

The Relationship Between Learning and Performance in a Total Enterprise Simulation: Revisited and New Data.

Jerry Gosenpud
University Of Wisconsin - Whitewater
gosenpuj@uww.edu

John Washbush
University Of Wisconsin - Whitewater
washbusj@uww.edu

ABSTRACT

This paper re-explores the relationship between learning and performance in the total enterprise simulation, cites recent research on the topic, and reports the results of a pilot study performed by this paper's authors. Research from twenty or so years ago reports mixed results as to whether the players who perform the best in a competitive simulation also learn the most, and scholars in the past have attempted to attribute negative results to imperfect simulations. Research from 2009 (Wolfe and Deloach, 2009), and the present study found positive results between simulation learning and performance, suggesting that data can be attained to demonstrate the validity of the simulation as a learning tool. This paper recommends that the field review the issue and initiate new research.

INTRODUCTION AND BACKGROUND

This study focuses on the relationship between learning and performance in the total enterprise simulation, and one of its purposes is to reopen a scholarly discussion on that relationship. In large part, this is an empirical paper. However, since the measurement of learning in the simulation has generated considerable discussion over the years among ABSEL simulation scholars, and since this study utilizes an instrument that measures learning that many simulation users employ in their courses, namely an analytic paper, this article will also focus on the measurement of learning.

Common sense suggests that in a simulation, those who perform the best learn the most, but the research attempting to verify this has been inconsistent. Anderson and Lawton (1992) found that performance correlated with only two of seven measures of learning and concluded that the relationship between learning and performance was weak. Some other studies (Teach, 1989; Wolfe and Deloach, 2009) show a positive relationship between learning variables and

performance in the simulation, but still others (Anderson and Lawton, 1992; Wellington and Faria, 1992; Washbush and Gosen, 2001) do not.

There are explanations in the literature as to why a linear relationship between learning and performance has not been consistently found. For example, Burns, Gentry and Wolfe (1990) and Thorngate and Carroll (1987) argued that performance could easily be influenced by luck, and Wolfe and Chanin (1993) hypothesized and Washbush and Gosenpud (1993) found that average performers who struggled a lot learned more than high performers who struggled less.

There are two additional reasons why research does not consistently show a positive relationship between learning and performance. The first is presumptive and involves a pragmatic understanding of the value of the game to many of its players. Played in a university setting, the game is a means of attaining grades. For these players, measurement of simulation learning probably involves taking a test or writing a paper. These tests or papers are usually administered near the end of a term when students have a good idea of how well they are performing in the simulation. Students doing well in the simulation not seeking high course grades are not likely to try hard on an exam while students doing poorly on the simulation will likely study hard for a simulation exam in an attempt to avoid low course grades. Grade-seeking motivation thus hides any relationship that may inherently exist between performance and learning.

The second reason involves the illusive nature of learning from a simulation. Wolfe and Deloach (2009) define this learning eloquently (and in our opinion accurately) as the ability to create, process, apply knowledge to useful ends required by the game. The problem is that most efforts described in the literature to measure simulation learning do not reflect Wolfe and Deloach's definition. Scholars have used a variety of ways to measure learning. Early scholars (see Gosen and Washbush, 2004 for a review) used examinations covering course learning to assess simulation learning, with Whiteley and Faria (1989), Faria and Whiteley

(1990), Wolfe (1976), and Wolfe and Guth (1975) dividing course learning into categories. Many, including Comer and Nichols (1996) and Herz and Merz (1998) and have used perceptions of learning as a measure for simulation learning, and Peach and Platt (2000) and Teach (1989) have used forecasting accuracy as a proxy for learning. Two sets of authors (Feinstein, 2001 and Gosen and Washbush, 2001) were more elaborate and perhaps came closer to Wolfe and Deloach's (2009) definition in that they developed instruments 1) designed to test the learning constructs inherent to the simulation the students were playing and 2) for which validity studies were developed. Feinstein (2001) concluded that his test was valid, as people taking the test scored significantly higher after playing the simulation than before. For Gosen and Washbush (2002) convergent validity studies did not show the test to be valid. Although both sets of authors did extensive work on developing these scales, neither of the focal scholars followed up on their work, and no other research study applied any of the findings or used the tests. Thus most of the above efforts do not seem to be consistent with Wolfe and Deloach's (2009) definition, and there is not enough follow up for the efforts that came closer to conclude that these were successful. A tentative conclusion may drawn, then, that early attempts to find valid measures for simulation learning failed.

