

# EMPIRICAL ANALYSIS OF THE RELIABILITY AND VALIDITY OF BALANCED SCORECARD MEASURES AND DIMENSIONS<sup>☆</sup>

Emilio Boulianne

## ABSTRACT

*For many years management accountants have been involved in the design of information systems for decision-making. To be effective in system design, accountants need pertinent and reliable performance measures within a valid framework. The Balanced Scorecard (BSC) has received a great deal of attention as a comprehensive model of performance that takes into account both financial and non-financial measures. This paper examines the empirical reliability and validity of the BSC framework and its associated measures. With reference to content validity, internal consistency reliability, and factorial validity, results show that BSC, with measures grouped into its four dimensions, is a valid performance model.*

*Previous studies have called for better reliability and validity of BSC measures. The present study may help in the design and implementation of BSCs in business units by adding robustness to the BSC framework, and by suggesting a set of valid measures associated with the four BSC*

<sup>☆</sup>This paper is based on my dissertation completed at HEC-Montreal, Canada.

*dimensions. The results may lead to reduced costs of BSC design and implementation, and enhanced consistency of future studies of the BSC.*

## 1. INTRODUCTION

A consensus has emerged among academics and practitioners that it is important to design and implement performance measurement systems that consider non-financial measures to obtain a better assessment of business performance (Chenhall & Langfield-Smith, 1998; Ullrich & Tuttle, 2004). The Balanced Scorecard (BSC), introduced by Kaplan and Norton (1992), has received a great deal of interest as a framework that takes into account both financial and non-financial measures to provide a comprehensive model of performance measurement. The BSC is one of the major topics examined in management accounting research during the past decade (Ittner & Larcker, 2001). The BSC proposes four dimensions to represent business-unit performance—Financial, Customer, Internal business processes, and Learning and Growth. Despite surveys reporting that a growing number of organizations use the BSC, little is known about the reliability and validity of the BSC's framework and its suggested measures (Ittner, Larcker, & Meyer, 2003; Chenhall, 2003). Surprisingly, little attention has been paid to the BSC as a valid performance measurement model as originally proposed by Kaplan and Norton.<sup>1</sup>

Kaplan and Norton (2001) report that a reason for delay in BSC implementation is that business units may not have developed reliable measures for the scorecards. Problems of valid and reliable measures also have an impact on the credibility and importance allocated to the BSC dimensions by managers (Lipe & Salterio, 2002). For example, Malina and Selto (2001) report that changes in importance are a function of how credible the BSC measures are. They report that for performance assessment of a particular unit, management initially allocated 20% weight to the Learning and Growth dimension, then the year after weighted it to just 4%, then finally eliminated the dimension because management felt the measures associated with this dimension were not reliable. Ittner et al. (1997) point out the importance of establishing the reliability and validity of measures before suggesting any business models. Surveys report that executives worry about the quality and reliability of non-financial performance measures in BSCs, which has an impact on BSC usage (Lingle & Schiemann, 1996; Ittner & Larcker, 2001; Reck, 2001). Moreover, the BSC is expensive to design and implement as it may mobilize management time for up to 2 years (Chow, Haddad, & Williamson, 1997; Lipe & Salterio, 2000).

The present study aims to help in the design and implementation of BSCs in business units by empirically examining the BSC dimensions and its suggested measures. To do this, we conducted a survey research among 90 Canadian business units. First, BSC common measures associated with each of the four BSC dimensions were selected. Measures can be unique to particular units or common to units, but as reported by [Lipe and Salterio \(2000\)](#) and [Dilla and Steinbart \(2005\)](#), it appears that only common measures count as evaluations of performance across units, as measures unique to individual units tend to be ignored. Second, we examined the reliability of these common measures. Finally, we examined the factorial validity of the four BSC dimensions; factorial validity refers to the degree of coherence between a theoretical expectation of dimensions and empirical results. The main question for the purpose of this research is: Does the BSC, with common measures along its four dimensions, represent a valid performance model?

The results show that the BSC represents a valid model; this is an important research contribution to the performance evaluation and management accounting literature. Because it is rather costly to develop, it is important for management to understand that BSC design can be enhanced and implementation issues mitigated by providing validity to the BSC framework.

To our knowledge, there has been no other study that has empirically investigated the BSC from a construct validity perspective, so the present study provides evidence in this area. [Chenhall \(2003\)](#) also points out the importance of spending more time to develop robust and reliable BSC measures to enhance consistency between studies on the BSC.

