

Developments in Business Simulation and Experiential Learning, Volume 29, 2002

A COMPARISON OF DISCRIMINATION-BASED VERSUS CONVENTIONAL SIMULATION GAME SCORING

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

ABSTRACT

Discrimination-based scoring is a potentially more philosophically and empirically compelling approach to scoring simulation game participants than the conventional cumulative performance approach. This study presents the rudiments of discrimination-based scoring and applies the approach to a typical simulation game competition. Results indicate that scores for many participants may be dramatically affected, providing both philosophical and empirical rationales for pursuing validation research.

INTRODUCTION

In 2001 Dickinson (2001a, 2001b) introduced the concept of discrimination-based scoring. Inspired by the work of Pogossian (1999, 1998, 1997), discrimination-based scoring repositions the typical longitudinal competitive simulation game. In the genre of marketing management and similar types of simulation games, a typical competition is among companies comprising an industry. That perspective is meaningful to managers of the competing companies. Where company performances can be at least ordered—which is normally the case—then the comparative performances of pairs of companies may be deduced from the overall ordering and for purposes of evaluating companies/managers those paired company “competitions” may provide a meaningful basis. Discrimination-based scoring considers each period to be a competition and evaluations of companies/managers are made on that basis. This is in contrast to the common practice of using some cumulative criterion. An apt analogy from sports may be drawn from two common scoring approaches in golf. The criterion under “medal play” is the total number of strokes taken; the criterion under “match play” is the number of holes won. Notably, despite thousands upon thousands of person-years of experience, neither approach has evolved to become the singular criterion. A stream of research comparing these two approaches in simulation gaming and parallel to the present one is being conducted by Thavikulwat (2001).

A theoretical rationale for a period-by-period base versus a cumulative base may be found in the continuing debate attending learning versus performance philosophies of evaluation. Several evaluation philosophies have been encapsulated by Gentry *et al.* (1996). Adherents of the performance philosophy maintain that performance outcomes

reflect learning; managers of a company that has earned greater profit have learned more than have managers of a company that has earned lesser profit. “Since gaming’s earliest years the literature has implicitly accepted the notion that teams that have performed well in the game have learned the most, but this basic relationship has not been investigated.” (Wolfe 1990, p. 293) Prefacing their empirical study that concluded, “Learning did not correlate with performance...” (p. 43), Gosenpud & Washburn (1996) stated, “Common wisdom suggests that people who perform best in simulations do so because they have learned how to play the game better than most.” (p. 43) Adherents of the learning philosophy do not accept that performance equates to learning. Bluntly, “Performance is not a surrogate for learning.” (Burns, Gentry, & Wolfe, 1990, p. 261). An example contention is that game participants may learn from the experience of a lesser profit strategy.

Several implications distinguish between the two philosophies. For example, a learning criterion might “grade on improvement,” while a cumulative performance criterion does not recognize improvement. A disastrously poor or extraordinarily good performance in one or a few periods may be sufficient to dominate cumulative performance, obscuring numerous superior or inferior, but more modestly so, performances in several periods. Numerous superior (inferior) period performances may reflect learning, but not sufficiently affect cumulative performance. Discrimination-based scoring contains elements of both philosophies, though it is not rooted in either of them. In that its basis is period-by-period, it does embody the noncumulative tenet of the learning philosophy. It can reflect improvement, extreme performances affect only single period criteria and, thus, do not dominate cumulative performance, and a majority of superior period performances will yield a higher score than a minority of superior period performances regardless of the degree of superiority. In these respects, discrimination-based scoring is akin to the learning philosophy. On the other hand, as applied in this study, the operational measure for evaluation is profit which is a tenet of the performance philosophy.

The purpose of this study is to investigate the potential impact of discrimination-based versus conventional scoring. It is designed to determine whether the two approaches in fact result in materially different scores for participants. If they do not, then motivation for future research

Developments in Business Simulation and Experiential Learning, Volume 29, 2002

of the approaches is reduced. Another implication of the study is that if the two approaches do not yield materially different scores, then the debate between the learning and performance philosophies may be somewhat mooted, at least from an empirical standpoint.

DATA

The simulation game utilized for this research was *The Marketing Management Experience (MME)* (Dickinson 2000). In the *MME*, participants assume the role of marketing manager for a manufacturer of both still and video digital cameras. Simulation managers formulated their marketing mixes for a total of nine periods plus an initial trial period. Participants in the research were 48 Masters of Business Administration students enrolled in an introductory marketing management course with each student managing his or her own simulation company. The 48 companies were grouped into eight industries of five companies each plus two industries of four companies each. Because companies compete only against other companies within their respective industries and that was the basis on which students were scored and also because different industries may and do grow at different rates across industries as a function of the collective sales-stimulating strategies of the member companies, analyses for this study are on a within-industry basis.

