Developments In Business Simulation Experiential Exercises, Volume 22, 1995
SIMULATION PERFORMANCE, LEARNING, AND STRUGGLE

John Washbush & Jerry Gosenpud
University of Wisconsin-Whitewater

ABSTRACT

The purpose of this study was to continue a line of research intended to evaluate learning and simulation performance. Using instructor designed instrumentation, learning scores were developed for three classes of undergraduate business-policy course students participating in simulations. One-way analysis of variance was performed on learning scores using top-middle-bottom simulation performance classifications. Consistent with previous findings by the researchers, analysis of these data found that simulation performance and learning did not co-vary. Using a series of questionnaires, the researchers also evaluated the concept that learning and the struggle to perform well in the simulation might be positively related. While this study did not find definitive proof of such an association, results indicated that this research should continue.

GENERAL

In recent research (Washbush and Gosenpud, 1993, 1994; Gosenpud and Washbush, 1993), we have argued that students' learning and their performances in total enterprise (TE) simulations do not co-vary. Anderson and Lawton (1992) have documented that most teachers who use simulations do grade on performance. While intuition suggest that this is reasonable, we believe that this intuition may be seriously misplaced. We have consistently found no linear relationship between learning and competitive performance, and, in some cases, poorer performers learned more than better performers.

The purpose of this paper is to continue the examination of the relationship between student learning and total enterprise (TE) simulation performance. Earlier papers (Washbush and Gosenpud, 1993, 1994), in addition to arguing that learning and simulation performance do not positively co-vary, have suggested that learning may be a function of the extent to which students struggle to improve performance standing or reverse negative performance trends. For the present study, we continued the use of a set of researcher-prepared pen and paper examinations which test knowledge, skills, and abilities relevant to the simulation environment. From these tests we derived learning performance measures which we compared with simulation performance. In addition, we developed and administered a series of four surveys which were designed to identify how and the extent to which students struggled during the course of the semester. The survey items were scored and evaluated against learning scores. The findings reported here continue to support our contention that, while learning occurs over the period of simulation play, learning and simulation performance are not positively related. However, the results of analysis of the struggle surveys did not definitively identify a positive linkage between struggle and learning.

BACKGROUND

There is extensive literature on student learning in the simulation environment, and studies have had widely divergent purposes with equally divergent results. For example, Greenlaw and Wyman (1973) have questioned the validity of simulations for teaching purposes, and Thorngate and Carroll (1987) have found that luck varies positively with the number of contestants.

Whiteley and Faria (1989) have concluded that simulation games are effective in improving quantitative skills but are not effective in improving the acquisition of applied or theoretical knowledge. Wellington and Faria (1991) have found no relationship between examination performance and simulation performance. They suggested that simulation play involves skills that may not be directly measurable by normal multiple-choice exams and further suggested that the pedagogical value of simulations should be focused on the development and acquisition of decision-making and interpersonal-communication skills as opposed to the acquisition of business principles and knowledge. On the other hand, Keys and Wolfe (1990) have concluded that games are internally valid for use in strategic management courses. Additionally, Wolfe and Roberts (1993) have argued that simulation games have external validity in predicting future career success of players.

The implication of these results is that performance is not a pure reflection of learning that other factors including luck are important determinants of performance, and that simulation performance may not be a valid reflection of learning.

Hypotheses

The formal hypotheses of this study were:

(1) Students who demonstrate superior performance in the simulation exhibit superior learning.

(2) Students who struggle to improve over the course of the simulation demonstrate superior learning.

METHOD

Subjects, Design and Procedure

The subjects of this study were students enrolled in the required under-graduate Administrative Policy course at the University of Wisconsin- Whitewater. We studied 105 students in three sections (n = 38, 34, 33) during the fall, 1993 semester. The first two sections were taught by the senior author and the third section was taught by the other author. The simulation used was Micromatic (Scott & Strickland, 1985). Each section contained one Micromatic industry and each was exposed to similar, but not identical, market characteristics. In all industries, play encompassed 12 quarters plus a practice round. All industries were identical with respect to decision factor weights and evaluation criteria. Game performance was 20% of the course grade, and 5% of the course grade reflected peer ratings of individual contribution. Play was on a team basis throughout each semester. Teams ranged in size from 2 to 4, but the majority contained 3 members.

