USING FORECASTING ACCURACY AS A MEASURE OF SUCCESS IN BUSINESS SIMULATIONS

Richard Teach, Georgia Institute of Technology

ABSTRACT

This paper describes the results of an experiment that investigated the link between the ability of simulation team participants to forecast the financial and/or market related outcomes and the actual results of their decision making. This experiment was repeated six times over the course of an academic year. Each replication involved 12 consecutive decision periods. The experiment required each member of the student teams to forecast either the expected market share and sales of the product for which they were making decisions or to forecast the cash flows and profits of the simulated firm. The managerial position a student held during the simulation determined the type of forecast (market share, cash flow, etc.) he or she would be assigned. The experiment showed a very strong link between the ability of the management team to forecast outcomes and their firms performance, as measured by profitability.

THE PROBLEM

A fairly large literature base exists on attempts to predict simulation performance using a priori variables. (For an excellent discussion and reference list, see Gosenpud, 1987, also Hornaday and Wheatley, 1986.) Most of these studies do not measure the learning that takes place while the game or simulation is in progress. Wolfe and Box (1988) studied the cohesion which takes place while playing a business simulation and linked this variable to performance, and Wolfe and Chacko (1983) researched team size as it affected simulation performance. Gosenpud and coauthors (1983, 1984 and 1985) have produced several papers on predicting simulation performance while Hornaday and Curran related decision styles to performance (1987 and 1986).

There is not common agreement on what should, or could, be used as success criteria when evaluating the results of a business simulation. Last year, House and Napier (1988) wrote on performance measures in business simulations and compared them to performance measures in 18 food/tobacco /cosmetic companies. Their measures were almost all profit derived. Anderson and Lawtons (1988) research concentrated upon assessing performance in business gaming. Their research centered around relating the learning objective and the assessment process. The methods of evaluating performance discussed in Anderson and Lawton’s paper were generally not profit measures. However, they did devote a paragraph to the possibility of using the ability to predict results as a measure of performance evaluation. They stated, “In a competitive business simulation, accurately forecasting market share, profit as a percentage of sales, and ROI reflects a teams ability to translate its decisions into simulation outputs in the midst of a dynamic environment. Without this ability, efforts by the team to plan its future activities are of little value." (page 242)

The evaluation of the participants in business simulations is frequently based upon the relative profitability (or derivatives of profitability) of the firms These different forms of prof. liability may include cash flow, market share, total sales, return on sales, stock price or any of a large number of measures. If various forms of profitability are to be used as performance criteria, then several very artificial and unrealistic conditions must exist. First and foremost, all firms start with exactly the same product(s), asset structures and economic histories. Second, the same parameters that guide the competition among the firms must equally apply to all firms. These conditions do not exist in competitive environments. The fact that profitability is used as an evaluative measure limits the richness of the competitive experience and reduces the knowledge that may be gained from participating in a business simulation (Teach, 1987).

This paper explores using the ability to forecast the outcomes of decisions made in a business simulation as a surrogate for performance as measured by firm profitability. If one is able to use measures such as forecasting, which are independent of firm size, asset composition competitive advantage, and efficiency of operations, then business simulations could be written which have multi-product firms competing with single product firms. Competition could exist between cash rich and cash poor firms and between firms with wildly different competitive structures. These structures are much more realistic and would enlighten students as to the very nature of actual competitive environments. In the soft drink industry, all firms are not Coca Cola. Royal Crown and even Double Cola coexist in this competitive market with Coke, in spite of the fact that they have very different asset structures and products which compete in different market segments or niches.

The central theme of the research described in this paper is the concept of linking the forecasting ability of the management team to the performance of the firms which the students manage. Pickett and Stell (1987) wrote of teaching forecasting but did not link forecasting to simulation results. In the same year, Chiesl (1987) linked forecasting to business games but did not tie the forecasting results to game performance.

There appears to be no strong consensus on a “best" performance criterion. However, the most common single criterion seems to be a firms relative profitability at the end of play. This forces the imposition of simplistic restrictions on business simulations and limits their realism and opportunities to create better learning environments. An interesting paper which initiated the discussion of measuring performance when teams operate in a business simulation where the starting conditions differ among firms was published by Pray and Gold (1987).