Trying to stay consistent with their own definition, Wolfe and Deloach (2009) utilized a way of measuring learning based on a diagnostic theory of organization learning developed by Choo (1998). This theory involves three stages of learning by organizations. The first is sense-making whereby organizational decision makers begin to interpret environmental signals, especially if there is a gap between the knowledge the organization has and the knowledge it needs to be effective. The second is knowledge making, as the organization either imports or develops the knowledge it needs to succeed. The third involves the action the organization undertakes given the knowledge it has acquired. Choo (1998) hypothesized that organizations that consciously go through this process will improve their effectiveness.

The instrument that Wolfe and Deloach (2009) used to measure learning was based on Choo's stage model and developed by Hansen (2004) and contains 30 Likert type items. Wolfe and Deloach (2009) treated the simulation company as a learning organization which could scan the environment for sense making, create knowledge about the simulation and its environment, and make decisions. Learning was the degree to which players said they scanned the environment and made sense, created knowledge, and made decisions based on the knowledge. Wolfe and Deloach (2009) found significant improvements in their groups' sense-making, knowledge and decision making, and

importantly for this study they found that end-of-simulation learning scores for all three stages of Choo's learning model correlated positively and significantly with company performance in their simulation. In other words, in their study those who performed the best learned the most.

Given this paper's purpose of revitalizing the exploration of the relationship between simulation learning and performance, it is important to point out the advantages and disadvantages of Wolfe and Deloach's model. Their way of conceptualizing and measuring learning in our opinion may have considerable advantages over other ways to measure simulation learning and perhaps two disadvantages. Key advantages come from the connection the theory has with the situation that Wolfe and Deloach's simulation students face. The simulation their study used, THE GLOBAL BUSINESS GAME (Wolfe, 2003) is very complex and in the study reported (Wolfe and Deloach (2009), the learning organization was indeed an organization with six simulation players per company. In addition, all (or at least most) total enterprise simulations, especially Wolfe's, require players to run a complex organization where players manage multiple functions including manufacturing, marketing, finance, inventory control and human resources. Thus, both the team and simulation milieus reflect the challenges inherent in real organizations, which is the kind of phenomenon Choo's learning theory focuses on. In addition, this instrument simply asks for a team member's perception of team's organization and behavior. Respondents are not judging what they've learned thus, minimizing the need to respond in a socially desirable manner. In addition, responses are not graded thus minimizing the 'grade-seeking motivation' problem, mentioned above. The disadvantages are that these constructs ignore the logic and principles of the actual simulation, which students must learn to be successful, and there is no extrinsic incentive to take the instrument seriously.

The present article introduces yet another method to measure learning, based on a student authored analytic paper, a paper discussing decisions and results in the simulation (and not vastly different from an annual report) that has been used in the past as a measure of learning (Anderson and Lawton, 1992). In the present study, players were asked to use their knowledge and understanding of game relationships to explain why they did as well (or as poorly) as they did in the game. Not only has this method been used in the past as a correlate of simulation performance but an analytic paper is probably often used methodology for supplementing performance to grade students in the simulation. The major advantages of this methodology are that it seeks to measure the learning of the actual principles that it takes to play the simulation

effectively, and it comes very close to measuring learning as defined by Wolfe and Deloach (2009).

METHOD

Subjects and Procedure. Thirty-four students taking the capstone policy course at the University of Wisconsin-Whitewater during the Fall of 2008 played the MICROMATIC (Scott et al., 2008) competitive game for sixteen quarters. Of those, twenty eight participated in the study in that they wrote the paper used to measure learning. (Of the six that did not write the paper, three did not need to write the paper for the grade they wanted in the course, two gave up on trying to pass the course as it ended and one paper was lost. All participants were seniors, 16 were graduating, and 11 were females.

