Theory-oriented evaluation for the design of and research in gaming and simulation

Willy Christian Kriz
University of Applied Sciences Vorarlberg
Jan Ulrich Hense
Ludwig Maximilians University

Jan Klabbers introduced the terms science of analysis *and* science of design. *The science of analysis uses games and simulations as research methods to test theories in various fields. Research in the science of design perspective emphasizes the design of the artifact, and testing its usability. However, in which way both communities can be of mutual benefit remains controversial. The authors discuss the potential contributions of theory-based evaluation in order to link both communities. Theory-oriented evaluation approaches are based on logic models which have the function of outlining how the simulation, its participants, and its environment interact with each other, and elicit the simulation's desired outcomes. Its primary aim is to gain evaluative knowledge from a particular gaming simulation which can then be used to improve the simulation and its implementation for practical purposes. The authors illustrate their viewpoints by discussing the theory-oriented evaluation of a business simulation game.*

KEYWORDS: *analytical science; design science; logic model; project example; theory-oriented evaluation*

There has been much criticism in the area of gaming and simulation research due to the disparity between conventional academic research and research based on practical experience. The conventional academic research carried out in this field has focused on developing and improving domain-specific knowledge by using simulation games in experimental environments (gaming and simulation laboratories). In contrast, the practical experience in this field has involved the transfer and dissemination of knowledge using specific simulation games with clearly defined designated audiences in a defined context of use. Klabbers (2004, 2006 [this issue]) uses the terms *science of analysis* and *science of design* to describe these two different approaches. The theory-driven science of analysis approach has used games and simulations as scenarios to test theories in various domains such as education, social psychology, politics, and economics. The main aim of the conventional science of analysis has been to develop generalized scientific concepts and context-independent knowledge. Accordingly, the external validity of findings is of primary importance to this approach. Research in the issue-driven science of design approach, on the other hand, puts the emphasis on the usability of the simulation game. In this case, games and simulations are studied with the aim of supporting and evaluating their development and use in practical contexts.

SIMULATION & GAMING, Vol. 37 No. 2, June 2006 268-283
DOI: 10.1177/1046878106287950

Here, it is necessary to focus on local knowledge and individual, unique circumstances to find customized solutions to practical problems. The science of design has two branches (Klabbers, 2003): *design-in-the-small*, which uses simulation games as artifacts to model processes of reality, and *design-in-the-large*, which uses the effects of simulation games to change existing dysfunctional situations, patterns of behavior, or systems structures into preferential ones. Therefore, it is essential that the artifact assessment takes the evaluation of a simulation game as a product into account, as well as its effect on the process of change. In the design science, the interplay of design-in-the-small and design-in-the-large is of primary importance.

Although we view both the science of analysis and the science of design as legitimate approaches to gaming and simulation, the question still remains how both approaches are to be interconnected and how each can benefit from the other's expertise. Both science traditions reach agreement in the area of gaming simulation, yet confusion and misunderstanding between representatives of both communities continue to cause conflict. It appears that representatives of both parties are reluctant to actively work together. As a consequence of this conflict of interests, a potentially fruitful exchange of ideas is failing to occur. One major cause for the obvious oppositional stance of both traditions—which Klabbers (2006) describes as a distinction between the communities of observers and communities of practice—has its roots in the different epistemologies and methodologies, as well as in practical terms as both communities use different criteria for success. Because we, as authors, belong to the scientific community of applied psychology, we sometimes experience this frustrating antagonism in our practice. In our case, the main criterion of success is in the number of new research findings and scientific publications that have to meet the specific standards of the science of analysis tradition, in particular, the rules of inference of empirical social research. On the other hand, we develop and evaluate concrete simulation games, which are solely regarded successful by our customers if they can realize well-defined learning effects and are able to trigger intended changes of human behavior and/or changes of organizational structures and work processes.

With regard to Klabbers's (2006) depicted framework, we take a position in between both traditions, as well as use the intersecting set of both of them. It is a position that requires the special commitment to both parties, which also puts us at risk of being misunderstood by both sides. However, especially in our discipline of educational and organizational psychology, the optimization of learning environments (e.g., in the vocational education processes) is a legitimate and common area of research. Here, applied (analytical) science and concrete requirements of practice encounter each other and are being linked by the so-called "design-based research." In this article, we discuss the potential contributions of theory-oriented evaluation to bridge the gap between both communities. We begin with a short overview of theory-based evaluation approaches.

Theory-based evaluation of gaming simulation with logic models

The main aim of theory-oriented evaluation approaches is to go beyond testing the outcomes of gaming simulations with regard to meeting their learning goals. The goal is not only to prove *whether* a simulation works but also to show *how* and *why* it works (or fails to work) in a given context. So-called logic models, which are also sometimes referred to as program models, program logics, or program theories, are important tools to accomplish these goals (Bickman, 1987). Logic models depict how relevant variables and their interaction lead to the desired and unintentional outcomes of gaming simulations. They normally differentiate at least between inputs (preconditions), processes (actions), as well as outputs/outcomes (effects) and describe the mutual relationships between these variables. Logic models rely heavily on theoretical considerations and knowledge in the domain of the artifact that is to be evaluated. Thus, in the case of evaluating gaming simulations, evaluators should have considerable knowledge about practical application, and it is often advisable for them to cooperate with the practitioners who develop and/or facilitate the game. Such preconditions imply that by using logic models, it is not enough to refer only to abstract, disembodied theories, as is common in the analytical science domain. Actually, we are also referring to "theories-in-use," which puts this approach at the same time into the domain of design science.

