THE FEASIBILITY OF THE BALANCED SCORECARD FOR BUSINESS GAMES

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

ABSTRACT

Many business simulation games offer a multiple-criterion scoring feature. Typically, the game administrator may select specific criteria from a menu and may also specify weights for each criterion. The resultant, then, is used as the basis for performance evaluation, in contrast to, say, simple cumulative earnings. In the business management literature, and to a substantial extent in business practice, a multiple-criterion approach to evaluation has been formalized as the Balanced Scorecard. This paper presents a review of the Balanced Scorecard and examines the integration of it for evaluation of simulation game performance.

INTRODUCTION

Evaluation of participants in business simulation games is vital and, accordingly, is an ongoing focus of debate and empirical research. Position statements and reviews of evaluation approaches have been compiled by Gentry *et al.* (1998), Thavikulwat *et al.* (1998), Anderson and Lawton (1997; 1988), and McDevitt (1997), among others. Two main philosophical schools have evolved. One school is premised on performance (Anderson and Leigh [1992] reported that 92.5% of simulation users in colleges used performance as a determinant of students' grades), the other school is premised on learning (Gentry *et al.* 2002). Within the performance philosophy, the most direct and simplest evaluation criterion is usually profit. However, several business simulation games offer more comprehensive bases for evaluating performance:

> *The Marketing Management Experience* (*MME*, Dickinson 2002)
> *The Multinational Management Game* (*MMG*, Keys and Wells 1997)
> *Micromatic* (Scott *et al.* 1992)
> *The Business Strategy Game* (*BSG*, Thompson and 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 and 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."

Outside the realm of simulation gaming, the multiple-item approach to measuring performance has received widespread endorsement by companies, specifically in the form of the Balanced Scorecard first introduced by Kaplan and Norton in 1992 (Kaplan and Norton 1992, hereafter K&N). The Balanced Scorecard has been adopted by numerous enterprises. It has also received at least passing consideration in the gaming literature (Sauaia 2001). The purpose of this paper is to examine the Balanced Scorecard with respect to its feasibility for application to business games and, with or without complete feasibility, posit implications of the Balanced Scorecard for evaluation in business games. The Balanced Scorecard represents a well-grounded framework that might serve as a unifying reference for measuring performance in business games.

DESCRIPTION OF THE BALANCED SCORECARD

The concept of the Balanced Scorecard is a model of simplicity. The Balanced Scorecard is but an extended (beyond traditional financial criteria) set of measures used to gauge enterprise performance. Its inauspicious concept notwithstanding, the Balanced Scorecard has evolved into an impetus fostering strategy-focused organizations. (K&N 2000, p. 7) The Balanced Scorecard has become a "strategic management system." (K&N 2000, p. 17) The linkage is that Balanced Scorecard measures are derived from the enterprise's vision and strategy. "The scorecard should tell the story of the strategy..." (K&N 1996a, p. 47) As the enterprise works to the performance evaluation scheme of the Balanced Scorecard, all organizational resources tend to align with and focus on that strategy. (K&N 2000, pp. 7-8) The major impetus for Kaplan and 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." (K&N 1992, p. 71)

Kaplan and Norton (1996a, pp. 38-40) delineate numerous dysfunctional effects of the overemphasis on financial

measures on the United States corporate system, though, again, they advocate "balancing" financial measures, not displacing them. The means by which the Balanced Scorecard purports to portend future success are:

- inclusion of a broad-based mix of operations measures from customer, internal, and learning and growth "perspectives" that should drive future performance in addition to traditional measures from the financial "perspective" and

- inclusion of some measures of trend, such as *change* in market share and not just current market share.

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." (K&N 1996a, 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. (K&N 1996a, pp. 67, 306) These customer core measures are supplemented with customer value proposition measures: product/service attributes, customer relationship, and image and reputation. (K&N 1996a, 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." (K&N 1996a, 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. (K&N 1996a, pp. 96, 104)

Finally, the *learning and growth perspective* focuses on the internal resources of the enterprise. Kaplan and 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). (K&N 1996a, pp. 28, 127) (The learning and growth perspective was originally labeled "innovation and learning." [K&N 1992])

It is noteworthy that the Balanced Scorecard does not yield a single index number designed to reflect performance. No mechanism is prescribed for combining the various measures into a single index. In business game scorecard applications, though, where performance evaluation must presumably be transformed into a unidimensional grading continuum, some procedure must be in place for mapping the multiple criterion measures into a univariate index. All of the business games listed above utilize a weighted sum approach. This imposes two requirements: specification of weight values and transforming original criterion measures to some common scale. Each of these requirements presents potential limitations to the use of the Balanced Scorecard in business games as described below.