Each student was placed in groups of three (or two), for practice and consulting purposes only, heretofore called 'home teams.' However, they played the competitive game as individuals in three separate industries. The markets for each of the separate industries were different enough so that one person from a different industry playing the game for another would be extremely difficult, although not impossible.

After the fifth and twelfth quarters, students were to turn in a paper assigning them to use their understanding of the game, its complexity and its variable relationships, to explain why they did as well or as poorly as they did in the game. Students had some choices as to how to and whether to do the first of these papers, because if the second paper earned a higher grade than the first, the first paper grade would not count. In addition the first paper could be done in the home teams if students chose. The second paper had to be done as individuals and would earn a 0 if not turned in. As for grading the game, a cafeteria grading scheme was used to determine the relative weights of game play and the 'learning' paper. If a student's grade for playing the game was higher than the 'learning' paper grade, than the game play grade was counted twice as much as the paper, but the paper counted twice as much for those whose paper grades were higher.

Variable Measurement. Performance in the simulation was measured at the end of play using the game's scoring procedure and was based on sales (15%) net income (40%), return on assets (10%), return on sales (25%), and stock price (10%). The actual resulting performance grade was not the index that the game produces. It was modified to fit an A, B, C, grading scheme with A=90 and above, B = 80-89, and C= 70-79. Performance grades were modified in cases of MICROMATIC performance improvement, and in two cases, results from a late semester round of Solo.

Learning was measured (or rated) by a one of this paper's authors, who has considerable experience as a MICROMATIC instructor, but was not acquainted with the focal class of this study, so in that way the assessor was blind to team performance. He graded the student papers on focus, strategy and analytic depth, on a 10 point scale. By

assignment, the student papers included references to how they did in the simulation. Those references were whited out before the coder read the papers.

RESULTS AND DISCUSSION

The Pearson correlation was .387, $p=.042$ (2 tailed, $df=26$). This correlation is both significant and positive suggesting that in this simulation, the people who performed the well on the financial indices of the simulation also wrote papers reflecting a high degree of understanding of game principles and the ability to apply these when playing the simulation. Coupled with the positive learning-performance result found by Wolfe and Deloach (2009), this data supports the argument that our field should begin again to explore the relationship between simulation learning and performance. This is important because, if the field can demonstrate that those who perform the best also learn the most, it will confirm that learning does emerge from the simulation and thus supporting the notion that the simulation is valid learning tool. A positive learning-performance relationship dispels the idea that performing well in the simulation is a random or lucky occurrence.

This is a pilot and for a variety of reasons, the data should be interpreted tentatively. First it is based on 28 students in one study. Second there was only one rater, which could cause rater unreliability. Third, although care was taken to rid the learning instruments of references describing and implying how well a student was doing in the simulation, there is no guarantee that the resulting paper was void of hints of the writer's relative success. Fourth, the rater and the instructor did not communicate well as to the rating criteria. Students were told that they were to be graded on their ability to apply game principles to game results and that the degree to which their understanding was complex would also count. The rater evaluated focus, depth of analysis, and the degree to which the paper reflected a coherent strategy. Perhaps the fact that these criteria are different might suggest that the findings underestimate the true relationship between performance and learning. More likely though, differences between the grading criteria and the rater's criteria contribute to the data's tentativeness.

The data is stronger for the sample's poorest performers. Of the ten students whose performance grades were below 70, all but two received two or below given the paper's 10 point rating range. In contrast, for the ten people whose performance scores were over 80, only five papers were rated 5 or above. This suggests weakly that when students do really poorly in MICROMATIC, they also are unable to express that they have learned anything (even when the vehicle of expression is worth 20% of the course grade).

The learning instrument deserves further discussion. As indicated in the introduction, this study's instrument asks students to express their understanding of game relationships (which include game principles and the complex consequences of decisions) and apply this understanding to their own decisions and performance. This reflects the definition of learning expressed by and Wolfe and Deloach (2009) and would seem to be the kind of learning most simulation learners expect. The paper is a comprehensive assignment and arguably reflects the comprehension, analysis, synthesis, and application components of Bloom's taxonomy (Bloom et al, 1956). With respect to those critics who sometimes argue that such instruments merely tap the writer's ability to memorize game rules, we believe and the data from the present study suggests otherwise.

