

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

PERFORMANCE ON A TE SIMULATION: WHAT DOES IT REPRESENT?

Philip H. Anderson. University of St Thomas
Leigh Lawton. University of St Thomas

ABSTRACT

Analysis was conducted to assess the relationship between financial performance of an individually managed company in a business simulation exercise and various measures of (a) student mastery of business skills and subject matter, (b) global measures of ability, and (c) time spent on the exercise. No significant relationship was found between the financial performance and global measures of ability. Only a weak association was found between financial performance and measures of business mastery. There was also an absence of *any significant* relationship between the variables measured and performance on the group-based simulation, including a lack of a significant relationship with performance on the individual simulation. While the validity of the measures of mastery of the business discipline is debatable, the results suggest that group performance scores on a simulation may reflect neither mastery of any of the measures of the business discipline nor of the individual's mastery of the game. The results highlight the need to develop and validate instruments assessing mastery of the skills and subject matter of the business discipline and to determine the relationship between performance on a business simulation and those instruments. Until a relationship is established, dependence on financial performance in a group simulation as an indicator of an individual's mastery of either the simulation or the business discipline should be minimized.

INTRODUCTION

Assessing what students learn on a business simulation exercise has been a focus in the literature since simulation exercises were first introduced (Greenlaw and Wyman, 1973; Keys, 1976; Wolfe, 1981; Wolfe, 1985; Wolfe, 1987; Whiteley and Faria, 1989; Burns, Gentry, and Wolfe, 1990; Wolfe, 1990; Gosenpud, 1990; Anderson and Lawton, 1992a; Gosenpud and Washbush, 1993, 1994; Anderson and Lawton, 1995; Washbush and Gosenpud, 1995). Attempts to measure learning have focused on Bloom's Taxonomy of Learning (Bloom et al. 1956) which classifies learning into six levels arranged in a hierarchical order to reflect progressively higher levels of learning. They are, in ascending order, basic knowledge, comprehension, application, analysis, objective synthesis, and objective evaluation.

Early attempts to measure learning focused on students' *perceptions* of what they learned (e.g.; Schellenberger, *et al*, 1989). More recently, paper and pencil tests have been used to assess lower levels of learning on Bloom's Taxonomy (e.g.; Gosenpud and Washbush, 1994). While instructors have used a variety of methods in attempting to determine the level of mastery a student has achieved from exposure to the exercise, financial performance has remained a key-measuring tool. A survey by Anderson and Lawton (1992b) found that all respondents, without exception, used financial performance as one of the determinants, and sometimes the sole determinant, of a student's grade for the simulation exercise. At present, there are few objective measures for assessing learning at the higher levels of Bloom's Taxonomy. In the absence of these measures, financial performance has been relied on as a proxy for student learning at all levels of Bloom's Taxonomy (Anderson and Lawton, 1992b).

Is this heavy reliance on financial performance as an indicator of mastery justified? Does financial performance demonstrate the mastery of skills and subject matter of the discipline? Anderson and Lawton (1992a) found the relationship between financial performance and other measures of student mastery to be weak or non-existent. However, that study relied on a *group's* financial performance on the simulation to assess the exercise's relationship with mastery. Basing individual performance on group work can unfairly reward or penalize individual members of the group; as they are "carried" or "held back" by other members of the group, respectively (Gosenpud and Washbush, 1994).

In an attempt to eliminate group effects, Anderson and Lawton (1995) assessed the relationship between financial performance on an "individualized" version of a competitive, total enterprise (TE) simulation exercise and a set of independent variables including: performance on the group-managed version of the simulation, a peer group assessment of the subject's strategic management skills; a self-assessment of managerial skills; the grade received on a case study write-up; the grade received for class participation during the course; and overall GPA

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

level No significant relationship was found between a subject's performance on the "individualized" version of the simulation and any of the measures except the subject's business GPA. Whether the lack of associations resulted from flaws in the "Individualized" model or the inability of the independent variables to assess learning was not determined.

This paper reports the results of a study conducted to clarify what financial performance on a simulation exercise reflects. In this study, the dependent variable was performance on a simulation exercise where the competing companies were managed by individual students, rather than student groups, thereby eliminating group effects. The independent variable's included performance on the group-run simulation, some of the variables assessed by Anderson and Lawton (1995), plus others that have been identified in the literature as possible predictors of simulation performance. The goal of this study was to isolate what performance on a simulation exercise measures.

