Developments in Business Simulation and Experiential Learning, Volume 26, 1999

A FRAMEWORK FOR ASSESSING THE COMPETENCIES REFLECTED IN SIMULATION PERFORMANCE

Hugh M. Cannon, Wayne State University
Alvin C. Burns, Louisiana State University

ABSTRACT

This paper is part of an on-going stream of research designed to determine the potential of using simulation games as instruments of assessment. Previous research has failed to validate games for this purpose, in part, because it has not specified what competencies the games should assess. This paper seeks to address the problem by proposing a three-dimensional taxonomy of performance competence. It then discusses how researchers might operationalize the three dimensions, thus developing alternative competency measures to which game performance might be compared to establish its validity.

INTRODUCTION

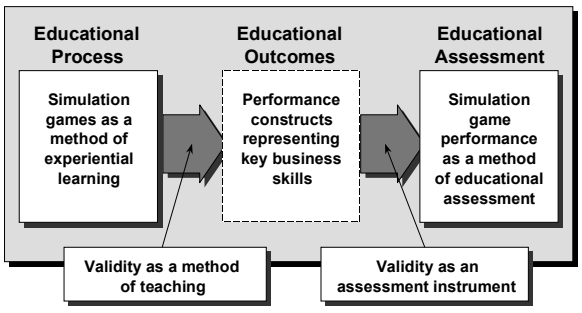
In 1990, Keys and Wolfe suggested that "management games will play a more significant role in management development and assessment efforts in business schools as part of the move toward competency-based outcome measure-ment" (p. 324). This provided an impetus for The Association for Business Simulation and Experiential Learning (ABSEL) to organize a Committee on Assessment (Thavikulwat, 1995) to pursue the use of games as measures of measures of managerial competence. The Committee drafted standards and registration procedure for assessment instruments (Cannon, Frazer, Goff, Markovich, Stevens, and Thavikulwat (1996), presented them at ABSEL's 23rd Annual Meeting in Orlando, Florida, and published the final version in *ABSEL News & Views* ("Standards and Registration Procedure," 1996). The standards applied to both simulation games and/or experiential exercises.

While several games were submitted for registration, the Committee found itself frustrated by a lack of evidence that games truly measured relevant managerial competence. Evidence would take the following forms:

1. Reliability in the instrument's measurements;
2. Discrimination by the instrument between individuals within a population with different types and/or degrees of learning;
3. Convergence between the instrument's measure and other reasonable measures of learning;
4. Normative scores for different populations.

In a paper presented to the 1998 annual ABSEL conference, Anderson, Cannon, Malik and Thavikulwat (1998) addressed the problem of validation from two perspectives. First, the traditional gaming literature has sought to determining the instructional validity of simulations by showing how they contribute to reliable and valid educational outcomes (Burns, Gentry, & Wolfe, 1990; Keys & Wolfe, 1990; Stanislaw, 1986). This is illustrated by the left-hand side of Exhibit 1. Second, the movement toward using games as tools of assessment seeks to determine validity of game performance as an indicator of managerial competence.

EXHIBIT 1
SIMULATION GAMES AS ASSESSMENT INSTRUMENTS VS. EDUCATIONAL TOOLS



Obviously, the two perspectives are related. They share a common set of performance constructs

that provides both the measure of educational success and of managerial competence. But this leads to a potential problem. But what if success in the simulation game does not represent managerial competence, but rather, some other construct – such as ability and motivation to play simulation games? If this is the case, teachers who use simulations as educational tools will naturally structure their teaching to promote game performance. Students who respond well to teaching activities will tend to do well and appear to be effective managers. This reinforces the notion that simulation performance is a valid measure of competence, even when it is not.

The real question, then, addresses the nature of simulation performance. Does it really represent managerial competence, competence as a student, or some other construct all together? The only way to answer the question is by studying the nature of simulation performance, comparing it to other measures of competence to see if it appears to be related.

The purpose of this paper will be to take another step in this line of research. Performance competence is not an easy thing to measure. Research suggests that it involves a number of different dimensions. Our job is to develop a theoretical framework which not only helps us understand the nature of performance competence, but which leads us closer to our ultimate goals of reliable and valid measurement.

TOWARD A FRAMEWORK FOR EVALUATING PERFORMANCE COMPETENCE

The quest for a valid set of performance constructs is by no means new to education. In the late 1940s, a committee of renowned educational psychologists sought to develop a comprehensive set of educational objectives. The result was a six-level taxonomy that suggested performance involved six increasingly demanding levels (Bloom, Englehart, Furst, Hill, and Krathwohl 1956). These ranged from simple knowledge to the ability to create new ideas and evaluate them according to their merits. They have come to be known commonly today as *Bloom's Taxonomy* of educational objectives.

Note that Bloom's taxonomy addressed *cognitive* learning, or learning that addresses how to consciously acquire and manipulate ideas. Subsequent work addressed the question of feelings – values and attitudes. It acknowledged the fact that success is more than thinking. It involves a great deal of socialization, molding to the culture in which business (or other types of) success takes place. The result was a second, *affective*, taxonomy (Krathwohl, Bloom, and Masia 1964) which involved such things as the propensity to pay attention to the appropriate cues, to the tendency to prioritize effectively, to the internalization of a set of appropriate values as a basis for governing one's behavior.

