Developments in Business Simulation & Experiential Exercises, Volume 14, 1987

GOAL SETTING AND PERFORMANCE EVALUATION WITH DIFFERENT STARTING POSITIONS - THE MODELING DILEMMA -

Thomas F. Pray, Rochester Institute of Technology
Steven C. Gold, Rochester Institute of Technology

ABSTRACT

This paper focuses on design and modeling issues for computerized business simulations which are primarily used to teach strategic planning. A review of ten popular micro and mainframe strategy-oriented simulations demonstrates that most do not formally integrate goal setting as a component of the simulation and that all simulations reviewed had each firm start with exactly the same financial and operation base.

The paper purports that even with the performance measurement problems associated with different starting positions, strategy-based simulations should permit each firm to set their owns goals and objectives and that the software should contain performance measures. These measure(s) should indicate not only if the firm reached its objective(s), but how well the team performed relative to other firms in the industry.

A statistical methodology is presented where firms with different starting positions can set and weight their own goals, and then be rated on whether they attained their goals. Two summary measures of goal difficulty are presented, one based on absolute difficulty and the other on relative difficulty. Chebyshev's inequality is then used to partition goals into three levels of difficulty: low, medium and high.

While the statistics of comparative performance put forth in the paper do not provide a strict ranking, they do segregate firm's performance and goal difficulty into different groups, and they should initiate a discussion of how to measure performance when teams have different starting scenarios.

INTRODUCTION

Over the past few years there has been an increase in the number of papers, presented at A.B.S.E.L. conferences, which have been directed specifically toward "improving the design and the modeling of computer-based simulations". This paper is also concerned with design and modeling issues of simulations, but confines itself to computer-based simulations which are integrative in design, and are primarily used to teach business policy and/or strategic management courses.

A review of existing strategy-oriented simulations showed a disturbing lack of formal goal setting and performance evaluation measures. It was also found that in every simulation reviewed all firms started with identical financial and operating bases. A reason often stated for this approach was "then no team would have an advantage over another." But is this realistic or the most appropriate design for teaching integrated business policy? aren't there pedagogical advantages to be gained by having firms start with diverse financial and operating bases?

When the learning objectives focus on strategy issues, a simulation model that is reasonably close to the "real world" would seem to be preferred. One way to make simulations more realistic is to have each firm start with different bases. But this approach raises a number of problems with evaluating goals and assessing the relative performance of competing firms in an industry. With different starting positions, one firm might have the "lion's share" of the market, but might not be as profitable as some of the smaller competitors. Another scenario might be an industry composed of two or three medium-sized firms, some profitable and some not, in a market composed of mainly small firms. But no matter what the scenario, each firm would be in a different position at the start of the simulation and would need to develop a different set of goals and objectives. With different staring scenarios, a number of difficult design and performance evaluation questions occur:

- How do we formally incorporate individual team goal setting into the computer simulation ?
- How will the algorithm assess whether the teams reached their desired goals?
- How will the algorithm know whether the goals were too difficult to attain, or whether the teams' goals were too easy?
- How will the teams' overall performance be summarized and ranked, after playing a number of periods?
- How will the administrator "stay on top's of each team's performance relative to their own goals and objectives, and relative to the other teams?

When simulations are used in a traditional context such as a semester course in business policy, the instructor may have time to answer

these questions after thoroughly studying each team's goals and objectives, the decisions each period of play, and the results obtained. But with different starting positions this can be both extremely time consuming and a challenging task. When simulations are used in two-day management development seminars or in residential week-long programs for graduate or undergraduate students, the time constraint prohibits such detailed administrator analysis. Thus, there appears to be a need for flexible and formal goal setting, and reliable summary measures of performance as a part of the simulation software.

PURPOSE

The purpose of this paper is fourfold:

(i) to review a representative number of mainframe and microcomputer-based simulations that are primarily used to teach business policy and/or strategic management issues, and see how they address team goal setting and what summary performance measures were embodied in their computer algorithms.