In the general field of evaluation research, there has been a call for a more explicit use of theory in evaluation for some time now. Early theorists criticized a tendency in evaluation to concentrate solely on the outputs and outcomes of a program. Suchman (1967), for example, proposed to conceptualize the processes and mediating factors in carrying out an evaluation. The term *program model* goes back to Weiss (1972), whereas Fitz-Gibbon and Morris (1975) were the first to introduce the term *theory-based evaluation*. Yet it took until the 1980s for theory-based approaches to come to life as fully fledged evaluation models (Chen & Rossi, 1983) and gain wide acceptance in the field (e.g., Bickman, 1987; Chen & Rossi, 1983; Rossi, Lipsey, & Freeman, 2004). Only recently have the authors suggested applying theory-oriented approaches for the evaluation of gaming and simulations in their context of use (Hense, 2004; Hense & Kriz, 2005; Kriz & Hense, 2004).

What was the motivation for these developments? Chen (1990) denounced common evaluation practice as "black-box evaluations." Similar to behavioristic learning theories, black-box evaluations observe only the inputs and outputs of an artifact, neglecting the intermediate relationships and mechanisms. This deficit goes back to restrictions of the classical quasi-experimental research design, which has its undisputed merits in determining the effects of programs, yet often fails to deliver information needed for improvement and implementation of a program. Moreover, in the domain of gaming and simulation—with its close ties to system dynamics—this kind of simplified view of social realities seems inappropriate.

To illustrate how the basic ideas of using logic models in evaluation are put into practice, one can refer to the comprehensive approach by Reynolds (1998). He distinguishes seven steps in conducting a theory-based evaluation, which can easily be adapted to the evaluation of gaming simulations. The first step is to develop a logic model for the game or simulation that is to be evaluated. Here, it is important to distinguish between two levels. In most cases, simulations are already based on a model that mimics a specific part of reality, dependent on the domain of the simulation. In the case of a simulation for the teaching of ecological awareness, for example, this model could depict an ecosystem, its various components, and their interdependencies. However, the logic model needed for evaluation purposes is a different one. It models the usage and application of the game or simulation for learning or other purposes in a given educational context such as biology lessons in schools. In the given example, this model would basically depict relevant antecedents, learning processes, and outcomes, which contribute to the desired goal of increased ecological awareness and competence. The second step is to measure the effects of participation in the simulation with regard to the logic model's outcomes. This is in accordance with traditional evaluations. In the third step, data are collected on mediating and background factors. They are also derived from the logic model. For the example used above, one would probably include learner characteristics such as previous knowledge and attitudes, or process variables such as joint strategic decisions and interactions affecting the ecosystem in the simulation. The fourth step includes estimating the main effects of participation (interactions and events), and the fifth step tests causal mechanisms of the logic model (as a dynamic artifact) to explain these effects. Here is where the theoretical approach can show one of its most important merits: In addition to learning about the effects of the game, the logic model can be used to analyze which factors contributed the most and which factors had a detrimental effect. In the sixth step, Reynolds's (1998) approach proposes to interpret the findings for the purposes of generalization and knowledge transfer, including generalizations that are relevant to other contexts of use for other audiences as well. The final step is the most important one for practical purposes because it contributes to improvement. Here, the results of the above steps are used for identifying possible areas of improvement of the gaming simulation or its implementation. This concerns mainly the design-in-the-large (e.g., in that the interactions between the players in the simulation game improve the effectiveness of the series of strategic decisions in an organization). Yet, also from a design-in-the-small perspective, the evaluation results could lead to improvements, such as including more accurate causal links in the dynamic game model.

The theory-oriented approach to evaluation is not the ultimate solution for every possible evaluation goal, because these goals can be different in different contexts. Two situations come to mind, where other approaches probably fit better (cf. Chen, 2005). The first one is in the very early stages of developing a gaming simulation. Here, the goal is to test prototypes of a game with only a small sample of stakeholders where purely qualitative, formative evaluations can be much more informative than a fully fledged quantitative design. Yet, also here, a theory-based approach can help to focus the evaluation, because it helps identify the key questions for interviews or observations referring to actions and intended and unintended effects. The second case where the theory-based approach is sometimes less well suited is within decision-making contexts, where one game or simulation is to be chosen from among a set of alternatives with the aim to serve a given goal in a fixed context of use with a clearly defined audience. To facilitate such a decision, a classical evaluation design using multiple experimental groups and measuring only the relevant outcome variables is often sufficient. Yet it has to be mentioned that although such decision-making situations are often cited as one of the most important reasons for conducting evaluations, in reality they seldom occur in a pure form (Alkin, 1985; Weiss, 1981). In the case of gaming simulation, for example, the result of a pure outcome evaluation is only applicable for decisions if one knows the conditions that brought forth these outcomes. If these are subject to variation (e.g., in cases where other audiences are to be addressed), knowledge of their influence is elementary, which again requires a consideration of mediating factors and causal mechanisms. These mediating factors are, among others, the interactions and consequences (effects) in their context of use for this intended audience with its particular history and interests.

To sum up the preceding arguments, logic models can fulfill multiple functions in evaluation, because they are representations of thorough theoretical and conceptual considerations in evaluation and game design (Hense, 2004). To mention only the most important ones—first, in an early stage, they can help identify the key factors in designing an evaluation. Second, they provide a frame of reference for interpreting the simulation's workings as a learning environment. Third, they help identify areas for improvement in the simulation's design or implementation. In the following sections, we exemplify these possible merits by introducing a concrete logic model and by discussing the relevance of logic models in the context of the traditions of design-in-the-large and design-in-the-small.

