THE INFLUENCE OF SCORECARD EVALUATION ON DECISIONS AND OUTCOMES

John R. Dickinson
University of Windsor
bjd@uwindsor.ca

ABSTRACT

The Balanced Scorecard represents a well-grounded and well-developed type of multiple-item index that potentially may be effectively adapted to performance evaluation in business games. A necessary condition for a scorecard-based approach, though, is that such an approach actually impact the strategy decisions of game participants. This study examines the effect of a scorecard approach on strategy decisions and outcomes compared with a simple cumulative profit criterion.

INTRODUCTION AND PURPOSE

Evaluation of performance in business games is obviously vital, that vitality being reflected in numerous theoretical and empirical studies. Among the published works providing overviews of evaluation are Anderson *et al.* (1998), Anderson & Lawton (1997, 1992), Gentry *et al.* (1998), and McDevitt (1997). Though there exist alternative philosophical schools (Gentry *et al.*, 2002), Anderson & Leigh (1992) reported that 92.5% of simulation users in colleges used performance as a determinant of students' grades.

While profit seems to be the predominant measure of performance, many contemporary business simulation games provide multiple-criteria bases for evaluating performance:

> *The Marketing Management Experience* (*MME*, Dickinson, 2002)
> *The Multinational Management Game* (*MMG*, Keys & Wells, 1997)
> *Micromatic* (Scott *et al.,* 1992)
> *The Business Strategy Game* (*BSG*, Thompson & Stappenbeck, 1999)
> *Web Marketplace* (Cadotte)
> *The Business Management Laboratory* (*BML*, Jensen, 1999)

These more comprehensive bases take the form of a multiple-item index; that is, a weighted sum of numerous criteria, rather than only a single criterion. Anderson & Lawton (1988, p. 243) observe that, "Perhaps the most common form of evaluating performance on a simulation exercise is to compare the ranking of a student's team to other teams on a number of predetermined measures generated by the simulation."

THE BALANCED SCORECARD

In parallel, a major development in the evaluation of real-world company performance has been the Balanced Scorecard first introduced by Kaplan & Norton in 1992 (Kaplan & Norton, 1992). Briefly, the Balanced Scorecard is an extended (beyond traditional financial criteria) set of measures used to gauge enterprise performance. The major impetus for Kaplan & Norton's balanced scorecard was to complement traditional financial measures of performance ("...measures that tell the results of actions already taken...") with "...operational measures that are the drivers of future financial performance." (Kaplan & Norton 1992, p. 71) The Balanced Scorecard comprises four perspectives, three in addition to the traditional financial perspective.

BALANCED SCORECARD PERSPECTIVES

The *customer perspective* enables companies, "...to identify and measure, explicitly, the value propositions they will deliver to targeted customers and market segments. The value propositions represent the drivers, the lead indicators, for the core customer outcome measures." (Kaplan & Norton, 1996, p. 63) Customer core measures, i.e., outcome measures that appear repeatedly on actual company scorecards, include market share, customer acquisition, customer retention, customer profitability, and customer satisfaction. (Kaplan & Norton, 1996, pp. 67, 306) These customer core measures are supplemented with customer value proposition measures: product/service attributes, customer relationship, and image and reputation. (Kaplan & Norton, 1996, pp. 73-77)

The *internal perspective* is something of a misnomer. Included in this perspective is the traditional measurement of operating processes, performance centers, and departments. However, "Conventional performance measurement systems focus only on monitoring and improving cost, quality and time-based measures of existing business processes. In contrast, the approach of the Balanced Scorecard enables the demands for internal process performance to be derived from the expectations of specific external constituencies." (Kaplan & Norton, 1996, p. 115) That is, two of the three core types of internal perspective measures derive from external influences. The internal perspective subsumes innovation (identifying the market and creating the product/service offering) and postsale service (servicing the customer) processes as well as operations. (Kaplan & Norton, 1996, pp. 96, 104)