(ii) to detail some of the performance measurement problems that have to be addressed if a simulation is going to start each team with a different financial and operating base.

(iii) to initiate a discussion on how to measure performance when simulations have teams starting with different financial and operating bases.

(iv) to present a "preliminary" quantitative and statistically-based approach which permits individual team goal setting and then measures team performance in both absolute and relative terms.

REVIEW OF LITERATURE

Ten general management and business policy-type simulations were selected to illustrate (i) how the designers measured and summarized team performance, (ii) whether or not individual teams could formally set goals and objectives in the simulation, and if goals could be set how they were tied to formal evaluation methods, and (iii) if different starting positions were a part of the software. summary of the survey is presented in Table I - Goal Setting and Performance Measures in Policy Games.

TABLE 1: GOAL SETTING AND PERFORMANCE MEASURES IN POLICY GAMES

SIMULATION	PERFORMANCE MEASURES	GOAL SETTING	STARTING VALUES
TEMPOMATIC IV	Multicriteria and ranked. Profits, sales, stock price, etc. Weights-instructor assigned	INFORMAL	IDENTICAL
DECIDE	Single criterion and ranked. Stock Value- Profit, debt, dividend policy.	INFORMAL	IDENTICAL
EXECUTIVE SIMULATION	No single measure. Reports: ROI, stock price, profits, sales, etc.	INFORMAL	IDENTICAL
MULTINATIONAL MANAGEMENT GAME	No single measure. Reports: ROI, stock price, profits, sales, etc.	INFORMAL	IDENTICAL
MICROMATIC	No single measure. Reports: profit, stock price, ROI.	INFORMAL	IDENTICAL
THE BUSINESS GAME	Multicriteria and ranked. ROI, divid, ratios, social conditions, quality, and personnel objectives. Winning standings based on largest weighted points.	FORMAL	IDENTICAL
THE BUSINESS MANAGEMENT LABORATORY	No single measure. Reports: Stock price, earnings, etc.	INFORMAL	IDENTICAL
MANAGING A DYNAMIC BUSINESS	Single criterion - profit. Other measures presented.	INFORMAL	IDENTICAL
BUSOP	Single criterion -profit Other measures presented.	FORMAL	IDENTICAL
ENTERPRISE: A SIMULATION	No single measure. Emphasis on profit.	INFORMAL	IDENTICAL

The review shows there is no one common method for measuring team performance. Some simulations used single measures such as cumulative profitability, stock values, or return on investment, while others opted for multicriteria methods where weights were assigned to the different criteria. A number of the business policy games did not recommend any one measure, instead left the task of performance evaluation to the individual teams and/or instructor. Nine out of ten of the simulations did not attempt to come up with a measure for "winning." Only one simulation, (6], presented a measure that could be used to determine relative position of each firm and the overall "winner" at the end of game play. Its performance measure used a multicriteria function, which included traditional financial and profitability measures, as well as, social conditions and personal objectives. Each criterion had assigned weights and points. During game play or at the end, the team with the largest number of weighted-points would be declared the winner.

Almost all of the simulations discussed the need for policy formulation and goal setting at the outset of the simulation. In some, such as [3] and [10], there were sections in the participant manual which provided direction on how to establish policy and set goals. One simulation provided questions and worksheets to be handed in to the instructor to insure that strategic planning was an integral part of simulation play. In general, however, most of the simulations reviewed did not formally integrate goal and objective setting as part of the simulation algorithm or as a formal part of the simulation process. Two of the ten simulations, [6] and [9], had a more formal role for goal setting in that they required teams to fill out business plans and/or submit reports concerning their goals and objectives. One simulation (6] allowed firms to set goals and formally weighted those goals in the performance evaluation.

One of the most interesting findings pertains to the starting positions of the firms. All ten of the business-policy simulations had the firms start with the same financial and operating values. However, to bring more realism into simulations by permitting firms

to set diverse goals and coupling that with different starting positions causes "a modeling dilemma".