Example: Logic model of SIMGAME

To illustrate the application of the theory-oriented evaluation approach, we describe the logic model of SIMGAME. The project SIMGAME was a Leonardo da Vinci program of the European Union and was carried out in 2003 and 2004. SIMGAME is a business simulation game for economy lessons in secondary schools and was evaluated in about 30 schools from six different countries (Kriz, Puschert, & Hense, 2004; Kriz & Hense, 2004). Until now, the project was the first in the German-speaking area, where theory-oriented formative evaluation for business simulation games was used.

The logic model of the evaluation of SIMGAME is based on several sources, among others, (a) current simulation game research (Faria, 2001; Garris, Ahlers, & Driskell, 2002; Hindle, 2002; Kriz & Brandstätter 2003; Wolfe, 1997); (b) approaches of situated learning (Brown, Collins, & Duguid, 1989; Gruber, Law, Mandl, & Renkl, 1995; Hense, Mandl, & Gräsel, 2001), especially the so-called problem-oriented learning (Mandl & Gerstenmaier, 2000); as well as (c) more general models for the quality of instruction and learning environments (Ditton, 2000, 2002).

The input variables included not only sociodemographic data of the participants but also the motivation for participation in the game, because motivation can be seen as an important factor that influences the learning results (cf. Deci & Ryan, 1993). Theoretically, the previous knowledge, or previous experience as the case may be, can be considered as input factors (Krapp, 1997). These variables are not only relevant for the part of the pupils but also on the part of the teacher. Furthermore, the preparation of the teacher is important. Even the quality of the simulation plays an important role.

It should be clear that the included outcome variables for the model were those that could be used as indicators for achieving the learning objectives of the simulation game, for example, the acceptance of the simulation game by teachers and pupils and the social, professional, and methodological skills of the pupils that they should improve. Organizational changes such as positive effects on the class climate were also included.

For the area of process variables, the QUAIT-model (quality of instruction, appropriateness, incentives, time) from Slavin (1996) can be used for adding several main factors, such as the usable time for learning ("time on task" in the logic model), the appropriateness of the level of difficulty of the learning input (in the logic model, among others described and recorded by "felt under- and overchallenged"), as well as the quality of the instruction of the facilitator, which has to include the debriefing process, especially for simulation games. Furthermore, the interaction of the participants is a classical process variable itself (Friedrich, Hron, & Hesse, 2001; see Figure 1).

All variables of the logic model should be based on research results and the corresponding theoretical concepts. In addition, all the described factors should be operationalized, rated, and because of the derived theory-based hypothesis, analyzed with regard to their interaction. The various results of this evaluation are not reported here (cf. Kriz, Hense, & Puschert, 2004), because in this article, we intend to point out aspects of the benefits that theory-based evaluation can achieve, using SIMGAME as an example.

Design-in-the-large, design-in-the-small, and logic models

Simon (1969) has pointed out that design means to conceive and to implement courses of action aimed at changing existing dysfunctional situations into preferred ones. This approach of design-in-the-large is the foundation of all forms of consulting work, training, and education in the attempt to foster new ways of thinking and acting and to develop organizations (Klabbers, 2003, 2006). Gaming simulation design as a design-in-the-small approach enhances a shift of existing organizational cultures and structures and in this way contributes to the design-in the-large process of social systems. Existing dysfunctional educational and organizational situations are changed and/or improved into preferred ones (design-in-the-large) through the design and use of simulation games (design-in-the-small).

Figure 2 shows how Greif and Kurtz (1996) describe organizational change as design-in-the-large with characteristic sequential phases and feedback loops: diagnosis, defining goals, development of change strategies, concrete planning, action, and evaluation. Gaming simulation methods support the diagnosis phase to determine the actual condition. Gaming simulation helps to understand existing organizational structures and work processes by designing a present state simulation game. When playing and debriefing such a simulation, existing advantages and disadvantages of these structures can be illustrated, by fostering discourse on ideas for potential change strategies. In an organization,

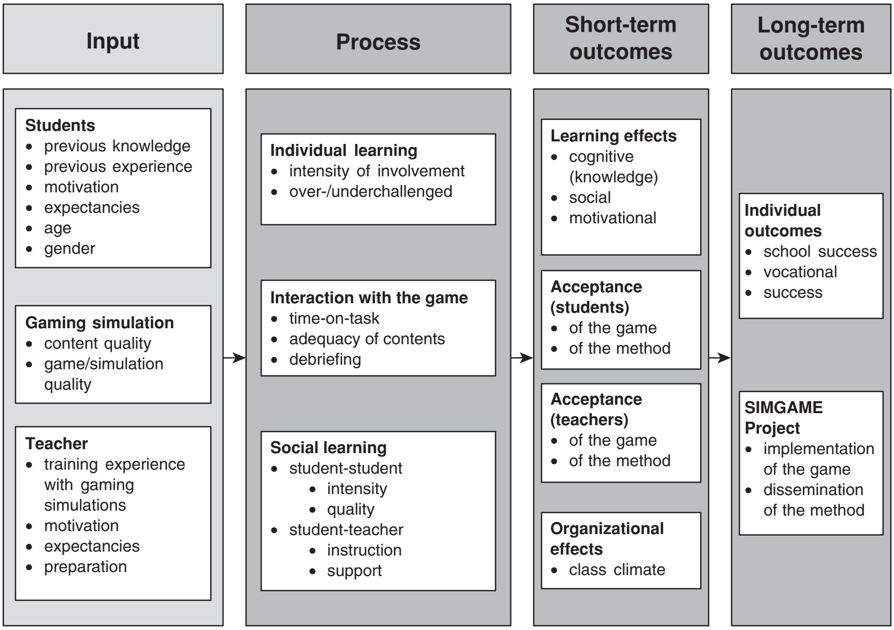

FIGURE 1: Logic Model of SIMGAME

desired changes can be illustrated in "vision/future state simulation games" (see Ruohomäki, 2002, 2003). The knowledge that is acquired and the conclusions that are drawn can be used to define goals and concrete planning of change measures.

