ANOTHER LOOK AT THE USE OF FORECASTING ACCURACY ON THE ASSESSMENT OF MANAGEMENT PERFORMANCE IN BUSINESS SIMULATION GAMES[1]

Moises Pacheco de Souza
Universidade Federal de Santa Catarina
mpsouza1980@yahoo.com.br

Ricardo R. S. Bernard
Universidade Federal de Santa Catarina
bernard@cse.ufsc.br

Hugh M. Cannon
Wayne State University
hugh.cannon@wayne.edu

ABSTRACT

Research in business games show that the reduction in forecast error can be used as a predictor of team performance. The objective of this study was to evaluate these findings with respect to high- versus low-level individual management functions. Using multiple linear regression, the study found that the set of independent variables (prediction error for indicators relevant to each management function) explained 40.75% of the overall company performance, while another portion is explained by external factors. As expected, the study found that the forecast accuracy for high-level functions (general management) had the greatest predictive impact and low-level functions (sales, human resources, and finance) the lowest. This supports the notion that Teach's (1990) forecast-accuracy approach to performance evaluation can be used in top-level strategic management simulations as well as lower-level functional simulations rather than limiting it functional simulations only, as Wolfe (1993a) suggests.
Key-words: Simulation Performance Evaluation, Managerial Performance, Forecasting Accuracy

INTRODUCTION

There is no clear consensus regarding the way we should evaluate learning in business games, either at the group or the individual level. Traditionally, instructors rely on the financial success of the simulated firm, reasoning that the best measure of what potential managers have learned is their ability to profitably manage their firms. This, however, tends to be a relatively poor indicator of learning (Thorngate and Carrol, 1987; Anderson and Lawton, 1990; Washbush and Gosen, 2001).

Teach (1987, 1989, 1990, 1993a, 1993b, 2007) has argued that a better measure would be the ability of managers to predict, or forecast, the consequences of their behavior. Their research suggests a positive correlation between forecasting accuracy and indicators of general management success in a simulation. That is, the better game participants are able to forecast, the better their performance, and conversely, the worse their forecast, the poorer their performance. In its most general form, the argument is that the ability to forecast, the impact of one's decisions is a necessary condition for good decision making, and hence, should be the basis for evaluating decision-making quality (i.e. student performance). Given Teach's strong advocacy of this position over the years, we will refer to it as Teach's forecasting-accuracy approach to student evaluation.

The logic behind Teach's approach is persuasive. How can a manager make effective decisions except by predicting their consequences? This, then, must be the essence of what students must learn to do in a simulation, or for that matter, in a real business organization. Teach (1987, 1990) goes farther, suggesting that in the absence of prediction, the success of a simulated firm might actually be misleading. In the real world of business, firms never begin with an even start. So, how can we evaluate management by which firm delivers better performance? In simulations, firms generally do have an equal start, but random errors early in the simulation quickly change this. By the time student managers have acquired the expertise to manage effectively, they no longer have an equal opportunity. And

---

even if this were not a problem, requiring firms to always begin on an equal footing severely limits the ability of game designers to address conditions that typify the real world of business.

Notwithstanding the logical appeal of Teach's approach, it has come under considerable criticism. Most notably, Wolfe (1993a) notes that the empirical support is mixed, and that, even if we accept its basic validity, it is more appropriate for more specialized "functional" games than for high-level top-management simulations. To address these items, we will construct a basic model of the learning process implicit in Teach's approach. We will then describe an experiment designed to address the two key issues raised by Wolfe. Finally, we will return to the conceptual model, and use it to interpret the results of our study, identifying key issues and directions for future research.

FORECASTING, LEARNING, AND MANAGERIAL PERFORMANCE: A CONCEPTUAL MODEL

Notwithstanding the prominence of financial measures as criteria for good simulation performance, as early as 1975, Hand and Sims argued for the use of forecasting accuracy. They empirically studied thirteen different performance criteria and concluded, "For the purposes of further research, the thirteen performance variables can be reduced to two: (a) Sales Forecasting Error, the primary 'driving variable,' and (b) Profit, the primary 'end result' variable" (p. 715). In a series of articles, Teach (1987, 1989; 1990, 1993, 2007; Teach and Patel 2007) expanded on this concept, arguing that, "The ability to adequately forecast the impact of changing key decision making variables must be learned before one can become a good practicing manager" (Teach 2007, p. 57).

To put this in perspective, consider the problem faced by students who are engaged in a business simulation game. This is portrayed in Figure 1. Students are confronted with a complex environment, consisting of both internal (company capabilities) and external (opportunities and constraints). In this context, they must marshal their business knowledge and conceptual abilities to conceptualize general causal relationships that will govern their decision making. Given these concepts, they formulate decision alternatives and seek to apply the general causal relationships to predict the specific consequences of their decisions.

Following the logic suggested by Cannon, Friesen, Lawrence and Feinstein (2009), we argue that much of the learning takes place at this stage of the process. Given the complexity of the links between specific decisions and their ultimate financial consequences, students must break down the system into understandable chunks, forecasting intermediate cause-and-effect relationships governing each one (Cannon 1995). This requires a great deal of knowledge and information processing, but more important, it require high-level reasoning ability to determine how to break down the system and which principles to apply in making the intermediate forecasts. The experiential decision process is a kind of intellectual "weight training," building the required thinking stills through practice. While we need not do so here, we can conceptualize the outcomes of this process in terms of the formation of knowledge structures and cognitive process abilities discussed in Bloom's revised taxonomy of educational objectives (Cannon and Feinstein 2005; Ben-Ziv and Carton 2008). The actual thinking patterns used to forecast decision consequences are stored as knowledge, but the ability to access and utilize the knowledge structures effectively in a specific decision-making domain constitutes both a general and a domain-specific reasoning ability. It takes the form of *tacit* knowledge – knowledge that is not easily taught or spoken -- but that enables people to grapple effectively with real-world problems (Gentry, Stoltman and Mehihoff 1992).