This paper is organized as follows: Section 2 presents a literature review and Section 3 describes the research methodology; Section 4 reports on the reliability of the BSC measures and the factorial validity of the BSC dimensions; and the last section discusses limitations and presents a conclusion.

## **2. LITERATURE REVIEW**

Organizational psychology literature has pointed out the importance of reliable measures for performance measurement ([Blum & Naylor, 1968](#)). For example, subjective measures for performance assessments are often considered less accurate and reliable than objective measures because they may be influenced by the rater's biases ([Heneman, 1986](#); [Campbell, 1990](#)). Reliability is also regarded as an important factor in the choice of performance

measures (Ittner et al., 2003). According to Malina and Selto (2001) “to be effective as a management control device, the BSC should result in evaluations that are accurate” (p. 75). To examine the reliability of BSC common measures and the validity of its four dimensions, we start our analysis with the content validity step.

### 2.1. Content Validity: Selection of the BSC Common Measures

*Content validity* refers to the use of relevant dimensions and measures to represent a construct.

First, the literature on the BSC clearly proposes four dimensions—Financial, Customers, Internal business processes, and Learning and Growth. These dimensions are considered essential to almost all organizations (Malina & Selto, 2001).

Second, the literature proposes a list of measures associated with each of the dimensions. Kaplan and Norton (1996) reported that across business units, a core set of common measures is found among BSCs observed, regardless of those units’ business objectives. This statement has been emphasized by Lipe and Salterio (2000), who report that experiment participants evaluated their divisions’ performance based solely on common measures, unique performance measures having no effect on the evaluation judgments. The above observations support the viewpoint that the BSC should include only critical performance measures that are mainly reflected in the common measures. As we examine several business units in the present study, the use of a set of common measures is considered appropriate.

We selected the BSC measures based on Kaplan and Norton (1996, 2001) and Kaplan and Atkinson (1998), where several scorecards are presented. Table 1 shows the selected measures associated with the respective dimensions. These measures aim to be representative of a typical BSC having (1) short- and long-term objectives, (2) drivers and outcome measures, and (3) objective and subjective measures.

The selection of measures takes also account of the availability of non-financial data from business units.<sup>2</sup>

For the Financial dimension, *return on assets* and *net profit margin* reflect the financial performance, and *working capital ratio* reflects asset utilization. For the Customer dimension, *marketing expenses to revenues* reflects marketing efforts to solicit new customers, and *revenue growth* is a *proxy* for *market share*. For this dimension, we selected two measures, although previous research has often used a single measure to represent the Customer

**Table 1.** Balanced Scorecard Measures Considered for Reliability and Validity Examination.

| Dimensions          | Measures                       | Sources  |
|---------------------|--------------------------------|--|
| Financial           | Return on asset                | K&N (2001), pp. 31, 82, and 172;<br>K&N (1996), p. 49                              |
|                     | Net profit margin              | K&N (2001), p. 31  |
|                     | Working capital ratio          | K&N (2001), pp. 100 and 172;<br>K&N (1996), p. 52                                  |
| Customer            | Marketing expenses to revenues | K&A (1998), p. 552   |
|                     | Revenue growth                 | K&N (2001), pp. 122 and 198;<br>K&N (1996), p. 80                                  |
| Internal business   | Number of new products         | K&N (2001), p. 37; K&N (1996),<br>pp. 26, 52, 99, 101, and 112                     |
|                     | Number of products offers      | K&A (1998), p. 553   |
|                     | R&D expenses to revenues       | K&N (1996), p. 52 and 99   |
| Learning and growth | Employee absenteeism rate      | K&N (2001), p. 19 and 248;<br>K&N (1996), p. 137                                   |
|                     | Employee turnover rate         | K&A (1998), p. 568; K&N<br>(2001), pp. 19, 99, 172, and<br>309; K&N (1996), p. 131 |
|                     | Training expenses to revenues  | K&N (2001), pp. 147, 248 and<br>309; K&N (1996), p. 29                             |
|                     | Revenue per employee           | K&A (1998), p. 568; K&N<br>(1996), pp. 52, 55, 131, and 154                        |

Note: K&N stands for Kaplan and Norton, while K&A for Kaplan and Atkinson.