Finally, the *learning and growth perspective* focuses on the internal resources of the enterprise. Kaplan & Norton categorize measures in this perspective into people (i.e., employee capabilities), systems (i.e., information systems capabilities), and organizational procedures (i.e., motivation, empowerment, and alignment). (Kaplan & Norton, 1996, pp. 28, 127) (The learning and growth perspective was originally labeled "innovation and learning." [Kaplan & Norton, 1992])

FEASIBILITY OF THE BALANCED SCORECARD IN SIMULATION GAMING

In the context of simulation gaming Dickinson (2003) systematically assessed the feasibility of the Balanced Scorecard. Also, the Balanced Scorecard was the focus of a panel discussion at the Association for Business Simulation and Experiential Learning (ABSEL) conference in 2003 and was recognized by Sauaia (2001) in his review of performance evaluation criteria. The Balanced Scorecard concept rests on the mission and vision of the enterprise, sometimes specifies target levels of performance rather than objectives of maximum (e.g., profit) or minimum (e.g., employee turnover) performance, and several common measures (e.g. productivity) lack construct validity in simulation games. Further, the Balanced Scorecard is not conceived to yield a single unidimensional continuum of performance index values; in no published Balanced Scorecard are the components combined in any fashion into a single value. All of these aspects of the Balanced Scorecard were viewed by Dickinson (2003) to be problematic, if not altogether infeasible, in the context of contemporary business simulation games. It is recommended, then, that the multiple-criteria indices used in conjunction with contemporary business games *not* be referred to as Balanced Scorecards, but simply as scorecards.

More positively, the "perspectives" aspect of the Balanced Scorecard clearly is adaptable to business simulation games. The Balanced Scorecard philosophy, then, does contribute a substantive and real-world foundation for the use of multiple performance criteria beyond simple cumulative profit, a feature as noted above already available in many business games. In addition, the Balanced Scorecard also contributes the substantive bases of its specific perspectives. That is, business game designers and users might specify criteria from across the four perspectives in contriving their multiple-criteria indices or scorecards.

USE OF SCORECARDS IN GAMING BASIC RESEARCH

Several researchers have made use of multiple-criteria indices or scorecards in their basic research of business gaming. Foremost are Gosen & Washbush (2002, & Scott 2000, 1996, 1995, 1994) in the research program to develop an instrument to measure learning. Anderson & Lawton (1995) and Wolfe (1993) also employed scorecards in their basic research.

Against this backdrop of the already common availability of multiple-criteria indexes and the substantive contributions of the Balanced Scorecard philosophy, no research to date has empirically investigated the effectiveness of multiple criteria indexes *vis-a-vis* the single criterion of cumulative profit. Below are reported some groundwork findings from an experiment designed to compare the predictive validities of these two approaches to evaluation.

EXPERIMENT DESIGN

The purpose of this experiment is to compare the effectiveness of a multiple-criteria performance evaluation approach with the use of simple cumulative profit as the sole basis for evaluation. Basically, a simulation game competition was administered in a single introductory marketing class. Companies were managed by individual students. Half of the students were instructed that their final score for the simulation competition would be determined solely by cumulative profit. The other half of the students were instructed that their final score for the simulation competition would be determined on the basis of a four-criteria index.

As validation criteria become available, e.g., cumulative grade point average at graduation, the predictive validity of each of the evaluation approaches will be determined. The present study examines the more immediate effects of the two evaluation approaches on students' strategy decisions and selected outcomes. That is, the present study examines the internal validity of the experiment design. If the two experimental treatments, i.e., the two evaluation approaches, do not evoke relevant, meaningful differences in the subjects' behaviors, then the predictive validities of the two approaches are moot.