For each period of the competition, each company realizes a profit (loss). However, that profit figure is not precisely reflective of the effectiveness of the manager's strategy for that period. An obvious example is fixed costs. In the *MME* these costs are not controllable by the manager and do not impact sales. Some costs stem from earlier periods. For example, inventory carrying costs for stock in inventory at the beginning of Period t that remains in inventory at the end of Period t are attributable to strategy decisions for periods prior to t . Some costs may be incurred in Period t whose express impact will be entirely in periods subsequent to t . For example, new sales force hires incur hiring and salary compensation costs in Period t , but the new hires do not actually sell until Period $t+1$. Where possible, then, expenses that are not specifically relevant to the performance of the company in a given period were subtracted from nominal profit for that period. No corresponding adjustments to revenue are feasible, though lagged, cumulative, and strategy consistency effects cause sales in a given period to be at least partly a function of strategies in prior periods. Complete isolation of period earnings is antithetical to the nature of a longitudinal game; it is the dependence of periods that defines

a longitudinal game.

Based on isolated period earnings as just described, for a given period companies may be rank ordered within their industry. From that ordering may be deduced pairwise comparisons indicating whether one company outperformed a second company or *vice versa*. In this manner, across the periods of the competition, the proportion of periods in which company k outperformed company j may be calculated. Resulting from this approach is an $n \times n$ matrix P whose elements are proportions, p_{jk} , where n is the number of companies in the industry. P_{jk} is the proportion of periods in which company k outperformed company j . That is, the P matrix is column-dominated. It follows that $p_{kj} = 1 - p_{jk}$

ANALYSIS

The basis for analysis, i.e., determination of each company's/manager's final score, is the P matrix as described above. A simple approach is to sum the respective columns of the P matrix. The performance value for participant k , then, would be the sum of the proportions in column k . More sophisticated methods, though, may be used to transform the proportions matrix into scale values, a seminal method being Thurstone's law of comparative judgment (Thurstone, 1927a, 1927b). In the context of competitive longitudinal simulation games, under Thurstone's Law a company's performance in a given period is viewed as a sample drawn from a distribution, the mean of which is the true measure of company/manager performance. Thurstone's law provides a model by which the P matrix may be used to estimate these true measures. Details of the application of Thurstone's Law may be found in Dickinson (2001a, 2001b). For a complete presentation, the reader is referred to Thurstone's original works (1927a, 1927b) and to Torgerson (1958). Both cumulative earnings and Thurstone scale values were linearly transformed so that the maximum score (within an industry) equaled 100 and the mean score equaled 70.

RESULTS

Results here mainly comprise absolute differences between scores determined in the traditional manner as a function of cumulative earnings and scores determined on a period-by-period basis via discrimination-based scoring using Thurstone's Law, both types of scores being transformed as described above. Results are presented in Table 1.

Developments in Business Simulation and Experiential Learning, Volume 29, 2002

TABLE 1: Discrimination-based Versus Cumulative Earnings Based Scores

Industry	Companies in Industry	Mean Absolute Difference ^a	Maximum Absolute Difference ^a	Minimum Absolute Difference ^a	Change in Ordinal Position ^b
1	5	8.17	18.84	0.00	0
2	5	6.01	9.38	0.00	2
3	5	1.49	3.73	0.00	2
4	5	12.71	23.70	6.62	3
5	5	2.64	5.81	0.00	0
6	5	9.11	20.69	0.00	2
7	5	7.00	12.45	0.00	3
8	5	5.12	12.80	0.00	0
9	4	19.99	39.98	2.62	2
10	4	17.99	35.97	0.00	0
Mean		9.02	18.34	0.09	

a Entries are absolute differences between discrimination-based and cumulative earnings based scores, each set of scores having been transformed to a maximum of 100 and a mean of 70.

b The number of companies whose respective discrimination-based and cumulative earnings based scores placed them in different rank order positions within their industry.

Across the ten industries, the mean absolute difference ranged from 1.49 to 19.99. For fully 12 companies/managers, or 25 percent, the absolute difference exceeded 10 points. For six companies/managers, the absolute difference exceeded 20 points. For substantial numbers of managers, the difference in score is marked. There was no systematic compression or expansion of the scores, with standard deviations for traditional scores exceeding those of discrimination-based scores in six of the ten industries. Ordinal positions within their industries changed for 14 of the 48 companies (29.2%).

Examining the details leading to some of the dramatic differences between traditional and discrimination-based scoring is informative. The most extreme mean absolute difference (=19.99) between the two scoring approaches occurred in the ninth industry comprising four companies. With conventional scoring, Company 4 had the highest earnings and would have received a score of 100. With discrimination-based scoring, Company 4 would have finished

third and received a score of 65.71, a reduction of 34.29 points. Company 2 with conventional scoring would have received a score of 60.02 and finished third. With discrimination-based scoring Company 2 would have placed first and received a score of 100, an increase of 39.98 points.