Learning

To measure learning, the researchers continued the use of two parallel forms of a multiple-choice and short-essay examination. The validity of these instruments is grounded in their content. The items were selected based on both the decision-making environment of Micromatic and the subject matter and analytic requirements of the policy course. The test developers used a

common scoring key for all questions to ensure uniformity of measurement. Statistical reliability estimates for the instruments range from .65 to .7. The examinations contained questions and situations routinely confronted by companies competing in Micromatic. These included manipulating and analyzing the marketing-mix, making operating decisions, determining costs, understanding financial and cashflow analyses, and understanding the relationship between capacity and marketing expenses. The questions tapped analytical, synthesis, and application skills of the Bloom Taxonomy (Bloom, 1956). We administered Form 1 as a pre-test at the beginning of the semester. Students completed Form 2 at the end of the semester (as an announced final examination which counted 10% of course grade). Learning over the period of play was defined as the difference in percentage score for Form 2 minus percentage score for Form 1 (Learning = %2 - %1). Therefore, a positive learning score indicated learning improvement.

Struggle

Previous research has caused us to suspect that learning may be related to the extent to which students struggled with game performance and put forth effort to improve game understanding, performance standing, or to reverse poor results. Accordingly, we developed a set of four self-report questionnaires designed to identify the presence of struggling. Questionnaires were administered as follows: #1 at about the third week of play; #2 at about the mid-point of play; #3 at about the ninth week of play; #4 at the end of play. In general, survey forms contained mixtures of open-ended and discrete responses, although each survey was, to at least some extent, unique. Typical information solicited included:

-Extent to which game was easy or difficult
-Extent to which the team was struggling or not
-Amount of time spent making decisions
-Degree of confidence
-Degree of confusion
-Amount of effort spent
-Time, effort, and struggle in relation to performance

Responses were marked according to predefined or interpretive scaling depending on item nature and structure.

ANALYSIS OF LEARNING AND PERFORMANCE

Our first hypothesis proposed that students who demonstrated superior performance in the simulation would exhibit superior learning. To establish a basis for analyses, at the end of each round of play we determined quarterly and cumulative simulation performance using the normalized scoring routine contained in the Micromatic software. Scoring factors used were after-tax earnings (40%), return-on-sales (30%), and return-on-assets (30%). We deliberately did not use leverage-manipulable factors (return-on-equity, earnings-per-share) to calculate performance.

Performance Categories and Analysis Results

At games end we categorized each company's relative performance within its industry by defining whether the team performed in the Top, Middle or Bottom of the performance scores in the respective industry. For analytic purposes, data were pooled for all industries. Table 1, displays the results of a one-way analysis of variance (Oneway ANOVA) of learning scores for individuals whose teams finished in the top, middle, and bottom of their industries. This table shows a mild positive relationship between learning and performance, but the differences are far from statistically significant. As in our previous research, these results led us to reject the first hypothesis-top simulation performers did not achieve significantly higher learning scores.

TABLE 1
ONEWAY ANOVA
LEARNING VS SIMULATION STANDING
RANKING BY TOP, MIDDLE, BOTTOM

SOURCE	DF	SS	MS	F	p
Finishing level	2	281	140	1.14	.324
Error	102	12577	123		
Total	104	12858			

LEVEL	N	MEAN	ST DEV
1 (Top)	37	14.24	12.83
2(middle)	32	13.97	8.47
3 (bottom)	36	10.68	11.25

It should be noted that our research has consistently demonstrated that players have learned, and they learned what the game intended to teach-namely, financial statement and data analysis, understanding the consequences of decisions, and applying that understanding to subsequent decision-making situations.

ANALYSIS OF THE STRUGGLE TO IMPROVE

Our second hypothesis proposed that students who struggled to improve over the course of the simulation experience would acquire superior learning. Learning scores for all subjects were categorized into quartiles. Data obtained from each of the items from the four questionnaires were compiled and analyzed against the quartile classifications using Chi-Square statistical procedures.

Significant Chi-Squares and Their Interpretation

Analyzing questionnaire items against learning quartiles produced only two statistically significant Chi-Squares.

1. After several periods of play, (first questionnaire) subjects who were least confident that they would do well in the game achieved learning scores in the lowest quartile. Those who were most confident achieved learning scores in the two middle quartiles (essentially average learning). However, those who were second lowest in confidence achieved top quartile learning scores (Chi-Square = 27.01 37, p = .02862, df = 15).

2. At the mid-point of play, subjects whose learning scores were in the bottom three quartiles indicated that they had not reduced game effort out of a sense that past efforts seemed fruitless. Curiously, those who achieved top quartile learning scores indicated that they had reduced effort slightly because of a perception that past efforts were fruitless (Chi-Square = 18.50423, p = .02975, df= 9).

Additional Analyses of Item Responses

A second set of analyses was conducted on the response data. Items were assumed to be a series of continuous-measure Likert scales. Each item was analyzed using Oneway ANOVA on item scores between those finishing in each of the learning score quartiles (defined above). While there were no statistically significant F's, three items approached significance (p>.05<.l).

1. Those who were least confident that they would do well at the third week of play achieved learning scores in the lowest quartile. However, The second least confident people were those who achieved highest quartile scores (F = 2.57, p = .059).