DATA COLLECTION

Subjects for this experiment were 40 students enrolled in an M.B.A. introductory marketing course. None of the students, by university admissions policy, had an undergraduate degree in business. Further, virtually all of the students were during that semester enrolled in the same core of introductory business courses, again by a "cohort" university M.B.A. program design. The business simulation game used was *The Marketing Management Experience* (MME, Dickinson, 2002). Each individual student managed his or her own MME company. The competition comprised nine periods in addition to an initial trial period.

In the MME companies are grouped into industries with a company competing against only the other companies in its same industry. Students, then, were assigned at random to companies with the companies then being assigned at random to industries. This plan yielded a total of 40 companies grouped into 10 industries, each comprising four companies. The 10 industries were assigned at random to one of the two experiment treatments. Specifically, companies/students in five industries were instructed that their scores for the simulation competition would be determined solely on the basis of cumulative earnings. Companies/students in the other five industries were instructed that their scores for the competition would be determined on the basis of a four-criteria scorecard as described below.

THE SCORECARD

Multiple-criteria indices are most usually linear combinations, which is to say they are weighted sums. With this, the indices are defined by two types of specifications: the specific criteria to comprise the index and the specific respective weights to be applied to the criteria. Though only these two types of specifications must be made to define a multiple-criteria index these are sufficient to yield a virtually limitless number of possible definitions. This, in turn, inevitably limits the mix of defined indexes that might be researched and, accordingly, inevitably limits the generalizability of the research results to the

"population" of multiple-criteria indexes. Recognizing this, research might still systematically and judiciously define specific indexes with generalizability left to judgment and not to classical statistical inference.

NUMBER OF CRITERIA

For this research, at least four criteria were to be specified in order for the scorecard to represent each of the four perspectives of the Balanced Scorecard. Early keystone applications of Balanced Scorecards described by Kaplan & Norton comprise 17 (1992, p. 76) and 20 criteria (1993, p. 136). However, as Dickinson (2003) noted, the responsibility for satisfying such large numbers of criteria is presumably delegated to numerous managers, not to single managers as would be the case in the present research. In the present research challenging individual subjects/managers to formulate strategies toward many criteria would be dysfunctionally complex or unmanageable. On this basis and on the basis of empirical pretest analyses described below, the number of components comprising the scorecard for this research was set at four.

PRETEST ANALYSES

Seventeen criteria available in the MME were considered for inclusion in the multiple-index criterion for this experiment: seven in the financial perspective, four in the customer perspective, four in the internal perspective, and two in the learning and growth perspective. Pretest data were drawn from an earlier administration of an MME competition. That earlier competition very much paralleled the present study in that it comprised 48 companies managed by individual students who were enrolled in an earlier offering of the same course, nine competition periods plus a single initial trial period, the competition accounting for the same percent of course grade, and so on. In the earlier competition cumulative profit was the single evaluation criterion.

Among financial perspective criteria, return on equity (or direct variations such as return on capital employed) appears most frequently in published Balanced Scorecards and is generally accepted as a universal measure of investment performance. In the pretest data, pairwise correlations among return on equity, cumulative earnings, and return on assets were all .999. On the basis of its real-world ubiquity and its virtually complete capture of statistical information with the other two common measures of profit, return on equity was included in the

scorecard for the present research.

Among the customer perspective criteria, customer loyalty and customer attraction were extremely highly correlated (.991) and, as such, each was equally correlated with return on equity (.663). Customer loyalty was slightly more highly correlated with both market share and stockouts (two other customer perspective criteria) and so captured slightly more of the statistical information in those two criteria. Too, customer loyalty, being essentially the percent of repeat brand purchasers, is more simple than is customer acquisition which involves percentages of customers attracted from each of several competitors. On these bases, customer loyalty was included in the scorecard of the present research.

