence:

1. From the *learning psychology* literature the relevant factors for individual learning processes were drawn. Its instructional theories based on cognitive and moderate constructivist views of learning were chosen because they emphasize authenticity and problem solving as the key factors in learning success (Brown, Collins, & Duguid, 1989; Hense, Mandl, & Gräsel, 2001). Additionally, Kolb's (1984) experiential learning cycle provided valuable insights because of its felicitous application to active learning situations (Kriz, 2003).
2. Additional information that described how learning takes place in gaming and simulations environments was obtained from the *gaming simulation* literature. The model proposed by Garris, Ahlers, and Driskell (2002) was helpful because it provided a basic input-process-outcome logic that mirrored the processes naturally associated with gaming simulations. Empirical findings on factors relevant for learning with business simulations were derived from overview articles, such as Faria (2001) and Wolfe (1997), as well as more recent research (Hindle, 2002). They

mainly provided insight into correlates of gaming performance such as the team's makeup and instructor involvement.

3. As a third source, the *instructional quality* literature was used to shape the broader classroom learning context created by SIMGAME (Brophy & Good, 1986). Parts of Ditton's (2002) model of school and instructional quality were adapted to extract the general factors associated with learning in school contexts. In a similar manner, parts of a framework model for learning in virtual seminars by Friedrich, Hron, & Hesse (2001) were adapted because the model emphasizes the role of interaction with technology in learning processes.

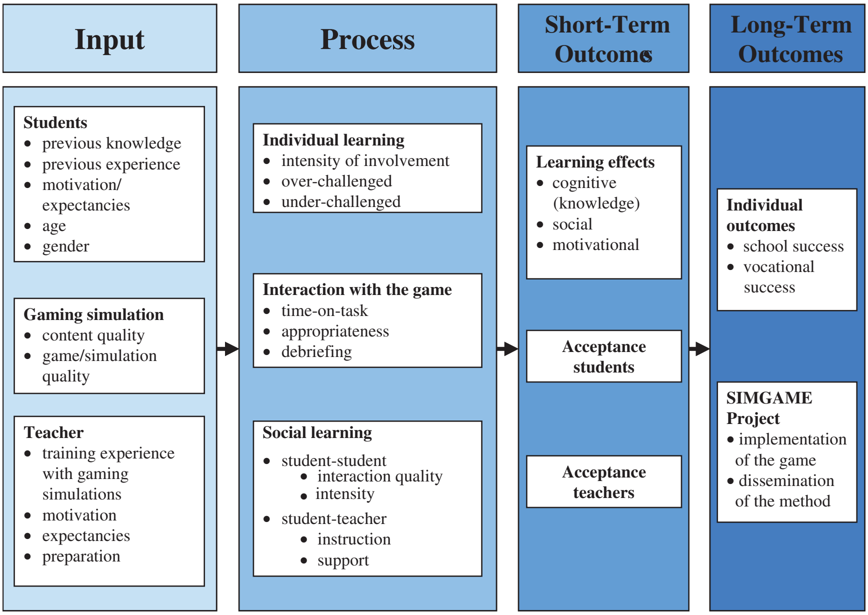
Additionally, to incorporate the stakeholders' perspective, a theoretical analysis of the program's materials was cross-validated by discussing them with its participants. The goals during this validation phase were twofold. The first goal was one of insuring that the theoretically derived model agreed with the stakeholders' experiences and expectancies about the program. The second goal was adjusting the model's goal element to guarantee a fair representation of the outcomes the program participants desired. This matching process led to minor changes in subaspects of the model. This was a further specification of learning goals on the student level.

Theoretical Background of the SIMGAME Logic Model

The program theory for SIMGAME, which emerged from the process that has been outlined above, is presented in Figure 3. It is basically an input-process-outcome model. The *outcome element* comprises the project goals and is divided into short-term and long-term outcomes. It was expected the operational goals on the student level (learning and acceptance) and teacher level (acceptance) contribute to accomplishing the project's overall goals. The *input and process elements* of the model comprise factors that were expected to contribute significantly to the attainment of the short-term outcomes. Thus it is here where the above mentioned sources of theoretical knowledge have been used as starting points to identify the different components of the model.

The *input* domain of the model comprises essential preconditions for learning in the game. On the *students'* side, it contains primarily cognitive, motivational, and demographic aspects. The importance of *previous knowledge* and *previous experience* is exemplified in Atkinson's (1974) model of cumulative learning, and has been confirmed repeatedly in empirical studies as one of the most powerful predictors of learning success (Helmke & Weinert, 1997). Affective preconditions of learning such as *motivation* and *expectancies* have long been neglected in research, yet they have also proven to exert substantial influence on learning processes and outcomes (Pekrun & Schiefele, 1996). Students' *age* has been included in the model because from the viewpoint of developmental psychology it can be expected to correlate significantly with learning. Finally, *gender* has been included to take into account potential gender-specific predispositions concerning the preferred learning environment.