Finally, the third dimension addressed the fact that neither conscious knowledge nor values and attitudes are sufficient to explain effective performance as we observe, or indeed, as we experience it. There is a third dimension – the ability to act quickly and effectively on an almost unconscious, instinctual level. This kind of ability, or skill, is incorporated in a psychomotor taxonomy (Simpson 1974). It addresses such things as the ability to carry out a specific sequence of guided activities to the ability to improvise appropriate sets of complex behavior. Exhibit 2 summarizes the three dimensions of learning.

EXHIBIT 2
HIERARCHICAL LEVELS THE THREE DIMENSIONS OF PERFORMANCE

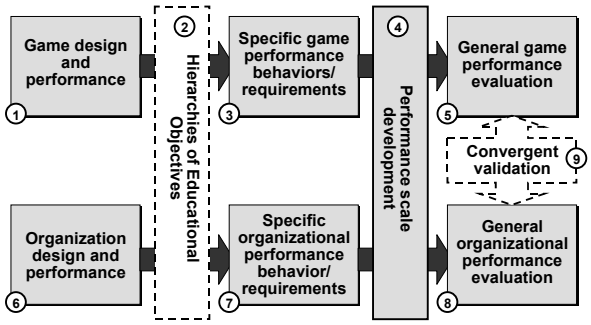
Cognitive Domain
Knowledge, or the ability to recall ideas such as facts, concepts and theories
Comprehension, or the ability to understand and make intellectual use of knowledge.
Application, or the ability to use abstract ideas in concrete situations.
Analysis, or the ability to break ideas down into their parts and logical premises.
Synthesis, or the ability to develop new ideas

from apparently unrelated parts.
Evaluation, or the ability to judge the merit of ideas for given purposes.
Affective Domain
Receiving, or the tendency to recognize and pay attention to important stimuli.
Responding, or the tendency to act in appropriate ways as a result of a stimulus.
Valuing, or the internalization of underlying motives that govern response.
Organization, or the arrangement of values into a coherent, stable system.
Characterization by a value, or the use of values to control one's behavior.
Psychomotor Domain
Perception, or the ability to sense objects, qualities and relationships via sensory organs.
Guided response, or the ability to perform a specific act under the guidance of a teacher.
Mechanism, or the ability to perform a learned act habitually without guidance.
Complex overt response, or the ability to perform a complex pattern of acts.
Adaptation, or the ability to alter an act to meet the demands of anew situation.
Origination, or the ability to develop new acts through the application of unrelated skills.

These three hierarchies provide a useful starting place for developing performance criteria. While they are too abstract for direct application, they can serve as a template for identifying the types of performance through which simulation games might be evaluated. Exhibit 3 illustrates the process.

The Exhibit suggests that the process of developing assessment criteria should proceed in parallel, building from both game performance and actual job performance. That is, evaluators would consider both the design of games and the design of the organizations the games are intended to simulate (boxes 1 and 6). In each case, the results of the analysis would be filtered through the lens of educational objectives (box 2). For instance, an evaluator might identify performance behaviors and requirements involved in sales forecasting by asking how forecasting relates to a set of questions growing out of Exhibit 2. What *knowledge* do students or workers need? What do they need to *comprehend*? To what extent do they have to *apply* general forecasting models to the specific requirements of a particular forecasting problem? To what extent must they *analyze* potential information sources to extract relevant data? Must they *synthesize* of potential information sources and forecasting principles to develop better approaches? To what extent must they *evaluate* how well the approaches are likely to work?

EXHIBIT 3
A MODEL FOR DEVELOPING PERFORMANCE ASSESSMENT CRITERIA



In similar fashion, how do the questions rising out of the *Affective* and *Psychomotor* domains translate into performance behaviors and requirements? What type of stimuli must they learn to notice? What kind of value system must they ultimately develop and incorporate into their application of the forecasting process? What kinds of patterns must they be able to recognize in the data? What kinds of complex behaviors must they learn to perform in order to carry out the forecasting process? To what extent must they develop new ones?

The result of this process is a pool of performance items that represents the ability to carry out an effective sales forecast, and by extension, any other skill or activity the simulation is designed to teach and evaluate. Given a pool of

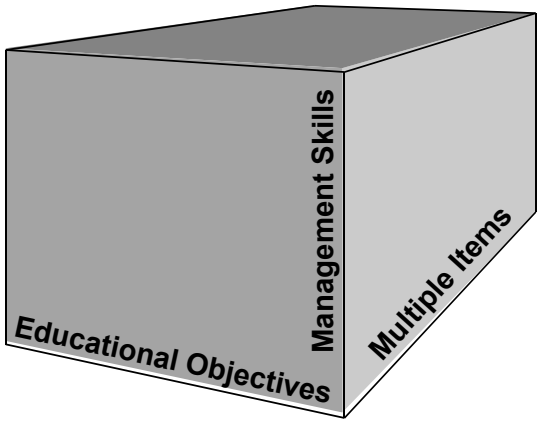
items, researchers can use standard techniques of psychometric scale development to develop multi-item performance indices (box 4).

