TOTAL ENTERPRISE SIMULATIONS AND OPTIMIZING THE DECISION SET: ASSESSING STUDENT LEARNING ACROSS DECISION PERIODS

E. Brian Peach, The University of West Florida
Richard G. Platt, The University of West Florida

ABSTRACT

The paper reviews the claim that the marginal value of learning ends after four to six decisions in a complex simulation by evaluating the level of student learning as a function of the number of decisions accomplished. Concern has been expressed over the lack of a relationship between learning and simulation performance, and a proxy for learning assessment is proposed and used to test learning. It was found that the first four to six decisions were used to learn how to play the game, and that desired conceptual learning occurred after this period. It was concluded that a minimum of 10 decisions was necessary to allow the student to apply concepts learned in class and to reinforce to the student the value of learning achieved through the simulation.

INTRODUCTION

Numerous studies attest to the widespread use of Total Enterprise Simulations in both the corporate world and business schools. The rise of corporate universities and the increasing sophistication of simulations available for PCs makes them attractive for use in the corporate environment. Faria (1989) estimated that more than 5,000 U.S. companies use business simulations in corporate training and development programs. Use of simulations is also widespread in Europe and other parts of the world (Joldersma & Guerts, 1998). In addition to the corporate world, simulations have become endemic to business schools. In a 1987 review, Faria estimated a relatively high number of AACSB schools to be using simulations (Faria, 1987); ten years later he increased the estimate from 95% to 97% (Faria & Nulsen, 1996).

The use of simulations does not come without cost. Rollier (1992) estimated it cost his corporation $14,000 per executive per course in executive wages, and emphasized that simulations must bring value to justify adding to this cost. Given the popularity of simulations, and the cost they add in terms of both dollars to the corporate world and complexity and workload in academia, it is reasonable that researchers should evaluate how many decision period cycles are required in a simulation to achieve the desired levels of learning.

Rollier concluded after observing a number of simulations that marginal learning ceased at about the fourth decision, and uncertainty was gone by the sixth (1992, p447). He cites a number of references to support his conclusion (Wolf, 1985 p277; Hogarth & Makridakis, 1981, p102; and Cangelosi & Dill, 1965).

OPTIMIZING THE DECISION SET

It is clear that each additional decision period required of simulation participants adds to the overall cost both for participants and for simulation administrators. Thus the decision period set would be optimized where the marginal learning benefit no longer exceeds the marginal cost of an additional decision. Although this is simple in theory, operationalizing a test encounters a number of difficulties. Primarily these would be identifying the Acost@ of an additional decision period and assessing the marginal learning benefit.

In the corporate world, the >cost= of an additional decision period could likely be formulated based on lost wages, instructor and administrator costs, computer time, etc. In academia, it is more qualitative as the cost essentially consists of requiring students to do an additional decision. Costs here are not incurred so much in dollars as they are in the time spent by students and game administrators. If the class instructor is also the game administrator, additional decisions are a very real >cost= in terms of workload as well as the psychic energy expended in counseling students and managing the game.

This conundrum is likely unsolvable in an absolute sense, but can be addressed in a qualitative sense. Typically, both for corporate as well as business school simulations, the game is integrated into a course with the purpose of enriching learning, or providing a vehicle for fulfilling mastery of a set of learning objectives. The instructor has certain learning goals for the course and the simulation's role is to support this learning. Thus for the instructor, the >cost= of an additional decision period is very similar to the >cost= of an additional homework exercise, quiz, report or project. The instructor knows that another repetition of an assignment will likely enhance learning, but there is the cost of the student doing the assignment and the instructor grading it. The instructor thus assigns sufficient quizzes, tests, assignments, etc., to reasonably assure that the student population achieves the desired level of learning. Through time and experience, the instructor ascertains the appropriate number and type of each pedagogical instrument such that the end performance meets the desired standard. Through extrapolation, the procedure for establishing the optimum number of decisions in a simulation is thus straightforward - additional decisions are required until the desired level of learning is achieved.