Figure 3
SIMGAME Program Theory (Logic Model)



According to Bloom (1976), learning is the result of the interaction of learner and instruction characteristics. Hence, in addition to student factors, aspects of the learning environment as well as teacher characteristics have been included on the input side of the model. Concerning the *learning environment,* research suggests that there are two important quality aspects of gaming simulations (Feinstein & Cannon, 2002; Gosen & Washbush, 2004). *Content quality* refers to the simulation model within the game, which has to be a valid representation of reality to enable students to acquire applicable knowledge. *Game/simulation quality* on the other hand refers to the features that turn the simulation into a learning environment. It includes features such as the game materials, student manuals for playing the game, and trainer manuals for facilitating the game.

Besides characteristics of the learner and the learning environment, *teacher* behavior is a third important input factor in explaining successful learning outcomes (Brophy & Good, 1986; Wright, Horn, & Sanders, 1997). The model incorporated their *training experience with gaming simulations,* because familiarity with a new instructional method can be expected to be crucial for organizing and managing an effective learning experience. *Motivation* and *expectancies* refer to important affective and attitudinal aspects of teacher behavior, while time for *preparation* of classes with SIMGAME is a behavioral component that indicates whether there was enough time spent in preparing learning with a complex and potentially new method (Berliner, 1984; Raths & McAninch, 1999).

In the model's *process* component, variables on three levels were incorporated: individual learning, social learning, and interaction processes with the learning environment. The specific variables included on each of these levels were mainly derived from the research on effective learning and teaching, and the simulation and gaming literature. *Intensity of involvement* stands for the individual learner's degree of participation in the game. It was included because in cooperative learning settings, individuals can engage in social loafing and free-rider behavior (Salomon & Globerson, 1989). Another condition of individual learning is the need for an optimal fit between task difficulty and learners' abilities (Brophy & Good, 1986). This aspect was covered by the *over-/underchallenged* variables, which in contrast to all other factors of the model, were expected to exert a negative influence on student outcomes. Within the "interaction with the game" component, two aspects of Slavin's (1994) well-known QUAIT model were included (*time-on-task* and *appropriateness*), together with *debriefing*. Debriefing refers to the phase after playing the game, when learners recapitulate what happened during the game, reflect on their learning experience, and thus are expected to gain a deeper understanding of the learning contents (Hill & Lance, 2002; Peters & Vissers, 2004; Thiagarajan, 1992).

Finally, because SIMGAME is a highly interactive learning environment, important aspects of social and cooperative learning were considered (Cohen, 1994; Slavin, 1990). *Student–student interaction* was divided into a quantitative element (*intensity*), which refers to the sheer amount of cooperation of an individual, and a qualitative element (*quality*), which represents the degree to which cooperation was productive and task directed. Similarly, the *teacher–student interaction* contained a quantitative and a qualitative component: *instruction* in the model meant the amount of additional instructional input teachers delivered during the game. *Support* mirrors whether this input was adequate for helping the students during the learning process.

Elements in the model's *output* domain were derived primarily from project materials and stakeholder input. The *short-term outcomes* comprised immediate cognitive, social, and motivational learning effects, as well as students' and teachers' acceptance of the game. Although the *short-term* outcomes referred to SIMGAME's operational goals, the *long-term outcomes* represented development goals that are not directly attainable during the project runtime.

Concerning the causal links contained in the model, it has to be noted they are relatively unspecific at the present stage of development. The model only implies a general causal influence from input to process and from process to outcomes, without further specifying which elements on one level exert substantial influence on which elements on the next. Also, there are no mutual interdependencies modeled within one level, such as between students' intensity of involvement and teacher support on the process level. Although a more detailed model would have been desirable, such as for the purposes of an action model in the sense of Chen (2005), further specifications of model assumptions were not attempted because of the following theoretical, methodological, and practical considerations.

At the present state of theory and research on learning with gaming and simulations, it is believed there is not clear evidence that would lead to particular priority of some causal links over others. The field of gaming and simulation is quite heterogeneous and comprises learning environments that differ considerably from each other. Considering the issue of external validity it is often questionable if strong causal chains found in one application can be transferred unchanged to another gaming simulation. When the perspective is broadened to the more general field of learning theories, it must be recognized that learning is commonly considered an extremely complex human activity that is influenced by numerous factors on different levels (Slavin, 1994). Characteristics of learners and instructors, aspects of the learning environment, and the interaction among these three elements have to be considered at the same time. Theories such as the Walberg's productivity model (Wang, Haertel, & Walberg, 1993) or aptitude treatment interaction (Cronbach & Snow, 1977) refer to this complexity by postulating a multiplicative relationship between characteristics of learners and the learning environment.