aspect (see Banker, Potter, & Srinivasan, 2000; Ittner et al., 1997). We have observed in some BSC studies that *revenue growth* appears in either the Financial dimension or in the Customer dimension, depending on the nature attributed to this measure. *Revenue growth* may be seen as an indicator of financial performance or as an indicator of competitiveness (with a customer focus) reflecting the relative market share and position. For example, growth in sales volume appears in the Customer dimension of Nova Scotia Power’s scorecard (see Kaplan & Norton, 2001, p. 122). Other performance measurement models have a similar classification. In Lynch and Cross’s (1991) Performance Pyramid Model, *revenue growth* is associated with the Market dimension instead of the Financial dimension, and in Fitzgerald, Johnston, Brignall, Silvestro, and Voss (1991) Determinants and Results Matrix, *revenue growth* is associated with the Competitiveness dimension instead of the Financial-performance dimension. For the Internal

business-process dimension, *number of new products* introduced over the last 3 years, *number of product offers*, and *R&D expenses to revenues* reflect innovation initiatives. Finally, for the Learning and Growth dimension, *employee absenteeism rate* and *employee turnover rate* reflect employee satisfaction, *training expenses to revenue* reflects employees' training efforts, and *revenue per employee* reflects employee productivity.

To examine the reliability of the common measures selected and the validity of the BSC dimensions, we collected the above measures among business units. The next section describes how we collected the data.

### 3. RESEARCH METHODOLOGY

Survey research was employed to collect the required data. As managers are reluctant to permit disclosure of information on their units, we worked with a professional accounting organization to support the study and used their members' directory to pre-select a set of units from both manufacturing and service industries.<sup>3</sup>

Members were contacted by telephone and first asked whether they were organized as a business unit, since the BSC literature indicates that the performance measures chosen should be tailored to this unit of analysis. Moreover, only business units of 100 employees or more were targeted as units with less than 100 employees that are unlikely to have clearly attributed fields of responsibilities (Brownell & Dunk, 1991). For those units that fulfilled these criteria, we explained the nature of the study and elaborated upon the information they would be asked to provide. To encourage participation, respondents were promised summarized outcomes of the study. Questionnaires were reviewed for clarity and forwarded to the units that agreed to participate.<sup>4</sup>

Respondents were asked to provide financial and non-financial data to calculate the return on asset, net profit margin, working capital ratio, revenue growth, marketing expenses to revenues, number of new products, number of product offers, R&D expenses to revenues, training expenses to revenues, and revenue per employee measures. For the employee absenteeism rate and employee turnover rate measures, respondents were asked to classify their business unit's compared with peers' using a 7-point scale (1 meaning a high rate, 7 a low rate). Respondents were also asked to provide annual revenues for size classification and Standard Industrial Classification (SIC code) for industry classification (the appendix shows how these measures were collected).

Five hundred firms were contacted, and the 380 that agreed to participate received questionnaires. We conducted three telephone reminders at intervals of two weeks, four weeks, and six weeks. We received the questionnaires from 128 units, although responses from 38 units were eliminated because the questionnaires were incomplete. The sample consequently consists of 90 questionnaires, for a response rate of 24%. From these 90 business units, 85 are stand-alone firms, and 5 are business units of two large firms. The main reasons mentioned for non-participation in the study were confidentiality concerns.

The profile of the average respondent is a comptroller who holds a bachelor's degree in commerce with an accounting designation and has an average age of 42 years. At the business-unit level, the average number of employees is 156, with average revenues of 22 million Canadian dollars. The sample of business units consists of 48 manufacturing (53%) and 42 services (47%).<sup>5</sup> A *t*-test on industry, including all variables, shows no significant differences between manufacturing versus services groups. To estimate the non-response bias, we compared *late respondents* vs. *early respondents* and results indicate that we do not have the presence of non-respondents bias. Table 2 provides descriptive statistics of measures collected, while Table 3 provides a correlation matrix showing some anticipated relationships between the measures. For example, strong correlations are observed for the measures associated with the Financial and the Learning and Growth dimensions. The next section examines the reliability of BSC measures and the validity of BSC dimensions.

#### 4. INTERNAL CONSISTENCY RELIABILITY OF MEASURES AND FACTORIAL VALIDITY ASSESSMENT OF THE BSC DIMENSIONS

Cronbach's  $\alpha$  is the most recognized estimation of reliability in management accounting research (Brownell, 1995). We used the Cronbach's  $\alpha$  coefficient to estimate the internal consistency reliability of measures. Coefficient  $\alpha$  is therefore calculated first for each dimension (Churchill, 1979). Table 4 presents the BSC measures along with Cronbach's  $\alpha$  coefficients for each dimension.