RESEARCH METHODOLOGY

The Subjects and the Course

Subjects for the study were seniors in their final year of study at a medium-sized, midwestern university. All were members of one of two sections of a business policy course taught by the same instructor that used a total enterprise simulation (described below) as a major component of the course pedagogy. The profile of the student was that of a typical, traditional college senior. All were majoring in various fields of business administration. A total of 44 students participated in the study.

The Simulation

The simulation used was *Micromatic: A Strategic Management Simulation*, 2nd ed. by Scott, *et al* (1992). *Micromatic* is a moderately complex simulation. Each decision set requires approximately 60 decisions in the areas of marketing, production, and finance. Each decision represents a three-month (one quarter) period.

The Group-based Simulation Exercise

Students were first exposed to the simulation as a member of a group. Each group was in competition with the other groups in its section of the course. Twelve simulated quarters of operation were run. The group-based simulation exercise comprised about one-half of the course in terms of

pedagogy and time, and determined 25% of the student's final grade.

The Individual-based Simulation Exercise

After completion of the group-based simulation exercise, students were required to manage a simulation company on their own. The simulation used was the same as that used by the groups. Because constraints in the software limited the number of companies in an industry to 16, students were randomly assigned to one of three industries. The number of companies in each of the three industries was 15, 15, and 14. Each industry was constructed with the same environment. That is, the demand, costs, and weightings of the marketing mix variables were identical. A total of 12 simulated quarters of operation were run. This portion of the course comprised 15% of the student's final grade.

The Simulation Performance Index (SPI)

Micromatic uses seven factors to determine current quarter and game-to-date rankings of company performance. These factors are sales revenues, net income, earnings per share, return on sales, return on assets, return on equity, and stock price. The percentage weights assigned to these factors was the same for both the group and the individually managed exercises.

The Dependent Variable

The dependent variable measured was the simulation performance index for the individual-based exercise (i.e.; the Individual SPI).

The Independent Variables

Ten independent variables were measured to check for relationships with the individual simulation performance measure. These were: (1) the simulation performance index for the group exercise (i.e.; the Group SPI), (2) an assessment of an individual's strategic management skills as judged by other members of the simulation group, (3) a self assessment of personal managerial skills, (4) the grade received on a quiz given at the end of the semester which assessed the subject's ability to interpret reports generated by the simulation, (5) the subject's business GPA, (6) the subject's overall GPA, (7) the subject's accuracy in forecasting unit sales on the individual-based exercise, (8) the self-reported average time spent by the subject making decisions on the individual-based

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

exercise, (9) the subject's performance on the Watson-Glaser Critical Thinking Appraisal test, and (10) the subject's performance on the Wesman PCT intelligence test.

There is a dearth of validated instruments for measuring mastery of the concepts and skills of the business discipline. The efficacy of existing instruments is limited by their weak validity and their frequent use of self-assessment measures. This lack of valid instruments inhibits our ability to relate simulation performance to actual managerial talents and skills, nor to the ability to succeed in a "live" business organization. While not pleased with the options available, we selected two of the better instruments from this small, weak pool -- a modified version of Stumpfs Strategic Management Skills Questionnaire and modified version of using the Managerial' Effectiveness Profile System.

A modified version of Stumpf's Strategic Management Skills Questionnaire (i.e.; the SMSQ) was used to measure an individual's strategic management skills. Each member of the group simulation exercise evaluated his/her teammates. Stumpf identified six strategic management skills: knowing the business, managing subunit rivalry, finding and overcoming problems, staying on strategy, being an entrepreneurial force, and accommodating adversity (Stumpf, 1988). Stumpf reports internal consistency estimates of reliability of .74 for his 20-item instrument. The Stumpf questionnaire was administered after the fourth and twelfth quarters of the group-based exercise. This allowed assessment of the relationship between the perceived improvement of the subject's strategic planning skills and performance on the simulation exercise.

Each student rated his or her personal managerial skills using the Managerial Effectiveness Profile System (i.e.; MEPS) developed by Human Synergetics (1983). A modified version of the MEPS (MMEPS) was designed for use with students (Anderson and Lawton, 1990). The MMEPS has 71 items that form fourteen scales such as planning effectively, managing conflict, making decisions, and demonstrating commitment. The Cronbach alpha for the MMEPS score, using a pool of 125 undergraduate students, was .95.