To make the problem even more challenging in this particular case, most gaming simulations are notably complex learning environments. They rely heavily on the mechanisms of system dynamics and social interaction. Although this has the advantage of creating a stimulating and authentic learning experience, it also makes the course of events in a game highly unpredictable. Finally, apart from these theoretical and methodological problems, according to the evaluation's goals, the main function of the model was to enable a systematic study of the factors relevant for learning with SIMGAME. It was neither meant for guiding the design of the game nor as an action model for controlling the project's implementation (Chen, 2005).

Taking into consideration these arguments, the consequences of neglecting possibly important links seemed more severe than the consequences of using a less detailed model. Methodologically, this implies that the model is more appropriate for exploratory analysis than for hypothesis testing. This corresponds to the main goal of this evaluation study—to identify plausible leverage points for improvement of SIMGAME and its classroom implementation.

Questions and Framework of the Evaluation

To approach the goals of the evaluation, two questions had to be asked (Reynolds, 1998):

1. Are there any main effects of participation in the game in terms of the short-term outcomes contained in the change model?
2. Which elements in the input and process parts of the model contribute substantially to predicting the outcomes of participation in SIMGAME?

Wholey (1979), one of the first to use the term "logic model" in connection with evaluation, differentiated the *logic* model from the *evaluable* model of a program. Although the logic model describes the entire program, the evaluable model is restricted to those aspects of the program that it is possible to assess under the conditions of the given evaluation context. The evaluation of SIMGAME was subject to several such practical and methodological constraints, which implied adapting the original change model (Figure 3) and resulted in an *evaluable change model* for SIMGAME. The following modifications were applied out of these external constraints.

The *long-term outcomes* were excluded completely from the evaluable model. The individual outcomes on this level (e.g. vocational success) become relevant only years after finishing the game, so that they were beyond the temporal scope of our study. Moreover, although SIMGAME was expected to contribute to these individual long-term goals, realistically a relatively short intervention like SIMGAME can only be expected to exert a very small influence on them, so that one would need a very large sample size to demonstrate such a small effect.

Within the *input-student* component, the factor *age* was eliminated from the model because nearly all students were either 17 or 18 years old, which resulted in a very low variance, thus rendering age an unsuitable factor for the analysis of possible predictors of learning outcomes. The same rationale applied for the factor *students' previous experience,* because practically no student had played an educational gaming simulation before SIMGAME*,* and for the *input-gaming simulation* component. The latter had to be excluded, because in analyzing one singular gaming application, the quality of contents and game components do not show any variance.[1] Analyzing these aspects requires comparing different games with different quality features.

Method and Data Sources

Evaluation data was collected in all participating schools in the Czech Republic, Slovakia, Austria, Italy, and Germany, which resulted in a sample of 25 classes with 37 teachers and 468 students. Immediately before and after playing the game a knowledge test (for students) and questionnaires (for students and teachers) were administered. Because of the organizational conditions in this large-scale project, the teachers were responsible for conducting the data collection on location, with the help of a previous face-to-face training and detailed written instructions.

The knowledge test covering SIMGAME's intended cognitive learning domains was developed by business education content experts. It consisted of 20 multiple-choice questions, with four answering alternatives each. Out of these four alternatives per question, a minimum of one and a maximum of four were correct, or in other words, each question had zero to three wrong answers (distractors) and at least one correct answer.

Three different scores were derived from the knowledge pre- and posttests: number of weighted correct answers (theoretical range 0 to 20), number of weighted incorrect answers (theoretical range 0 to 20), and a total score, which equaled the difference of correct minus incorrect answers (theoretical range –20 to 20). Weighing of answers was performed on a per-question basis to compensate for the different number of distractors in each question.

Although the test results were used to operationalize the elements of previous *knowledge* and *cognitive learning* in the model, all other components were operationalized by items and test scales derived from the student and teacher questionnaires. Data on the input variables were collected before and all other data after playing the game. Table 1 shows an overview of all items and scales used in the analysis.

Results and Discussion

Evaluation question 1 concerned the main short-term effects of participation in SIMGAME. Table 2 shows the results of the content knowledge test. A significant increase in correct answers with an effect size of nearly 1/2 standard deviation ($d' = 0.46$) was detected. Similarly, there was a significant increase in total scores of nearly the same effect size ($d' = 0.42$). Also, a small significant decrease of incorrect answers was observable. On the whole, these results indicate that SIMGAME increased students' correct content knowledge and that this increase was statistically significant.