With gaming simulation, consequences of alternative scenarios in a changed organizational structure can be tested, scored, and discussed (test scenario games). Gaming simulation can further be used as an intervention tool for human resource management and organizational development in the training of specific department-related (local) knowledge and required skills (training simulation game) and, in this way, support the execution of concrete change actions.

The whole process of gaming simulation is illustrated in Figure 3. A part of the existing situation of reality is selected as a reference system for the designed simulation game (Kriz, 2003). The final aim is to change organizational structures and processes of social systems and/or to use the simulation game for education and training to change mental models and to foster skills of individuals within the real world ("reality"). To carry out this design-in-the-large, a game as an artifact is created and used as a dynamic model of reality (design-in-the-small). The game is applied through play and facilitation. By inviting stakeholders and opinion leaders to participate in the design process, it becomes natural to have them contribute as agents as well as actors. Participating in designing, playing, and debriefing allows the players to take part in the design-in-the-small process, while ultimately contributing to the next phase of the

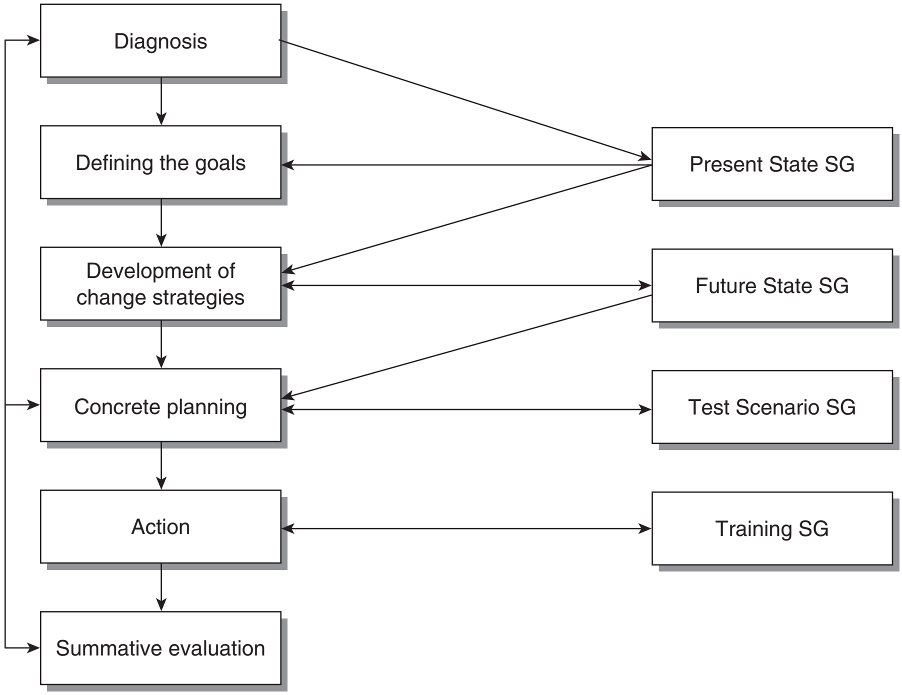

FIGURE 2: Greif and Kurtz (1996) Model of Organizational Change (design-in-the-large) and Possible Support With Gaming Simulation (design-in-the-small)

social system processes' design-in-the-large. Debriefing is facilitated to enhance the learning process and to apply newly gained insights, knowledge, and skills within the design-in-the-small aimed at changing reality (design-in-the-large). The secondary phase of debriefing, referred to as meta-debriefing and summative evaluation, is required to encourage further reflection on the linkages between design-in-the-small and design-in-the-large and to measure profits of changes in reality.

Logic models are helpful for the whole design process in various ways. By designing a game, the logic model can support the necessary process of reducing complexity. Logic models focus on achieving certain defined learning and change objectives. Therefore, they conduct the selection of the reality factors depicted in the simulation and their interdependencies. For facilitating, debriefing, and meta-debriefing of the game, they provide knowledge for the interaction and the behavior of the players and facilitators on the basis of the selected process variables. Especially simulation game didactic, which has to be included in designing the facilitation of simulation games and its reflection to use the whole potential of the learning environment simulation game perfectly, can be derived from the logic model. On one hand, the logic model sets a frame for operationalizing and measuring the variables as success indicators of the simulation game,