Among available internal perspective criteria, two–sales per retail outlet and sales per salesperson–are of questionable construct validity. As explained by Dickinson (2003), in real-world applications, measures such as these are surrogates for differential sales-affecting traits of specific outlets or salespeople. However, typically in business games, outlets and salespeople have no such differential traits. For example, individual retail outlets in business games do not have differential locations, ambiances, layouts, etc. A third internal perspective criterion, inventory management, is a combination of inventory carrying costs and opportunity costs of lost sales. While this measure is available to the MME administrator, in only limited real-world situations is the extent of lost sales known. Accordingly, inventory turnover was included in the scorecard of the present research.

Of the two available learning and growth perspective criteria, sales force morale had lower correlations with return on equity (.238 vs. .324) and customer loyalty (.131 vs. .299) than did sales force turnover, the two having nearly equal correlations with inventory turnover (.206 vs. .200). On the basis of its lower statistical redundancy with other components included in the scorecard, sales force morale was also included.

The four criteria comprising the scorecard for this experiment were, then, return on equity (representing the Balanced Scorecard financial perspective, Kaplan & Norton, 1992, p. 76), customer loyalty (consumer perspective, Kaplan & Norton, 1996, pp. 77, 80), inventory turnover (internal perspective, Kaplan & Norton, 1996, p. 111), and sales force morale (learning and growth perspective, Kaplan & Norton, 1996, p. 44). All four of the criteria are described and operationally defined in the *MME: Manager's Position Description*, i.e., player's manual. Intercorrelations among the four included scorecard components are presented in Table 1.

TABLE 1: Scorecard Component Intercorrelations				
	Return on Equity	Customer Loyalty	Inventory Turnover	Sales Force Morale
Return on Equity	1			
Customer Loyalty	0.663	1		
Inventory Turnover	0.463	0.482	1	
Sales Force Morale	0.238	0.131	0.206	1

CRITERIA WEIGHTS

As noted above, one of the infeasible aspects of the Balanced Scorecard for adaption to business games is that it is not conceived to yield a unidimensional continuum of performance index values (Dickinson, 2003). Specifically for the present research, real-world applications of the Balanced Scorecard do not provide any guidance as to weightings of criteria.

Business simulation games offering a multiple-criteria evaluation feature leave it to the administrator to specify weights for criteria. Where multiple-criteria indices have been used in basic research of business gaming, weights have been published, though no particular rationales for the particular weights are presented. Further, these criteria are invariably entirely within the financial perspective. Finally, as discussed above, potential sets of weights constitute an infinite population from which drawing a meaningful random sample is not feasible. For the present research, then, the most simple scheme of equal component weights, i.e., each scorecard component being weighted 25%, was used. (It may be tempting to conjure the financial perspective as more important, i.e., deserving of greater weight, than the other perspectives. That view, however, is antithetical to the Balanced Scorecard philosophy. That philosophy eschews short-run profit as an evaluation criterion and holds that the nonfinancial perspectives are portenders of long-run profit. Specification of component weights and the components themselves ultimately warrants theoretical, managerial, and empirical consideration.)

SCORECARD ADMINISTRATION AND FEEDBACK

The standardized (within industry) value of each criterion was calculated for each competition period. At the end of the competition the mean of each standardized criterion was determined, i.e., the sum of the nine period values divided by nine. The mean-across-periods standardized values for each criterion, then, were multiplied by weights of .25 and the products summed.

Subjects in the cumulative profit experiment group were apprised each period of the cumulative earnings of all companies in their respective industries. That is, they were given sufficient information for them to determine their exact evaluation standings within their respective industries. Students in the multiple-criteria experiment group were given comparable feedback. Specifically, they were informed of the mean (across competition periods to date) original values of each of the four criteria, the mean standardized values, and the scorecard values, i.e., weighted sums of the four mean standardized values. Thus, subjects in both experiment groups received the same sufficient information to determine their evaluation standings for the relevant evaluation approach.

As noted above, in the MME companies compete only against the other companies in their industries. Both experiment groups were evaluated accordingly. That is, regardless of their experiment group, subjects were evaluated relative to only their actual competitors. Specifically, depending on experiment group, within each industry either final cumulative earnings values or final scorecard scores were transformed to have mean equal to 75 and standard deviation equal to 12.