At the end of the competition Company 4 had earned \$7,057,749.13, while Company 2 had earned \$6,173,390.12. Stated differently, Company 4 had earned \$884,359.01 or 14.33% more than Company 2. The P matrix for the ninth industry is presented in Table 2. Compared to Company 3, Companies 2 and 4 are equal, each having outperformed Company 3 in four of the nine periods (44.4%). Company 2 outperformed Company 1 in seven of the periods, while Company 4 outperformed Company 1 in six of the periods. Head-to-head, Company 2 outperformed Company 4 in twice as many periods (6) as Company 4 outperformed Company 2 (3 periods). With this diagnosis, it is eminently clear why Company 2 improved its score under discrimination-based scoring and why Company 4's score was reduced.

TABLE 2: Column-dominated Proportions Matrix ^a

Company	1	2	3	4
1		0.778	0.667	0.667
2	0.222		0.556	0.333
3	0.333	0.444		0.444
4	0.333	0.667	0.556	

a Entries are the proportions of periods (out of nine) that the company in column k earned more than the company in row j in Industry 10.

Developments in Business Simulation and Experiential Learning, Volume 29, 2002

CONCLUSION AND LIMITATIONS

The results of this study are clear: evaluating simulation game participants using discrimination-based scoring on a period-by-period basis can yield scores for participants that vary materially from scores based on a cumulative end-of-game criterion. Too, considering the diagnosis of those material variations, the discrimination-based approach has logical, conceptual merit. This finding may be generalized to the context of the learning versus performance philosophies; taking into account interim data can yield remarkably different scores compared with the single datum of cumulative performance. The controversy between learning and performance approaches to scoring simulation games poses an interesting amalgam of philosophy and empiricism. This study has shown that the potential for improving the scoring of simulation games is great. Complementing philosophical rationales, there is now an empirical rationale for continuing research into the comparative validity of discrimination-based scoring *vis-a-vis* more conventional approaches.

REFERENCES

- Burns, Alvin C., Gentry, James W. & Wolfe, Joseph (1990). A cornucopia of considerations in evaluating the effectiveness of experiential pedagogies. In Gentry, James W. (Ed.), *Guide to Business Gaming and Experiential Learning* (Statesboro, GA: Association for Business Simulation and Experiential Learning).
- Dickinson, John R. (2001a). An alternative view to scoring selected simulation games and experiential exercises. Association for Business Simulation and Experiential Learning, presentation.
- Dickinson, John R. (2001b). A new perspective on the scoring of objective tests. In Krishnan, Ram & Viswanathan, Madhu (Eds.), *Marketing Theory and Applications*, (Chicago: American Marketing Association), 347-348.
- Dickinson, John R. (2000). *The marketing management experience*. Windsor, Canada: Management Experiences.
- Gentry, James W., Anderson, Phil, Burns, Alvin C., Cannon, Hugh, Faria, Anthony J., Frazer, Ron, Gosenpud, Jerry, Lawton, Leigh, Nulsen, Ray, Teach, Richard & Wellington, Bill (1996). How do we measure the 'learning' in experiential learning and how do we best stimulate it? In Patz, Alan L. & Butler, John K., Jr. (Eds.), *Developments in Business Simulation & Experiential Exercises*, Volume 23 (Statesboro, GA: Association for Business Simulation and Experiential Learning), 197.
- Gosenpud, Jerry & Washbush, John (1996). Correlates of learning in simulations. In Patz, Alan L. & Butler, John K., Jr. (Eds.), *Developments in Business Simulation & Experiential Exercises*, Volume 23 (Statesboro, GA: Association for Business Simulation and Experiential Learning), 43-48.
- Pogossian, Edward M. (1999). Increasing efficiency of management skill assessment. In Morgan, Sandra & Page, Diana (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 26 (Statesboro, GA: Association for Business Simulation and Experiential Learning), 80-81.
- Pogossian, Edward M. (1998). Development of management skill assessment. In Butler, John K., Jr, Leonard, Nancy H., & Morgan, Sandra W. (editors), *Developments in Business Simulation and Experiential Learning*, Volume 25 (Statesboro, GA: Association for Business Simulation and Experiential Learning), 29-30.
- Pogossian, Edward M. (1997). Ability of efficient evaluation of knowledge-based management strategies. 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), 128-129.
- Thavikulwat, Precha (2001). The tournament concept: Extending the idea of using simulations as instruments of assessment. In Pittenger, Khushwant & Vaughan, Mary Jo (Eds.), *Developments in Business Simulation and Experiential Learning*, Volume 28 (Statesboro, GA: Association for Business Simulation and Experiential Learning), 225-233.
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, Vol. 4 (July), 273-286.
- Thurstone, L. L. (1927b). Psychophysical analysis. *American Journal of Psychology*, Vol. 38, 368-389.
- Torgerson, Warren S. (1958). *Theory and methods of scaling*. New York: John Wiley & Sons, Inc.
- Wolfe, Joseph (1990). The evaluation of computer-based business games: methodology, findings, and future needs. In Gentry, James W. (Ed.), *Guide to Business Gaming and Experiential Learning* (Statesboro, GA: Association for Business Simulation and Experiential Learning).