SOME SCORECARDS IN CURRENT GAMES

Menus of scorecard options for six current simulation games are presented in Table 1. Criteria available in these six games are very predominately in the traditional financial perspective; very few options may not be determined from the income statement and balance sheet.

TABLE 1: Some Current Game Scorecard Menu Options

MME	*MMG*	*Micromatic*
cumulative earnings	market share	income after taxes
return on equity	return on sales	sales revenue
return on assets	return on assets	stock price
return on sales	debt to total assets	return on assets
dollar sales	quality/technology	return on sales
unit market share	return on equity	return on equity
inventory turnover	stock price	earnings per share
asset turnover		
operating expense ratio		
inventory management		
sales force turnover		

BSG	*Marketplace*	*BML*
revenues	market share	stockholder wealth
profits	customer satisfaction	stock price
return on equity	profitability	dividends
firm value	investments in future	company stability
bond rating	return on investment	quality control
strategy rating		order delivery rate
		employee compensation
		steady work

The most prominent use of a scorecard approach in basic research has been by Gosen and Washbush (2002) in their development of an instrument to measure learning. Their research has used the *Micromatic* simulation with the scorecard usually being specified as net income weighted 40%, return on sales weighted 30%, and return on assets weighted 30%.

Kaplan and Norton make very clear that financial measures should not be scrapped. To the contrary, "The financial objectives serve as the focus for the objectives and measures in all the other scorecard perspectives." (1996a, p. 47) They soundly dismiss the theory that, "...as companies make fundamental improvements in their operations, the financial numbers will take care of themselves." (1996a, p. 32) Thus, current scorecards in simulation games are very consistent with one of the Balanced Scorecard perspectives, but also very lacking with respect to the remaining three perspectives. Should the Balanced Scorecard prove meritorious, the addition of criteria from those three perspectives is a basic implication for business games.

INFEASIBLE CORNERSTONE

For all but the most complex and sophisticated (and unwieldy) business games, the very cornerstone of the Balanced Scorecard is infeasible. The Balanced Scorecard evaluation approach commences with managers conceiving the specific nature of the Scorecard for their companies.

"The objectives and measures for the Balanced Scorecard are more than just a somewhat ad hoc collection of financial and nonfinancial performance measures; they are derived from a top-down process driven by the mission and strategy of the business unit. The Balanced Scorecard should translate a business unit's mission and strategy into tangible objectives and measures." (K&N 1996a, pp. 11,12)

"Since every strategy is unique, every scorecard should be unique and contain several unique measures." (K&N 1996a, p. 306)

Kaplan and Norton (1996a, pp. 300-308) outline a 10-task process for building a Balanced Scorecard and provide several case study descriptions of the process (1993, pp. 138-139, 143-147; 1996b, pp. 78-79). The process throughout requires custom input from managers and consensus building among them. The feasibility of building a scorecard by this process in a simulation game context (as opposed to invoking a scorecard specified by the simulation administrator) is very questionable. In business games, a number of factors mitigate against this creative process. As a practical matter, students expect to be graded on a common basis. This, in turn, suggests a common scorecard imposed by the simulation administrator and not idiosyncratic scorecards created by students.

Might not students be allowed to compose their own scorecards? For example, possibly students could themselves select components from scorecard menus and also specify weights for the respective components. Perhaps, but within what framework would resulting scores be transformed into grades? The *Micromatic* and *MMG* scorecards first rank companies on each component and then applies weights to the rank positions, not the original scorecard component values. But if not all competing companies have included a given component in their scorecards, then the rankings of competing companies will be not be comparable across components. A company might rank first on a component simply because it was the only company to include that component, while a company might rank first on another component due to its having bested several competing companies on that component. The *MME* provides additional component normalizing options, but in all of them competing companies play a role.

An approach of how-well-did-the-company-accomplish-what-it-set-out-to-do would presumably not involve competing companies. In this scenario, though, against what maximum value or other standard are realized scorecard scores to be compared to arrive at grades? Many potential scorecard components, e.g., profit, have no meaningful maximum value so a percentage-of-maximum transformation would also not be meaningful.