THE PROBLEM OF MEASURING PERFORMANCE

In executive training seminars using business simulations, a frequently asked question is:
"How well are we doing?" or" Who did the best?" Participants not only expect a response but generally expect an immediate response. But as discussed earlier, it is not always clear how to judge performance among competing firms, especially in a realistic business setting in which:

(1) Each firm begins at a different starting point;
(ii) Each firm develops its own goals.

To illustrate the performance measurement problem, lets assume there is an industry in which two firms (say A and B) have the following initial market shares and goals:

Now assume both firms achieve their respective goals, that is, firm A increases market share from 5% to 10%; and firm B increases market share from 50% to 55%. In this case, if one were asked to assess the performance of firm A compared to firm B, what could one say? Did both firms

FIRM	INITIAL MARKET SHARE	GOAL
A	5.0%	10.0%
B	50.0%	55.0%

perform equally well since each achieved their respective goals? But were both goals of equal merit? In an absolute sense, both goals were equal, and market share increased by 5% for each firm. But is it a greater achievement to go from a market share of 5% to 10%, or is it relatively more difficult to increase market share beyond 50%? A case can be made for either position. An increase in market share from 5% to 10% represents a doubling of sales volume, certainly an impressive achievement. However, increasing market share from 50% to 55% could be equally as difficult even though it only represents a ten percent increase in sales volume. After all, for firm B to increase market share by this amount, it may have to eliminate a small firm, like firm A in this example, completely from the market. Small firms may fight very hard to maintain their shares of the market. A decline of 1% in market share for firm A represents a 20% decline in sales volume.

The dilemma deals with absolute versus relative changes. Is it more appropriate to use absolute or relative changes to measure performance? Perhaps both measures should be utilized since a case can be made for either approach. But what can one say about a firm that achieves a greater relative change but a smaller absolute change in a performance measure, or vice-versa? For instance, modifying the example above, suppose firm A increased market share by only one percent, from 5% to 6%; while firm B again increased market share from 50% to 55%. In this case firm A still has a larger relative change in market share (i.e 20% for firm A versus 10% for firm B); but firm A now has a smaller absolute change in market share compared to firm B (i.e. 1% versus 5%, respectively). Consequently, the relative measures contradict the absolute measures of performance.

There are additional difficult questions in evaluating and comparing performance among different firms in an industry. Some of these include: How does one deal with multiple objectives? Generally, firms try to achieve a set of objectives. Some objectives may be realized while others may not. How does one compare performance in this case? For example, suppose firms have as their objectives both market share and ROI. What if firm A achieves its market share objective but not its ROI objective; while firm B achieves it ROI objective but not its market share objective. What can we say in this case? Did both perform equally well since each achieved one objective out of a set of two? Complicating matters further, what if the firms had different objectives? What if firm B did not set a market share objective, but rather looked at ROI and debt? What type of comparative assessment of performance could one make in this circumstance?

A SUGGESTED EVALUATION CRITERION

Although there are a number of difficult questions that may not have answers, comparative measures of assessment are still necessary; especially when using business simulations in executive training programs. Therefore, we will propose a preliminary approach for evaluation, recognizing the difficulties mentioned previously and realizing there are shortcomings with any measure of comparative assessment.

The evaluation method is composed of three parts. First, firms set their own quantifiable goals. These end-result goals may be expressed in a variety units such as dollars, inventory turnover, share of market percentage, etc. In addition to establishing goals, firms then assign "weights of importance" to each of their goals. The second part of the method then calculates two statistically-based "measures of difficulty." One is based on relative goal difficulty and the other in terms of absolute difficulty. The third step uses Chebyshev's Inequality to partition the performance objectives and results into three groups, (low, medium and high) based both, on the difficulty of goals set and performance results obtained.