ASSESSING LEARNING

How to assess learning for any venue has been a research topic probably for millennia. Certainly the topic of assessing learning from simulations has been of consistent interest by simulation users and developers (Gentry, 1990). Most Total Enterprise (TE) simulations provide some performance measure that game administrators typically use as a measure of (or proxy for) learning. However, recent research on TE simulations questions whether learning and simulation performance are related (Gosenpud & Washbush, 1993; Washbush & Gosenpud, 1993, 1994, 1995).

Academics generally view the overall grading process as a seriously flawed measure of student learning, but historical precedents, administrative requirements plus the fact no alternative system has gained ascendancy mandate their continued use. General dissatisfaction with grading procedures can thus be compounded with simulations. Because teachers using TE simulations generally use performance measures provided by the TE simulations as the basis for grading (Anderson & Lawton, 1992), students who believe they learned a substantial amount from the simulation may perceive a 'grade' based on simulation performance measures as unfair or irrelevant.

If a lack of correlation between simulation-provided performance measures and actual learning was definitively established, it would require that conscientious instructors develop alternative measures. It has been recommended that to assess learning researchers must develop supplemental tests (e.g., Washbush & Gosenpud, 1993, 1994, 1995). Unfortunately, the nature and type of acceptable alternatives remain subjects of debate. Gentry, Stoltman, and Mehloff (1992) suggest instructors develop measures of what they are trying to teach. Gosenpud and Washbush (1994), citing Anderson and Lawton's (1992) survey results, state it is implicit in what instructors measure as to what they expect students to learn. But if, as Anderson and Lawton (1992) report, 93% of instructors grade solely on performance, either those instructors must believe overall performance is an adequate proxy for their desired learning outcomes, or they cannot or will not define their desired learning outcomes. Assuming 'winning is everything' is not the sole desired learning outcome for 93% of instructors, devising additional measures of desired learning outcomes should continue to be a goal for simulation users.

Gosenpud and Washbush (1994) acknowledge that learning measures, such as exams, may focus students' attention too narrowly. However, they conclude that because it is not known from overall simulation performance alone whether students are learning, instructors should define the types of learning that are most valuable and assess them. Gosenpud & Washbush (1994) polled instructors to ascertain what users think players should be learning. Strategic decision making ranked highest, but cash, inventory and production management also ranked high.

Responding to the recommendations of Gentry, Stoltman, and Mehloff (1992) and Washbush and Gosen (1994) that instructors develop measures of what they are trying to teach, one such approach to assessing learning for the BSG was proposed by Peach (1996).

DEVELOPMENT OF LEARNING OBJECTIVES FOR THE BSG TE SIMULATION

The TE simulation used for this research was the Business Strategy Game (BSG) (Thompson & Stappenbeck, 1999b, 1998, 1997, 1995), which uses shoes as the product. For a more complete description of the game, see Snyder (1995). A number of variables are used in making decision inputs. These include production, pricing, marketing, and financial decisions. As a global simulation, decisions also deal with tariffs, exchange rates, and foreign markets. The simulation provides the team with projected income and balance sheet data, and a complete projection of typical performance data including inventories, sales, profitability, etc. Data are provided on a geographic basis as well as in corporate summaries.

There are six performance measures provided by the simulation: revenues, profits, return on equity, firm value, bond rating, and a game specific measure called strategy rating. This measure provides the players with the Apower of each company's strategy and the distinctiveness of the resulting market position@ (Thompson & Stappenbeck, 1999a, p76). The scores for each of these measures are summed and the total scores are used to rank teams.

In this paper, the authors are proposing that it is possible to use performance measures provided by the simulation to assess whether learning is taking place. The total scores provided by the TE simulation were one of the variables used in the analysis to determine levels of learning because the individual measures do represent variables that are not only of interest to managers, but capture learning about critical areas of the course objectives. The authors have found that actual manipulation of ROE through debt and equity decisions is the single most powerful learning tool for students to internalize understanding of such concepts as the Dupont formula and the equity multiplier.