EXPLANATION TO AND ASSIMILATION BY PARTICIPANTS

A set of material limitations on the use of the Balanced Scorecard simply derives from the Scorecard's or any scorecard's greater complexity compared with the more global univariate criterion of cumulative profit. This complexity requires greater explanation to participants and greater assimilation by them. This complexity subsumes purpose, scope, specification, and data normalization.

Using any multiple-criterion scorecard approach rather than the single criterion of profit raises, or should raise, with both participants and administrators the fundamental question of Why? Why do not good strategies, including operations management, simply result in higher earnings? If the scorecard components do impact earnings, then why include the individual components in the performance evaluation procedure? If the components do not impact earnings, then why not exclude them from the evaluation procedure on that basis?

The Balanced Scorecard four perspectives' greater number of criteria clearly do serve a diagnostic purpose and diagnosis might well be informative for improving strategy formulation, i.e., a means toward better performance. But the objective here is evaluation, not diagnosis. A diagnostic purpose may be accomplished by simply making the scorecard information available to participants. To a limited extent it may be argued that some real-world managers are responsible for specific operational procedures, e.g., inventory management or production, and not for business unit

profitability. Thus, evaluation on more specific criteria is valid. Such specific criteria, though, have only a very small presence in core Balanced Scorecard measures (Kaplan and Norton 1996a, p. 306) and an even smaller presence in the scorecard menus of the games listed earlier.

The four perspectives and the several specific measures among them require explanation and assimilation of just that many more constructs and their operational definitions than simple cumulative profits. *Per se*, explanation/assimilation of the more numerous constructs requires additional time and effort, presumably subtracting from time and effort available for other uses, e.g., lectures, in the course. In addition to this logistical consideration, a pedagogical consideration attends this greater mix of constructs. Some of the specific scorecard components may be at too sophisticated a level or be otherwise inappropriate for the business game context. For example, cash flow may be too advanced a concept for the participation cohort. Or cash flow may not be appropriate for introduction until some time in the course well after the simulation competition has commenced. Introducing all of the components at an early stage in a course may be inappropriate.

In some instances, specific knowledge of the measure may simply not be available. The explanation for "company stability" in the BML is:

> "One [surrogate] might be how well they have managed their financial matters (e.g. how many and how large are the special loans, if any, and do they have room for financial expansion or are they >maxed out' and approaching a bottleneck). Another factor might be what has been happening to their market share (e.g. has it been at least steady or has it been declining to where they may be unable to function adequately?). A performance index for each firm is calculated on these factors and the firm with the highest score is rated the >best'." ("Update for version 5.2," p. 5)

This explanation is not sufficient to allow students to replicate their evaluations.

As with "Purpose" above, where introduction of the numerous scorecard components is appropriate for the participation cohort and consistent with course content progression, then the use of a scorecard may be beneficial rather than detrimental.

A given scorecard configuration comprises specific components and specific weightings of the respective components. The simulation administrator presumably exercises his or her judgment in selecting components and in setting weight values. This, in turn, entails the need for additional explanation/assimilation not only of the components themselves (see above), but also why specific components were selected and others not selected and the rationale for their relative importance in the evaluation scheme. A possibly beneficial exchange, but also possibly an unproductive one.

As explained earlier, the Balanced Scorecard prescribes no mechanism for combining its multiple measures into a

unidimensional index, a mechanism that is necessary for transforming multiple measures a business game into a unidimensional grading continuum. The *MME*, *Micromatic*, and the *MMG* all employ a weighted sum approach. Specification of weights is addressed immediately above. It is generally not feasible to apply weights to original criterion values. Some common criterion measures, e.g., profit, may be in thousands or millions, other common measures, e.g., inventory turnover, may be in the teens, and yet other common measures, e.g., return on equity, may be fractions. Applying weights to measures of such disparate scale levels would very much distort the intended weighting scheme. Instead, the original criterion measure values must be normalized in some way and then the weights applied to the normalized values.