REFERENCES

- Anderson, P.H. and Lawton, L. (1992). A survey of methods used for evaluating on student performance on business simulations. *Simulation & Gaming*, 23, 490-498.
- Choo, C.W. (1998). *The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge, and Make Decisions*. NY: Oxford University Press.
- Comer, L.B. and Nichols, J.A.F. (1996). Simulation as an aid to learning: How does participation influence the process? *Developments in Business Simulation & Experiential Exercises*, 23, 8. *The Bernie Keys Library*.
- Faria, A. J. and Whitely T. R. (1990). An empirical investigation of the pedagogical value of playing a simulation game in a Principles of Marketing course. *Developments in Business Simulation & Experiential Exercises*, Vol. 17, 53-57. *The Bernie Keys Library*.
- Feinstein (2001) "An Assessment of the Effectiveness of the Simulation as an Instructional System for Food Service." *Journal of Hospitality and Tourism Research*, Volume Twenty Five, 2001, 421-443.
- Gosen, J. and Washbush, J. (2002). Total enterprise simulation learning compared to traditional learning in a business policy course. *Developments in Business Simulation & Experiential Exercises*, Vol. 29, 247-251. *The Bernie Keys Library*.
- Gosen, J. and Washbush, J. (2004). A review of the scholarship assessing experiential learning effectiveness. *Simulation & Gaming*. Vol. 35, 270-293
- Hansen, J.H. (2004). Learning, Sense Making, Action, and Firm Performance: An Empirical Test of Choo's Theory of the Knowing Organization. Unpublished Doctoral Dissertation, University of Nebraska-Lincoln.
- Herz B. and Merz (1998) Experiential learning and the effectiveness of economic simulation games. *Simulation & Gaming*. Vol. 29. 238-250.
- Kolb D. A. (1984). Learning and problem solving: On management and the learning process. In D.A. Kolb (ed.), *Experiential Learning*. Englewood Cliffs NJ: Prentice-Hall, 27-41.
- Peach, E. B., & Platt, R.G. (2000). Total enterprise simulations and optimizing the decision set: Assessing student learning across decision periods. *Developments in Business Simulation & Experiential Exercises*, Vol. 27, 242-247. *The Bernie Keys Library*.
- Scott., T., Kaliski, J., and Anderson P. (2008). MICROMATIC. <http://www.oaktreesim.com/micromatic/>
- Teach (1989) Using forecasting accuracy as a measure of success in business simulations. *Developments in Business Simulation & Experiential Exercises*, Vol. 16, 103-107. *The Bernie Keys Library*.
- Thorngate, W. and Carroll, B. (1987). Why the best person rarely wins. *Simulation & Gaming*. Vol. 18, 299-320.
- Washbush & Gosen (2001) Learning in total enterprise simulations *Simulation & Gaming*. Vol. 32, 3, 281-296.
- Wellington, W. J. and Faria, A. J. (1992). An investigation of the relationship between team cohesion, player attitude, performance attitude on player performance. *Developments in Business Simulation & Experiential Exercises*, Vol. 19, 184-189. *The Bernie Keys Library*.
- 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. (1976). The effects and effectiveness of simulations in business policy teaching applications. *Academy of Management Review*, 1 (2): 47-56.
- Wolfe, J. (2003). THE GLOBAL BUSINESS GAME: A STRATEGIC MANAGEMENT & INTERNATIONAL BUSINESS SIMULATION. Mason OH: South-Western.
- Wolfe, J. and Deloach, G. (2009). Individual and organizational learning in a top management game. *Developments in Business Simulation and Experiential Learning*, Vol. 36, 231-237. *The Bernie Keys Library*.
- Wolfe, J. and Guth, G. R. (1975). The case approach vs. the simulation in teaching Business Policy. *Journal of Business*. Vol. 48, 349-364