CRITERIA

The purpose of this study is to compare selected decisions and selected outcomes of companies in the two experiment treatment groups. The general operant theory is that subjects in the two groups in formulating their strategy decisions managed their companies differently in order to maximize their final competition evaluation; that is, in order to maximize either their cumulative profits or their scorecard index values. The impetus for the former experiment group is generic: maximize revenue while minimizing costs. Selectively, the impetus for the latter group is more comprehensive and more specific. Differences between the two groups, then, may be anticipated in the context of the respective performance evaluation experiment treatments.

The two experiment groups were compared on 15 criteria, representing all four of the Balanced Scorecard perspectives in keeping with the scorecard specified for this research. The two criteria from the financial perspective–profit and return on equity–are equivocal. It may be that the profit experiment group, having but the one evaluation objective, would be expected to realize greater profit than the scorecard group whose managerial skills were obliged toward multiple evaluation objectives. Against this, one of the scorecard components is return on equity, a factor of which is profit. Equity in the MME comprises only common stock, which is common to all companies and constant throughout the competition, plus retained earnings, i.e., cumulative profit. Cumulative profit and return on equity, then, tend to be highly correlated (recognizing that this need not absolutely necessarily be the case). In the pretest analyses described above, cumulative profit and return on equity had a correlation of .999. It might also be reasoned that the scorecard experiment group is more highly motivated to attend to various key areas of their companies which, in turn, may lead to more effective management and, thus, greater profit.

The scorecard criterion drawn from the customer perspective is loyalty. Operationally, the MME generates a brand (actually company) switching matrix the diagonal elements of which indicate loyalty and with the off-diagonal elements indicating switching between competing companies. The scorecard loyalty criterion was the diagonal element relevant to the particular company. It was anticipated that company loyalty for scorecard companies would be greater than for profit companies. MME managers may select from a menu of six messages for their broadcast, print, and trade advertising. The pioneering message is "designed to introduce customers to the basic concept of digital cameras..." (Dickinson, 2002, pp. 6-4) This message appealing to first-time purchasers would not serve the interest of scorecard companies evaluated on loyalty, i.e., repeat purchasers. Accordingly, it was anticipated that scorecard companies would select the pioneering message less often than would profit companies.

Scorecard companies were each period provided with all of the switching matrix diagonal or loyalty percentages. Thus, subjects were continually informed of their own loyalty measures and those of their competitors. (This put this

experiment group on a comparable basis with the profit experiment group which was informed each period of the cumulative profit of all competitors.) However, scorecard companies were not provided with the off-diagonal information. The complete switching matrix is available to all competitors through the purchase of a marketing research report. It was anticipated that scorecard companies would have a greater interest in this off-diagonal company switching information than would profit companies and, thus, that the former would purchase the relevant research report more often.

The internal perspective scorecard component was inventory turnover. It was, therefore, anticipated that the scorecard experiment group inventory turnover would be greater than that for the profit experiment group. Related to inventory, it was anticipated that the scorecard group inventory carrying cost would be less than that for the profit experiment group. One way to increase inventory turnover, of course, is to maintain low inventory levels. Thus, it was anticipated that both stockouts and the opportunity cost of lost sales due to stockouts would be higher for the scorecard experiment group than for the profit experiment group.

The fourth scorecard evaluation component was sales force morale. In the MME morale is indicated by an index defined so that the average morale for all companies in an industry equals 100. So, for example, where a company's morale index equals 110 the company manager knows that his or her sales force has a morale level 10 percent above the industry average. With this index, then, the mean morale index values for companies in each experiment group are by definition equal. However, competition participants are informed in the *MME: Manager's Position Description* that morale is dependent on sales force compensation and turnover. For this research, then, sales force compensation and turnover were suitable criteria representing the learning and growth perspective with the anticipation that compensation would be greater and turnover lower for the scorecard experiment group compared with the profit experiment group.