If one examines the evaluation of actual managers or management teams, the competence or ability of the managers is rarely based upon the relative profitability of their firm within its industry. This is because the corporate cultures, asset structure, and market niches in which each firm operates tends to be unique. Therefore, forget about trying to find a best success criterion and focus on the proficiency with which managers (or student simulation participants) execute a critical management process which is highly associated with firm success, regardless of how success is define. This *critical managerial process* is forecasting.

THE EXPERIMENT

This paper reports the results of an experiment run three times during three, eleven week academic terms in which junior level students participated in a marketing simulation. In addition to the regular simulation decisions, the students were required to provide forecasts of the expected outcomes based upon their decisions and the anticipated decisions of their competitors. The accuracy of the forecasts were compared to the competitive performance of the firms in the simulation.

The 5th edition of *Marketing in Action* simulation (Day and Ness. 1986) was used throughout the three terms of this experiment. This particular simulation uses a soft drink scenario, where six or fewer firms, each producing up to three soft drinks, compete in an industry. If more than six teams are required during a term, additional regions are run, each independent of the other. If a team produces all three products, the simulation requires a minimum of 27 decisions per period. There are eight decisions associated with each product and three company-wide decisions. In addition to the required marketing and manufacturing decisions, a wide variety ot marketing research may be requested. One decision period equates to a simulated month.

During this experiment, the students were divided into four-person teams. Twenty-nine teams were formed over the academic year time period and these teams competed in six separate competitions. During the fall term, there were twelve teams divided into two regions of six firms each. In the winter term there were ten teams, divided into two, five-firm regions. The spring term had two regions of three and four firms, respectively. Each competition consisted of a trial set of four decisions and a competitive run of 12 decisions, simulating one years activities in the soft drink industry.

The organizational structure of these teams was imposed by the researcher. Each team was composed of a president and three brand managers. The brand managers were responsible for their respective brand decisions. The president was to settle internal disputes, monitor the cash flows along with the profitability of the firm, and be responsible for having the decisions turned in on time.

Each member of the team was required to produce two forecasts for each decision period. Each brand manager was to forecast the expected market share of his/her brand (the sales of the student's brand divided by the total sales of all competing brands, including all flavors) and the total market's expected unit sales of the flavor for which the brand manager was responsible. The president was required to forecast the net cash flows during each decision period and the profit or loss expected to occur as of the end of the period.

In order to get the attention of the players and to assure that reasonable efforts were made in producing the forecasts, a grade was assigned to the forecasting accuracy. This grade assigned 10 forecasting accuracy was weighted as one-sixth of the total grade in the course. However, no part of anyone's grade was based upon their teams performance as defined by the individual firms profitability. While, on the surface, this may seem to reduce the pressure to maximize profits, the natural competitiveness of the students to run the most profitable firm was paramount in their minds. Whenever the profits by team were posted, every team had a representative on hand to either cheer or jeer the firm that had the greatest profits during the particular posted period. The remaining five-sixths of the students grades were based upon both oral and written reports and a peer evaluation. The total profits of the simulated firm was not used as an evaluation criterion.

FORECASTING ACCURACY

In the *Marketing in Action* simulation, under cateris paribus conditions, the three flavors of soft drinks have inherently different market shares. The parameters for the cola flavor initially give it about 50 percent market share, diet cola has 30 percent and lemon lime obtains the remaining 20 percent with a long-run tendency for the diet cola to gain market share at the expense of the regular cola flavor. These percentages, however, may differ fairly widely, depending upon the decisions the competitive teams make during the periods the simulation is run.

In order to compare the forecast accuracy of the brand managers across all the brands, a relative error term was calculated. Relative error was defined as:

$$Re = \left| \frac{Fv - Av}{Av} \right| \quad (1)$$

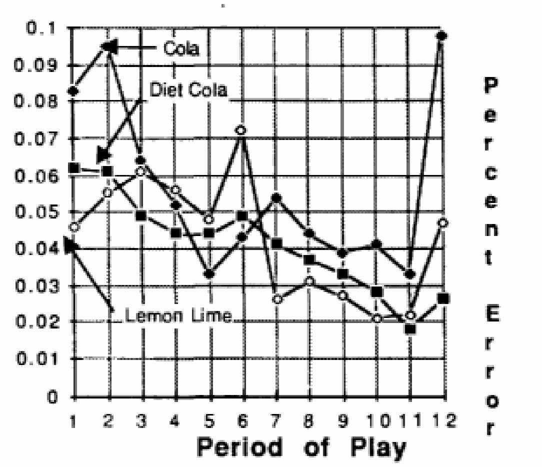
Where: Re = Relative Error (a positive value)
Av = Actual Value
Fv = Forecast Value