It is acknowledged that both the SMSQ and the MMEPS are based on perceptions, either subjective assessments made by team-mates or self-assessments. Thus, neither of these instruments may provide a particularly good measure of mastery of managerial skills. Nonetheless. In the absence of any better instruments, they were used in an effort to test for

associations between the simulation exercise and managerial aptitude, however tenuous.

The Watson-Glaser Critical Thinking Appraisal' test (i.e.; the CTA) was administered at the beginning and the end of the semester. The Wesman PCT intelligence test (i.e.; the WI) was administered at the beginning of the semester. Both of these instruments report high reliability and validity measures, making them useful' assessments of general ability. The CTA reports validity coefficients of .56 to .79 and reliability coefficients of .73 (Watson and Glaser, 1980). Facione (1986) reported a Cronbach alpha of .90 overall (and .72 to .80 on the subscales) for the CTA. The WI reports validity coefficients of .57 to .80 and reliability coefficients of .84 to .91 (Wesman, 1965).

RESEARCH HYPOTHESES

The hypotheses tested in this study were:

- H1: The Individual SPI will be positively related to the Group SPI.
- H2: The Individual' SPI will be positively related to the SMSQ score.
- H3: The Individual SPI will be positively related to the MMEPS score.
- H4: The Individual SPI will be positively related to the simulation quiz.
- H5: The Individual SPI will be positively related to accuracy of forecasting unit sales on the individual-based exercise.
- H6: The Individual SPI will be positively related to the average time spent on making decisions on the individual-based exercise.
- H7: The Individual SPI will be positively related to the business GPA.
- H8: The Individual SPI will be positively related to the overall GPA.
- H9: The Individual SPI will be positively related to Watson-Glaser Critical Thinking Appraisal' scores.
- H10: The Individual SPI will be positively related to Wesman PCT intelligence scores.

RESULTS

Measurements for two of the independent variables, the Stumpf Strategic Management Skills Questionnaire (i.e.. the SMSQ) and the Watson-Glaser Critical Thinking Appraisal questionnaire (i.e., the CTA), were taken at both the beginning and the end of the semester. Matched

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

pair t-tests were conducted to determine if scores improved between the first and second administrations of these two instruments. Table 1 shows that no significant change occurred for either of the instruments.

The reason for the lack of change is not clear. It is possible that participating in the simulation exercise may have had little effect on the subjects. It is also possible that the instruments lacked the sensitivity or validity to detect changes. While the validity of the SMSQ is questionable, extensive testing of the CTA suggests that the lack of improvement in CTA scores reflects a failure of the simulation to produce measurable change. However, given the global nature of the CTA and the multiplicity of elements which influence the score received, it is entirely possible that participation in the simulation was beneficial, but that the time duration between the two assessments was too short for a measurable change to be observed. Since no significant differences were found between the two assessment periods, the second measurement was used in the analysis of the subjects' performance on the simulation exercise.

| Table 1 Tests for Significance Between Pre and Post Measures | | | | |
|---|-------------|-----------------|-----------------|-----------------|
| <u>Inst.</u> | <u>Mean</u> | <u>St. Dev.</u> | <u>t score*</u> | <u>p-value*</u> |
| SMSQ | 0.000 | 17.40 | 0.00 | NS |
| CTA | 0.727 | 6.30 | 0.77 | NS |
| * one-tailed tests were performed | | | | |

Analysis was also conducted to test for relationships among the independent variables. This was done to determine the extent to which two variables might be measuring the same characteristic. Table 2 shows the Pearson Correlation analysis conducted to assess these possible associations. While a significant relationship was found between the SMSQ and the MMEPS and CTA, the adjusted r-squared for both was less than 6%, indicating very little of the variation in the SMSQ score is explained by its relationship with those instruments. The relationship between the CTA and the WI is also quite weak. One would expect a relationship between an individual's intelligence and his or her critical thinking skills. However, the adjusted r-square of 17.1% shows only a small portion of the variation in an individual's critical thinking skills can be explained by intelligence. Thus, it appears that each independent variable is measuring a different construct.