The second variable was devised to capture students' ability to predict environmental responses to their decisions. One of the fundamental premises of strategic management is that organizations act as open systems in the environment. Thus, the challenge for managers is to position the firm within the environment such that it can pursue its objectives consistent with organizational capabilities and in response to environmental demands (Ansoff, 1990, pp242-232). This was operationalized as the ability to accurately predict sales, or more specifically, the ability to anticipate the effects of various marketing decisions by accurately predicting shoe sales volume.

METHODOLOGY

The Sample

The sample consisted of senior business students enrolled in the capstone business policy course at an AACSB accredited regional university in the southeast United States. The students had completed all other business core courses with a minimum grade of C in every course, and their demographic profile was typical of the university's College of Business demographic profile. Each semester, students form teams and are grouped into one or two industries depending on the number of teams as the simulation software limits an industry to sixteen teams. Each student is part of a team for the purposes of playing the BSG. Company decisions are the result of group decision making. Group decision making allows the team members to synergistically share the knowledge gained from previous courses and life experiences.

Data Collection

For any industry, the simulation consisted of a number of decision periods, typically scheduled to be twelve. For some industries, computer problems, hurricanes and other events limited the total to a lower number. Most industries had eleven or twelve decision periods; the lowest number of decision periods was eight. For each decision period, each team submitted decisions about how its company was to operate during the next simulation iteration (e.g., simulated year of operations). One of the values submitted for each decision period was the predicted sales volume in pairs of shoes.

Data were available for thirty-three industries over a period of nine years. Over this period, the simulation was released in a series of new versions, and a significant change in the game limited usable industries to the last sixteen industries. For each industry, the game administrator recorded decisions and performance results. All classes were taught by the same instructor who also acted as the game administrator.

VARIABLES

Team performance is a value calculated each decision period based on the total score provided by the simulation. This total score is the sum of six individual factors under the control of the players. The algorithm used by the simulation scores teams on each factor relative to the highest score for any team on that factor and gives 100% of the factor score to the highest team for that factor. Other teams then get scores proportional to performance relative to the highest team. Because it is unusual for a team to be the top at all six factors, the highest total score is typically less than the 100 points possible. The objective, therefore, is to achieve a total score as high as possible relative to the highest performer. To allow comparisons of team relative performance across industries, a difference score was calculated for each team relative to the highest scoring team in that industry for each decision, and for each decision period a mean difference score was calculated. A higher mean difference score was interpreted as the industry doing less well relative to the leader. A decreasing mean difference score was interpreted to mean that the industry as a whole was doing better (learning) relative to whichever team was the best at that point.

As a second measure of learning, this project used the ability of teams to accurately predict its sales. The accuracy of demand forecasted for their product was expressed as a percentage; and was calculated using actual demand divided by projected demand. One hundred percent would indicate perfect accuracy, but because actual demand could be either greater than or less than projected demand, these values could be greater or less than one hundred percent. Numbers closer to one hundred percent were interpreted as indicating more learning.

PROPOSITIONS

Rollier (1992) contends learning effectively occurs in the first four decision periods. If this is true, then there should be a difference in our measures for decision periods one and four. Rejection of these hypotheses would be support for Rollier.

- **H_{1A}:** *There will be no difference in the mean difference scores for decision periods one and four.*
- **H_{1B}:** *There will be no difference in the means of the team decision accuracy percentages for decision periods one and four.*

If learning stops at decision period four, then the measures of learning for decision period four and decision period eight should be the same.

- $ **H_{2A}:** *There will be no difference in the mean difference scores for decision periods four and eight.*

- $ **H_{2B}:** *There will be no difference in the means of the team decision accuracy percentages for decision periods four and eight.*

And finally, if the learning stopped with decision period four, then the measures of learning for decision periods eight and twelve should be the same.

- $ **H_{3A}:** *There will be no difference in the means of the difference scores for decision periods eight and twelve.*

- $ **H_{3B}:** *There will be no difference in the means of the team decision accuracy percentages for decision periods eight and twelve.*

RESULTS

Means of the difference scores were computed by decision period across the teams for all terms (e.g., simulation industry).