In *Micromatic* and the *MMG* competing companies are ranked on each criterion and weights are applied to the resulting ordinal position values and the resulting products summed. The practice is tenuous from a data scale level standpoint, i.e., multiplying and summing ordinal data, and participants may (properly) question it. It may also be difficult to rationalize to participants why a close second on one criterion carries the same contribution as a distant second on a second criterion. In the *BSG* and the *BML* the "best" company on a given factor is assigned a score of 100 or 10, respectively, with other companies receiving scores relative to those maxima. In the *MME* a choice from three normalization procedures may be made by the administrator. While two of the procedures may be more tenable than the weighting-of-ranks approach, they still require explanation/assimilation. In any event, the normalizing of data extends the issue of required explanation/assimilation resources.

FEASIBILITY OF KEY BALANCED SCORECARD MEASURES

The essence of the Balanced Scorecard concept lies in the four perspectives it prescribes and the multiple measures within each perspective. While there is no set list of measures, "...certain core outcome measures appear repeatedly on scorecards." (K&N 1996a, p. 306) Across their various publications, several actual measures incorporated by major companies are identified.

Among Rockwater's measures are project profitability (financial perspective) and project performance index and project closeout cycle (internal business). (K&N 1993, p. 136) The notion of project connotes uniqueness and dynamism from conception to implementation. Such uniqueness and dynamism are generally infeasible in a competitive business game where an environment common to all competitors, i.e., a common playing field, seems necessary.

On the other hand, Rockwater's profit forecast reliability measure in its financial perspective already has precedent in the gaming literature. Teach (1992; 1990) proposed that forecasting accuracy is a suitable evaluation criterion. Among the items that might be forecast, Teach notably suggests several that would fall into the internal business processes perspective-manufacturing costs, raw material inventory levels, cash flow, and so on-as well as market share which falls into the customer perspective. Akin to Rockwater, Gosen and Washbush employed forecasting accuracy of profit (2002), as well as market share (2002) and units sold (2001), as criteria for validating their learning instrument. Peach and Platt (2000, p. 245) used the accuracy of forecasting demand as a measure of learning. Wolfe (1993a, 1993b), though, has taken exception to the forecasting accuracy criterion. Incorporating forecasts into business games seems feasible. The *BML* does not include forecasting accuracy among its scorecard options, but does make provision for processing forecasts of either net income, units sold, or market share.

Rockwater's customer ranking survey and customer satisfaction index measures in its customer perspective could easily be incorporated into a game's algorithms. Just as sales is a functional outcome of a company's strategy, so could be these criteria. The perceptual maps in *Markstrat* (Larreche and Gatignon 1990) exemplify this possibility for business games generally. The staff attitude survey in Rockwater's innovation and learning perspective could be implemented likewise. The *MME*, for example, regularly generates a sales force morale index in this fashion. Rockwater's rate of improvement index in the innovation and learning perspective is generic enough to be adapted to many game contexts.

As presented earlier, the internal perspective includes innovation. Also as discussed earlier, true innovation or creativity is generally not feasible within the typical competitive business game structure. The Balanced Scorecard's internal perspective also includes, however, measurement of operations processes and the postsale service process. Regarding the former, the *MME* scorecard menu offers an inventory management option (in addition to inventory turnover). But typical total enterprise simulations such as the *BSG* and *Micromatic* do not offer any operations-based measures, though this seems feasible. Likewise, there is no apparent reason why business games cannot incorporate postsale strategy decisions and specific market responses to these decisions.

In the learning and growth perspective, the category of employee capabilities seems particularly feasible for incorporation into business games. Core measures include satisfaction, retention, and productivity. As uncommon examples, the *BML* scorecard includes a "steady work" criterion reflecting layoffs and overtime and the *MME* scorecard includes a sales force turnover criterion. Though not available in its scorecard, the *MME* also generates a sales force morale index (dependent mainly on compensation) and consequent resignations of salespeople. Productivity might be a simple matter of dividing some desired outcome by the number of employees or employees of certain types though, as discussed below, this construct may have inadequate validity in some game designs.

Since information systems capabilities are largely a matter of design and since information generally is highly defined/constrained in business games, that category within learning and growth does not seem to hold promise for business games.

The final category includes motivation and empowerment. Strategy decisions in both areas would be eminently feasible, including job rotation, training, compensation, delegation, span of control, and so on. The strategy decision input-measure output relationship would seem no different in kind than common sales response algorithms or cost functions. This final category also includes alignment which refers to the objectives of departments being consistent with organizational objectives. As discussed earlier, at the present time it seems necessary that evaluation systems be imposed by the game administrator and it is not feasible for participants to devise their own objectives. However, what *seems* infeasible may not, in fact, be infeasible. The Balanced Scorecard paradigm does the service of at least bringing this prospect into consideration.