A summary of the criteria on which the two experiment groups were compared and the anticipated direction of their differences is presented in Table 2.

TABLE 2: Comparison Criteria

		Treatment Group	
Perspective	Operational Measure	Cumulative Profit	Scorecard
Financial	Profit Return on equity	equivocal equivocal	equivocal equivocal
Customer	Loyalty Pioneering advertising message, % Purchases of company loyalty research	lower higher lower	higher lower higher
Internal	Inventory turnover Carrying cost Stockouts, % Opportunity loss	lower higher lower lower	higher lower higher higher
Learning & Growth	Sales force total compensation Sales force turnover, %	lower higher	higher lower

ANALYSIS AND RESULTS

Analysis for this research comprised comparisons of mean values of the two experiment groups and simple t-tests of those means. For each company, total profit, inventory carrying cost, and opportunity loss dollar amounts were determined summed over all competition periods. Return on investment, inventory turnover, and sales force turnover were calculated for each company each competition period and then the mean per period determined. Loyalty was the percent figure from the diagonal of the switching matrix for each competition period averaged over competition periods. Use of the pioneering advertising message was operationalized for each company as the percent of all advertising messages that were the pioneering message. Purchase of the brand/company switching matrix marketing research report was operationalized as the total of those purchases over the course of the competition.

Since there are four potential market segments in the MME

(two products x two regions, though participants are not required to market in all four segments), generally there are four inventories for each competition period or a total of 36 potential inventories across the nine competition periods. Stockouts were operationalized as the percent of all potential inventories that stocked out.

Sales force compensation was measured as the average dollar amount per salesperson per competition period.

The parameters for all industries in this study were identical. Even so, industry sales (and profits) may vary across industries depending on the collective strategies of the companies comprising the industries. For this reason, the four criteria measured in dollar amounts–profit, inventory carrying cost, opportunity loss, and sales force compensation–were also expressed as percentages of sales and were analyzed. Mean values of the 11 original criteria plus four of those criteria expressed as percents of sales were calculated for each of the two experiment groups. Those mean values are presented in Table 3.

TABLE 3: Means by Experiment Group

Criterion	Profit Group	Scorecard Group	
Profit	792,444.10	2,810,097.00	n.s.
Return on Equity	-0.44	0.71	n.s.
Loyalty	37.49	42.89	*
Pioneering Adv. Message	4.05	1.59	*
Loyalty Research Report	2.00	3.70	*
Inventory Turnover	3.74	21.22	*
Inventory Carrying Cost	1,318,853.00	409,992.00	*
Stockouts	11.39	30.43	*
Opportunity Loss	524,512.00	1,565,527.00	*
Sales Force Compensation	25,655.05	32,754.22	*
Sales Force Turnover	6.37	4.44	n.s.
Profit as Percent of Sales	-1.44	0.80	n.s.
Carrying Cost as Percent of Sales	2.61	0.64	*
Opportunity Loss as Percent of Sales	1.00	2.31	*
Compensation as Percent of Sales	0.42	0.48	n.s.

* = p < .10, one-tailed test
n.s. = not significant

The scorecard experiment group realized greater profit, return on equity, and profit as a percent of sales, though for all three criteria the difference was not significant (p>.10, two-tailed t-test). As explained earlier, there is no clear *a priori* theory as to which experiment group would realize the greater profit or return. The results of this study in this respect are inconclusive, though they do suggest that use of a multiple-criteria evaluation approach does not necessarily lead managers to formulate less profitable strategies.

For all of the 12 remaining criteria, the differences between mean values for the two experiment groups were in the direction anticipated and for 10 of these 12 the differences were statistically significant (p<.10, one-tailed t-test). The preponderance of statistically significant differences is all the more imposing in light of the small sample sizes and consequent lower power of the tests.