In addition to these objective test results, the learning effects in the cognitive, social, and motivational domains were measured by subjective ratings of students and teachers on a 5-point Likert scale (Table 3). Although analysis of these ratings is equivocal because of an unspecific reference point, the following interpretations are evident:

1. Students' and teachers' ratings are favorable regarding the three learning domains.
2. The learning effect reported by both groups was about the same across the three domains.
3. The teachers were more optimistic about learning results than were the students.

Table 3 also presents students' and teachers' game acceptance ratings. Both ratings are more favorable for the game than the subjective ratings for learning effects. Students, again, were less approving than their teachers and were less uniform in their opinion ($SD = 1.23$).

Table 1
Overview of Constructs and Scale Characteristics

	Source[a]	k	α	Exemplary Item/Operationalization
Input—students				
Previous knowledge	S – t1	40	n.a.	(knowledge pretest scores)
Motivation/expectancies	S – t1	4	.76	“I expect to raise my interest for business studies.”
Gender	S – t1	1	n.a.	(1 = male, 2 = female)
Input—teachers				
Training experience	T – t1	2	.84	(Total number of simulations games played/facilitated)
Teacher motivation	T – t1	5	.80	“I am looking forward to playing SIMGAME in class.”
Expected problems	T – t1	4	.82	“I expect problems in organizing the class.”
Preparation	T – t1	1	n.a.	(Hours spent preparing SIMGAME)
Process—individual learning				
Intensity of involvement	S – t2	3	.74	“During game play I tackled the contents of the game.”
Overchallenged	S – t2	1	n.a.	“Playing SIMGAME expected too much from me.”
Underchallenged	S – t2	1	n.a.	“Playing SIMGAME expected too little from me.”
Process—interaction w/game				
Time-on-task	T – t2	1	n.a.	(Hours spent playing the game)
Game appropriateness	T – t2	4	.78	“The game is suited for its target group.”
Debriefing	T – t2	1	n.a.	(Hours spent with debriefing)
Process—social learning				
Interaction quality	S – t2	4	.79	“I worked effectively with other students.”
Teacher instruction	S – t2	4	.83	“My teacher facilitated SIMGAME well.”
Teacher support	S – t2	2	.72	“My teacher explained a lot during the game.”
Outcome—student learning				
Cognitive learning	S – t2	40	n.a.	(knowledge posttest scores)
Cognitive learning	S – t2	12	.91	“SIMGAME increased my knowledge of operating figures.”
Cognitive learning	T – t2	12	.87	“SIMGAME increased students’ knowledge of operating figures.”
Social learning	S – t2	3	.73	“SIMGAME increased my teamwork skills.”
Social learning	T – t2	3	.80	“SIMGAME increased students’ teamwork skills.”
Motivational learning	S – t2	3	.83	“SIMGAME raised my interest in economics.”
Motivational learning	T – t2	3	.80	“SIMGAME raised students’ interest in economics.”
Outcome—acceptance				
Acceptance students	S – t2	4	.78	“I would like to play SIMGAME again in class.”
Acceptance teachers	T – t2	5	.77	“I will recommend SIMGAME to other teachers.”

a. S = Student questionnaire; T = Teachers questionnaire; t1 = before playing; t2 = after playing *k* = number of items; α = Cronbach’s alpha (if applicable).

Table 2
Cognitive Learning Effects (N = 409)

	$\bar{x}$	SD	t	p	d'
Pretest correct answers	8.97	2.31	6.52	.00	0.46
Posttest correct answers	9.85	2.88			
Pretest incorrect answers	4.52	1.35	−1.92	.03	−0.13
Posttest incorrect answers	4.39	1.38			
Pretest total score	4.44	2.88	6.00	.00	0.42
Posttest total score	5.46	3.56			

Note: One-tailed *t*-tests for dependent samples; degree of freedom = 408; *d'* = Cohen's *d* for dependent samples.

Table 3
Subjective Learning Effects and Acceptance Rated by Students and Teachers

	N	$\bar{x}$	SD
Student ratings			
Cognitive learning	420	3.51	0.77
Social learning	406	3.54	0.94
Motivational learning	421	3.27	0.96
Acceptance	395	3.74	1.23
Teacher ratings			
Cognitive learning	35	3.89	0.60
Social learning	34	4.09	0.57
Motivational learning	33	4.02	0.80
Acceptance	24	4.44	0.53

Note: Theoretical range for $\bar{x}$: 1 = *no learning effect*; 5 = *large learning effect*.

As an overall conclusion for evaluation question 1, SIMGAME succeeded in eliciting substantial learning effects. Also, students' and especially teachers' acceptance ratings for the game indicated that the attitudinal goals of SIMGAME were successfully met.

Evaluation question 2 asked which of the model's input and process parts contributed the most to the prediction of outcomes associated with SIMGAME. Because *cognitive learning* was the only outcome variable measured by an objective test, and because it can be considered the single most important student-level outcome, the following analyses concentrate on explaining this factor. The relationship between cognitive learning and its potential predictors within SIMGAME was investigated via regression analysis.

