

Developments In Business Simulation & Experiential Exercises, Volume 23, 1996

HOW DO WE KNOW WHERE WE'RE GOING IF WE DON'T KNOW WHERE WE HAVE BEEN: A REVIEW OF BUSINESS SIMULATION RESEARCH

**S. Dolly Malik State University of New York, College at Geneseo
Barbara J. Howard, State University of New York, College at Geneseo**

ABSTRACT

One of the themes of the 1996 ABSEL meetings is to focus on the future of business simulations. Often, in order to examine where the future is going to take you, it is necessary to examine the past. With this in mind, we undertake a review of literature relevant to understanding where the field currently stands with respect to business simulations. Our examination of the literature from 1985 to 1995 focused on articles related to business simulations that fell into the following categories: reviews on different aspects of business simulations; research on both the external and internal validity of business simulations; research examining the effectiveness of simulations as a learning tool; the examination of the individual factors and group processes involved in playing the simulation. We will review the literature pertaining to each category.

THE REVIEWS

Since the seminal review by Greenlaw and Wyman (1973), examining the teaching effectiveness of business simulations, there have been a number of reviews on various aspects of business simulations. This review did not find much support for the relative effectiveness of simulations over other teaching methods. Wolfe (1985) did a follow-up to the Greenlaw and Wyman review while Keys and Wolfe (1990) examined the role that simulations play in education and research. These two reviews, plus two others by Wolfe (1990) and Klein and Fleck (1990), conclude that, while there is some controversy over the effectiveness of simulations in increasing learning of theoretical concepts over pedagogical methods, the majority of the studies support the general validity of simulations as a classroom technique.

Gosenpud (1987) reviews the research on predicting simulation performance, as measured by quantitative measures, based on factors such as; academic ability, major, personality, motivation, team cohesion and organizational formality. His review finds that none of the factors are consistent predictors, nor are they totally unrelated to simulation performance.

A slightly different approach was taken by Keys (1987) who reviewed the ten most popular total enterprise simulation

games. A review of the most popular functional games was done by Biggs (1987). Klein and Fleck (1990) compare and contrast management games that might be used to teach international business concepts.

Another review topic that has generated interest is the actual use of simulations in business and academia. Recent surveys of the field have found that the usage of business simulations in academia has actually decreased (Faria, 1987; Keefe, Dyson and Edwards, 1993). According to the respondents in Williams' (1993) study, simulation users stopped including simulations in their business policy courses because they did not find a significant learning impact. An intriguing finding across all three surveys was that those educators with more teaching experience are more likely to use simulations. It would be interesting for these researchers to do a follow-up with their respondents to see if simulation use increases among the less experienced educators over time. Taking a more futuristic viewpoint, Wolfe (1993) used the Delphi technique to have experts estimate the future use of Business gaming in 2010. An increase in the use of business gaming, particularly simulations, at the graduate and corporate training levels was predicted.

EXTERNAL AND INTERNAL VALIDITY ISSUES

External Validity

External validity is defined as the extent to which the model matches its real-world counterpart (Carvalho, 1991). Based upon his review of the literature, Norris (1986;1987) called for more rigorous investigation of the external validity of business simulations. Wolfe and Roberts (1986) investigated the longitudinal relationship between students' simulation performance and their subsequent real-world performance. Their results found that an individual's simulation performance was related to salary level five years after graduation but not to salary changes or to promotion. In an extension of this study, the authors had peers evaluate simulation group members on interpersonal attributes, like decision making, leadership and value to the team, which have been shown to be important to managerial success (Wolfe and Roberts, 1993). The results showed a relationship between peer evaluations of these attributes and subsequent income levels and salary changes.