The relevant purpose of the Balanced Scorecard approach in the present study is evaluation. (It is recognized that in other contexts the Balanced Scorecard may serve other purposes such as diagnosis.) This purpose imposes another limitation on the types of components that might be appropriated from actual Balanced Scorecard implementations. For most components across all four Balanced Scorecard perspectives, it is clear that the most favorable evaluation will attend either the maximum (e.g., profit) or minimum (e.g., employee turnover) value of the measure. Some components, however, are targeted at specific levels other than maximum or minimum. The Kenyon Stores application, for example, specifies average unit retail price as a customer perspective component on the basis of the company wishing to avoid price discounting (Kaplan and Norton 1966, p. 78). In most business simulation games, though, such "targeting" is infeasible. If price is among a game's decision mix then imposition by the game administrator of some target price serves to remove price from the discretion of the participants. It is conceivable that some *average* target price be imposed, e.g., average over the course of the competition or average across sales territories. As a general observation, though, the imposition of target levels by administrators is problematic from a learning standpoint and, to some degree, from a logistical standpoint.

INADEQUATE CONSTRUCT VALIDITY

As discussed above, the feasibility of key Balanced Scorecard measures for business games is mixed. An additional caveat is that some constructs may be only nominally feasible. Productivity types of measures, for example, are commonly defined as ratios of sales or profits per employee or outlet. In the business realm sales per retail outlet is presumably a surrogate for numerous sales-affecting strategies that vary across outlets (as well as idiosyncratic environments, etc.). Store layout, cleanliness, in-store promotions, staff characteristics, and so on presumably impact the sales of outlets and sales per outlet reflects these impacts. Similarly, employee productivity (a core measure, Kaplan and Norton 1966, p. 306) is presumably a function of the skill, knowledge, work habits, and other such traits that vary across

employees. However, it is often the case in business games that a retail outlet is a retail outlet or an employee is an employee. The simulation strategy decision is often just the number of outlets or employees and not any more specific strategies attending those outlets or employees. Outlets and employees are by definition of the game design of identical productivity. Sales per retail outlet, then, is in fact a reflection of the aggregate of all other strategy decisions that affect sales. It is not a reflection of any effectiveness of the retail outlets other than their sheer number. Though sales per outlet, for example, may be identically calculated in the business realm and in a business game, the respective underlying constructs may not be similar.

COMPLEXITY IN INTEGRATION BY PARTICIPANTS

Whatever the evaluation format invoked for participants in simulation games, there is an obvious incentive for participants to play or manage to that format; participants have an imperative to formulate their strategies to maximize their evaluations. The multiple-criteria Balanced Scorecard or any scorecard is presumably much more complex and difficult to integrate into strategy formulation by game participants. Rather than simply concerning themselves with whether their strategies will yield earnings, they must concern themselves with the impact of their strategies on each component comprising the scorecard. Depending on the game context, this may be pedagogically beneficial. But it is more complicated and a potentially more confusing framework in which to formulate strategy. For some game contexts, e.g., high school students or introductory level courses, this confusion may serve to compromise learning.

Reflecting the aim of Balanced Scorecards to incorporate many facets of the enterprise, Balanced Scorecards in use by corporations may comprise large numbers of components. For example, early keystone applications of Balanced Scorecards described by Kaplan and Norton comprise 17 (1992, p. 76) and 20 criteria (1993, p. 136). Typical business simulation games are not so all encompassing. However, the number of potential criteria for specifying a scorecard may still be large; *The Marketing Management Experience* menu presents 11 criteria.

The large number of components in real-world Scorecards is generally not replicable in business simulation games. The responsibility for satisfying 17 or 20 criteria in the real-world is presumably delegated to numerous managers. Simulation game enterprises are typically managed by individual participants or management teams of a limited number of participants. Challenging individual participants to formulate strategies toward achieving numerous criteria is probably dysfunctionally complex.