The suggested evaluation criterion will be illustrated with a simplified example. For this example, it is assumed that firms can select among a set of three possible

objectives: market share, ROI, and debt. (It is assumed firms want to minimize debt.) Firms then have the discretion to determine the importance or weight of each objective and the value of the goal it wants to achieve in relation to each objective. Table 2 -MEASURING PERFORMANCE illustrates the choices and weights selected.

TABLE 2: MEASURING PERFORMANCE

	Goal Setting & Evaluation			
Objective	Weight	Goal	Result	Rating
Share	.40	10%	7%	70%
ROI	.50	60%	20%	33%
Debt	.10	30%	25%	100%
	1.00			
OVERALL				55%

As demonstrated by the weight, the highest priority objective established by the firm was ROI, with a weight of 50%, while the lowest priority was the Debt ratio which was given a weight of only 10%. This implies that the firm views the ROI as being five times as important as the Debt objective. The Goals for each of the objectives, the Results, a weighted average Rating are also listed. The firm has set a Goal of 10% for the Market Share but according to the results has only achieved a 7% Market Share. The Goal for the ROI was set at an optimistic 60% but the results indicate that only a 20% ROI was realized.

The Rating measures how closely the Objective is being meet. The Rating for Market Share indicates that 70% of the Goal has been meet; whereas the Rating for ROI indicates that only 33% of the Goal was meet. The rating for Debt was constrained to 100%, since the Goal was exceeded.

The Overall Rating of 55% is the weighted average of the Ratings for each Objective (e.g. 55% = 0.40x70% + 0.5x33% + 0.10 x 10% The interpretation of the Overall Rating is that 55% of all objectives are being meet (taking into consideration, of course, the relative importance of each objective).

The difficult question is how this measure of Overall Rating can be used in a comparison sense with other firms in the industry. Some firms may choose goals that are easy to achieve and, consequently, obtain very high Overall Ratings; whereas other firms may choose goals that are quite difficult to achieve and obtain very low Overall Ratings, despite the fact that they may have advanced significantly in market share, ROI, and debt. Therefore, some consideration of the difficulty of the goals must be incorporated into the Overall Rating.

A set of Difficulty Indices are included in Table 3 - PERFORMANCE EVALUATION WITH DIFFICULTY INDICES. The Difficulty Indices measure the number of standard deviations the firm's Goal varies from the mean Goal for the industry, both in an Absolute and Relative sense. For the Absolute Difficulty Index, the absolute difference between each firm's starting point and it's Goal is calculated. The mean difference and the standard deviation for the industry are then calculated. Each firm's absolute difference is then expressed in standardized units. In the absence of outliers, these difficulty indices should take on values between -3.00 and +3.00. Equations 1-4 summarize

$$ADi = Xi - Yi \quad (1)$$

$$MAD = \sum_i (Xi - Yi)/N \quad (2)$$

$$S = \sqrt{\sum_i (ADi - MAD)^2/(N-1)} \quad (3)$$

$$DIi = (ADi - MAD)/S \quad (4)$$

Where i = Firm number
Yi = Starting value of the Objective for firm i
Xi = Goal for the Objective for firm i
N = Number of firms
ADi = Absolute Deviation for firm i
MAD = Mean Absolute Deviation for the industry
S = Standard Deviation of absolute deviations
DIi = Difficulty Index for firm i in standardized units

the calculations:
For the Relative Difficulty Index the "percent" difference between the starting point and the Goal is calculated and compared to the mean percent difference for the industry. The relative difficulty index would then be calculated in same fashion as the absolute difficulty index, but using proportions. The Overall Difficulty Indices are simply a weighted average of the Indices for each Objective, where the weights used were "weights of importance" set by each firm.