Developments In Business Simulation & Experiential Exercises, Volume 23, 1996

Both Mehrez, Reichel and Olami (1987) and Napier and House (1990) compared simulated industries and firms with real industries and firms to examine the external validity of simulations. Mehrez et al (1987) found some similarities between the performance of simulated firms within a simulated detergent industry and a real firm within the Israeli detergent industry. They noted that the biggest differences were a result of the constraints built into the simulation. Napier and House (1990) compared data from 21 simulated and 20 actual firms in the food processing industry on the basis of a financial ratio matrix. The simulated firms had more significant relationships among ratios than did the actual firms. This finding is not surprising considering that simulations are built through the use of algorithms that specify relationships among variables while the real world is not so perfectly and consistently structured.

While there is some more research examining the external validity of simulations since Norris' (1986) article, this is an area that needs more research. The longitudinal design used by Wolfe and Roberts (1986; 1993) should be replicated in other external validity studies. Similarly, research should continue to investigate the relationships between the performance of simulated and real firms along the lines of Mehrez et.al. (1987) and Napier and House's (1990) studies.

Internal Validity

Most of the validation work done between 1985 and 1995 seems to have been on the internal validity of the simulations. Snyder (1994) argues that this is exactly the way it should be. Pray and Gold (1982) and Gold and Pray (1984) were among the first to call attention to the question of the internal validity of simulations. To date, most internal validation efforts have followed two research designs (Dickinson and Faria, 1994). The first involves comparing the simulation performance of students with superior academic performance to that of students with poorer academic performance. The second method tries to evaluate whether or not students' decisions, over time, adapt to the simulation's environment (Dickinson and Faria, 1994). These two methods have been used extensively to investigate the learning effectiveness of simulation participation (e.g., Washbush and Gosenpud, 1994; Wellington and Faria, 1991).

Several researchers have proposed alternative methods to examine the internal validity of simulations. Snyder (1994) proposes a framework for assessing the validity of total enterprise simulations (TES) based upon the criteria of user-friendliness, comprehensiveness, theoretical grounding and adaptability. Stanislaw (1986) contends that simulations

consist of deriving a theory of real world behavior, modeling the theory and translating the model into a computer program. There is a different type of validity associated with each one of these steps; theory, model and program, and computer simulations should be tested for these.

Two new ways of examining internal validity were investigated by Dickinson and Faria (1994) and Wolfe and Jackson (1989). Wolfe and Jackson propose that simulation designers need to examine algorithmic validity to assess their realism. In their study, the majority of simulation players were unable to identify the illogical algorithms programmed into the simulation. Wolfe and Jackson contend that, since simulation participants have their own sense of reality, the simulation's program is only one aspect of perceived realism. However, they suggest further testing of this type of validity since their study only manipulated one algorithm. Defining validity as meaningfulness of experience, Dickinson and Faria (1994) investigated the use of a random strategy criterion to test the validity of participation in a marketing simulation. The results of their study found that students' simulation performance was better than a random strategy 60 to 64 percent of the time. On the basis of these results, the researchers conclude that the simulation had validity because it was a meaningful experience for participants. Since a primary concern of educators is the meaningfulness of their classroom techniques, the random-strategy criterion could easily be used to test the validity of participation in TES.

When simulations were introduced 35 years ago (Watson 1981), there was an assumption of their superiority over pedagogical techniques as a learning method. Based upon extensive reviews of the literature by Greenlaw and Wyman (1973) and Wolfe (1985) that raised questions about these assumptions, research investigating the external and internal validity of simulations has increased in the past ten years. As is evident by the research reviewed in the next section, the debate over the effectiveness of simulations continues. Therefore, further empirical research on the external and internal validity of simulations is needed.

LEARNING ISSUES

Most simulations are used in business policy, management or marketing courses (Anderson and Lawton, 1992). Furthermore, they usually last 8 to 16 quarters and performance on the simulation comprises an average of 30% of the final course grade (Anderson and Lawton, 1992). These researchers also found that simulation performance grades consisted of some

Developments In Business Simulation & Experiential Exercises, Volume 23, 1996

combination of overall relative performance in the actual game, a written strategic plan, an analysis of performance results and other items like peer evaluations and tests.