Note that this term has a lower bound of zero, but no upper bound. In a few cases in early periods of several simulations, mistakes were made in the term market share. A few brand managers thought that market share was to be estimated based upon individual flavors and not total market sales. In addition, some presidents got carried away with profit and/or cash flow projections. As a result, a few of these relative error terms exceeded 1 .0. None of the relative error terms exceeded 1 .0 after the third round of forecasting. For the purposes of this report, all the error terms above 1.0 were rounded down to 1.0

It was anticipated that the relative errors in forecasting small market shares (primarily the lemon lime brands) would be much higher than the relative errors in estimating the high market shares. This was expected simply because the denominator, or the actual market shares of lemon lime, was a smaller value than the large market share brands. As can be seen in Figure 1, this was not the case. The relative error terms of all three brands had strikingly similar patterns. The initial, high relative errors can be attributed to

FIGURE 1

Average Relative Errors in Forecasts of Market Share
(Each brand averaged over 29 brand managers)



almost all the firms, assuming they would, or could, gain a substantial market share advantage the first few rounds of the competition.

One should note the jump in error for the last period of the simulation. It was assumed that "end play led to this dramatic change in the ability to forecasts the outcome of decisions. It has been the author's experience that students frequently make rash decisions for the last period, the reasons for which are either WAPs or SITTOs and known only to themselves.

While market share lakes all flavors and all competitors in consideration, the forecasts of case sales was only for the flavor in question. The brand manager was to forecast the total number of cases of his/her flavor that was expected to be sold. Figure 2 shows the relative errors in forecasting case sales.

The average error rate declined almost every period. By the eleventh period, the average error in forecasting case sales was below five percent for all three brands. As shown in the previous figure, the end play affected the ability to forecast the final period. It is evident in Figures 1 and 2, that the ability to forecast gets better with practice. The more practice. the better the forecasting results.

The presidents had to forecast cash flows and profits. Because both of these values are dependent upon the abilities and forecasts of the brand managers, the presidents' errors tended to be larger. Information had to be exchanged between the brand managers and the presidents. In order to enhance this upward passing of information, the evaluations of the brand managers' forecasting abilities were independent of the evaluations for the presidents forecasting accuracy. Figure 3 displays the average accuracy of the 29 presidents in forecasting cash flows and profits over the to rounds of the six simulation runs. Again, the forecasting abilities increased with practice, with the exception of the final period.

FIGURE 2

Average Relative Errors in Forecasts of Case Sales
(Each brand averaged over 29 brand managers)

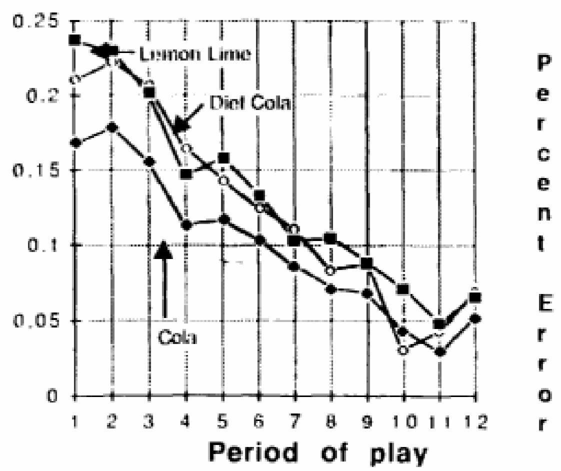
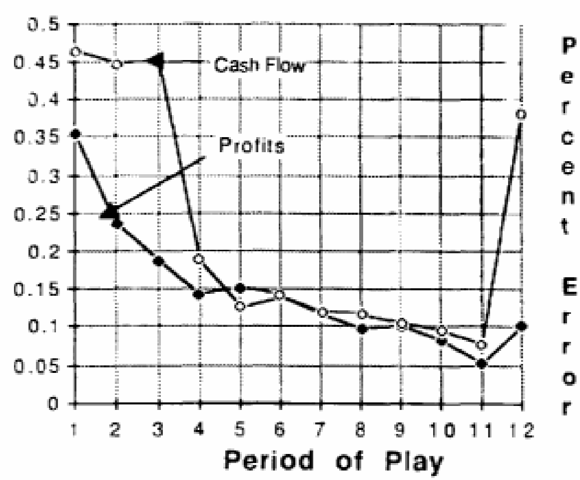


FIGURE 3

Relative Errors in Forecasts of
Cash Flows and Profit
(By 29 Presidents)



PROFITABILITY AND FORECASTING ERRORS

From the previous discussion, it can be seen that the participants were earning how to forecast. The question still remained, however, 'Was forecasting accuracy related to the firms performances during and at the end of the simulation?"