TABLE 3: PERFORMANCE EVALUATION WITH DIFFICULTY INDICES

	Goal Setting & Evaluation				DIFFICULTY INDICES	
Objective	Weight	Goal	Result	Rating	Absolute Difficulty	Relative Difficulty
Share	.40	10%	7%	70%	0.80	3.00
ROI	.50	60%	20%	33%	2.20	4.00
Debt	.10	30%	25%	100%	0.30	0.10
OVERALL				55%	1.45	3.21

Examining the Difficulty Indices in Table 3, we see that the Market Share Goal was 3 standard deviations from the mean in relative

terms. This says that the percent increase in the Goal for market share from the starting position of the firm was high (statistically) compared to the industry. But in an absolute measure the goal that was set by the firm was not high in relation to the industry, as noted by an Absolute Difficulty Index of only 0.80. The Goal set for the ROI, however, was significantly higher than the mean for the industry, in both an absolute and relative sense. In contrast, the Debt Goal implies a low level of difficulty in both an absolute and relative measure. The Overall Difficulty Indices imply that the Goals set by the firm, based on relative and absolute changes from its starting position, were "difficult" to achieve when compared to the average for the industry.

IMPLICATIONS OF THE EVALUATION CRITERIA

What can one conclude about comparative performance using the evaluation criteria suggested in this paper? If firms in an industry have different measures of difficulty, can performance be compared? For example, suppose a computerized business simulation consisting of 6 firms yields the Overall Ratings and Difficulty Indices given in Table 4 - INTERFIRM PERFORMANCE EVALUATION.

Firm B, with the highest Overall Rating, has low measures of difficulty. In this case, the high Overall Rating can be attributed to the firm setting Goals fairly close to its initial

TABLE 4: INTERFIRM PERFORMANCE EVALUATION

Firm	Overall Rating	Absolute Diff.	Relative Diff.	Group
A	55	.95	3.21	middle
B	95	-2.20	-2.70	low
C	85	2.92	2.10	high
D	45	3.20	4.50	high
E	75	.15	1.10	middle
F	60	-2.70	-3.30	low

starting position. Per contra, Firm C has the second highest Overall Rating but scores high in measures of difficulty. What does this imply about the performance ranking of firms B relative to C? What does it imply about the performance ranking of the other firms?

Our recommendation is NOT to attempt to get a strict ranking. We suggest, as a first step, separating the firms into three groups based on the difficulty indices: firms with high difficulty (group H), firms with middle difficulty (group M), and firms with low difficulty (group L). The use of Chebyshev's inequality will aid in separating the firms into three distinct groups. According to the inequality at least 75% of the firms should have indices between -2.00 and +2.00, with the remaining 25 percent distributed in the tails.

Therefore as a starting point, high difficulty is defined as + 2.0 or more in both the absolute and relative measures. Low difficulty is defined as less than -2.0 for both the absolute and relative measures. All other firms are placed in the middle, group M. The purpose of this scheme is to separate firms that we can say, confidently, have set high goals from those firms that have set :relatively easy goals. The firms in the middle cannot be differentiated statistically. Figure I-CHEBYSHEV AND DIFFICULTY GROUPINGS portrays the decision rule.

Figure I: CHEBYSHEV AND DIFFICULTY GROUPINGS

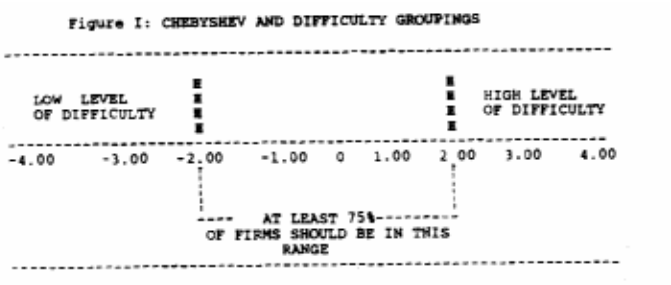

Table 5 - THREE GROUPS is based on the use of the inequality and the decision rule. Within each group, the firm with the highest Overall Rating is ranked above the other firms in performance. In Group H, firm C is ranked above firm D; in Group M, firm E is ranked above firm A; and in Group L, firm B is ranked above firm F. Intergroup comparisons are not as straightforward, but some conclusions can be reached. It seems reasonable to assume that any firm in Group H that has a Rating equal to or greater than a firm in the other groups has a better performance ranking. Similarly, any firm in Group M, that has a Rating equal to or greater than a firm in Group L, has a better performance ranking. Given this scenario, we can conclude