The second level of experiential learning comes through feedback regarding forecasting accuracy. In this sense, the simulation provides a laboratory in which students can experiment, testing their forecasting hypotheses to see if they were correct. If so, the knowledge structure that describes how the forecast was made is reinforced and will used again in similar situations. If the forecast turns out to be wrong, the feedback triggers additional high-level reasoning to determine why and how the process can be fixed for the next decision.

Finally, forecasting accuracy – actually, the forecasting accuracy for each of the various analytical chunks into which the company decision-making process was divided -- is folded into an overall evaluation of company performance. For all the emphasis on forecasting accuracy, the fact is that profit and other financial measures of simulation success are the ultimate test for a firm. Indeed, they are the criterion against which tests of the forecasting-accuracy approach are validated. That is, researchers seek to determine whether forecasting accuracy is a valid measure of performance by testing the relationship between forecasting accuracy and the traditional financial measures of performance, such as profit.

Drawing on Figure 1, we see why this would be the case. Company performance is a function of the actual consequences of good decision making, and short of highly effective intuition (or tacit knowledge), we assume that positive consequences result from the analytical process described in the figure, where seek to estimate the consequences of their decisions (i.e. forecast) and choose the decision that produces the most desirable consequence. The problem is that company performance is also a function of other factors that are not under the control of the student managers (what the figure refers to as "non-decision-related factors"). If the forecasting approach has any validity at all, its contribution should show up statistically over a large number of observations, where the "noise" effect of non-decision-related factors tend to cancel each other out. In any specific study, however, we are left to wonder whether the results are truly a reflection of the forecasting approach's validity or simply an expression of excessive noise.

Figure 1:
A Model for Understanding the Role of Forecasting in
Simulation-Based Experiential Learning
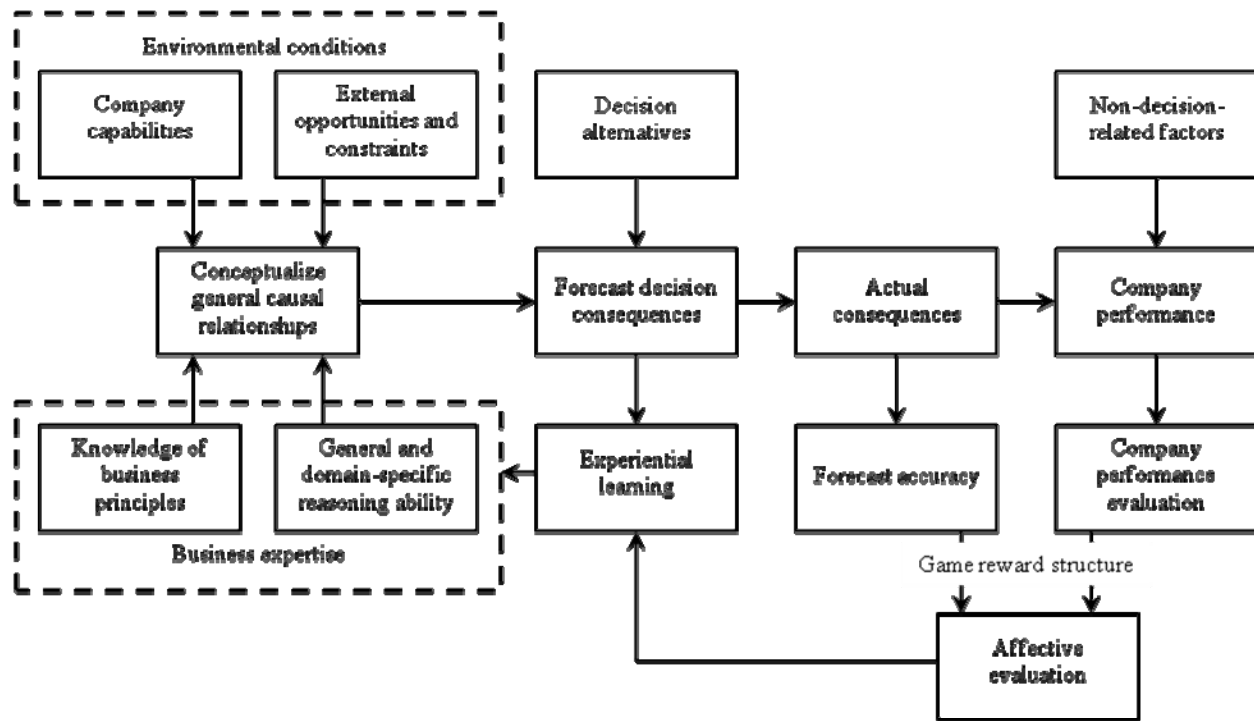
The fact that company performance is driven by both controllable and uncontrollable factors is significant. For instance, Teach and Patel (2007) argue that, given the structure of most simulation game designs, once a company becomes "dominant," other teams cannot effectively compete, regardless of what they do. To the extent that company performance is used to reward student participants (e.g. making company performance an important part of the grade) it as a key objective in student decision making. To the extent that this performance is out of students' control, they will be highly motivated with no logical bases for decision making. That is, they will receive random reinforcement. This tends to promote student irrationality, which, of course, is precisely the opposite of what we would hope the game would accomplish.