A Pearson Correlation analysis was conducted to test for

| Table 2 Tests for Relationships Among Questionnaires | | | |
|---|----------------|--------------|----------------|
| <u>Instrument</u> | <u>SMSQ</u> | <u>MMEPS</u> | <u>CTA</u> |
| Managerial Effectiveness (MMEPS) | 0.278 0.034 | | |
| Critical Thinking Appraisal (CTA) | 0.282 0.032 | 0.096 NS | |
| Wesman Intelligence (WI) | 0.167 NS | -0.246 NS | 0.439 0.003 |
| top number = r Bottom number = p-value for one-tailed tests (NS = $p > .05$) | | | |

significant associations between performance on the individual-based simulation and the ten independent variables. In an attempt to measure the individuals' performance at the top of their learning curve on the individual-based simulation, assessment of the performance over the last four quarters was used as a second dependent variable. Similarly, performance over the last four quarters for forecasting accuracy and for the group-based simulation were also added to the independent variables assessed. Table 3 reports the results of the analysis for all these variables.

Table 3 also shows the relationship between the Group SPI and the independent variables. The purpose was to determine whether the same significant relationships that exist between the *individual* simulation SPI and the independent variables also exist between the *group* simulation SPI and the independent variables. This side-by-side comparison of the group and individual performance indexes provides insight into what financial performance on the group-based simulation represents. That is, if a particular variable is significantly related to both the individual and group performance measures, it provides evidence that the group-based performance score is reflecting individual mastery on that variable.

The independent variables were grouped into three categories; (1) Assessments of Mastery of the Discipline. (2) Assessments of Global Measures of Ability and (3) Other Measures. Utilizing these categories provides insight into just what it is that the performance on a simulation measures. The Business GPA was included as a measure of the mastery of the discipline since it

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

indicates an intellectual understanding of business concepts. Forecasting accuracy was also included with this group. It is assumed that forecasting accuracy is a reflection of the simulation participant's understanding of the relative importance of the marketing mix variables and the competitive nature of the environment in which the company is operating.

| Table 3 Correlation Analysis - All Subjects | | | | |
|--|--------------------|-----------------|--------------------|-----------------|
| | Individual SPI | | Group SPI | |
| | Last 4 quarters | All quarters | Last 4 quarters | All quarters |
| Assessments of Mastery of the Discipline | | | | |
| 1. Group SPI (Last 4 quarters) | .012 NS | .168 NS | | |
| 2. Group SPI (All quarters) | -.003 NS | .126 NS | | |
| 3. SMSQ | .143 NS | .331 .028 | .000 NS | .000 NS |
| 4. MMEPS | .227 NS | .236 NS | -.094 NS | -.007 NS |
| 5. Quiz | .405 .001 | .461 .001 | .137 NS | .058 NS |
| 6. Business GPA | .160 NS | .380 .006 | .046 NS | -.048 NS |
| 7 Forecasting Accuracy (Last 4 quarters) | -.416 * .003 | -.496 * .001 | -.043 * NS | .066 NS |
| 8. Forecastin g Accuracy (All quarters) | -.153 * NS | -.413 * .003 | -.128 * NS | .064 NS |
| Assessments of Global Measures of Ability | | | | |
| 9. Overall GPA | .055 NS | .300 .024 | .088 NS | .015 NS |
| 10. Watson Glaser Critical Thinking Appraisal | .074 NS | .164 NS | -.010 NS | -.142 NS |
| 11. Wesman Intelligence | .093 NS | .113 NS | .002 NS | -.203 NS |
| Other Measures | | | | |
| 12. Average Time Spent | .475 .001 | .278 .034 | .040 NS | .161 NS |
| * The negative correlation indicates that the smaller the median forecasting error made by a subject, the better his or her financial performance on the simulation. | | | | |
| top number = r Bottom number = p-value for a one-tailed test (NS = p > .05) | | | | |

Individual SPI and Mastery of the Discipline

For financial performance over the last four quarters, only quiz scores and forecasting accuracy (measured by the median error of the forecast) were significantly related to individual SPI. Financial performance over the course of all quarters of the game fared a bit better. Individual SPI for the entire simulation was significantly related again to the quiz scores and forecasting accuracy, plus the Business GPA and scores on the SMSQ. The lack of a relationship between individual SPI and MMEPS and the weak relationship between SPI and SMSQ (adjusted $r^2 = 8.8\%$) may be due to the weakness of these independent variables or to the failure of individual SPI to reflect mastery. There was *no* significant relationship between the individual SPI and the group SPI, for either the last four quarters or the entire simulation. In fact, the relationship for the last four quarters is slightly *negative*.