The relative errors across all eight forecasts (three market share, three case sales forecasts and a cash flow and a profit forecast) made by a loam were added together This sum of relative errors was associated with the firm's profit, period by period. In this way, a two-tuple, relating a firm's profit to the errors in forecasting *existed* for each firm, every period.

It was first thought that plotting these values would reveal a relationship between forecasting accuracy and profits that showed smaller errors were associated with higher profits This was not the Marketing in action has a significant seasonal pattern and a small general economic cycle built into it. these exogenous economic variables caused the companies profits to peak about halfway through the simulation In addition. the total profitability varied dramatically by region and by academic term, due to the fact that decision making and competitive pressures differed as the student teams differed. Thus, across simulation runs, profits and the ability to forecast were not related. However, if the teams held constant and if comparisons were made within single region, then the profits and forecasting abilities were very closely related. The relationship between forecasting abilities of the student management teams and firms' profits was replication dependent.

Determining how to measure the relationship

Since simple and direct measures of the relationship between profitability arid forecasting errors is not possible, because it is replication dependent, a different method of comparing profitability and forecasting error had to be devised one that accounted for the different competitive simulations. Both the summed forecasting errors and the firms profitability by region and academic term were rank-ordered. Thus, each term and region was treated as a replication. The rank orders were aggregated into a matrix, where the rows and columns both equaled largest number of teams in the six replications. For any one simulated period, the rows represented the rank order of the summed forecasting errors and the column represented the rank Order of the firms profit during that period for all replications. Table 1 provides an example, displaying the results of the fifth simulated run for all 6 replications.

TABLE 1

A Matrix of Ranked Summed Forecasting Errors
and Firm Profits
for period 5

Rank order 1 Summed Forecasting Errors	Rank order of profits					
	1	2	3	4	5	6
1	0	2	2	1	1	0
2	1	1	1	1	1	1
3	3	1	2	0	0	0
4	2	0	1	1	0	1
5	0	1	0	1	2	0
6	0	1	0	0	1	0

Read this cell as: Out of the six replications for period five, no firms whose summed forecasting errors were the smallest also produced the highest profits.

There were six, first, second and third place firms (one for each replication), five, fourth place firms, four, fifth place firms and two, sixth place firms. A region with only four firms can have no firms ranking lower than fourth.

Note that by producing a matrix as shown above, comparisons can be made across simulation replications without the same number of teams per replication. In the above example, only two of tie six replications had six teams.

The next step was to determine how to compare the matrix from one period to another. If the relationship between forecasting abilities and firms' profits were perfectly monotonic, that is, the rank order of the summed forecasting errors was identical to the rank order of the profits, all the observations would fall on the diagonal. The farther from the diagonal an observation is, the worse the fit. It was decided to weight the deviations from the diagonal by the square of the number of places the observations deviated from the diagonal: an analogy to the "squared error" concept.

$$Df = \sum_{i=1}^{\#\text{ of rows}} \sum_{j=1}^{\#\text{ of Columns}} N_{ij} * (i-j)^2 \qquad (2)$$

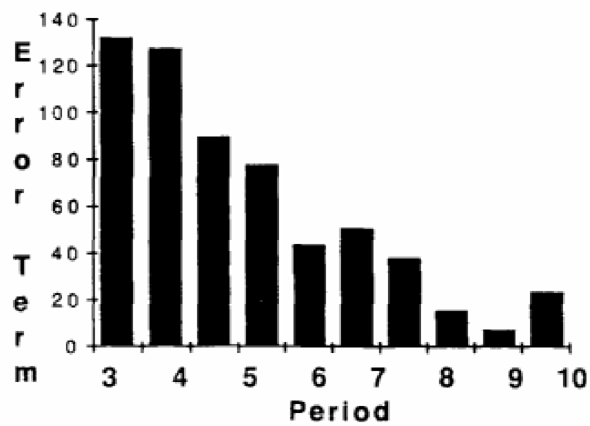
Where:

Df = Degree of fit
Nij = The number of observations in the i^{th} row and j^{th} column.
(i-j) = the difference in the subscripts or the number of places away from the diagonal position.