EVALUATING THE UNDERLYING THEORY

The concern regarding random reinforcement is relevant, because the empirical results in support of Teach's forecasting-accuracy approach are equivocal. That is, it is not clear from the literature that being able to correctly forecast the consequences of key decisions leads to company success. Our discussion will proceed in two parts: First, we will consider the relationship between forecasting accuracy and company performance, seeking to determine whether the basic premise of the forecasting accuracy approach makes sense. Second, we will Wolfe's (1993a)

argument that the forecasting accuracy approach is less appropriate for "top-management" than for "functional" simulations.

The Relationship between Forecasting Accuracy and Overall Company Performance

Figure 1 illustrates the relationship between forecasting decision consequences and company performance. The traditional approach to evaluating the forecasting-accuracy approach to student performance has been to test the strength of this relationship. On the positive side, Teach (1989, 1993a, 2007) and Washbush (2003) offer positive empirical evidence in support of the approach. Opposing this, Wolfe (1993a,b,c) offers a host of countervailing evidence. For instance, he reinterprets the Hand and Sims (1975) findings, suggesting that they do not support the relationship between forecasting accuracy and performance (Wolfe 1993a). He then cites an additional unpublished study by Smith and Golden (1991) where forecasting accuracy showed no significant relationship to financial performance. Turning to related issues, he notes that the logic of forecasting-accuracy approach would suggest that teams who do careful planning would deliver better financial results than teams that do not. He cites two studies by Curran and Hornaday, noting that one (Curran and Hornaday 1987) showed no relationship between planning and company performance. In a second study (Hornaday and Curran 1988), where the market size was substantially greater, they found that there was a relationship.

Wolfe (1993a) goes on to conduct his own study. First, he tested for increases in forecasting accuracy over time, an effect we would expect if students were actually learning, as the approach suggests. He found mixed results. For two key indicators (input costs and profits), forecasting accuracy actually decreased dramatically over the course of the game. Second, he found that, while high-performing companies did more forecasting than low-performing ones, the accuracy of forecasts was higher for the low-performing companies.

Given our discussion of Figure 1, none of these results should surprise us. Consider the inconsistent results regarding the relationship between forecasting error and company performance. If a simulation is well designed, there is no question that the quality of student decisions, and by extension, forecasting accuracy will have an impact on company performance. The question is how much? The more complex the simulation, the less effect any given decision is likely to have on overall company performance, thus attenuating its relationship to forecast-error.

Note that this is exactly what Cannon's (1995) "complexity paradox" would lead us to expect. According to Cannon, the paradox is as follows (p. 96):

> *"On one hand, the purpose of a simulation game is to provide a realistic laboratory in which students can learn business decision making, experimenting with various decisions and getting feedback regarding their relative level of success (Gentry 1990). On the other hand, the more faithfully a game portrays the true complexity of an actual situation, the more decisions there are to make and the more phenomena there are to model. This increases the potential for obscuring the linkage between cause and effect, thus defeating the purpose of the simulation (Fritzsche and Cotter 1990)."*

In other words, when simulations are complex enough to provide a reasonable representation of an actual firm, they are so complex that managers cannot relate the consequences of their individual decisions to the firm's performance. To draw on an old adage, no one questions that there will be a "straw that broke the camel's back," but straw is never likely to show up in a regression equation explaining the variance in camels' spinal integrity! So it is with management decisions. They are all important, but there are so many, and their incremental contribution can be so relatively small, that we often cannot measure their effect on the firm's overall success.

One response to this is to continue on our present course of research, replicating the basic studies across an ever-broadening set of situations. We would not expect dramatic results, again consistent with what we currently find in the literature. However, we should continue to see a relationship, even in individual studies. In the end, we should be able to apply meta-analytical techniques to establish a much more definitive confirmation of the validity of Teach's approach. This leads us to our first hypothesis. Given a broad range of functional forecasts:

H1: The combined accuracy of functional forecasts will be positively related to overall company performance.

The Relationship between Forecast Level and the Predictive Ability of Forecast Accuracy

Wolfe (1993a) offers a second, more damning criticism of the forecasting-accuracy approach. This grows out of two pieces of research. First, Wolfe and Richards (1993) found that, consistent with earlier studies, various aspects of overall company performance predict career success, years after the gaming experience. Of course, this does not demonstrate a causal relationship between company performance in a simulation game and career success, but at very least, it suggests that the two have something in common. If company performance is an unreliable measure of student skills, why does it predict future success?