As a check on the presence of extreme cases, median values were calculated for each of the 15 criteria. All of the differences in medians between the two experiment groups were in the same direction as the differences in means.

CONCLUSION

The general conclusion from this study is that under different performance evaluation schemes participants in a simulation game competition do, indeed, manage their companies differently with resultant meaningfully different outcomes. This is a preliminary and important finding in the context of a longer term research project that will investigate the concurrent and predictive validities of the two performance evaluation approaches. Had the anticipated differences in selected decisions and outcomes not materialized then a necessary condition for the validity research would not be satisfied.

This general conclusion also has an important implication in its own right. That implication is that simulation competition participants are able to translate given objectives (in the form of the evaluation scheme) into strategies to achieve those objectives. This is an important finding. Simply dictating an evaluation scheme does not necessarily mean that participants can assimilate that scheme into their management strategies. This research has demonstrated that they can which, in turn, attests to the meaningfulness of the simulation participation experience itself.

It is interesting to note that the scorecard companies actually realized greater profits, on average, than did profit companies. It may be tempting to speculate that the greater profit realized by the scorecard companies is attributable to the more broad-based managerial evaluation of the scorecard and, thus, participants' greater attention to key areas of their companies.

The consistency of the results of this study is compelling. Of the 12 hypothesis tests where there was a clear theoretical basis for direction, 10 were statistically significant. (As explained above, there are sound rationales for theorizing either group would realize greater profit and return on investment.) In light of the small sample sizes in each experiment group, this consistency of direction and statistical significance is imposing.

REFERENCES

Anderson, Philip H., Cannon, Hugh M., Malik, Dolly & Thavikulwat, Precha (1998). "Games as Instruments of Assessment: A Framework for Evaluation," in Butler, John K., Jr., Leonard, Nancy H. & Morgan, Sandra W. (Eds.) *Developments in Business Simulation and Experiential Learning*, Volume 25. Statesboro, GA: Association for Business Simulation and Experiential Learning, 31-37.

Anderson, Philip H. & Lawton, Leigh (1997). "Demonstrating the Learning Effectiveness of Simulations: Where We are and Where We Need to Go," in Butler, John K., Jr. & Leonard, Nancy H. (Eds.) *Developments in Business Simulation and Experiential Learning*, Volume 24. Statesboro, GA: Association for Business Simulation and Experiential Learning, 68-73.

Anderson, Philip H. & Lawton, Leigh (1995). "The Problem of Determining an 'Individualized' Simulation's Validity as an Assessment Tool," in Overby, John D. & Patz, Alan L. (Eds.), *Developments in Business Simulation & Experiential Exercises*, Volume 22. Statesboro, GA: Association for Business Simulation and Experiential Learning, 43-48.

Anderson, Philip H. & Lawton, Leigh (1992). "A Survey of Methods Used for Evaluating Student Performance on Business Simulations," *Simulation & Gaming*, Vol. 23, December, 490-498.

Anderson, Philip H. & Lawton, Leigh (1988). "Assessing Student Performance on a Business Simulation Exercise," in Sanders, Patricia & Pray, Thomas (Eds.), *Developments in Business Simulation & Experiential Exercises*, Volume 15. Statesboro, GA: Association for Business Simulation and Experiential Learning, 241-245.

Cadotte, Ernest R., *Web Marketplace*.

Dickinson, John R. (2003). "The Feasibility of the Balanced Scorecard for Business Games," in Pillutla, Sharma & Feinstein, Andrew Hale (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 30. Statesboro, GA: Association for Business Simulation and Experiential Learning, 90-98.

Dickinson, John R. (2002). *The Marketing Management Experience*. Windsor, Ontario: Management Experiences.