Table 1 - Difference Scores

Decision Period	t-value	Null Hypothesis
1 - 4	-.078	Accept
4-8	-4.50*	Reject
4 – 12	-3.07*	Reject
8-12	0.65	Accept
1-12	-3.57*	Reject

* = p=.01

Table 2 - Sales Forecast Accuracy

Decision Period	t-value	Null Hypothesis
1 - 4	9.63*	Reject
4-8	-2.43	Accept
4 – 12	-5.07*	Reject
8-12	-3.21*	Reject
1-12	1.96	Accept

Values for Student's t-test for the mean difference scores are shown in Table 1 and t-values for Student's t-test for means of sales forecast accuracy are shown in Table 2. H_{1A} was not supported but H_{1B} was supported ($p = .01$). H_{2A} was not supported but H_{2B} was supported ($p = .01$). H_{3A} was not supported but H_{3B} was supported ($p = .01$).

DISCUSSION

At first glance, the results seem mixed and inconclusive. However, they actually indicate a clear pattern of learning that extends beyond the fourth period. The contention that learning does not extend past four decision periods was based on the stability of performance during the first four periods. This contention is confirmed by the failure to reject H_{1A}, as the mean performance scores for decision period one and decision period four were not statistically different. The correct interpretation of this is that during the first couple of periods, players are learning the rules of the game and are basically imitating each other's strategies for lack of a better guide. This leads to similarity of strategies and thus similarity of performance. Although there are apparent differences between teams by the fourth decision, as an industry, the differences in strategy have not had time to reflect themselves in significant differences in the mean difference score. At the same time, the failure to accept H_{1B} indicates players are beginning to experiment with strategies but have not gained sufficient understanding of their environment to accurately predict sales.

The failure to accept H_{2A} indicates that significant differences in industry performance occur after the fourth decision period (t=-4.5). In addition, there is significance between decision period four and decision period 12 as well (t=-3.07). Decision periods 1-4 are relatively peaceful as players learn the rules of the simulation. However, no real achievement of the course learning goals occurs during this period. In fact, if anything students despair at the notion that planning has any positive effect because what they perceived as planning was leading to disaster. The first six decision periods consistently indicate a combination of seat-of-the-pants decisions by some teams and partially flawed application of course principles by others leading to severe flux as the industry is beset with many conflicting strategies. These strategies are typically less predicated on environmental realities (overcapacity, declining demand, incipient tariffs, exchange rate fluctuations) than the students' preconceived notions if not pure hopefulness. As declining performance scores raise the specter of failing the course, students are forced to resort to a more thoughtful and consistent use of the strategic analysis tools provided in the course. Long range planning, assessment of competitor capabilities and demonstrated intentions, and financial management all become useful tools.

By Decision period 8, teams have stabilized their performance relative to the lead team, and there is no significant difference in the industry means (t=0.65). The sales forecast accuracy, however, shows significant improvements in accuracy between decision periods four and twelve (t= -5.07) as well as between decision periods eight and twelve (p=-3.21). Thus, students continue to improve in their ability to assess their environment and better sales forecast accuracy confirms to them their improved skills. But the industry has now reached a point where the skill improvement is across all teams and although individual performance metrics (sales forecast accuracy) improve significantly, mean industry performance does not (t=0.65).

CONCLUSION

Policy (Strategic Management) courses typically use complex TE simulations that require the proper application of a variety of analytical tools, techniques, and models for teams to compete effectively. The first few decision periods, rather than accomplishing all desired learning tasks, set the stage for learning as students create and then (typically) resolve a turbulent environment that demands rigorous application of relevant tools and techniques. Although marginal learning becomes insignificant subsequent to decision period 8, the next few decision periods are critical in affirming to students that the tools they are using work.

Ending the simulation at decision period seven or eight leaves the student unconvinced that recent success was the result of planning rather than random luck. After six or seven decision periods that seemingly were unrelated to resulting performance, it takes more than one effective decision to cement belief in the strategic process. Ten to twelve decision periods are optimal for achieving learning and then reinforcing belief and acceptance. TE simulations are effective pedagogical tools when properly utilized, and ensuring a sufficient number of decisions is a critical step in effective utilization.

REFERENCES

Available upon request