Second, Wolfe (1993a) discusses another study he conducted (Wolfe and Chanin 1993). In his analysis, he notes that the study supported the proposed relationship between forecasting accuracy and company performance. However, it also found that adopting and effectively implementing an effective strategy played a critical role in company success. The implication, of course, is that the contribution of students to simulated company performance involves more than just an ability to forecast accurately. He concludes that (p. 59):

> *"Using a different perspective than that taken by Teach, one might ask under what learning/playing conditions could one employ his suggested intermediate criteria. Certainly his criteria would be relevant within a company's functional areas of isolation, or associated with judging effectiveness within the many functional games available (Biggs 1987). For the large crop of general, top-management simulations in use today (Keys 1987), substantive, bottom-line criteria may be appropriate criteria, and intermediate criteria may be applicable to functionally related simulations such as marketing games, which comprise the field's second highest number of applications (Faria 1987)."*

This drives to the heart of the problem. Wolfe distinguishes between "intermediate" versus "bottom-line" criteria, relating to the objectives of "functional" versus "top-management" simulations, but Teach's approach is not about the level at which the measurement is taken; it is about the type of measure. The approach would not preclude to the use of high-level variables such as market share, sales, profit, return on assets, or return on equity. In fact, these would be the kinds of variables most appropriate for measuring the effects of high-level strategies. However, the measurement would be on forecasting accuracy rather than results. The key would be to evaluate students on their

Table 1:
Characteristics of the Participant Groups

Group Code	Size of Teams	Program	Institution	Schedule
AD	8	Accounting	A	Day
AE	8	Accounting	A	Evening
BD	8	Administration	B	Day
BE	8	Administration	B	Evening

ability to anticipate the consequences of their strategic decisions. The higher level the variables, the higher their correlation with company performance is likely to be, since they would not be subject to the mediating effects of other decisions, but only of environmental conditions over which the managers would have no control.

In Wolfe's (1993a) critique of the forecasting-accuracy approach, he presents results of a study that appear to disconfirm this expectation. He notes that forecast accuracy was a worse predictor later in the game rather than earlier, while the effect of learning would lead us to the opposite conclusion. Furthermore, forecasting accuracy was higher for low-performance teams than for high-performance teams, which is just the opposite of what the theory behind Teach's approach would lead us to expect.

These results are troubling, because they run contrary to the theory underlying the forecasting-accuracy approach, and they do not suggest any alternate theory to guide further testing. This suggests a need for further exploration. In the meantime, however, we can propose an alternative test based on extant theory, evaluating the proposition that forecasting accuracy for high-level variables will correlate more strongly with company performance. Given the fact that the chief executive (CEO) is responsible for strategic decisions, we would expect:

H2: Forecasting accuracy at the CEO level will have the greatest impact on overall company performance.

If supported, this hypothesis addresses Wolfe's (1993a) recommendation that the forecasting-accuracy approach be relegated to use in lower-level, "functional" simulations. Contrary to his recommendation, it would suggest that the logic of the method apply at all levels of simulated company decision making.

METHODOLOGY

The study drew on students from two separate universities in Brazil, drawn from two different majors: accounting and management, forming four sets of eight teams each. (See Table 1). The students were selected from a convenience sample, drawing on classes for which simulation games were being used. Given the limited number of classes available, we were not able to develop a truly random sample. Table 1 summarizes the characteristics of the sample.

The unit of analysis was the team of students representing a company in the simulation. Each student in a team assumed responsibility for one or more roles within the company. Each team consisted of between three and five students, with the selection for team membership made at random. The principal differences among the teams were their status as day or evening students, their major (accounting or administration), and the fact that teams came from either a public or a private university.

The student teams were evaluated using data gathered in five consecutive periods. Professors in the discipline were responsible for evaluating the accuracy of the forecasts on a scale of 0-10, indicating the least to the greatest level of error using the formula proposed by Teach (1989):

$$Re_i = |(Fv_i - Av_i)/Av_i| \qquad (1)$$

Where:

Re_i = relative forecasting error for each of *20* dimensions of performance

Fv_i = forecast value for each of *20* dimensions of performance

Av_i = actual value for each of *20* dimensions of performance

The actual data were gathered through the use of two instruments: First was a survey addressing student forecasted values (Fv_i). It consisted of twenty (20) questions regarding expected results in various areas of management responsibility (the 20 dimensions of performance enumerated in Table 2). It was completed following each period of play by each of the teams participating in the simulations. The questions were taken from the evaluation model built into the "Bernard Websimulator: Virtual Environment" (Bernard 2004, 2007) used in numerous locations throughout Brazil. The survey was provided automatically at the end of each period of play through the regular processing of the simulation results.

The second set of data collected consisted of the actual performance results (Av_i) along same dimensions as were used for the forecasts (Table 2). These were obtained directly from the Websimulator, and were, in turn, generated

Table 2:
Variables Used for Evaluating Performance and Forecasting Accuracy

Function	Dimensions of Performance (and forecast)	Performance measurement key	Pm_i	Av_i	Fv_i	Pe_i
Marketing	Market share	HB	1..10	%	%	%
	Sales growth	HB	1..10	%	%	%
	Sales	HB	1..10	$	$	%
	Demand of Product A	HB	1..10	units	units	%
	Demand of Product B	HB	1..10	units	units	%
	Demand of Product C	HB	1..10	units	units	%
	Demand of Product D	HB	1..10	units	units	%
	Demand of Product E	HB	1..10	units	units	%
Finance	Total inflow cash	HB	1..10	$	$	%
	Total outflow cash	LB	1..10	$	$	%
	Cash flow balance	LB	1..10	$	$	%
	Current liquidity ratio	HB	1..10	%	%	%
Personnel	Operating employee productivity	HB	1..10	#	#	%
	Salespersons productivity	HB	1..10	#	#	%
	Operating employee balance	NOB	1..10	ratio	ratio	%
	Salespersons balance	NOB	1..10	ratio	ratio	%
CEO	Share value	HB	1..10	$	$	%
	Share value	HB	1..10	rank	rank	%
	Net profit	HB	1..10	$	$	%
	Net profit margin	HB	1..10	%	%	%

Note: HB = higher is better; LB = lower is better when evaluating performance
         NOB = near one is better (negative and positive values are possible)

by the simulator's algorithm in response to 44 decisions the students were required to make while playing the game.

