

Evaluating Instructional Games and Simulations

Ted Ward

Evaluation of instructional games and simulations has two purposes. First, evaluation can provide a basis for making modifications and improvements in the learning exercise. Evaluation for this purpose has become known as *formative* evaluation. The second set of purposes is the more classical—i.e., determining the effectiveness of the learning exercises. Evaluation for this purpose, *summative* evaluation, is concerned that an instructional material or system has accomplished the learning objectives for which it was intended. Summative evaluation is essentially *product* evaluation or *outcome* evaluation, where the outcome is in terms of changed learners or accrued learning.

Evaluation of instructional games and simulations should give account of three aspects of the instructional materials and experience:

1. As would be the case with any educational evaluation, there must be a concern for what has been learned. *Learning* is typically evaluated in terms of the change in the learner which can be credited to the given instructional experience. This first aspect of evaluation is concerned with the gain or changes in content knowledge (and all other significant aspects of learning) which can be seen as a consequence of the instructional game or simulation. A taxonomic approach to learning—concern for the specific kinds and levels of cognitive, affective and psychomotor learnings—is required.
2. One of the major reasons for the use of instructional games and simulations is to increase the interest level of the learners while moving toward the desired learning outcomes. *Motivation*, in this sense, is concerned with the degree to which an instructional experience stimulates and sustains the interest of the learner.
3. Although much of formal education evades the test of demonstrated positive *transfer of learning* to the assumed equivalent tasks of life, concern for transfer is often the primary reason for using instructional games and simulations. Thus it is incumbent upon designers and users of instructional games and simulations to be concerned about transfer of learning when they are planning an evaluation program.

Evaluating Learning

Although it is more characteristic to evaluate learning by giving a test at the end of a learning experience, any valid claims about change in the learners must be based on comparisons of data from tests at the end of the experience and comparable tests at the beginning. Thus a pretest/post-test approach to evaluating the learning increment is indicated. There should be a high degree of consistency between the stated objectives for the experience and the items to be examined in the pretest/post-test routine. Typically, in a well-designed instructional program it is possible to infer the objectives by examining the tests or to project the tests by examining the statement of objectives.

One of the commonly overlooked concerns in evaluation of content gain is the possibility that learners may be regressing in some important area while they are apparently gaining on the intended learning. A comprehensive evaluation of learning should give some attention to the sorts of “bad learnings” or unintended accomplishments that may accompany a learning experience. This can be accomplished by providing for a monitoring (in the pre/post testing) of crucial concomitant learnings that should not be allowed to regress or develop in the wrong direction as a byproduct of the instructional simulation.

Evaluating Motivation

Evaluating motivation can be accomplished best by observing learners of different sorts while they are engaged in a learning experience. Often the developer or leader can pay attention to the levels of interest and the points in the game or simulation at which interest lags or the students become disengaged from one another or from the experience itself. It is helpful to make a chart, blocking off on the horizontal the various phases of the instructional experience, and to mark a series of points representing highness and lowness of interest at many points in time. Thus it is possible to derive a relatively reliable picture of the rising and falling of apparent interest during each phase of the experience. A simple graph can be produced using several colored pencils to represent the interest level of several students under observation. Another commonly used technique for evaluating motivation is to ask participants to write their level of interest or to respond to questions that ask specifically about their interests in, as well as enjoyment and satisfaction derived from, various parts of the instructional exercise. It has been our experience that, in general, students are able to add significant information through this sort of questionnaire or open-ended response to questions about motivation, interest and enjoyment. But such information is no substitute for direct observation of a group of students engaged in the experience under conditions for which it was designed.

Evaluating Transfer of Learning

In fairness, there would seem to be no strong reason why instructional simulations and games should be expected to pass a more critical test of transferability from the school environment to the real world than are other school experiences. Nevertheless it has been accepted among designers and users of instructional games and simulations that their concern for evaluating transfer of learning is required because of the claims that are made about simulation and instructional games. If instructional games and simulations are used in order to increase the relevancy of learning to the real world, designers and users should determine that, in fact, such a claim is valid.

The most valid tests of transfer would necessarily be done in longitudinal studies. Testing the learners as they encounter the real situation for which the simulation had supposedly trained them would be required. Since in many instances the moment or encounter with the real world equivalent may be months or years later, a longitudinal approach, while ideal, may be impractical. (It is just this impracticality that has been used historically to exempt the school from answering the tough questions about the relevancy and the transferability of learnings gained within formal education.) Thus instructional simulation designers and users often use a second approach when longitudinal studies are impractical; a second simulation is designed in which the major purpose is to test the application of skills gained in the first simulation. In other words, the transfer of learning gained in a particular simulation or game is assessed in terms of the degree to which the student can competently deal with the problems in a new but related situation. Thus some simulations have as their primary purpose *training* while others have as their primary purpose the *evaluation* of the learning from a previous simulation.

Distinctions Between Formative and Summative Evaluation

With reference to the *evaluation of learning* (content acquisition), both formative and summative evaluation is indicated. Of primary concern during the development of or the refinement of an instructional game or simulation is evaluation that indicates which of the learnings are being accomplished and which are not and, to some degree, what sorts of learners are learning and what sorts are not. Such information should be made available from a *formative* evaluation so that

designers can take steps to make corrections or adjustments as indicated by the patterns of strengths and weaknesses in the learning exercises. From a summative point of view, evaluation of learning is necessary in order to establish the validity of the game or simulation in terms of its potentialities for achieving its claimed learning objectives. For the users, whether or not to use the particular instructional experience again is decided by summative evaluation. Summative evaluation can also be useful to determine what auxiliary experiences must be provided in order to broaden the learnings and to clear up the less well accomplished objectives.

Summative evaluation, especially if derived at several different points during an experience, is also useful as an instructional input to the learner. As learners gain information based on the results of their actions and decisions, they become more fully a party to the whole learning process, making changes and corrections as needed to more fully accomplish their own learning objectives.

Most *evaluation of motivation* of an instructional game or simulation is done for formative purposes. Such evaluations are helpful in the refinement of an instructional experience, assuming that the designer is free to alter the characteristics in such a way that the phases of an activity in which interest wanes can be rebuilt or redesigned so as to keep the interest level higher. As a summative evaluation, indications of the observed motivation level of the game or simulation can be useful in making decisions about whether or not to use it for particular groups.

Evaluations of transfer of learning can serve formative and summative purposes. In both instances they serve as the crucial indicator of whether or not the experience in its present form will be justified in terms of cost and time as the educational intervention leading toward a certain competency in the real world. For the designer, this sort of information—usually derived from a research study of the relationship of learning acquired in the exercise to the learning demanded in the real world—is extremely useful in the redesign of a faulty game or simulation. For the user of the exercise, data concerning the degree of transfer can be a basis for making decisions about whether or not to use the experience.

Summary

Instructional games and simulations are a product of the technological approach to educational design. As such they are necessarily subject to the major criteria of effective educational technology: precisely defined objectives and precisely measured learning outcomes.