As shown in Table 4, we obtained a Cronbach's  $\alpha$  coefficient of 0.64 for the three measures of the Financial dimension. This coefficient would be higher if we deleted the *working capital ratio*, but we kept it because of its sound content validity and because at early stage, a coefficient of around

**Table 2.** Descriptive Statistics of Measures Collected.

| Measures:                            | Data Obtained from Respondents |         |         |           |                                   |
|--------------------------------------|--------------------------------|---------|---------|-----------|-----------------------------------|
|                                      | Mean                           | S.D.    | Minimum | Maximum   | Theoretical Range                 |
| Return on assets (ROA)               | 9.49                           | 6.52    | -4.70   | 31.40     | Does not apply for these measures |
| Net profit margin (NPM)              | 5.05                           | 5.36    | -3.00   | 35.00     |                                   |
| Working capital ratio (WC)           | 1.73                           | 1.25    | 0.19    | 10.90     |                                   |
| Marketing expenses to revenues (MRK) | 0.02                           | 0.04    | 0       | 0.32      |                                   |
| Revenue growth (REVGR)               | 6.81                           | 18.04   | -38.5   | 71.2      |                                   |
| Number of new products (NEWP)        | 15.29                          | 19.64   | 0       | 50        |                                   |
| Number of products offers (POFF)     | 20.28                          | 35.36   | 1       | 90        |                                   |
| R&D expenses to revenues (R&D)       | 0.0095                         | 0.0158  | 0       | 0.12      |                                   |
| Training expenses to revenues (TRAI) | 0.0026                         | 0.0056  | 0       | 0.02      |                                   |
| Revenue per employee (RPE)           | 254,300                        | 465,233 | 25,000  | 4,277,992 |                                   |
| Employee absenteeism rate (ABS)      | 5.41                           | 1.25    | 2       | 7         | 1-7                               |
| Employee turnover rate (TURN)        | 5.24                           | 1.63    | 1       | 7         | 1-7                               |

Note:  $n = 90$ .

0.60 is considered reasonable (Nunnally, 1967, p. 226). For the Customer dimension the  $\alpha$  coefficient is 0.51, which shows that the two measures are compatible enough for purposes of reliability. For the Internal business dimension, we have to delete the *R&D expenses to revenues* measure to obtain an  $\alpha$  coefficient of 0.55. Finally, for the Learning and Growth dimension, two iterations are necessary; first, we must delete the *revenue per employee* measure to obtain an  $\alpha$  of 0.43, then we must delete the *training expenses to revenue* measure to obtain an  $\alpha$  coefficient of 0.58.

There are theoretical arguments to support this iterative process of Cronbach's  $\alpha$  coefficient computation, deletion of items, and recomputation until an acceptable coefficient is achieved for each dimension (see Churchill, 1979, p. 69). Factor analysis can then be used to validate whether the four dimensions as proposed by Kaplan and Norton can be observed empirically, which would permit the examination of the factorial validity of the BSC.

**Table 3.** Correlation Matrix.

|       | ROA         | NPM         | WC          | MRK         | REVGR       | NEWP        | POFF        | R&D         | TRAI        | RPE         | ABS         | TURN        |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| ROA   | <i>1.00</i> |             |             |             |             |             |             |             |             |             |             |             |
| NPM   | 0.71**      | <i>1.00</i> |             |             |             |             |             |             |             |             |             |             |
| WC    | 0.07        | 0.26*       | <i>1.00</i> |             |             |             |             |             |             |             |             |             |
| MRK   | -0.16       | 0.04        | 0.45**      | <i>1.00</i> |             |             |             |             |             |             |             |             |
| REVGR | 0.11        | 0.04        | -0.04       | 0.29*       | <i>1.00</i> |             |             |             |             |             |             |             |
| NEWP  | 0.01        | 0.08        | -0.12       | 0.08        | 0.08        | <i>1.00</i> |             |             |             |             |             |             |
| POFF  | 0.06        | -0.01       | -0.15       | -0.02       | -0.11       | 0.39**      | <i>1.00</i> |             |             |             |             |             |
| R&D   | -0.10       | 0.00        | 0.17        | 0.22        | -0.12       | 0.11        | 0.03        | <i>1.00</i> |             |             |             |             |
| TRAI  | 0.01        | -0.04       | -0.17       | -0.05       | -0.24*      | 0.07        | 0.17        | 0.13        | <i>1.00</i> |             |             |             |
| RPE   | -0.27**     | -0.14       | 0.02        | 0.14        | -0.04       | 0.06        | 0.02        | 0.01        | -0.03       | <i>1.00</i> |             |             |
| ABS   | 0.06        | 0.12        | 0.04        | 0.02        | 0.10        | -0.11       | 0.03        | 0.02        | -0.13       | 0.21*       | <i>1.00</i> |             |
| TURN  | 0.04        | 0.13        | 0.27**      | -0.01       | -0.06       | -0.07       | -0.01       | -0.04       | -0.15       | 0.02        | 0.43**      | <i>1.00</i> |