Greenlaw and Wyman (1973) questioned the effectiveness of simulations as a classroom technique. However, Keys and Wolfe (1990) argued that simulations were effective teaching tools. This conclusion has been supported by subsequent studies examining participants' reactions toward simulation play. A survey of Australian users and participants found very favorable reactions toward simulations as a learning tool (McKenna, 1991). Washbush and Gosenpud (1991) also found favorable attitudes toward simulations among their student subjects. Leonard and Leonard (1995) found that graduates surveyed felt that simulations or simulations with cases did a better job of preparing them than cases alone for their current positions. Williams (1987) surveyed AACSB schools and found that 69% of policy professors that used simulations felt that they were the best way to teach policy and surprisingly almost 50% of those not using a simulation felt the same way. Ninety-eight percent of the students involved in the simulations rated them favorably. Yahr (1995) gathered student opinions of a particular simulation and found that students perceived the simulation as both enjoyable and a learning experience. McLaughlin and Bryant (1987) asked game participants to write a paper detailing the advantages and disadvantages of their experience. Although the participants agreed with the literature on the advantages, students ranked the relative importance of the advantages differently. Students did not perceive the disadvantages of simulations as indicated by the literature. Hemmasi and Graf (1991; 1992) surveyed seniors in an undergraduate business program and alumni who had been working for five years about the perceived effectiveness of simulation play. They report that practitioners actually rated the simulations as being more effective in enhancing managerial skills than did students. In fact, Teach and Govahi (1988; 1993) found that, of the five managerial skills rated most important by the business people in their sample, simulation play was rated as highly effective in enhancing four of these: the ability to adapt to new tasks, make decisions, organize, and assess a new situation quickly. Soukup and Whitney (1987) report on exploratory research on the effectiveness of simulation in meeting specific learning objectives. Students perceived a difference in the effectiveness of teaching methodologies and that effectiveness was a function of the learning objective. Simulations were most useful in gaining communication, decision-making and group behavior skills.

Several studies have examined independent variables, which might influence simulation performance. Curran and Hornaday (1987) originally found no difference in performance or satisfaction between groups that were required to write a formal long-range plan and those that didn't. A replication of that study by Hornaday and Curran (1988) found that planning teams achieved significantly higher earnings, stock prices, EPS & ROI. Non-planners had larger firms in total assets; higher D/E, more loans and unnecessary assets. There was no difference between planners & non-planners in terms of satisfaction with the simulation or with regard to their teammates. Anderson and Lawton (1992) also found that a group's annual plan was a predictor of financial performance in a simulation. Gosenpud and Washbush's (1991) study found that group factors such as cohesion and organizational formality influenced group performance whereas an individual's academic interests (primarily finance and accounting) predicted individual performance. Participant's motivation toward the game also predicts individual performance (Gosenpud & Miesing, 1992). These affective variables, even when measured prior to simulation play, relate to performance (Gosenpud, 1989). Wolfe and Box (1987) determined that team cohesion was both directly and indirectly related to economic performance on the simulation. Hornaday and Wheatley (1986) found that GPA had no relationship to group performance, but that all female teams, teams with accounting majors and non-marginal teams did perform better than their opposites.

The biggest controversy associated with business simulations is their effectiveness as learning tools, a more rigorous criteria than participant reactions. A number of empirical tests have failed to find a significant relationship between simulation performance and actual learning. Whiteley and Faria (1989) found no impact of simulation performance on a midterm or a final exam performance. However, when they examined individual questions, there was a relationship between performance and quantitative questions. They suggest that, since the game allows participants to implement and use financial knowledge, the quantitative questions reflected the learning that occurred through simulation play. Wellington and Faria (1991) found no relationship between recency of simulation play and performance level on either a midterm or final exam performance. One of the problems of examining the learning impact of simulation performance is that most studies cross units of analysis since simulation play is at a group level and exam performance is measured at an individual level (Wellington and Faria, 1991). However, when Faria and Whiteley (1990) examined the relationship between individual performance on a simulation and