Given these indices as tools, researchers can evaluate targeted performance in both a simulation and actual organizational performance environment (boxes 5 and 8). They can then compare these to establish a measure of convergent validity (box 9).

CASTING THE PROCESS IN AN ON-GOING PROGRAM OF RESEARCH

Note that the process described in Exhibit 3 represents a research program of enormous proportions. A single aspect of the game playing process – forecasting, for instance – can constitute a major research study. It involves the development of a host of items, several representing each of the various aspects of the forecasting process. And how many skills are there along the order of forecasting?

EXHIBIT 4
THE THREE FACES OF PERFORMANCE-ITEM DEVELOPMENT



Ultimately, the research program must proceed not only from a taxonomy of educational objectives, but also from an overall taxonomy of the behaviors and requirements that constitute effective management in the kind of organization the simulation is trying to address. These, in turn, should all be represented by multiple items to ensure scale reliability, as suggested in Exhibit 4.

One of the advantages in pursuing this kind of research in a simulation game environment is that gamers ply their trade by identifying the critical elements of organizational performance that need to be simulated. In many ways, the ultimate test of one's understanding of management processes is one's ability to simulate them.

SUMMARY AND CONCLUSIONS

This paper is an outgrowth of an assignment made by ABSEL's Committee on Assessment during the 1998 annual conference in Hawaii. Its purpose is to extend ABSEL's program of research regarding the use of simulation games as instruments of assessment. While the program has been more problematic than originally anticipated, the research has continued. The Committee has retrenched from a premature call for immediate registration of games to a more systematic, theory-based approach (Anderson, Cannon, Malik and Thavikulwat 1998).

This paper addresses the seemingly overwhelming problem of identifying measurable criteria by which gaming experience can be evaluated against the demands of actual management situations. Even so, the task is daunting. It promises to involve an enormous amount of painstaking research. Nevertheless, by developing a specific, theory-based program, the research can progress at a reasonable pace. The program itself would proceed as follows:

1. The process would begin for gamers, as it would for management system analysts, with the identification of the specific behaviors and requirements that constitute performance success.

2. In order to structure this search, and anchor it in an underlying theoretical framework, gamers and system analysts would view their be-

haviors and requirements through the lens of established taxonomies of educational objectives. Consistent with the broader view of education that characterizes the ABSEL approach, these taxonomies would include *affective* and *p sychomotor* as well as *cognitive* objectives.

3. The output of the search and conceptualization process would be a set of performance items which will provide the basis for developing specific performance scales. These would be developed by using standard psychometric scale development techniques.

4. Once developed and validated, the performance scales provide the missing link in the game validation process. They provide a common instrument for comparing performance in actual versus simulated decision environments.

This paper provides the basic framework. However, the real work is yet to come. We anticipate that the next step in the research program will be to begin developing specific validation studies. These would provide actual experience with the practical problems researchers are likely to encounter in game validation.

REFERENCES

Anderson, P., Cannon H. M., Malik D. & Precha Thavikulwat (1998). Games as Instruments of Assessment: A Framework for Evaluation. *Developments in Business Simulation and Experiential Learning*, 25, 31-39.

Bloom, B. S., N. D. Englehart, E. J. Furst, W. H. Hill, & D. R. Krathwohl (1956). *Taxonomy of Educational Objectives -- The Classification of Educational Goals, Handbook I: Cognitive Domain* (New York: David McKay Company).

Burns, A. C., Gentry J. W., and Wolfe J. (1990). A Cornucopia of Considerations in Evaluating the Effectiveness of Experiential Pedagogies, *in Guide to Business Gaming and Experiential Learning*, J.W. Gentry (ed.), New York: Nichols/GP Publishing, 253-278.

Cannon, H. M., J. R. Frazer, J. Goff, D. R. Markovich, R. Stevens, & P. Thavikulwat (1996). Draft standards and registration procedure for assessment instruments. *Developments in Business Simulation & Experiential Exercises*, *23*, 94.

Keys, B. & Wolfe J. (1990). The Role of Management Games and Simulations in Education and Research. *Journal of Management*, 16:1 311.

Krathwohl, D. R., Bloom B. S., & Masia B. B. (1964). *Taxonomy of Educational Objectives -- The Classification of Educational Goals, Handbook II: Affective Domain* (New York: David McKay Company).

Simpson, E. J. (1974)). The Classification of Educational Objectives in the Psychomotor Domain, in Robert J. Kibler, Donald J. Cegala, Larry L. Barker, and David T. Miles. *Objectives for Instruction and Evaluation* (Boston: Allyn and Bacon) 107-112.

Standards and registration procedure for assessment instruments. (1996). *ABSEL News & Views, 16* (Spring/Summer), 7.

Stanislaw, H. (1986). Tests of computer simulation validity. *Simulation & Games, 17*, 173-191.

Thavikulwat, P. (1995). Report from the program chair. *ABSEL News & Views, 15* (Spring/Summer), 1.