Individual SPI and Ability

Table 3 also shows that the only global measure of ability significantly related to performance on the individual simulation was the subject's overall GPA. There was no significant association with either critical thinking skills or intelligence.

Individual SPI and Time Spent

The average time spent on the exercise showed the strongest significant relationship with performance over both the last four quarters and all quarters of the individual simulation. The more time spent making decisions on the simulation, the better the financial performance. The strength of this relationship is surprising given that the measure of time spent is probably quite unreliable. It is a simple self-reported estimate by the subjects and thus prone to mis-estimates. It is very possible that the relationship may be even stronger than that found here, since it seems likely that those subjects who spent the *least* amount of time on the simulation would be the most inclined to *over* estimate the time spent.

Group SPI

As shown in Table 3, the analysis found no significant relationship between performance on the individual-based simulation exercise and performance on the group-based simulation. This was the case regardless of whether using performance over the last four quarters or

Developments In Business Simulation & Experiential Learning, Volume 24, 1997

for all quarters. In addition, none of the independent variables were significantly related to performance on the group-based simulation exercise. High financial performance by the group managed simulation company was not related to any measures of mastery of business skills, any global measures of ability, nor to time spent on the exercise.

DISCUSSION AND CONCLUSIONS

The results of this study rather strongly suggest that general ability does not play a major role in determining financial performance on a simulation. No significant relationship was found between measures of intelligence, critical thinking skills, or overall GPA and performance on either the individual or group simulations. This lack of a significant relationship held true, except for overall GPA and performance on the individual simulation, regardless of whether assessing performance over all quarters of the exercise, or just the last four quarters.

These results support a rather weak association between financial performance on the simulation and mastery of the discipline. While the Business GPA was significantly related to individual performance for all quarters, it is not clear whether Business GPA measures knowledge and mastery of the business discipline or the dedication and effort put into one's studies. If it is the former, we would have expected to find significance with performance over the last four quarters and not just with performance over all quarters. Performance on the quiz, a measure of the ability to interpret the simulations business reports, *does* appear to reflect mastery of the discipline. So does the ability to forecast accurately.

Perhaps the most striking result of this study is the absence of *any* significant relationship between the independent variables and performance on the *group*-based simulation. This includes the lack of a significant relationship with performance on the individual simulation. While the relationship between the variables and performance on the individual simulation may be mixed, and there may be questions regarding the ability of the instruments used to measure their intended dimension, the totality of the absence of *any* significant relationships with the group-based simulation is impressive. Important questions remain unanswered. What does financial performance on a group-based simulation exercise represent? To what extent, if any, should it influence the final grade an individual receives for the course in which a simulation is used? What does a grade

for a course represent? Mastery of the discipline, or just mastery of the game? The evidence from this study strongly suggests these may be two different masteries.

Regardless of any weaknesses of the measures of mastery used in this study, the fact still remains that performance on the group-based simulation showed *no* significant relationship with *any* of the measures. The group performance scores do not reflect mastery of any of the measures of the business discipline *nor of the individuals mastery of the game* (as indicated by performance in the individual simulation). This lack of significant relationships leads to the following conclusions. Until such time as we can demonstrate a link between performance on a group-based simulation exercise and measures of mastery of the discipline, its use as a means for assigning grades should be minimized.

The limitations of this study also indicate the need to develop valid instruments for measuring mastery. While the merits of the SMSQ and the MMEPS are debatable, both were used because valid instruments for measuring the upper level's of Bloom's Taxonomy are simply not available. Instruments are needed to assess mastery of (a) the course and/or business discipline and (b) mastery of the game.

While pursuing the development of these instruments, it is important to recognize that performance in a simulation is the result of an active, dynamic series of events. By contrast, assessment instruments yield scores that are the result of a passive, static singular event -- the filling out of the questionnaire. The simulation is inherently behavioral in its focus, while a questionnaire is intellectual. Consequently, questionnaires designed to capture the underlying factors driving performance on a simulation must be constructed with these differences in mind.

References available upon request