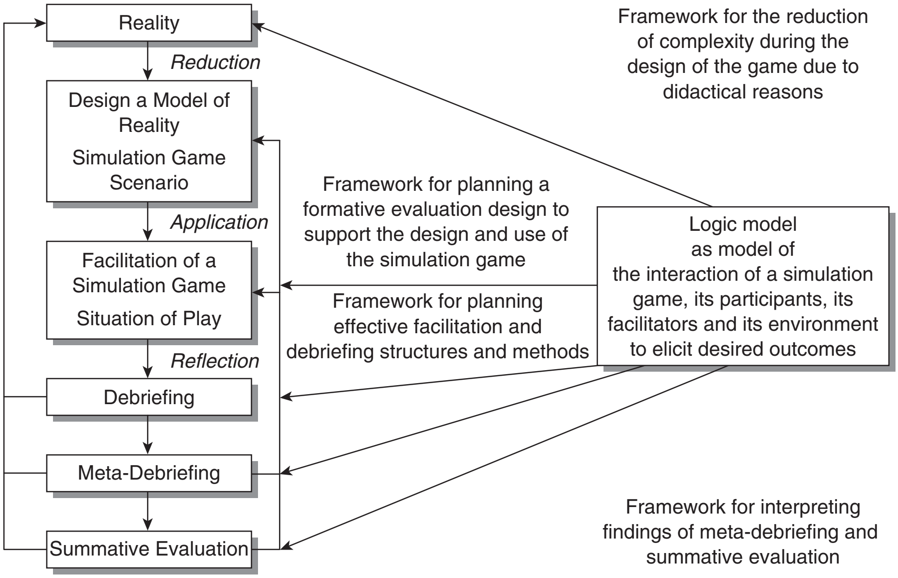

FIGURE 3: The Design-in-the-Small Process of Gaming Simulation and Its Support by Logic Models

as well as for the interpretation of results in the sense of a *summative evaluation*. On the other hand, the logic model can also be used for ongoing optimization of development and implementation of the simulation game in the sense of *formative evaluation*.

Therefore, the logic model has to be designed simultaneously within the first steps of a simulation model and, in an ideal case, by the same project team that is also responsible for designing the simulation game. Only this can ensure that the logic model leads to a target-oriented workflow and to a commonly shared vision of all designers, right from the beginning of the communication among all persons involved. Naturally, the logic model—as well as the simulation game itself—can change during the design process, because the logic model and the model of the simulation game are deeply interlocked with each other.

For SIMGAME, right from the beginning this was used to weight the project objectives transparently and with regard to its importance. In a later phase, interdependencies between the postulated variables could be better interpreted on the basis of the logic model and the corresponding hypotheses. Therefore, it was possible, for example, to properly testify further optimizations for the simulation game process or to answer the cost output question, Which of the several designed versions of SIMGAME leads under which conditions for which target group to the best effects?

The advantages and practical values of the theory-based evaluation are now described with another concrete example. It has already been said that the quality of the simulation game is an important input variable for achieving the learning objectives. It should be clear that it is not enough only to name this variable. In a further step, every variable of the logic model has to be operationalized and measured. In the case of SIMGAME, experts (in game design as well as in the game-content-related field of economy), stakeholders (e.g., representatives of federal economic chambers and of business associations), and teachers and pupils have given feedback several times on the quality of the simulation game in the design and implementation phase, using a questionnaire. Therefore, an appropriate list of quality criteria, based on gaming simulation research, was defined. The list of criteria shown in Table 1 (Kriz & Hense, 2005) supported and guided the design of the simulation game, to achieve all these 50 criteria as well as possible. In this way, the theory-based approach that uses knowledge from analytical science adds practical value to the design-in-the-small of the game and later contributes to the design-in-the-large of our educational systems by applying the simulation game in the classrooms.

Bridging the gap between design and analytical science

Simulations are used as a method for experimenting in the framework of research in the analytical sciences. To be able to study complex structures and processes, the participants have defined positions and limited room for maneuvering in these experiment-simulation games (e.g., Kriz, Hettinger, Nerl, & Gräsel, 2001). In this context, in principle, the simulation is not different from the experimental/quasi-experimental method, even if it is not set up as a quasi-experimental study. Nevertheless, for example, observations and qualitative research are carried out in the context of traditional academic research. Moreover, in the single discipline of educational psychology, which puts the optimization of learning environments in the center of its research interests, gaming simulation experiments can lead to new knowledge and new theories about instruction and learning processes. These new theoretical approaches are of highly practical value in the field of gaming simulation design and support the definition of new logic models, which are based on scientific research and founded correlations of input, process, and output variables.

Game designers must use knowledge from the analytical sciences as input into their design practice, because this knowledge can contribute to building and using the designed artifacts (the game). To do so, they must, of course, translate the universal context-independent knowledge of analytical science research into their concrete local circumstances. As an example, there is a typical science of analysis study by Stark, Gruber, Graf, Renkl, and Mandl (1996), who used the business simulation game JEANS-COMPANY as an experimental setting to examine, among other things, the influence of different instructional conditions, with forms of guided versus nonguided problem solving by the facilitator in relation to different groups of participants. The knowledge gained from this experiment is mainly important for the definition of relevant process variables in logic models dealing with simulation games. Conclusions from this study provide practical input to the design of the facilitation of simulation

TABLE 1: SIMGAME: List of Criteria for the Quality of a Simulation Game as Example for the Operationalization of an Input Variable of the Logic Model