\*\*Pearson correlation is significant at the 0.01 level.

\*Significant at the 0.05 level, (2-tailed),  $n = 90$ .

*Factorial validity* refers to the degree to which an empirical factor analysis is coherent with a priori theoretical expectations (Kerlinger, 1986). We therefore performed a principal components analysis, Varimax rotation, with the remaining measures (measures in italic in Table 4).

Table 5 presents the results that confirm the four BSC dimensions proposed by Kaplan and Norton, results that are also consistent with Hoque and James’ (2000) study. Only one measure, *working capital ratio*, does not clearly fit the BSC dimensions, with a loading of 0.268 for the Financial dimension and a loading of 0.275 for the Learning and Growth dimension. Kerlinger (1986, p. 572) indicates that in some studies, low-factor loadings have already been retained. We therefore maintain for now the *working capital ratio* measure for the Financial dimension because of its sound content validity and weak association with the Learning and Growth dimension. As a reminder, the previous reliability analysis shows that Cronbach’s  $\alpha$  coefficient of the Financial dimension could be improved from 0.64. to 0.82 by deleting the *working capital ratio* measure; this will be kept in mind during analysis.

To increase robustness to the above results, we also ran a factor analysis with the BSC measures, but without reference to Kaplan and Norton’s dimensions (see Table 6). The first rotation provided five factors, with two measures not loading on any factors—*training expenses to revenue* and *revenue per employee*. We deleted these two measures and the second rotation also provided five factors. We then calculated Cronbach’s  $\alpha$  coefficient for each factor (dimension). Results obtained are the same as in Table 5 for the

**Table 4.** Balanced Scorecard Measures with Cronbach's Alpha Coefficients per Dimension ( $n = 90$ ).

| Dimension           | Measures                              | Cronbach Alpha Coefficient | Alpha if Item Deleted after First Iteration |                  | Final Cronbach Alpha |
|---------------------|---------------------------------------|----------------------------|---|------------------|----------------------|
| Financial           | <i>Return on asset</i>                | 0.64                       | 0.21  |                  | 0.64                 |
|                     | <i>Net profit margin</i>              |                            | 0.05  |                  |                      |
|                     | <i>Working capital ratio</i>          |                            | 0.82  |                  |                      |
| Customer            | <i>Marketing expenses to Revenues</i> | 0.51                       | 0.23  |                  | 0.51                 |
|                     | <i>Revenue growth</i>                 |                            | 0.03  |                  |                      |
| Internal business   | <i>Number of new products</i>         | 0.42                       | 0.00  |                  | 0.55                 |
|                     | <i>Number of products offers</i>      |                            | 0.01  |                  |                      |
|                     | <i>R&amp;D expenses to revenues</i>   |                            | 0.55  |                  |                      |
| Learning and growth | <i>Employee absenteeism rate</i>      | 0.00                       | First Iteration                             | Second Iteration | 0.58                 |
|                     | <i>Employee turnover rate</i>         |                            | 0.00  | 0.00             |                      |
|                     | <i>Training expenses to revenues</i>  |                            | 0.00  | 0.58             |                      |
|                     | <i>Revenue per employee</i>           |                            | 0.43  | deleted          |                      |

*Note:* Descriptive statistics for the measures above are available in Table 2. Due to the high kurtosis index, Internal business measures have been transformed using the square foot for use in reliability analysis.

The nine measures in italics will be examined in further analysis.