Gentry, James W., Burns, Alvin C., Putrevu, Dickinson, John R., Sanjay, Chun, Seungwoo, Hongyan, Yu, Williams, Laura, Bare, Thomas & Gentry, Ruth Ann (2002). "Managing the Curiosity Gap Does Matter: What do We Need to do About It?" in Vaughan, Mary Jo & Pillutla, Sharma (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 29. Statesboro, GA: Association for Business Simulation and Experiential Learning, 67-73.

Gentry, James W., Commuri, Suraj R., Burns, Alvin C. & Dickinson, John R. (1998). "The Second Component to Experiential Learning: A Look Back at How ABSEL has Handled the Conceptual and Operational Definitions of Learning," in Butler, John K., Jr., Leonard, Nancy H. & Morgan, Sandra W. (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 25. Statesboro, GA: Association for Business Simulation and Experiential Learning, 62-68.

Gosen, Jerry & Washbush, John (2002). "The Validity Investigation of a Test Assessing Total Enterprise Simulation Learning," in Vaughan, Mary Jo & Pillutla, Sharma (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 29. Statesboro, GA: Association for Business Simulation and Experiential Learning, 247-251.

Gosenpud, Jerry & Washbush, John (1996). "Correlates of Learning in Simulations," in Patz, Alan L. & Butler, John K., Jr. (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 23. Statesboro, GA: Association for Business Simulation and Experiential Learning, 43-48.

Gosen, Jerry, Washbush, John & Scott, Tim (2000). "Initial Data on a Test Bank Assessing Total Enterprise Simulation Learning," in Page, Diana & Snyder, LT (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 27. Statesboro, GA: Association for Business Simulation and Experiential Learning, 166-171.

Jensen, Ron (1999). *The Business Management Laboratory*, Revised Fifth Edition. Little Rock, AR: Micro Business Publications.

Kaplan, Robert S. & Norton, David P. (1996). *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press.

Kaplan, Robert S. & Norton, David P. (1993). "Putting the Balanced Scorecard to Work," *Harvard Business Review*, Vol. 71, September-October, 134-147.

Kaplan, Robert S. & Norton, David P. (1992). "The Balanced Scorecard–Measures That Drive Performance," *Harvard Business Review*, Vol. 70, January-February, 71-79.

Keys, J. Bernard & Wells, Robert A. (1997). *The Multinational Game*, Fourth Edition. Little Rock, AR: Micro Business Publications.

McDevitt, Paul K. (1997). ΑMeasuring Student Learning Using Business Simulations: A Theory Based Perspective,@ in Butler, John K., Jr. & Leonard, Nancy (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 24. Statesboro, GA: Association for Business Simulation and Experiential Learning, 247-252.

Sauaia, Antonio Carlos Aider (2001). "Evaluation of Performance in Business Games: Financial and Nonfinancial Approaches," in Pittenger, Khushwant S. & Vaughan, Mary Jo (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 28. Statesboro, GA: Association for Business Simulation and Experiential Learning, 210-214.

Scott, Timothy, Strickland, Alonzo J., III, Hofmeister, David L. & Thompson, Michael D. (1992). *Micromatic: Instructor's Resource Manual*, Second Edition. Boston: Houghton Mifflin Company.

Thompson and Stappenbeck (1999). *The Business Strategy Game*.

Washbush, John & Gosenpud, Jerry (1995). "Simulation Performance, Learning and Struggle," in Overby, John D. & Patz, Alan L. (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 22. Statesboro, GA: Association for Business Simulation and Experiential Learning, 1-4.

Washbush, John & Gosenpud, Jerry (1994). "Simulation Performance and Learning Revisited," in Thavikulwat, Precha & Overby, John D. (Eds.), *Developments in Business Simulation & Experiential Learning*, Volume 21. Statesboro, GA: Association for Business Simulation and Experiential Learning, 83-86.

Wolfe, Joseph (1993). "On the Propriety of Forecasting Accuracy as a Measure of Team Management Ability: A Preliminary Investigation," *Simulation & Gaming*, Vol. 24, March, 47-62.