TABLE 5: THREE GROUPS

Group L		Group M		Group H	
Firm	Rating	Firm	Rating	Firm	Rating
B	95	E	75	C	85
F	60	A	55	D	45

that Firm C is ranked above Firms D, E, A and F. But we are not saying that Firm C exceeded Firm B in performance, since Firm B has a higher Rating. What we can say, based on this analysis, is that although Firm B had a Rating of 95, the firm set Goals that were relatively easy to achieve; and consequently, cannot be ranked above Firm C without further investigation. Furthermore, it seems likely that once the investigation is done, Firm C will be judged superior in performance relative to Firm B.

CONCLUSION

Strategic management focuses attention on issues of goal setting and performance evaluation. It is curious, therefore, that almost no integrated business policy simulations, oriented towards strategy, include formal goal setting and performance evaluation. The reasons for such an omission appear to be related to the modeling problems associated with quantifying goals and assessing comparative performance. Although there are a number of difficult technical problems centered around this issue, it does not provide adequate justification for omission owing to its critical importance to strategic management. Goal setting and performance evaluation criteria need to be embodied directly in the simulation algorithm. The purpose of this paper has been to present and illustrate such a model.

The evaluation criterion presented in this paper offers a number of advantages over the approaches currently utilized in existing business simulations. Firms are allowed to begin the simulation with different financial and operating characteristics. Given these different starting characteristics, firms select their own objectives and set different goals for each objective. Furthermore, the relative importance of each objective is specified by each firm through a "weighting" f actor. The simulation algorithm then evaluates: (1) the extent to which each firm achieves its goals; (2) the relative difficulty of the goals set by each firm, and (3) the comparative performance of each firm in relation to the competition. The measures of comparative performance do not provide a strict ranking but do segregate firms into different groups based on difficulty indices.

The type of information provided by this model focuses attention on some of the key components of strategic management and gives the user and administrator of the simulation useful feedback measures. This does not mean that the measures are perfect reflections of performance, they are not. Difficult questions of interpretation still remain, but avoiding the problem through omission or by have all firms start with the same financial and operating characteristics is certainly not the appropriate response. Rather, further research in the design and modeling of integrated business policy simulations, centering on the issues of goal setting and performance evaluation, are needed.

REFERENCES

[1] O. Embry, A. Strickland, and C. Scott, Tempomatic IV: A Management Simulation Houghton Mifflin Co., Boston 1974.

[2] T. Pray and D. Strang, DECIDE- A Management Decision Making Simulation, Random House Inc., New York 1980.

[3] B. Keys and H. Leftwich, The Executive Simulation, Kendall/Hunt Publishing Co, Iowa 1977.

[4] A. Edge, B. Keys, and W. Remus, The Multinational Game, Business Publications Inc., Dallas 1980

[5] A. Strickland and T. Scott, Micromatic–A Management Simulation, Houghton Mifflin Co., Boston 1985.

[6] L. Mills and D. McDowell, The Business Game, Little Brown and Company, Boston 1985.

[7] R. Jensen and D. Cherrington, The Business Management Laboratory, Business Publications Inc., Dallas 1977.

[8] J. Estes, C. Honess, and A. Corbett, Managing A Dynamic Business, Eagle Publishing and Graphics Company, Lexington 1980.

[9] J. Smith, BUSOP - A General Management Simulation of Business Operations, Management Simulations, Cambridge 1975.

[10] J. Smith, Enterprise II A Simulation, Houghton Mifflin Co., Boston 1985.