Learning and Growth (F3), Internal business (F4), and Customer (F5) dimensions. The Financial dimension (F1) still includes *return on assets* and *net profit margin* measures, but not the *working capital ratio* measure, which loads highly (0.729) on another dimension (F2), leading to an increase of the  $\alpha$  coefficient for the Financial dimension from 0.64 to 0.82. This analysis indicates again that the  $\alpha$  coefficient could be improved by deleting the *working capital ratio*; this measure is therefore finally deleted from the Financial dimension. The F2 dimension includes the *working capital ratio* and *R&D expenses to revenues* measures; we calculated the  $\alpha$  coefficient for these two measures but the  $\alpha$  was only 0.34, which reveals reliability issues.

**Table 5.** Factor Analysis of BSC Measures with Reference to Kaplan and Norton’s Dimensions.

| Measures                       | Factor Loadings |                           |                         |                |
|--------------------------------|-----------------|---------------------------|-------------------------|----------------|
|                                | F1<br>Financial | F2<br>Learning and Growth | F3<br>Internal Business | F4<br>Customer |
| Return on asset                | 0.914           |                           |                         |                |
| Net profit margin              | 0.917           |                           |                         |                |
| Working capital ratio          | 0.268           | 0.275                     | -0.349                  | -0.505         |
| Marketing expenses to revenues |                 |                           |                         | 0.487          |
| Revenue growth                 | 0.127           |                           | -0.144                  | 0.850          |
| Number of new products         |                 |                           | 0.777                   |                |
| Number of product offers       |                 |                           | 0.842                   |                |
| Employee absenteeism rate      |                 | 0.836                     |                         | 0.229          |
| Employee turnover rate         |                 | 0.823                     |                         | -0.240         |
| Eigenvalues                    | 1.948           | 1.456                     | 1.341                   | 1.088          |

*Note:* Extraction method: Principal component analysis.  
 Rotation method: Varimax with Kaiser normalization.  
 Variance explained with the four factors: 72.912%.  
 Absolute values less than 0.10 have been suppressed.

The above results combined (Tables 4, 5, and 6) demonstrate internal consistency reliability of eight BSC common measures. These measures associated with the four BSC dimensions represent a valid core set of measures that may be used as a starting point for BSC design. Results also support the specific BSC structure of four dimensions as proposed by Kaplan and Norton as showing factorial validity (i.e., coherence between theoretical expectations and empirical results). These results support the [Lipe and Salterio \(2002\)](#) study, which demonstrates that the four BSC dimensions are important to managers for performance evaluation.

## 5. DISCUSSION, LIMITATIONS, AND CONCLUSION

The objective of this paper was to examine the reliability of BSC measures and the validity of its framework. [Chenhall \(2003\)](#) points out the importance of developing robust and reliable BSC measures to enhance consistency between BSC studies.

Referring to the concepts of content validity, internal consistency reliability, and factorial validity, results indicate that the BSC four dimensions with a set of common measures represent a valid performance model. The

**Table 6.** Factor Analysis of BSC Measures, with no Reference to Kaplan and Norton's Dimension, and Cronbach's alpha Coefficients.

| Measures                      | Factor Loadings |        |        |        |        |                 |        |        |        |                 |
|-------------------------------|-----------------|--------|--------|--------|--------|-----------------|--------|--------|--------|-----------------|
|                               | First Rotation  |        |        |        |        | Second Rotation |        |        |        |                 |
|                               | F1              | F2     | F3     | F4     | F5     | F1              | F2     | F3     | F4     | F5              |
| Return on asset               | 0.899           | -0.113 |        |        | 0.102  | 0.922           | -0.137 |        |        |                 |
| Net profit margin             | 0.863           | 0.156  | 0.154  |        |        | 0.914           | 0.124  |        |        |                 |
| Working capital ratio         | 0.193           | 0.755  | 0.218  | -0.208 |        | 0.232           | 0.729  | 0.211  | -0.240 |                 |
| Revenue growth                |                 | -0.119 |        |        | 0.796  |                 |        |        |        | 0.964           |
| Marketing expenses to revenue | -0.127          |        |        |        | 0.412  |                 |        |        |        | 0.468           |
| Number of new products        |                 |        | -0.117 | 0.835  |        |                 |        | -0.102 | 0.822  | 0.126           |
| Number of product offers      |                 | -0.116 | 0.116  | 0.774  | -0.206 |                 | -0.113 |        | 0.799  | -0.175          |
| R&D expenses to revenue       |                 | 0.556  |        | 0.204  | -0.291 |                 | 0.591  |        | 0.260  | -0.104          |
| Employee absenteeism rate     |                 |        | 0.797  |        | 0.124  |                 |        | 0.826  |        | 0.183           |
| Employee turnover rate        |                 |        | 0.836  |        |        |                 |        | 0.841  |        | -0.156          |
| Training expenses to revenue  |                 |        | -0.191 | 0.174  | -0.693 |                 |        |        |        | measure deleted |
| Revenue per employee          | -0.472          |        | 0.186  | 0.150  | 0.208  |                 |        |        |        | measure deleted |
| Eigenvalues                   | 2.056           | 1.723  | 1.511  | 1.390  | 1.125  | 1.968           | 1.678  | 1.395  | 1.321  | 1.001           |
| Final Cronbach alpha          | n.a.            | n.a.   | n.a.   | n.a.   | n.a.   | 0.82            | 0.34   | 0.58   | 0.55   | 0.51            |