1. The learning objectives are clearly defined.
2. The target groups are clearly defined.
3. The possible areas for implementation are clearly defined.
4. The schedule and the structure of the game are clearly defined.
5. The spatial game setting is clearly defined.
6. The simulation game can be completely played in the defined time schedule.
7. The rules of the game are clearly defined.
8. The roles of the players are clearly defined.
9. The scenario of the game and the events occurring in the game are clearly defined.
10. All materials, game manuals, resources, and so on needed are available for successfully playing the game.
11. The facilitation skills that are needed are clearly defined.
12. The skills (preknowledge) of the players that are needed to play the game are clearly defined.
13. The understandability of the written materials (manual, facilitator and player guide, etc.) is very high.
14. The written materials provided are adequately comprehensive.
15. An adequately formulated theoretical model exists that shows how and what can be learned during the game.
16. The simulation has very good visualizations of the simulated processes and structures.
17. The simulation has attractive materials.
18. The simulation has sustainable materials.
19. The materials of the simulation can be used easily.
20. The design of the game supports an easy and intuitive usage of the simulation for facilitator and players (in addition, computer-based simulation games have a technically perfect and user-friendly interface).
21. The simulation can be used with adequate effort.
22. The simulation results can be analyzed with adequate effort.
23. The simulation includes a good reporting and recording system (decision processes, changes of simulated systems, and achievement of learning objectives can be derived from adequate indicators all the time).
24. The simulation offers adequate flexibility in the workflow (e.g., it is possible to go one step back and decline decisions).
25. The simulation offers adequate adaptability for changed framework conditions (e.g., for smaller/larger number of participants or for longer/shorter schedule, etc.), and the facilitator guide offers suggestions and hints for a flexible usage under changed framework conditions.
26. The simulation offers a motivating and interesting game scenario.
27. The simulation offers the players uncertainty to an adequate extent.
28. The simulation activates the participants to think about interconnections of simulated systems elements.
29. The simulation activates the participants to develop strategies.
30. The simulation activates the participants to rate sequences and side effects of problem-solving alternatives.
31. The simulation offers a variety of interactions between participants.
32. The simulation encourages a variety of perspectives and change of perspectives.
33. The simulation encourages an understanding of different interest groups.
34. The simulation offers an adequate link to reality for the target group; rules, roles, and simulated resources correspond to real, authentic situations.
35. Main processes and interconnected factors of reality are translated into the game model correctly.

(continued)

TABLE 1 (continued)

36. The simulation has an adequate level of complexity for the target group (no permanent under- or overchallenge).
37. The simulation offers several different alternatives of acting and deciding.
38. There is a realistic scope of acting and deciding for the players.
39. Highly skilled players/teams achieve better game results (with regard to learning objectives) in comparison with lower skilled players/teams.
40. The facilitator guide contains explicit hints for briefing the simulation game (e.g., role-taking processes, basic information, guidelines for tolerated and not-tolerated behavior of the participants, etc.).
41. The instructions in the facilitator manual for gaming simulation didactic contribute to a perfect workflow (the tasks of the facilitator—e.g., the roles the facilitator has to take—during the game are clearly expressed).
42. The guidelines in the facilitator manual about debriefing ensure the learning objectives that should be achieved (i.e., there are hints about topics, structure/schedule, and methods of debriefing).
43. The guidelines about gaming simulation didactic ensure the realization of desired learning objectives in practice (e.g., there are explicit hints about connecting the simulation with the real work processes of the target group).
44. There are explicit hints in the manual about embedding the simulation game in a whole teaching/learning context (e.g., with regard to the curriculum).
45. Beside the simulation game, there are complementary learning modules (i.e., in addition to the debriefing modules), which are target-group oriented and help link the experience of the simulation game with important knowledge and competence components in the sense of a higher qualification concept (e.g., case studies, texts for teaching, professional teaching videos, etc.).
46. The game is evaluated continuously and improved if needed.
47. The main learning targets are achieved by the majority of the players.
48. The game offers an adequate cost outcome relation (price, time amount, compared with suggested intervention or learning effects).
49. The game sticks to usual ethical guidelines (e.g., human dignity is not injured, no sustaining and unjustified discrimination, etc.).
50. The participants have the freedom to stop the game whenever they want (challenge by choice) if they are anxious about personal limits of dignity or unreasonable stress; in the case of their leaving the game, these participants have the choice to take the role or task of an observer.

games, particularly by avoiding "under- and overchallenge" (notice the variable in the logic model shown in Figure 1 and a quality criteria—no. 36—in Table 1). In the SIMGAME project, this knowledge had a substantial effect on the design of the game, as well as on the further development of a special teacher-training course.

In return, once a game has been designed as an artifact of a design-in-the-small process, by taking knowledge from analytical science into account, it can be used both for the design-in-the-large and/or as a method or model in the analytical science tradition to develop and test theories (Klabbers, 2004, 2006). In addition, the design, the application, and the evaluation of a simulation game as a training method can be a subject of research. The formative and theory-based evaluation offers results that contribute to the evolution of the analytical science in the sense of applied psychological-educational evaluation research. The list of quality criteria of simulation games, named above, is a good example for knowledge gained about simulation games, viewed as learning environments. That knowledge is based on the results of applied analytical science, which can also be of relevance for the design science and, more specifically, for the design of simulation games. Using this list of criteria in the evaluation of the SIMGAME project as a realization of the science of design in return offers data that can be used for the theory testing in analytical science. Recently, this linkage has been highly valued especially in the educational sciences, and it is required in the conceptualization of design-based research (Hoadley et al., 2003).

> Design-based research, which blends empirical educational research with theory-driven design of learning environments, is an important methodology for understanding how, when and why educational innovations work in practice. Design based research . . . help[s] us understand the relationships among educational theory, designed artefact, and practice. Design is central in efforts to foster learning, create usable knowledge, and advance theories of learning and teaching in complex settings. (p. 5)

Especially by connecting the science of analysis and the science of design, which is the objective of our current and future integrated research and design projects, generalized logic models for whole types of simulation games can be defined. Some elements of the SIMGAME can and will surely be generalized for all rigid-rule board-based simulation games, when used in schools. In every simulation game, actors, rules, and resources are interrelated (Klabbers, 1999). The interaction of the players acting on the game model within the constraints of the game rules, and the behavior of the game facilitator, are therefore mediating factors that must be included in the process variables in the logic models. Knowledge both from the communities of observers and practice and from academic analytical research and experiences from designing practice can thus be integrated in the logic model, which in return supports the designing process of simulation games.