To maximize the precision of the analysis, only cases without missing data in the variables used for regression were included (listwise deletion), which resulted in

Table 4
Regression Analysis to Predict Cognitive Learning (*N* = 207)

	B	SE B	β	p
(intercept)	.97	.91		.29
Input factors—students				
Previous knowledge	.59	.06	.53**	.00
Expectancies/motivation	–.11	.22	–.03	.61
Gender	1.10	.36	.15**	.00
Input factors—teachers				
Training experience	1.50	.20	.42**	.00
Expected problems	–.69	.26	–.16*	.01
Teacher motivation	1.16	.29	.30**	.00
Preparation	–.03	.04	–.05	.44
Process factors—individual learning				
Intensity of involvement	–.61	.24	–.14*	.01
Overchallenged	–.15	.18	–.04	.41
Underchallenged	.15	.17	.04	.40
Process factors—interaction with game				
Time on task	.04	.03	.08	.13
Game appropriateness	1.90	.37	.54**	.00
Debriefing	.41	.24	.09	.08
Process factors—social learning				
Interaction quality	.60	.23	.14*	.01
Teacher instruction	.22	.22	.05	.32
Teacher support	–.10	.20	–.03	.62

Note: $R_{corr}^2 = .56$; $F(16, 190) = 17.25$, $p < .01$.
*$p < .05$. **$p < .01$.

a reduced sample size of $N = 207$. All variables derived from rating-scales were *z*-standardized prior to the analysis, the results of which are shown in Table 4. Overall, the combined factors in the model's input and process components explained 56% of the variance in students' posttest scores: $R_{corr}^2 = .56$; $F(16, 190) = 17.25$, $p < .01$.

Looking at the *student* component of the input level of the model, students' *previous knowledge* is the most important factor for student learning (β = .53). In the context of the model's other factors, *gender* played a minor role in the analysis. Female students scored somewhat better than male students (β = .15). *Expectancies and motivation,* on the other hand, did not play a significant role in the analyses as one would have expected from the model's assumptions.

Turning to *teacher* input factors, the results suggest their *training experience* (β = .42) and *motivation* (β = .30) contributed significantly to predicting student learning. *Expected problems* reported by them exerted a small negative influence in the prediction (β = –.16). Although all of these findings are in accordance with the expectations contained in the model, teachers' time spent on *preparation* was not significantly correlated with student learning within the context of a multiple correlation analysis.

Within the model's *process* factors, on the individual learning level only *intensity of involvement* played a small but negative role in predicting student outcomes (β = –.14). Reports of being challenged too much or too little, on the other hand, were not significantly correlated with learning.

In the *interaction with the game* component, the factor *game appropriateness* shows a high influence on the prediction of students' learning (β = .54). This can be interpreted to mean that students learned more if teachers felt the game was well suited for the students' learning preconditions. *Time-on-task* and *debriefing* on the other hand had no significant contribution in this analysis.

Concerning the process factors on the level of social learning, only *interaction quality* among students (β = .14) was relevant for predicting student learning. The more students interacted with other students during game play the greater their own learning scores increased. In contrast, *teacher instruction* and *teacher support* did not show a significant contribution.

Summarizing the results for evaluation question 2, the analyses for predictors of students' cognitive learning showed, after controlling for previous knowledge, the most important factors were found at the teacher level. Their training experience and motivation were the most important input factors. In the process of the game, the propriety of the game for students' learning prerequisites was even more substantial in explaining learning results. Other factors of minor relevance were expected problems by the teachers and interaction quality, both of which were in accordance with the model's assumptions. The findings for gender and intensity of involvement in the game, however, are unexpected and deserve further discussion.

Gender was included in the model as it is generally an important demographic factor in learning research. Previous research on business simulations in educational settings, however, has not suggested a significant influence direction. Most gender research in games has been within the context of computer games (Brunner, Bennett, & Honey, 1998; Carr, 2005). The negative contribution of intensity of involvement was a second unexpected finding. This negative result might indicate that a student's delving too deeply in SIMGAME's details causes them to lose sight of the overall picture. An alternative explanation might be that an artifact is being observed because low-performing students naturally need a more detailed involvement with the game's materials. This, in turn, may cut them off from other game play learning-relevant processes.

For the improvement of the implementation of SIMGAME, the findings suggest that increased emphasis should be put on preparing teachers for SIMGAME. If they are not already experienced in facilitating gaming simulations, compensatory measures would appear to be necessary and beneficial. Also, the findings for teacher motivation are important. Care has to be taken to select teachers for simulation game facilitation who are open and willing to teach with this new learning method, or that additional preparation is necessary to raise teacher motivation for the game.