Figure 4 displays the values for the degree of fit over the twelve simulated periods for the six replications involving 29 teams.

FIGURE 4

The Degree of Fit Between Forecasting Errors and Profitability by Period



From the results shown in Figure 4, it is apparent that forecasting errors get smaller over time and the relationship between profits and forecasting ability gets stronger as the number of simulations runs increase. But the question remains, "Does the firm that forecasts better make more profits over the entire period of play?"

One could simply sum up all the forecasting errors over the number of plays 10 obtain a total forecasting error and compare the sum of the forecasting errors to the sum of profits. However, this would weigh errors made during the early part of the simulation the same as errors made during the final run. To prevent equal weighting, an exponential smoothing technique was applied to the summed errors. Exponential smoothing is a recursive technique, which differentially weighs the terms based upon when they were estimated. This process can be summarized as:

$$\sim E_t = W * E_t + (1-W) * (\sim E_{t-1}) \qquad (3)$$

Where:

$\sim E_t$ = Estimated error term for period t.
$\sim E_{t-1}$ = Estimated error term for period t-1.
E_t = Actual error in period t.
W = A weight between one and zero.

This recursive relationship can only be established at the beginning of the second period. The ~Et-1 for period one is equated to the actual error made in the first period. Thus, ~Et-1 equals Et, and since the weight W must be between one and zero, the actual error in the first period becomes the estimated error for the second period. If W is set equal to one, then the estimated error for the period t is always equal to the actual error in the period
t-1. As W declines from 1, the errors in estimation made during prior periods are used in producing the estimate for the current period. If W ever equals zero, the estimated error for the current period always equals the actual error made during the period when t=1.

Table 2 displays the ranked smoothed forecasting errors when W was set equal to .5. Several values of W were tried and the resulting tables remained quite stable as long as the values were greater than or equal to, .3 and less than or equal to, .9. It was originally thought that using a W value of one, which places all the weights on the last error term, would produce the same result, but the "end play", as shown in Figures 1 2 and 3, affected both the profitabilities of the firms and the teams' forecasting abilities during the last round to such a degree that the errors in the last period did not accurately reflect the forecasting abilities of the participants.

TABLE 2

Ranked Smoothed Forecasting Errors vs the Ranks of the Sum of the Firms Profits (at the End of the Simulation)

Rank order of Smoothed Forecasting Errors	Rank Order of Sum of Profits					
	1	2		4	5	6
1	5*	1	0	0	0	0
2	1	4	1	0	0	0
3	0	1	5	0	0	0
4	0	0	0	4	1	0
5	0	0	0	1	2	1
6	0	0	0	0	1	1

Read this cell as: "For the six replications of the simulation experiment, five out of the six firms which had the smallest smoothed error terms also produced the greatest profits"

Table 2 clearly indicates that in a simulated environment, there is a very strong relationship between the ability of a simulation team to forecast and the relative profitable of the firm which the team manages. If one assumes that comparative firm profitability is an effective measure of comparative managerial competence, then measuring forecasting abilities of the management team may be substituted for direct profitability measures. In evaluating these results, it is important that one notes that the above analysis reduces relationships to monotonic ones. If the first and second place teams were only separated by a few dollars or by a large amount the rankings were the same, one first place firm and one second place firm. The same holds true for the size of the forecasting error.

THE CONCLUSIONS

This paper reported the results of a single set of experiments using a single business simulation *(Marketing in Action),* one that concentrates in the marketing function only. These experiments were conducted in marketing classes on a campus which has very bright, analytical students. (Students in this university have SAT scores which average just under 1200.) In this limited case, it is evident that measuring the forecasting accuracy across eight separate forecasts per period would yield very similar results to directly measuring the total profitability of the firms. This relationship between forecast accuracy and firm performance was true *for* both the period by period results and for the results summed over the total 12 simulated periods.

If these results hold for total enterprise simulations, other functional area games and for students with a wide range of abilities, then much more diversity can be designed into business simulations. Foe example, business simulations could have both small and large firms, each with different amounts of resources, competing in the same market place. Product differences, some with unique market segments, could be devised for some teams. while others in the same simulation could offer generalized products with mass appeal. In short, if these results can be generalized, gaming can represent the complex nature of competition in a more realistic manner than is currently the case.