The overall evaluation of company performance measurement (which we shall designate Pm) was drawn from these same 20 dimensions. The "performance measurement key" column of Table 2 provides a general scoring key. The course professors scored each team in each period of play, assigning a score from "0" to "10" reflecting the quality of the team's performance along each of the 20 dimensions, as shown in the performance measurement (Pm_i) column of Table 2.

In order to test our two hypotheses, we performed two multiple linear regression analyses. The first regressed the 20 measures of forecasting error (Pe_i) on overall company performance (Pm), as depicted in equation (2):

$$Pm = f(Pe_i) \qquad (2)$$

Where:
Pm = overall company performance, operationalized by averaging the 20 individual measures of performance (Pm_i)
Pe_i = forecasting error for each of the 20 dimensions of performance

The second analysis grouped forecasting error by the four managerial functions shown in Table 2 and regressed

them on overall company performance (Pm), as depicted in equation (3):

$$Pm = f(Pe_m, Pe_f, Pe_p, Pe_x) \qquad (3)$$

Where:
Pm = overall company performance, operationalized by averaging the 20 individual measures of performance (Pm_i)
Pe_m = forecasting error for the marketing function, operationalized by averaging the 8 individual measures of forecasting error (Pe_i) that relate to the marketing function
Pe_f = forecasting error for the finance function, operationalized by averaging the 4 individual measures of forecasting error (Pe_i) that relate to the finance function
Pe_p = forecasting error for the personnel function, operationalized by averaging the 4 individual measures of forecasting error (Pe_i) that relate to the personnel function
Pe_x = forecasting error for the executive (CEO) function, operationalized by averaging the 4 individual measures of forecasting error (Pe_i) that relate to the executive function

Table 3 presents the principal characteristics of the five periods in which measurements were taken. These were created by the researchers to ensure that they would be

Table 3:
Principal Characteristics of the Five Periods of Performance Measurement

Measurement Period	Corresponding Period of the Simulation	Characteristics of the Period
1	2	- No readjustments of the macroeconomic conditions - Seasonal demand of product D
2	3	- Suppliers' prices readjustment - Seasonal demand of product D
3	4	- Suppliers' prices readjustment - Seasonal demand of product B, C and E
4	5	- Suppliers' prices readjustment - No seasonal demand
5	6	- No readjustments of the macroeconomic conditions - Seasonal demand of product D

standardized over all the teams participating in the experiment.

The experiment used a retailing simulation with the companies having their shares listed in a simulated stock exchange. The values of these shares were determined by the interaction of company performance and the macroeconomic aspects of the simulation. In SIMCO 5.0 (software used in the experiment), the decisions were made by members of the teams, taking one or more functions (depending on the number of pupils for team). Basically, the functions were chief executive officer (CEO), human resources, finance, and sales (including purchasing and inventory). Responsibility for the 44 decisions was assigned to the various functions, depending on the function to which they were most relevant.

We conducted a pre-test of the experiment during 2007 in an accounting course. During the daily pre-test, students identified problems with some forecast variables (basically, difficulty of interpretation and complexity). We replaced excluded the problem variables prior to the final experiment. The excluded variables were "abnormal loans," "liquidity," "indebtedness," "employees' turnover,", "Return on Investment - ROI" and "accumulated dividends". The experiment was conducted during the first semester of 2008.

RESULTS

Hypothesis H1 posited that the combined accuracy of functional forecasts will be positively related to overall company performance. To test this, we ran the multiple linear regression analysis defined by equation 2 above. The result was an R^2 of 0.4075, indicating the accuracy of the forecasts for the 20 variables accounted for 40.75% of the variance in company performance. This was significant above the .0001 level, strongly supporting the hypothesis.

Hypothesis H2 posits that forecasting accuracy at the CEO level will have the greatest impact on overall company performance. We tested this with the second multiple linear regression analysis defined by equation 3. As suggested by Table 4, this hypothesis was also supported. Forecasting accuracy at the CEO level explained 24.49% of the variance in overall company performance, while the next closest (marketing) explained 17.86%.

Table 4:
Multiple Regression of Company Performance by Management Function

Management Function	R^2
CEO	0.2449
Marketing	0.1786
Personnel	0.1341
Finance	0.0839

CONCLUSION

The results of our study were positive. Forecasting accuracy explained 40.75% of the variance in company performance. , This is particularly significant when we consider the fact that forecast accuracy's effect is expressed through a team's decisions, whereas actual performance is moderated by external factors over which game players have no direct control. We would expect the variance explained by forecasting error to vary across studies, depending on the nature of the simulation algorithms, expressed through the degree of control student decisions had over results versus to the influence of uncontrollable factors, such as competition and macrconomic considerations.

Additionally, we would expect the results for specific management functions to vary according to the complexity of the overall game, the relative influence of other decisions, and the level of the decisions in the overall hierarchy of decision making. In our study, for instance, we found that forecasting accuracy had the greatest impact for the CEO function. This makes sense, because, as the chief executive is responsible for ensuring that all the lower-level decisions work together in support of an overall strategy, the efficacy of which should be determined by the expected results. In our study, the expected results are operationalized as the CEO's forecast of company-level financial outcomes – in

this case, the probably impact it will have on share value, share value ranking, net profit, and margin (see Table 2).