Two more findings highlight the importance of careful selection within a SIMGAME context. It is important that students meet the preconditions for successful participation in the game. The negative predictor for expected problems by teachers indicates that temporal, spatial, organizational, and curricular preconditions of game play also need to be met to maximize SIMGAME.

Conclusion

By using a framework change model for the analysis of learning SIMGAME, several factors were identified that could be shown to be most important in explaining relevant learning outcomes. Because of the novelty of this method for evaluating a gaming simulation, this study's results should be considered to be exploratory in nature and may not be generalized for other games. Based on this study's experiences with the theory-based approach, several questions need to be addressed.

Methodologically, the simultaneous analysis of many potential influences needs considerably large sample sizes. In the context of a large European innovation project, it was possible to collect the necessary data for the chosen method of analysis. In other evaluation contexts this might be much harder to accomplish. Another methodological problem is that most data gathered in simulation games is normally hierarchically nested. In theory, such data would best be subjected to multilevel analysis, which takes into account variation both on the individual and the classroom level (Hox, 1995). Unfortunately, multilevel analysis has even stronger demands on sample sizes, which renders it unfeasible in many evaluation contexts. Thus, if theory-based evaluations are often associated with methodologically demanding procedures (Hennessy & Greenberg, 1999), it remains to be seen how these can be implemented in further simulation gaming research.

Theoretically, it is an open question if all factors relevant for determining learning outcomes in simulation games can be comprised in a single model. In the more general field of classroom learning, a plethora of studies were needed to arrive empirically at a set of generalizable high-impact factors (Wang et al., 1993). Such a process remains to be undertaken in the field of learning with gaming simulations. This present study is only one small step in this direction.

From a distinct gaming simulation perspective, the theory-based approach promises a better understanding of *learning within games.* Important antecedent and process conditions can be analyzed, as this study aimed to illustrate. However, within the evaluation of one particular game, it is not viable to investigate the role that distinctive *features of the game,* such as its game and content quality, have in learning. Usually, these features cannot be subjected to systematic manipulation within one single evaluation study, which shows that experiments in natural and laboratory settings are not rendered useless by the theory-based evaluation approach.

Bearing these restrictions in mind, the theory-based approach chosen in this study had significant advantages. In contrast to a conventional black-box evaluation, several key factors for further improvement and implementation of SIMGAME could be identified. From an educational-psychology perspective, it is important to investigate the factors that are associated with effective simulation game play and participation. Further studies with other games may show if this kind of approach is able to contribute to a general theory of gaming simulation learning.

An overall question this article addressed was “What is the right way to conduct evaluations of gaming simulations?” By raising this question it is not implied there is a single “right” way. After all, designing an evaluation study is heavily dependent on many contextual factors, and above all on the intended evaluation study’s purposes (Chelimsky & Shadish, 1997; Stufflebeam, Madaus, & Kellaghan, 2000). Yet, for the evaluation of gaming simulations, as well as for the evaluation of other educational procedures or social programs, the theory-oriented approach possesses many potential benefits. The approach proposes a frame of action that directs evaluators’ attention to a more holistic understanding of the how and why of learning within gaming simulation learning environments.

Note

1. The quality aspects of SIMGAME have been evaluated by a panel of gaming simulation and content experts in a forthcoming study.

References

Anderson, P. H., & Lawton, L. (1992). A survey of methods used for evaluating student performance on business simulations. *Simulation & Gaming, 23,* 490-498.

Atkinson, J. W. (1974). Motivational determinants of intellectual performance and cumulative achievement. In J.W. Atkinson & J. Raynor (Eds.), *Achievement and performance* (pp. 389-410). Washington, DC: Winston.

Berliner, D. (1984). The half-full glass: A review of research on teaching. In P. Hosford (Ed.), *Using what we know about teaching.* Alexandria, VA: ASCD.

Bloom, B. S. (1976). *Human characteristics and school learning.* New York: McGraw-Hill.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18,* 32-42.

Brunner, C., Bennett, D., & Honey, M. (1998). Girl games and technological desire. In J. Cassell & H. Jenkins, (Eds.), *From Barbie to Mortal Combat: Gender and computer games* (pp. 72-89). Boston: MIT Press.

Carr, D. (2005). Contexts, gaming pleasures, and gendered preferences. *Simulation & Gaming, 36,* 464-482.

Chelimsky, E., & Shadish, W. R. (Eds.). (1997). *Evaluation for the 21st century. A handbook.* Thousand Oaks, CA: Sage.

Chen, H. T. (1990). *Theory-driven evaluations.* Newbury Park, CA: Sage.

Chen, H. T. (2005). *Practical program evaluation.* Thousand Oaks, CA: Sage.

Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7,* 283-302.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1-35.

Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., et al. (1980). *Toward reform of program evaluation.* San Francisco: Jossey-Bass.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. *New Directions for Evaluation, 87,* 5-13.

Ditton, H. (2002). Unterrichtsqualität. Konzeptionen, methodische Überlegungen und Perspektiven [Quality of classroom learning and teaching. Concepts, methodological considerations and perspectives]. *Unterrichtswissenschaft, 30,* 197-212.

Donaldson, S. I., & Gooler, L. E. (2003). Theory-driven evaluation in action: Lessons from a $20 million statewide Work and Health Initiative. *Evaluation and Program Planning, 26,* 355-366.

Dorn, D. S. (1989). Simulation games: One more tool on the pedagogical shelf. *Teaching Sociology, 17,* 1-18.

Faria, A. J. (2001). The changing nature of business simulation/gaming research: A brief history. *Simulation & Gaming, 32,* 97-110.

Feinstein, A. H., & Cannon, H. M. (2002). Constructs of simulation evaluation. *Simulation & Gaming, 33,* 425-440.

Friedrich, H. F., Hron, A., & Hesse, F. W. (2001). A framework for designing and evaluating virtual seminars. *European Journal of Education, 36,* 157-174.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33,* 441-467.

Gosen, J., & Washbush, J. (2004). A review of scholarship on assessing experiential learning effectiveness. *Simulation & Gaming, 35*(2), 270-293.

Greenblat, C. S. (1989). Extending the range of experience. In D. Crookall & D. Saunders (Eds.), *Communication and simulation: From two fields to one theme* (pp. 269-283). Clevendon, England: Multilingual Matters.

Greenlaw, P. S., & Wyman, F. P. (1973). The teaching effectiveness of games in collegiate business courses. *Simulation & Games, 4*(3), 259-294.

Helmke, A., & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen [Determinants of school performance]. In F. E. Weinert (Ed.), *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Pädagogische Psychologie, Bd. 3)* [*Psychology of teaching and school psychology (Encyclopedia of psychology, educational psychology, Vol. 3)*] (pp. 71-176). Göttingen, Germany: Hogrefe.

Hennessy, M., & Greenberg, J. (1999). Bringing it all together: Modeling intervention processes using structural equation modeling. *American Journal of Evaluation, 20,* 471-480.

Hense, J., Mandl, H., & Gräsel, C. (2001). Problemorientiertes Lernen. Warum der Unterricht mit Neuen Medien mehr sein muss als Unterrichten mit neuen Medien [Problem oriented learning. Why learning with new media has to be more than instructing with new media]. *Computer und Unterricht, 11*(4), 66-82.

Hill, J. L., & Lance, C. G. (2002). Debriefing stress. *Simulation & Gaming, 33*(4), 490-503.

Hindle, K. (2002). A grounded theory for teaching entrepreneurship using simulation games. *Simulation & Gaming, 33,* 236-241.

Hox, J. (1995). *Applied multilevel analysis.* Amsterdam: TT-Publikaties.

Huebner, T. A. (2000). Theory-based evaluation: Gaining a shared understanding between school staff and evaluators. *New Directions for Evaluation, 87,* 79-89.

Keys, J. B., & Wolfe, J. (1990). The role of management games and simulations in education and research. *Yearly Review of Management, 16*(2), 307-336.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Kriz, W. C. (2003). Creating effective learning environments and learning organizations through gaming simulation design. *Simulation & Gaming, 34,* 495-511.

Leeuw, F. L. (2003). Reconstructing program theories. Methods available and problems to be solved. *American Journal of Evaluation, 24,* 5-20.

Mark, M. M., Hoffman, D. A., & Reichardt, C. S. (1992). Testing theories in theory-driven evaluations: (Tests of) moderation in all things. In H. T. Chen & P. H. Rossi (Eds.), *Using theory to improve program and policy evaluations* (pp. 71-84). New York: Greenwood Press.

Pekrun, R., & Schiefele, U. (1996). Emotions und motivationspsychologische Bedingungen der Lernleistungen [Emotional and motivational determinants of learning performance]. In F. E. Weinert (Ed.), *Psychologie des Lernens und der Instruktion* (pp. 153-180). Göttingen, Germany: Hogrefe.

Peters, V. A. M., & Vissers, G. A. N. (2004). A simple classification model for debriefing simulation games. *Simulation & Gaming, 35*(1), 70-84.

Petrosino, A. (2000). Whether and why? The potential benefits of including program theory evaluation in meta-analysis. *New Directions for Evaluation, 87,* 59-69.

Raths, J. D., & McAninch A. C. (1999). What counts as knowledge in teacher education? In J. D. Raths & A. C. McAninch (Eds.), *Advances in Teacher Education (vol. 5).* Stanford, CT: Ablex.