FEEDBACK

It is important for anyone to understand the basis on

which they are being evaluated. Because of information requirements and the extensive calculations involved, it would be difficult for participants to replicate scorecard index values produced by even modest scorecards. This extends the "Explanation to and Assimilation by Participants" issues presented above. Even if the scorecard approach is conceptually assimilated by participants, they may not be able to very well interpret their and their competitors' actual evaluations. With the univariate profit criterion, one company receives a higher or lower score than a second company because it earned more or less profit, respectively. This is a very comprehensible performance evaluation. It has nowhere near the learning potential, though, of a scorecard approach. Lacking is any indication as to *why* one company earned more than another.

Complete scorecard feedback would provide much greater insight explaining why one company receives a higher or lower score. The *MME* scorecard menu contains 11 criteria, the *Micromatic* and *MMG* menus contain seven, the *BML* menu contains eight, and so on, though all criteria need not be present in any given scorecard. It may be an overload of information to provide participants with, say, both the rank or proration of each competing company on each criterion and the product of that rank position or proration times the criterion's weight. A compromise might be to provide rank or proration information only. Participants, then, could compare their companies with their competitors in terms of performance on the various components. Too, they would have sufficient information to calculate weight x rank or proration products, though with some effort. Regarding evaluation, it is the product that in fact reflects the contribution of the criterion to the final scorecard value. The tradeoff is between participants understanding (via ranks or prorations) the performance of their companies *vis-a-vis* competitors and participants understanding (via products) their evaluation scores.

Under the single criterion of cumulative profit, the custom of informing all participants of the cumulative profit of all competitors is consistent with disclosure practices by publicly held companies. A scorecard approach, however, may require disclosure to participants of two types of information of which they would normally not be aware.

First, meaningful criteria from a normative company performance point of view may not normally be available to managers. The inventory management option of the *MME* mentioned under "Feasibility of Key Balanced Scorecard Measures" presents just such a dilemma. That option is defined as the sum of inventory carrying cost plus the gross margin of lost sales. Realistically, managers usually do not know the amount of sales lost, granting exceptions such as backorders. Informing game participants of this figure toward their understanding their evaluation would diminish the realism of the game experience. At the same time, not informing participants of this figure would bring uncertainty to the basis for their evaluation.

The use of stocking out itself as a criterion bears a similar dilemma. Stocking out is positioned in the customer perspective (K&N 1996, p. 80). Informing game participants

of the extent of stocking out is tantamount to informing them of the attendant opportunity loss as addressed above and is generally unrealistic. Rather, stocking out might better be operationalized as a count or percent of times that a stockout occurred over the course of a simulation competition.

A more general feedback disclosure issue is that the Balanced Scorecard is used for internal evaluation of the enterprise; most of the typical performance measures are not disclosed publicly, much less actively made available to competitors. Compositions of Balanced Scorecards pointedly draw from a wide mix of areas across the enterprise. Adapting Balanced Scorecard tenets for comparative evaluation purposes, i.e., grading, is a mutation of the Scorecard's intended purpose. On one hand, then, continuing feedback to participants would seem to involve disclosing to competitors information that would normally be proprietary and not available to them. On the other hand, such feedback would seem to be necessary. A middle ground might be found in providing to participants only scorecard scores and not the specific values underlying those scores.

SUMMARY AND CONCLUSION

The Balanced Scorecard is a sensible and straightforward performance measurement concept. It offers several positive potential contributions to both substantive learning and performance evaluation in simulation games. Its essence of multiple criteria encompassing many aspects of management is meaningful. Too, administration of the Balanced Scorecard may provide a vehicle for learning in many respects. At the same time, several considerations call into question the feasibility of applying the Balanced Scorecard *per se* to business games or, at least, require material adaptation of the approach.

- Since its specifics in any given organization derive from the organization's vision, mission, and corporate strategy, it seems infeasible for the Balance Scorecard to be adapted *in toto* to competitive business games.

- The Balanced Scorecard does not accommodate the weighting of Scorecard components and does not yield a unidimensional continuum that seems to be necessary for class grading purposes.

- The need to rationalize the use of criteria other than simple cumulative profit.

- The need to communicate and to rationalize specification of particular components to comprise the scorecard.

- The need to rationalize particular weights assigned to scorecard components.

- The need to rationalize a particular data normalization approach.

- The infeasibility of some common Scorecard components due to their being idiosyncratic to companies or targeted to something other than maximum or minimum value.