The difference in predictive value of forecasting accuracy for high- versus low-level managerial functions is especially important. It demonstrates the feasibility of operationalizing the forecast-accuracy approach for evaluating abstract, high-level decision making. For lower-level decisions – pricing, for instance – a decision maker is trying to forecast the impact a price increase or decrease would have on demand. This is no easy task, but at least the forecast represents a clear, easy-to-conceptualize relationship. A high-level decision to adopt a broad, high-quality strategy involves fitting every decision the firm makes into an overall strategic pattern, and then relating this pattern to specific financial consequences for the firm. This is a daunting task. If it can be done, however, as suggested by our results, it addresses Wolfe's (1993a) concerns that Teach's forecast-accuracy approach for evaluating student performance might not be appropriate for top-management simulations.

Cast in the light of Figure 1, this is good news. If we rely on company performance to evaluate student performance, we sensitize our students to an element of random reinforcement, in the sense that they are being rewarded for something over which they only have partial control. By contrast, their ability to forecast, as crude as it might be, is something that is totally under their control. From a learning perspective, by using the forecasting-accuracy approach, we are saying, "Work on your analytical skills, and make the best decision available to you." Evaluating them based on company performance is simply saying, "Perform!" Which is most likely to stimulate their development of management skills?

Directions for Future Research: Cognitive Processes

For all our enthusiasm, we must acknowledge that we are still closer to the beginning than the end of the story. While our results are encouraging, they are by no means definitive. In the case of CEO's forecasting accuracy, for instance, we do not really know whether the forecasts were the cause or the result of the strategy-selection process. Figure 1 suggests that they are the cause. But what if managers were making decisions intuitively, based on some kind of *tacit* knowledge (as we discussed earlier in our discussion of Gentry, Stoltman and Mehihoff, 1992), and then inferring results from prior company performance? This is not hard to imagine. `When asked for a forecast, the CEO would simply look at past share values, rankings, earnings, and margins, estimating whether they would go up or down, without ever linking them to strategic decisions. This could account for the Wolfe's (1993a) puzzling finding, that forecasting accuracy actually decreased rather than increased in the second half of the game. If managers were forecasting the consequences of their decisions, and learning in the process, their accuracy should have increased. However, if they were forecasting as an afterthought, based on past results, changing conditions in

the simulated environment could decrease the accuracy of their forecasts.

A similar line of reasoning could explain the fact that low-performing teams forecasted more accurately than high-performing teams in Wolfe's experiment. If the success of high-performing teams resulted from the application of superior *tacit* knowledge, they would have no forecasts available to draw upon and would treat the forecasting assignment as an unrelated and relatively unimportant task, to be dispensed with a minimum of effort. Lacking the requisite *tacit* knowledge to make effective decisions, lower performing teams might have worked more diligently at their analytical tasks, producing better forecasts, but falling short in actual performance.

Another possibility is that the difference between high- and low-performing teams was, as Wolfe (1993a) suggests, in large part the ability of students to select appropriate strategies. What if students can forecast the accuracy of the strategies they select, but not of the strategies they should have selected? Indeed, what if they can't even conceptualize the nature of winning strategies? This would explain why Wolfe found that low-performance might forecast accurately, but it would not explain why high-performance teams did not. Again, our assumption is that successful students select appropriate strategies by anticipating their financial consequences – i.e. by forecasting. However, it is also possible that this is not what happens, but rather, that students use short-cut heuristics, telling them what strategies work in what kinds of situations without ever linking them to financial principles. In a paper addressing this issue, Cannon, Friesen, Lawrence, and Feinstein (2009) cited a paper by Wolfe and Castrogiovanni (2006) in which MBA students were confronted with a simulation to which Duncan's (1972) environmental uncertainty framework provided a potential key for successful strategy development. The students might well have applied the framework without ever translating their strategy into financial projections. Indeed, the nature of these projections would involve a dramatically different skill set.

The point of these scenarios is not to suggest that they represent what happened. We have no way of knowing. But they might have happened. This suggests that we still have a lot to learn. Clearly, one of the areas that needs study is the role of *tacit* versus *explicit* knowledge. Another is the extent to which successful student decision makers use financial projections or qualitative decision heuristics to make decisions. While these address the general theory behind the forecast-accuracy approach to student evaluation, they involve very different thinking processes and call for a very different model of research. The research would begin by looking at students' thinking processes, as portrayed in Figure 1, rather than the outcomes of those thinking processes, as traditionally portrayed in traditional forecasting-accuracy-approach research (Teach Teach (1989, 1993a, 2007; Wolfe 1993a; Washbush 2003).

We have precedents for this kind of research. Indeed, the literature is too vast to review here. Some of the most

promising approaches center on efforts to apply of Bloom's taxonomy of educational objectives (Bloom, Englehart, Furst, Hill, and Krathwohl 1956), and more recently, his revised educational taxonomy (Anderson and Krathwohl 2001; Cannon and Feinstein 2004; Ben-Zvi and Carton 2008; Cannon, Friesen, Lawrence, and Feinstein 2009). This too is a virgin field. While we have made progress in conceptualizing the problems, we have not yet developed operational procedures for applying our conceptualizations in empirical studies, such as would be needed to evaluate the thinking processes that students use when making actual decisions in a simulation game environment.