Remus, W. E. (1981). Experimental designs for analyzing data on games: Or, even the best statistical methods do not replace good experimental control. *Simulation & Games, 12,* 3-14.

Reynolds, A. J. (1998). Confirmatory program evaluation: A method for strengthening causal inference. *American Journal of Evaluation, 19,* 203-221.

Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsi, T. A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation, 87,* 5-13.

Rossi, P., Lipsey, M. W., & Freeman, H. (2004). *Evaluation. A systematic approach.* Thousand Oaks, CA: Sage.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13,* 89-99.

Scriven, M. (1973). Goal-free evaluation. In E. R. House (Ed.), *School evaluation: The politics and process* (pp. 319-328). Berkeley, CA: McCutchan.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation. Theories of practice.* Newbury Park, CA: Sage.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice Hall.

Slavin, R. E. (1994). *Educational psychology. Theory and practice.* Boston: Allyn & Bacon.

Spector, J. M. (2000). System dynamics and interactive learning environments: Lessons learned and implications for the future. *Simulation & Gaming, 31,* 528-535.

Spector, J. M. Christensen, D. L., Sioutine, A. V., & McCormack, D. (2001). Models and simulations for learning in complex domains: Using causal loop diagrams for assessment and evaluation. *Computers in Human Behaviour, 14,* 517-545.

Stame, N. (2004). Theory-based evaluation and types of complexity. *Evaluation, 10,* 58-76.

Stufflebeam, D. L., Madaus G. F., & Kellaghan T. (Eds.). (2000). *Evaluation models—viewpoints on educational and human services evaluation.* Boston: Kluwer.

Thiagarajan, S. (1992). Using games for debriefing. *Simulation & Gaming, 23*(2), 161-173.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63,* 249–294.

Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation, 76,* 41-56.

Weiss, C. H. (2000). Which links in which theories shall we evaluate? *New Directions for Evaluation, 87,* 35-45.

Wholey, J., S. (1979). *Evaluation. Promise and performance.* Washington, DC: The Urban Institute.
W. K. Kellogg Foundation. (2004a). *W. K. Kellogg Foundation evaluation handbook.* Battle Creek, MI: W.K. Kellogg Foundation.
W. K. Kellogg Foundation. (2004b). *W. K. Kellogg Foundation logic model development guide.* Battle Creek, MI: W. K. Kellogg Foundation.
Wolfe, J. (1990). The evaluation of computer-based business gaming. In J. W. Gentry (Ed.), *Guide to business gaming and experiential learning* (pp. 279-300). London: Kogan Page.
Wolfe, J. (1997). The effectiveness of business games in strategic management course work. *Simulation & Gaming, 28,* 360-376.
Wolfe, J., & Crookall, D. (1998). Developing a scientific knowledge of simulation/gaming. *Simulation & Gaming, 29,* 7-19.
Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11,* 57-67.
Yampolskaya, S., Nesman, T. M., Hernandez, M., & Koch, D. (2004). Using concept mapping to develop a logic model and articulate a program theory: A case example. *American Journal of Evaluation, 25,* 191-207.

Jan Hense, PhD (University of Munich, Germany), is assistant professor of education and psychology at the Ludwig Maximilians University Munich (Germany). His research interests include evaluation theory and research, entrepreneurship education, and innovative learning methods with new media. He is an active member of the Swiss Austrian German Simulation and Gaming Association.

Willy Christian Kriz, PhD (University of Vienna, Austria), Professor of Human Resources Management and Gaming Simulation Methods at the University of Applied Sciences Vorarlberg (Austria), Guest Professor at Universities in Austria, Germany, Italy, Peru and India. Chairman of the Swiss Austrian German Simulation and Gaming Association, Past-President of the International Simulation and Gaming Association.

Joseph Wolfe received his PhD from New York University's Stern School of Business. He is a professor emeritus at the University of Tulsa, and is president of Experiential Adventures LLC as well as a fellow and a past president of the Association for Business Simulation and Experiential Learning (ABSEL).

ADDRESSES: JH: Ludwig Maximilians University, Department Psychology, Leopoldstr. 13, D-80802 Munich, Germany; tel.: +49 (89) 2180-3257; fax: +49 (89) 2180-99-3257; e-mail: jan.hense@psy.lmu.de. WCK: University of Applied Sciences Vorarlberg, Hochschulstrasse 1, 6850 Dornbirn, Austria; tel.: +43 5572 792 3218; fax: +43 5572 792 9509; e-mail: willy.kriz@fhv.at. JW: Experiential Adventures LLC, 500 West Roy Street W208, Seattle, WA 98119, USA; telephone: +1 206-829-8028; e-mail: jwolfe8125@aol.com.