- Inadequate construct validity of some common Scorecard components.

- Unrealistic responsibility of one or a few managers for numerous Scorecard criteria.

- Quantity of feedback information required.

- Disclosure issues attending feedback.

The above list of limitations would seem to suggest a gridlock that is contradicted by multiple-criterion evaluation approaches already in place and in practice. As reviewed above, several contemporary games offer scorecard evaluation features and some basic research has employed multiple-criterion indices. The reconciliation is that currently available and used scorecard-type evaluations are almost entirely within the traditional financial perspective. On this operational basis and also on several important philosophical bases, they are only minimally within the Balanced Scorecard concept.

A major implication of the Balanced Scorecard for the design of business games, then, may be found in its three nonfinancial perspectives. Within each of the three there are numerous examples of measures that are feasible for business games. Incorporation of such measures into games' algorithms, foremost, might enhance the validity of evaluation. To the extent that the Balanced Scorecard represents a parameter of real-world enterprises, incorporation of its tenets would serve to enhance the validity of the games themselves.

REFERENCES

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

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

Cadotte, Ernest R., *Web Marketplace*.

Dickinson, John R. (2002), "A Comparison of Discrimination-Based Versus Conventional Simulation Game Scoring," *Developments in Business Simulation and Experiential Learning*, Volume 29, 22-25.

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, and Gentry, Ruth Ann (2002), "Managing the Curiosity Gap Does Matter: What do We Need to Do About It?" in Vaughan, Mary Jo and Pillutla, Sharma (editors), *Developments in Business Simulation and Experiential Learning*, Volume 29 (Association for Business Simulation and Experiential Learning), 67-73.

Gentry, James W., Commuri, Suraj R., Burns, Alvin C., and 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., and Morgan, Sandra W. (editors), *Developments in Business Simulation and Experiential Learning*, Volume 25 (Association for Business Simulation and Experiential Learning), 62-68.

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

Gosen and Washbush (2001), "An Initial Validity Investigation of a Test Investigating Total Enterprise Simulation Learning," in Pittenger, Khushwant S. and Vaughan, Mary Jo (editors), *Developments in Business Simulation and Experiential Learning*, Volume 28 (Association for Business Simulation and Experiential Learning), 92-95.

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

Kaplan, Robert S. and Norton, David P. (2001), *The Strategy-Focused Organization* (Boston: Harvard Business School Press).

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

Kaplan, Robert S. and Norton, David P. (1996b), "Using the Balanced Scorecard as a Strategic Management System," *Harvard Business Review*, Vol. 74 (January-February), 75-85.

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

Kaplan, Robert S. and Norton, David P. (1992), "The Balanced ScorecardBMeasures That Drive Performance,"

Harvard Business Review, Vol. 70 (January-February), 71-79.

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

Larreche, Jean-Claude and Gatignon, Hubert (1990), *Markstrat 2* (Redwood, CA: The Scientific Press).

"Letters to the Editor" (1996), *Harvard Business Review*, Vol. 74 (March-April), 168-176

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

Peach, E. Brian and Platt, Richard G. (2000), "Total Enterprise Simulations and Optimizing the Decision Set: Assessing Student Learning Across Decision Periods," in Page, Diana and Snyder, LT (editors), *Developments in Business Simulation & Experiential Learning*, Volume 27 (Association for Business Simulation and Experiential Learning), 242-247.

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

Teach, Richard D. (1992), "Using Forecasting Accuracy as a Measure of Success in Business Simulations," in Gosenpud, Jerry and Gold, Steven C. (editors), *Developments in Business Simulation & Experiential Exercises*, Volume 19 (Association for Business Simulation and Experiential Learning), 103-107.

Teach, Richard D. (1990), "Profits: The False Prophet in Business Gaming," *Simulation & Gaming*, Vol. 21 (March), 12-26.

Thavikulwat, Anderson, Cannon, and Malik (1998), "Games as Instruments of Assessment: A Framework for Evaluation," in Butler, John K., Jr., Leonard, Nancy H., and Morgan, Sandra W. (editors) *Developments in Business Simulation and Experiential Learning*, Volume 25, 31-37.

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

Wolfe, Joseph (1993a), "Forecasting and Management Ability: Rebuttal to Teach's Response," *Simulation & Gaming*, Vol. 24 (March), 73-75.

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