Through the implementation of the theory-oriented evaluation approach in several unique design projects in practice, again, universal concepts are created, which—by being tested—are finally the objectives of science of analysis. Furthermore, this approach can also increase the usability of simulation games in specific situations, which helps to achieve the objectives of the design science. Despite the term *theory*-oriented evaluation, like any other evaluation model, this approach nevertheless belongs also to the science of design tradition: Its primary aim is to gain evaluative knowledge on one particular gaming simulation, which in turn can be used to improve the simulation or its implementation for practical purposes. In this context, the term *theory* should be understood as "meta-artifact" (Klabbers, 2006). Yet, due to the intensive use of theoretical considerations, it has considerable potential for producing findings that can be transformed and accumulated into generalized knowledge. The gap of both science traditions could be bridged here.

Conclusions

Our detailed analysis shows that we sometimes use research terminology of the analytical science (e.g., when we talk about input and output variables) to underpin our theory-based approach for the design sciences. This might well lead to criticism that possible differences may become blurred in reality through different existing methodologies and terminologies of both communities. As a matter of fact, we think that by staying with the theory-oriented approach within the standards of the analytical science, we stay recognized and more visible by our peers in the scientific community. Performing evaluations of gaming simulation, it is definitely important to us to examine classical theory-driven problems of our field (e.g., for the evaluation of learning environments), using the classical methodical approach of the quasi-experimental research method, based on the variable approach of causality, performing our evaluation studies with common statistical techniques, and interpreting the results by the scientific standards of our peer group. At the same time, within the framework of formative evaluations in real issue-driven projects (e.g., such as SIMGAME; see above), in practice our approach yields concrete results with regard to usability and optimization of the designed simulation games. In this way, the participating stakeholders accept the parallel use and development of the logic model along with the simulation model. We find ourselves in a double role. We are researchers who claim to develop generalized concepts that are part of the analytical science tradition. We are also change agents in the design science tradition, offering domain-specific knowledge and expertise in their context of use. We are only accepted in practice if we can meet the expectations of the sponsors with regard to practical outcomes. In our view, the logic model, which emphasizes theory-based evaluation, is well suited as a link to both communities and research traditions in the field of gaming simulation.

References

Alkin, M. C. (1985). *A guide for evaluation decision makers*. London: Sage.

Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation*, *33*, 5-18.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-42.

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.

Deci, L. E., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik [Theory of self-determination and its impact on pedagogics]. *Zeitschrift für Pädagogik, 39*(1), 223-238.

Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. [Quality assurance in school and education.] *Zeitschrift für Pädagogik*, 41(Suppl.), 72-92.

Ditton, H. (2002). Unterrichtsqualität. Konzeptionen, methodische Überlegungen und Perspektiven [Quality of education. Concepts, methodical considerations and perspectives]. *Unterrichtswissenschaft*, pp. 197-212.

Faria, A. J. (2001). The changing nature of business simulation/gaming research: A brief history. *Simulation & Gaming*, *32*(1), 97-110.

Fitz-Gibbon, C. T., & Morris, L. L. (1975). Theory based evaluation. *Evaluation Comment*, *5*(1), 1-4.

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. *Journal of Simulation & Gaming*, *33*(4), 441-467.

Greif, S., & Kurtz, H.-J. (1996). *Handbuch Selbstorganisiertes Lernen* [Handbook of self-organized learning]. Göttingen: Verlag für Angewandte Psychologie.

Gruber, H., Law, L.-C., Mandl, H., & Renkl, A. (1995). Situated learning and transfer. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 168-188). Oxford: Pergamon.

Hense, J. (2004). Theory-oriented evaluation of gaming simulations—The case of Simgame. In W. C. Kriz & T. Eberle (Eds.), *Bridging the gap: Transforming knowledge into action through gaming & simulation* (pp. 339-351). München: SAGSAGA.

Hense, J., & Kriz, W. C. (2005). Theoriebasierte Evaluation und Bildungscontrolling. In M. Gust & R. Weiß (Eds.), *Praxishandbuch Bildungscontrolling für exzellente Personalarbeit. Konzepte—Methoden—Instrumente—Unternehmenspraxis* (S.) (pp. 231-242). München: USP Publishing.

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 should there be education with new media more than just instruction with new media?]. *Computer und Unterricht*, *11*(4), 66-11.

Hindle, K. (2002). A grounded theory for teaching entrepreneurship using simulation games. *Simulation & Gaming*, *33*(2), 236-241.

Hoadley, C., Baumgartner, E., Bell, P., Brophy, S., Hsi, S., Joseph, D., et al. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5-8.

Klabbers, J. H. G. (1999). Three easy pieces. In D. Saunders & J. Severn (Eds.), *The international simulation & gaming research yearbook. Volume 7: Simulation and games for strategy and policy planning* (pp. 16-33). London: Kogan Page.

Klabbers, J. H. G. (2003). Gaming and simulation: Principles of a science of design. *Simulation & Gaming*, *34*(4), 569-591.

Klabbers, J. H. G. (2004). *Saga about ISAGA: A retrospect and a prospect*. Keynote lecture at the 35th ISAGA Conference, Munich.