Note: Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser normalization.

Absolute values less than 0.10 have been suppressed.

present study may therefore help the design and implementation of BSC in organizations by suggesting a set of measures associated with the specified BSC structure of four dimensions. Business units adapt their BSC measures to changes in strategy and/or the availability/development of reliable measures (Malina & Selto 2001). Simons (2000) reports that a well-designed BSC should permit a balance between short-term and long-term objectives, drivers and outcome measures, and objective and subjective measures; when examined the common measures reflect this.

In the future, researchers should examine the reliability of the BSC measures analyzed here with other units in different business settings. Churchill (1979) states that if a construct is more than a measurement artifact, it should be reproducible with a new sample when using reliable

measures: reliable and valid measurement is the *sine qua non* of science. Doing this will enhance the robustness and reliability of BSC studies and offer a stronger base for BSC theory development. Lipe and Salterio (2000) report that accounting research should be conducted with relevant theories, but the theory is not yet developed for performance assessment. The present study is an initiative toward a theory-building perspective in examining the validity of the BSC as a performance model.

The present study has limitations and we note the most important. First, we agree that a larger sample would increase confidence in the results, but we had to deal with the difficulties of obtaining financial and non-financial data at the business-unit level, which also limited the number of BSC measures when examined. Second, although we carefully developed questionnaires to be concise and clear, some respondents may have misunderstood the instrument; this is a limit of this method. Third, we referred to and applied rigorous reliability and validity concepts, although these notions have limits. For example, reliability is rarely fully measured, but always estimated (Peter, 1979). Finally, as reported by Ittner and Larcker (1998), BSC measures developed for planning/management, compensation, or performance evaluation, are most likely not appropriated for the three contexts. The present results therefore apply to performance evaluation only.

For many years, management accountants have been involved in the design of information systems for decision-making. With the advent of integrated information systems such as the BSC, the “information producer” function of the accountant has become more challenging. To be effective in the design of BSCs, accountants need pertinent and reliable BSC measures within a valid framework—otherwise, measures used will not reflect business-unit performance.

Rigorous research on the BSC is only beginning to emerge. The present study aims to be one of them.

## NOTES

1. Kaplan and Norton (2001) stated that “several years ago, we introduced the Balanced Scorecard. At that time, the Balanced Scorecard was about performance measurement, not about strategy” (p. 3). The reader should see the BSC as a construct aiming to assess business unit performance. This is the original aspect of this paper, since previous studies on the BSC took for granted the suggested measures and the four quadrants/dimensions.

2. Discussions with business unit managers, before we developed the questionnaires, provided us indications on the performance measures we could obtain from

them. As we were not interested in asking for measures not available from respondents, those discussions helped us to define information we could ask for.

3. This professional accounting organization is the Certified General Accountants (CGA). CGA-Canada is a Canadian professional accounting association representing 62,000 members and students. We worked with CGA-Quebec, an affiliate of CGA-Canada, which represents 10,000 members and students. The study follows an initiative by the author and CGA-Quebec on a project called Performance Indicators. Respondents were aware of the BSC approach.

4. Two academics and an adviser in linguistics reviewed the questionnaires.

5. The business units were in pulp and paper, textile, transformation, construction, industrial products, food products, retailer, wholesaler, leasing, and dealers. The percentage per industry is similar to the five hundred units contacted.

## ACKNOWLEDGEMENT

I thank the members of my dissertation committee, Suzanne Rivard, Michel Vezina, Claude Laurin and Michel Guindon from HEC-Montreal, and Alain Pinsonneault from McGill University, Canada. I acknowledge the comments of two anonymous reviewers. Thanks to CGA-Canada and the Lawrence Bloomberg Chair in Accountancy at Concordia University for their financial support.