Before this leap of faith to measure simulation results by forecasting accuracy takes place, more experiments similar to this one must take place. The author hopes that after this trial, others will attempt lo replicate the process using total enterprise business simulations and new forecasting variables. A lot of additional experimentation needs to be done with confirming results before there is a generalized equation between forecasting accuracy and simulated firm profitability.

REFERENCES

Anderson, Philip and Leigh Lawton' "Assessing Student Performance on a Business Simulation Exercise", Developments in Business Simulation and Experiential Exercises, Vol 15. Edited by Patricia Sanders and Tom Pray, Oklahoma State University, Stillwater, Oklahoma,. 1988, pp. 241-245.

Chiesl, Newell, "The Use of Simple Forecasting Techniques During an Interactive Computerized Business Game", Developments in Business Simulation and Experiential Exercises, Vol. 14. Edited by Lane Kelly and Patricia Sanders, Oklahoma State University, Stillwater, Oklahoma, 1987, pp. 39 - 42.

Gosenpud, Jerry. "Research on Predicting Performance in the Simulation", Developments in Business Simulation and Experiential Exercises, Vol. 14. Edited by Lane Kelly and Patricia Sanders, Oklahoma State University, Stillwater, Oklahoma, 1987, pp. 75 - 79.

Gosenpud. Jerry and P. Miessing, "Determinants of Performance in the Simulation', Developments in Business Simulation and Experiential Exercises, Vol. 10, Edited by Lee Graf and D. M. Currie, 1983, pp. 53 - 65.

Gosenpud, Jerry, c. Milton and A. Larson, "Predicting Performance over the Course of the Simulation", Developments in Business Simulation and Experiential Exercises, Vol 12, Edited by James Gentry and Alvin Burns, 1985 pp. 5 - 10.

Hornaday, Robert and Kent Curran, "Decision Styles and Student Simulation Performance". Developments in Business Simulation and Experiential Exercises. Vol. 14. Edited by Lane Kelly and Patricia Sanders, Oklahoma State University, Stillwater, OK, 1987. pp. 95 - 99.

Hornaday, Robert and W. Wheatley, "Four Factors Affecting Group Performance in Business Policy Simulations" in Developments in Business Simulation and Experiential Exercises, Vol 13. Edited by Alvin Burns and Lane Kelly, Oklahoma State University, Stillwater, Oklahoma,. 1986. pp. 17 - 21.

House, William and Herman Napier, "A Comparative Study of Strategic Performance Factors in Actual and Simulated Business Environments' In Developments in Business Simulation and Experiential Exercises, Vol 15. Edited by Patricia Sanders and Tom Pray, Oklahoma State University, Stillwater, Oklahoma,. 1988. pp. 50 - - 55.

Pickett, Gregory and Roxanne Stell, "Teaching Forecasting to the Masses: Little Decision Support Training. But a Lot of Answers in Developments in Business Simulation and Experiential Exercises, Vot. 14. Edited by Lane Kelly and Patricia Sanders, Oklahoma State University, Stillwater, Oklahoma, 1987, pp. 165 - 188.

Pray, Thomas and Steven Gold' "Goal Setting and Performance Evaluation with Different Starting Positions The Modeling Dilemma in Developments in Business Simulation and Experiential Exercises, Vol. 14. Edited by Lane Kelly and Patricia Sanders. Oklahoma State University, Stillwater, Oklahoma, 1987, pp. 169 - 174.

Ness, Thomas and Ralph Day, Marketing in Action, 5th Edition (The microcomputer version), 1986. Richard D. Irwin, Inc., Homewood, IL.

Teach. Richard D., Profits: The False Profit in Developments in Business Simulation and Experiential Exercises, Vol. 14. Edited by Lane Kelly and Patricia Sanders. Oklahoma State University, Stillwater, Oklahoma, 1987. pp. 205 - - 207.

Wolfe, Joseph and Thomas Box, "Team Cohesion Effects on Business Game Performance", Simulation and Games, Vol 19, Number 1 March, 1988, pp. 82 - 98.

Wolfe, Joseph and T. L Chacko, "Team-Size Effects on Business Game Performance and Decision Making Behaviors, Decision Science, Vol 14, 1983. pp. 121 133.