The problem of *tacit* versus *explicit* knowledge is more difficult to work with than the use of decision heuristics rather than optimizing decision alternatives by forecast outcomes. We tend to assume that *tacit* knowledge results from an internalization of *explicit* knowledge through practice (experiential learning). However, what if students acquire *tacit* knowledge directly, by-passing the conscious processing of formal principles and procedures? Again, this possibility begs more study and understanding.

Directions for Future Research: Affective and psychomotor aspects

A second major area for future research relates to less analytical aspects of decision making in a simulation game environment. One of the most disconcerting critiques of the forecasting-accuracy approach comes from Washbush (2003), who reports research that supports Teach's covariation between forecasting accuracy and company performance. Having provided this evidence, he goes on to question its relevance. He argues that there many other kinds of learning to be gained from simulation game participation beyond the ability to forecast. These include: "taking responsibility for outcomes of decisions; problem finding; identification of key strategy concerns through analysis of performance data; testing aptitude for and desire to manage; assessing personal performance under risk-stress conditions; developing other abilities relevant to an organizational career (p. 252).

Washbush cites additional learning objectives from Wolfe and Rogé (1997), noting that they "have identified a number of important elements of learning common to most popular total enterprise simulations. These include strategy, environmental analysis, forecasting, market development and penetration, cost and differentiation strategies, and performance measures."

If our study truly reflects the kind of student decision process described in Figure 1, we have addressed the link between forecasting accuracy and strategy. By extension, the same argument could be used to address environmental analysis, market development and penetration, cost and differentiation strategies. It does not address issues such as taking responsibility, problem finding, desire to manage, assessing personal performance under risk, and career development. These are all potentially important learning outcomes of simulation game participation, but they are side-effects of trying to win the game. Many of them are driven by the energy, or motivation, playing a game can engender. While the literature discusses the "energy factor," we are in need of an analytical framework for dealing with it (Yakonich, Cannon, and Ternan 1997).

We may dismiss some of these – career development, for example – as important, but not relevant to the decision process described in Figure 1. At the other extreme, however, problem finding is central. How does the decision maker decide what alternatives to consider? What criteria to use when forecasting their consequences? Answering this kind of question requires prioritization and motivation to respond, based on judgments regarding importance, propriety, values, consistency with a larger system meaning, and the relevance of these issues to one's actual behavior. This approach is not addressed by Bloom's *cognitive* taxonomy, either old or revised. Rather, it is addressed in Krathwohl, Bloom, and Masia's (1964) *affective* taxonomy. Depending on the nature of the decision, it also might involve instinctive behavioral responses, as described in Simpon's (1974) *psychomotor* taxonomy. Cannon and Burns (1999) suggest that all three taxonomies should be used when evaluating competencies involved in simulation performance.

As with adaptation of the cognitive taxonomies for evaluating learning, our work has yet to go beyond the conceptual stage. This suggests a second major area for future work. We need to develop empirical studies to better understand how affective, and perhaps, psychomotor processes fit into learning and performance in a simulation game environment, and more particularly, how they relate to students' assessment of the impact their decisions will make, as viewed in the context of Figure 1.

REFERENCES

Anderson, Lorin W. and David R. Krathwohl (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Anderson, Philip H. and Leigh Lawton (1990). The relationship between financial performance and other measures of learning on a simulation exercise. *Developments In Business Simulation e Experimental Exercises*, Vol. 17: pp. 6-10 (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Ben-Zvi, Tal and Thomas C. Carton (2008). Applying Bloom's revised taxonomy in business games. *Developments in Business Simulation and Experiential Learning*, Vol. 35: pp. 265-272.

Bernard, Ricardo R. S. (2004). Assessing individual performance in a total enterprise simulation. *Developments in Business Simulations and Experiential Learning*, Vol. 31: pp. 197-203. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Bernard, Ricardo R. S. (2007). Students' perceptions on the individual managerial performance. *Developments in Business Simulation and Experiential Learning*, Vol. 34: pp. 161-169. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Biggs, William D. (1987). Functional business games. *Simulation & Games*, 18:2 (June), pp. 242-267.

Bloom, Benjamin S., Max D. Englehart, Edward J. Furst, Walker H. Hill, & David R. Krathwohl (1956). *Taxonomy of Educational Objectives -- The Classification of Educational Goals, Handbook I: Cognitive Domain* (New York: David McKay Company).

Cannon, Hugh M. (1995). Dealing with the complexity paradox in business simulation games. *Developments in Business Simulations and Experiential Exercises*, Vol. 22: pp. 96-102. (Reprinted in the Bernie Keys Library, available from http://absel.org).

Cannon, Hugh M. and Alvin C. Burns (1999). A framework for assessing the competencies reflected in simulation performance. *Developments in Business Simulation and Experiential Learning*, Vol 26: pp. 40-44. . (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Cannon, Hugh M. and Andrew H. Feinstein (2005). Bloom beyond Bloom: Using the revised taxonomy to develop experiential learning strategies. *Developments in Business Simulations and Experiential Exercises*, Vol. 32: pp. 348-356. (Reprinted in the Bernie Keys Library, available from http://absel.org).

Cannon, Hugh M., Daniel P. Friesen, Steven J. Lawrence, and Andrew H. Feinstein (2009). The simplicity paradox: Another look at complexity in design of simulations and experiential exercises. *Developments in Business Simulations and Experiential Exercises*, Vol. 36: pp. 243-250. (Reprinted in the Bernie Keys Library, available from http://absel.org).