Developments In Business Simulation & Experiential Exercises, Volume 23, 1996

performance on a multiple choice exam, there still was not a significant relationship. Anderson and Lawton (1990) found no relationship between simulation performance and a final essay exam. Studies by Washbush and Gosenpud (1993; 1995) and Gosenpud and Washbush (1993) also supported these findings. In fact, the results from Washbush and Gosenpud's (1994) study suggest that poorer students actually learn more during the early part of simulation participation than do better students. Along a similar line, a study by Wolfe and Chanin (1993) found that low skill players increased performance and skill acquisition over time more than high skill players even though their economic performance was still relatively low. However, Wellington and Faria (1995) examined the relationship over two rounds of play of the same simulation but under different environmental and competitive conditions. Simulation performance remained consistent, i.e. good performers stayed good performers.

Klein and Fleck (1990), on the basis of their review of computer games, raise the point that the learning issue may be confounded since simulations are primarily used in a capstone course like Business Strategy and that initial learning of concepts may have taken place prior to this course. This, however, does not explain the results of the Faria studies since these were done with a functional marketing simulation. It is possible that the dimensions of learning as currently operationalized do not tap the dimensions actually enhanced by simulation play. Furthermore, the empirical tests of the learning effectiveness of simulations contradict the reports of the widespread positive reactions participants and users have toward simulations. Therefore, it may be in the best interest of researchers to continue to examine this relationship by using different ways to operationalize learning.

There are a number of researchers who have proposed several new methods of operationalizing the learning criteria. Biggs, Miles and Schubert (1993) have developed an instrument to measure perceptions of effectiveness. These authors compare their instrument with Anderson and Lawton's (1990) use of a similar instrument as an indication of the construct validity of the measure. Burns (1993) proposes that researchers use a grid, which compares teams' effectiveness (defined as the team's achievement of market share goal) and efficiency (achievement of the team's per unit profitability). A unique performance assessment tool is presented by Sackson (1992) who suggests that cluster analysis of performance be used to analyze both game and student performance.

One of the major problems of testing for the learning effectiveness of a technique like simulation gaming is

that the objectives of the user must be specified (Hsu, 1989). In fact, it is these objectives that should be the basis of the selection of a learning method (Gentry and Burns, 1981). Hsu (1989) also notes that simulation performance is the function of so many factors that it is difficult to isolate and test the impact of these on learning. Instead of performance, Hsu (1989) proposes that researchers and users focus on the use of simulations for skill building. Teach (1987; 1990) echoes this call for a decreased emphasis on profits as a measure of superior simulation performance. In a subsequent research study, Teach (1993a) investigated the use of a financial forecasting accuracy criteria to judge learning instead of financial performance. His study found a strong relationship between ability to accurately forecast financial factors and the firm's relative profitability. Wolfe (1993a) challenged these results by replicating Teach's study. He actually found that poorer performers had more accurate forecasts than better performers. This challenge resulted in a spirited debate between the two researchers over the technique and the respective conclusions (Teach, 1993b; Wolfe, 1993b).

The research, to date, has not been unanimously encouraging with respect to the effectiveness of simulations as a learning tool. There are a number of issues that researchers need to focus on before simulations are completely dismissed in terms of their effectiveness in enhancing learning. The first issue is that, given most games are played an average of 8-16 times, how feasible is it that learning does take place in this short a timeframe? Second, as pointed out by Hsu (1989) and Wellington and Faria (1991), most tests of learning cross units of analysis and this is likely to confound the results. Third, in order to examine learning specific to simulation performance, the learning measure has to tap the concepts the simulation actually forces players to use. When Whiteley and Faria (1989) and Gosenpud and Washbush (1993) examined the test items that related to simulation play, both studies found support for the simulation as a learning tool. Malik and Howard's (1995) study indicates that the simulation is effective in teaching specific concrete principles of a theoretical model, but that students were less able to apply the underlying theory. This suggests that researchers address all three of these issues when investigating the learning effectiveness of simulations.