2. Similarly, those who found the game most difficult at the mid-point of play achieved learning scores in the lowest quartile, and those who were just behind in the perception of difficulty achieved highest quartile scores (F = 2.71, p = .053).

3. Those who, at mid-game, reported that they were putting forth slightly less effort, because they perceived past efforts to have been fruitless, achieved highest quartile learning scores (F = 2.21, p=.092).

The results of these analyses of learning and questionnaire responses failed to indicate any clear relationship between struggle in playing the simulation and learning performance. Accordingly, the second hypothesis was not accepted.

DISCUSSION

This study continued research efforts intended to evaluate the relationship between learning and performance in the total enterprise simulation. We found that there was no direct, positive linear relationship between the two variables. Those who performed best did not learn the most; those who performed the worst did not necessarily learn the least. On the other hand, here, as in our previous related studies (Washbush & Gosenpud, 1993, 1994), students did learn. The fact that learning occurred supports the validity of the simulation as learning experience.

Additionally, in those two previous studies, we found evidence that suggested to us that some teams appeared to surrender and that surrender seemed to correlate negatively with learning. The phenomenon's repetition across those two studies suggested to us the presence of a complex set of variables that could exert a strong negative influence on some teams. We continue to believe that simulation team effort should be stronger when team members collectively believe that they have something to gain, intellectually or emotionally, by purposeful striving (i.e., struggle to compete, learn, or improve) to perform and that such efforts should enable players to better understand techniques of analysis and decision making, and positive learning should result.

Unfortunately, in the present study, the analytic results obtained from responses to a series of questionnaires, specifically designed to solicit information on the nature and role of struggle, have not provided us with concrete evidence that our prior interpretations are correct.

Despite the detailed nature of the questionnaires and their sequential administration over the period of play, we suspect that there may exist a number of methodological flaws:

1. Many students doubtless perceived that filling out the questionnaires during class time was a nuisance and a requirement to be dispensed with by exerting minimal effort.

2. Although multiple questionnaires were administered, each covered at least several weeks of decisions. Student attitudes and perceptions were likely to have changed from week to week, and responses may not have adequately captured these changes.

3. Data analysis of struggle data in the present study suggest, however weakly, that some people may find the game challenging, but are not able to translate that into learning, while others find the game challenging and are able to respond in ways that may positively influence learning. Perhaps there exists a set of intervening variables, which interact with struggle in positive ways. However, this study did not incorporate a level of sophistication to permit us to identify such variables.

4. Data analysis in the present study suggested the existence of a possible contaminating effect from using the post-test as a graded examination. Students on teams who performed in the lowest learning score quartile were generally on teams that were improving over the last 4-5 periods of play, while those scoring in the higher three quartiles were on teams that were generally exhibiting static or declining performance over that period. This suggests that post-test performance may have reflected grade compensation effort by those who feared that their simulation grade was in jeopardy. Perhaps, therefore, the post-test should not be graded.

The consistency of results over three separate studies, having been conducted over eight industries, using two instructors, and over three different periods of time, continues to send the message that learning and simulation performance do not positively co-vary. Therefore, we again caution instructors who desire to grade solely on learning to not incorporate simulation performance standing in final grades.

In this paper, we have attempted to define and examine the conditions under which supportive struggle occurs and is sustained, but the present results are, at best, disappointing. We believe, however, that our studies support continued efforts to evaluate the possible existence and nature of this struggle as a type of learning-related behavior. It is clear that the matter of the potential positive implications of struggle is obscured by great complexity.

REFERENCES

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

Bloom, B. (ed.) (1956). The Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay.

Gosenpud, J. & Washbush, J. (1993) Comparing the simulation with the case approach: again! But this time using criteria appropriate for the simulation. Developments in Business Simulation & Experiential Exercises, 20, 126.

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

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

Scott, T.W. & Strickland A.J. (1985) Micromatic: A Management Simulation. Boston: Houghton Muffin.

Thorngate W. & Carroll, B. (1987) Why the best person rarely wins Simulation & Games, 18, 299-320

Washbush, J. & Gosenpud, J. (1993) The relationship between total enterprise simulation performance and learning Developments in Business Simulation & Experiential Exercises, 20, 141.

Washbush, J. & Gosenpud, J. (1994) Simulation performance and learning revisited. Developments in Business Simulation & Experiential Exercises, 21, 83-86

Wellington, W.J. & Faria, A.J. (1991) An investigation of the relationship between simulation play, performance level and recency of play on exam scores. Developments in Business Simulation & Experiential Exercises, 18, 111-115

Whiteley, T.R. & Faria, A.J. (1989) A study of the relationship between student final exam performance and simulation game participation. Simulation & Games, 20, 44-64

Wolfe, J. & Roberts, C.R. (1993) A further study of the external validity of business games: five-year peer group indicators. Simulation & Gaming, 24, 21-33