Curran, Kent E. and Robert W. Hornaday (1987). An investigation of the relationships between formal planning and simulation team performance and satisfaction. *Developments In Business Simulation e Experimental Exercises*, Vol 14: pp. 43-46 (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Duncan, Robert B. (1972). Characteristics of perceived environments and perceived environmental uncertainty. *Administrative Science Quarterly* 17:3 (September), 313-327.

Faria, Anthony J. (1987). A survey of the use of business games in academia and business. *Simulation & Games*, 18:2 (June), pp. 207-224.

Fritzsche, David J. and Richard V. Cotter (1990). Guidelines for administering business games. In James W. Gentry (ed.), *Guide to Business Gaming and Experiential Learning*. East Brunswick: Nichols/GP Publishing, pp. 74-89. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Gentry, James W. (1990). What is experiential learning? In James W. Gentry (ed.), *Guide to Business Gaming and Experiential Learning*. East Brunswick: Nichols/GP Publishing, pp. 9-20. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Gentry, James W., Jeffrey J. Stoltman, and Carol E. Mehihoff (1992). How should we measure experiential learning. *Developments in Business Simulation and Experiential Exercises*, Vol. 19: pp. 54-57. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Hand, Herbert H and Henry P. Sims (1975). Statistical evaluation of complex gaming performance. *Management Science* 21:6 (February): pp. 708-717.

Hornaday, Robert N. and Kent E. Curran (1988). Formal planning, simulation team performance, and satisfaction: A Replication. *Developments In Business Simulation e Experimental Exercises*, Vol. 15: pp. 6-10 (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Keys, J. Bernard. (1987). Total enterprise business games. *Simulation & Games*, 18:2 (June), pp. 225-241.

Krathwohl, D. R., Bloom B. S., & Masia B. B. (1964). Taxonomy of Educational Objectives -- The Classification of Educational Goals, Handbook II: Affective Domain (New York: David McKay Company).

Smith, Jerrold R. and Peggy A. Golden (1991). An unexpected relationship between forecasting and performance. Working paper, University of Louisville. (Cited by Wolfe 1993a).

Teach, Richard D. (1987). "Profit: The false prophet". *Developments in Business Simulation and Experiential Exercises*, Vol. 14: pp. 205-207. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Teach. Richard D. (1989). Using forecast accuracy as a measure of success in business simulations. *Developments in Business Simulation and Experiential Exercises*. Vol. 16: pp. 103-107. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Teach, Richard D. (1990) "Profits: the false prophet in business gaming." *Simulation & Gaming*, 21:1 (March): pp. 12-26.

Teach, Richard D. (1993a). Forecasting accuracy as a performance measure in business simulations. *Simulation & Gaming*, 24.4 (December): pp. 476-490.

Teach, Richard D. (1993a). Forecasting and management ability: A response to Wolfe. *Simulation & Gaming*, 24.1 (March): pp. 63-72.

Teach, Richard D. (2007). Forecasting accuracy and learning: The key to measuring simulation performance. *Developments in Business Simulation and Experiential Exercises*, Vol 34: pp. 57-66. . (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Teach, Richard D. and Vishal Patel (2007). Assessing participant learning in a business simulation. *Developments in Business Simulation and Experiential Learning*, Vol 34: pp. 76-84. (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Thorngate, Warren and Barbara Carroll *1987). Why the Best Person Rarely Wins. *Simulation & Games,* 18:3 (September): pp. 299-320.

Washbush, John B. (2003). Simulation performance and forecast accuracy – is that all? *Developments in Business Simulation and Experiential Exercises*, Vol 30: pp. 250-253. . (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).

Washbush, John B. and Jerry Gosen (2001). An Exploration of Game-derived Learning in Total Enterprise Simulations. *Simulation & Gaming*, 32:3 (September): pp. 281-296.

Wolfe, Joseph (1993a). An Exploration of Game-derived Learning in Total Enterprise Simulations. *Simulation & Gaming*, 32:3 (September): pp. 281-296.

Wolfe, Joseph (1993b). On the propriety of forecasting accuracy as a measure of team management ability: A preliminary investigation. *Simulation & Gaming* 24:1 (March), pp. 47-62

Wolfe, Joseph (1993c). Forecasting and management ability: Rebuttal to Teach's response. *Simulation & Gaming* 24:1 (March), pp. 73-75.

Wolfe, Joseph and Gary Castrogiovanni (2006). Business games as strategic management laboratories. *Developments in Business Simulations and Experiential Exercises*, 33, 31-40 (Reprinted in the Bernie Keys Library, available from http://absel.org).

Wolfe, Joseph and Michael Chanin (1993). The integration of functional and strategy management skills in a business game learning environment. *Simulation & Gaming* 24:1 (March), pp. 34-46.

Wolfe, Joseph and C. Richard Roberts (1993). A further study of the external validity of business games: Five-year peer group indicators. *Simulation & Gaming* 24:1 (March), pp. 21-33.

Wolfe, Joseph and Joseph N. Rogé (1997). Computerized general management games as strategic management learning environments. *Simulation & Gaming* 28:4 (December), pp. 423-440.

Yakonich, David Anthony, Hugh M. Cannon, Amy Ternan (1997). The energy factor: Building motivation in the simulation gaming environment. *Developments in Business Simulation and Experiential Learning*, Vol 24: pp. 329-335. . (Reprinted in *The Bernie Keys Library*, available at http://www.absel.org).