Klabbers, J. H. G. (2006). A framework for artifact assessment and theory testing. *Simulation & Gaming*, *37*(2), 155-173.

Krapp, A. (1997). Selbstkonzept und Leistung: Dynamik eines Zusammenspiels [Self-concept and performance: Dynamics of an interaction]. In F. E. Weinert & A. Helmke (Eds.), *Entwicklung im Grundschulalter* (pp. 325-339). Weinheim: Beltz.

Kriz, W. C. (2003). Creating effective interactive learning environments through gaming simulation design. *Simulation & Gaming*, *34*(4), 117-134.

Kriz, W. C., & Brandstätter, E. (2003). Evaluation of a training program for systems-thinking and teamwork-skills with gaming and simulation. In F. Percival, H. Godfrey, P. Laybourn, & S. Murray (Eds.), *The international simulation & gaming research yearbook. Volume 11: Interactive learning through gaming and simulation* (pp. 243-247). Edinburgh: Edinburgh University Press.

Kriz, W. C., & Hense, J. (2004). Evaluation of the EU-Project "Simgame" in business education. In W. C. Kriz & T. Eberle (Eds.), *Bridging the gap: Transforming knowledge into action through gaming & simulation* (pp. 352-363). München: SAGSAGA.

Kriz, W. C., & Hense, J. (2005). Evaluation und Qualitätssicherung von Planspielen [Evalution and quality assurance of simulation games]. In U. Blötz & Bundesinstitut für Berufsbildung BIBB Bonn (Eds.), *Planspiele in der beruflichen Bildung* (Hauptartikel im Fachbuch; 4. erweiterte Auflage). Bielefeld: Bertelsmann.

Kriz, W. C., Hense, J., & Puschert, M. (2004). Endbericht zur Evaluation des EU-Projektes "Simulation betriebswirtschaftlicher Entscheidungsprozesse—Simgame [Final report on evaluation of the European Union project "Simulation of economic decision processes-Simgame"]."

Kriz, W. C., Hettinger, S., Nerl, A., & Gräsel, C. (2001). Sustainable resource management in the simulation game Fish Banks Ltd.: An experiment on the influence of systemic and social knowledge. In A. Villems (Ed.), *Bridging the information and knowledge societies* (pp. 224-234). Tartu: Tartu University Press.

Kriz, W. C., Puschert, M., Karl, J. & Dufter-Weis, A. (2004). Effective learning through gaming simulation design. In R. Shiratori, K. Arai, & F. Kato (Eds.), *Gaming, simulation & society* (pp. 217-227). Tokyo: Springer.

Mandl, H., & Gerstenmaier, J. (2000). *Die Kluft zwischen Wissen und Handeln* [The gap between knowledge and action]. Göttingen: Hogrefe.

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

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

Ruohomäki, V. (2002). *Simulation game for organisation development. Development, use and evaluation of the Work Flow Game*. Unpublished dissertation, Helsinki University of Technology, Industrial Management and Work and Organizational Psychology.

Ruohomäki, V. (2003). Simulation gaming for organizational development. *Simulation & Gaming*, *34*(4), 531-549.

Simon, H. (1969). *The sciences of the artificial*. Boston: MIT Press.

Slavin, R. E. (1996). *Education for all*. Lisse: Swets & Zeitlinger.

Stark, R., Gruber, H., Graf, M., Renkl, A., & Mandl, H. (1996). Komplexes Lernen in der kaufmännischen Erstausbildung: Kognitive und motivationale Aspekte [Complex learning in basic education of merchants]. *Zeitschrift für Berufs—und Wirtschaftspädagogik*, *13*, 23-36.

Suchman, E. A. (1967). *Evaluative research. Principles and practices in public service and social action programs*. New York: Russell Sage.

Weiss, C. H. (1972). *Evaluation research: Methods of assessing program effectiveness*. Englewood Cliffs, NJ: Prentice Hall.

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

Wolfe, J. (1997). The effectiveness of business games in strategic management course work. *Simulation & Gaming: An International Journal*, *28*(4), 360-376.

Willy Christian Kriz, PhD, MSc, is a professor of human resources management at the University of Applied Sciences Vorarlberg, a Steering Committee member, the former president of ISAGA, the director of the ISAGA Summerschool, and the chairman of the Swiss Austrian German Simulation and Gaming Association (SAGSAGA). He was the head of the evaluation of gaming simulation in a Leonardo da Vinci Program of the European Union. He is the author of five books and more than 60 papers. He teaches organizational psychology, project management, and organizational and personnel development. His further research interest is training of systems thinking and team management with gaming simulation, gaming simulation design, and evaluation. He is also working as a consultant and trainer in continuing education and business.

Jan Ulrich Hense is an assistant professor of education and psychology at Ludwig Maximilians University. His research interests include evaluation theory and evaluation research, entrepreneurship education, innovative learning methods, and learning with new media. As an evaluator and evaluation consultant, he has been involved in several local, national, and transnational evaluation projects. He teaches instructional design, statistics, research methods, and evaluation and is an active member of the Swiss Austrian German Simulation and Gaming Association (SAGSAGA).

ADDRESSES: WCK: University of Applied Sciences Vorarlberg, Department of Management and Research Methods, Hochschulstr. 1, 6850 Dornbirn, Austria; telephone: +43 (0) 5572-792-3218; fax: +43 (0) 5572-792-9274; e-mail: willy.kriz@fhv.at. JUH: Ludwig Maximilians University, Department of Psychology, Leopoldstr. 13, 80802 Munich, Germany; e-mail: hense@emp.paed.uni-muenchen.de.