SIMULATIONS AND GROUP DYNAMICS ISSUES

One of the suggestions made by Keys and Wolfe (1990) was that researchers should investigate the group dynamics and process issues that result from simulation play. This charge has been undertaken by a

Developments In Business Simulation & Experiential Exercises, Volume 23, 1996

number of researchers in the past decade who have looked at various process and group dynamic issues.

Wolfe, Bowen and Roberts (1989) explored the impact of a 20-hour team-building training session on group performance for a simulation that lasted the equivalent of two years. Their results found that trained teams reported more cohesiveness for the first year than did the control teams but there was no significant difference between groups on learning or satisfaction with group learning. In another experiment, which involved training, Fandt, Richardson and Conner (1990) trained a group of their participants on goal-setting techniques. While goal setting led to better group dynamics (i.e., more cohesion and less conflict) there was no significant difference between the control and trained group on simulation performance. Jaffe and Nebenzahl (1990) compared the group dynamics of successful teams versus the losing teams. They found that successful teams initially established group cohesion and then moved to a task orientation while losing teams did the opposite.

Along another line of group dynamics issues, Patz and Milliman (1992) investigated the impact of levels of confidence of individuals in-groups on simulation performance. Their research found that when there were confidence extremes within the group, the quality of performance decreased. In an interesting twist, Wheatley, Amin, Armstrong and VanderLinde (1991) found that individuals' cognitive styles and an interaction of cognitive information processing style and leadership style significantly impacted peer evaluations of simulation performance.

In 1990 and 1992, Patz reported studies that examined participants' personality bias as measured by the Myers-Biggs Type Indicators (MBTI). For each group, he averaged the members' score on the MBTI for two personality biases, intuition (N) and thinking (T) and then examined the relationship of this measure with simulation performance. Patz reported a strong relationship between groups dominated by N and T personalities. Replications by Anderson and Lawton (1991; 1993) and Gosenpud and Washbush (1992), however, found no support for such a relationship. Washbush (1992) tried using another measure of personality bias, the Learning Style Inventory, but, again, found no significant impact on performance. Anderson and Lawton (1993) argued that group dynamics may be the reason for Patz's significant results; not individual personality biases. Gosenpud and Washbush (1992) did find a stronger relationship between TES performance and group dynamics (i.e., cohesion, motivation and organization) than individual personality biases. Results from a study by Wellington and Faria (1992) support this conclusion. They found that

participant attitudes did not have a significant impact on simulation performance whereas group cohesion, by impacting performance expectations, did.

IN CONCLUSION

This article is an attempt to review the literature from 1985 to 1995 in an effort to identify where the research on business simulation currently stands. Based upon this review, it is possible to make some general statements about where the field needs to head in the future.

In order to establish the role of business simulations as a learning tool, researchers must do a more rigorous job of examining how simulations increase learning. The relationship of different dimensions of learning to simulation play must be examined. If the field does not do this, the current trend of a decrease in the classroom use of simulations will continue. Given the predictions of the experts in Wolfe's study (1993), more research is also needed to examine the use of simulations at graduate and corporate training levels.

Another area that needs more attention is the external validity of simulations. One way that simulation experience may have a more generalizable effect would be to increase the time that the simulation is played. Longitudinal designs like the one used by Wolfe and Roberts (1986) should be replicated with different TES and timeframes. Another method to address external validity might be to model algorithms that exist in actual industries and then examine the similarity of these with those that structure TES.

Finally, there has been some progress in examining group dynamics and simulation process issues. However, there is considerable room for additional research. There are a number of factors in-group research that could be adapted to simulation research further. For example, does diversity within a group increase or decrease performance? Are there gender differences in-group performance? How does leadership style relate to simulation performance?

As ABSEL looks to the future of business simulations and their effective use in the classroom, there are a number of issues still to be decided. Research has progressed this past decade, but simulation writers and researchers are not dinosaurs, YET!

REFERENCES

References are available by contacting the authors.