## REFERENCES

- Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review*, 75, 65–92.
- Blum, M. L., & Naylor, J. C. (1968). *Industrial psychology*. New York, NY: Harper & Row.
- Brownell, P. (1995). *Research methods in management accounting (Monograph no. 2)*. Australia and New Zealand: Coopers & Lybrand Accounting Research Methodology.
- Brownell, P., & Dunk, A. S. (1991). Task uncertainty and its interaction with budgetary participation and budget emphasis: Some methodological issues and empirical investigation. *Accounting Organization and Society*, 16(8), 693–703.
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In: M. D. Dunnette, L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1, 2nd ed., pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.
- Chenhall, R. H. (2003). Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organization, and Society*, 28, 127–168.
- Chenhall, R. H., & Langfield-Smith, K. (1998). The relationship between strategic priorities, management techniques and management accounting: An empirical investigation using a system approach. *Accounting Organization and Society*, 23(3), 243–264.

- Chow, C. W., Haddad, K. M., & Williamson, J. E. (1997). Applying the balanced scorecard to small companies. *Management Accounting*, 79(2), 21–27.
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64–73.
- Dilla, W., & Steinbart, P. (2005). Relative weighting of common and unique balanced scorecard measures by knowledge decision makers. *Behavioral Research in Accounting*, 17, 43–53.
- Fitzgerald, L., Johnston, R., Brignall, S., Silvestro, R., & Voss, C. (1991). *Performance measurement in service businesses*. London, UK: CIMA.
- Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology*, 39(4), 811–826.
- Hoque, Z., & James, W. (2000). Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research*, 12, 1–17.
- Ittner, C. D., & Larcker, D. F. (1998). Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research*, 10, 205–238.
- Ittner, C. D., & Larcker, D. F. (2001). Assessing empirical research in managerial accounting: A value based management perspective. *Journal of Accounting and Economics*, 32, 349–410.
- Ittner, C. D., Larcker, D. F., Meyer, M. W. (1997). *Performance, compensation, and the balanced scorecard*. Working Paper, The Wharton School, University of Pennsylvania.
- Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, 78(3), 725–758.
- Kaplan, R. S., & Atkinson, A. A. (1998). *Advanced management accounting* (3rd ed). New Jersey, NY: Prentice Hall.
- Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action*. Boston, MA: Harvard Business School Press, 322 pages.
- Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston, MA: Harvard Business School Press, 400 pages.
- Kerlinger, F. N. (1986). *Foundations of behavioral research*, New York: Holt, Rinehart & Wilson, 667 pages.
- Lingle, J. H., & Schiemann, W. A. (1996). From balanced scorecard to strategic gauges: Is measurement worth it. *Management Review*, 85(3), 56–61.
- Lipe, M. G., & Salterio, S. E. (2000). The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, 75(3), 283–298.
- Lipe, M. G., & Salterio, S. E. (2002). A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organization and Society*, 27, 531–540.
- Lynch, R. L., & Cross, K. F. (1991). *Measure up! Yardsticks for continuous improvements*. Cambridge, MA: Blackwell.
- Malina, M. A., & Selto, F. H. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research*, 13, 47–90.
- Nunnally, J. (1967). *Psychometric methods*. New York: McGraw-Hill Books, 640 Pages.
- Peter, J. P. (1979). Reliability: A review of psychometric basics and recent marketing practices. *Journal of Marketing Research*, 16(1), 6–17.

- Reck, J. L. (2001). The usefulness of financial and nonfinancial performance information in resource allocation decisions. *Journal of Accounting and Public Policy*, 20, 45–71.
- Simons, R. (2000). *Performance measurement and control systems for implementing strategy*. Upper Saddle River, NJ: Prentice-Hall.
- Ullrich, M. J., & Tuttle, B. M. (2004). The effects of comprehensive information reporting systems and economic incentives on managers' time-planning decisions. *Behavioral Research in Accounting*, 16, 89–105.

## APPENDIX

Based on the definitions provided, calculate the following measures:

**Return on asset:**

$$\frac{\text{net profit} + \text{interest expense}}{\text{total assets}}$$

For your unit, the  
**return on asset** is:

---

**Net profit margin:**

$$\frac{\text{net profit}}{\text{total revenue}}$$

For your unit, the  
**net profit margin** is:

---

**Revenue growth:**

$$\frac{\text{sales current year (less)} \\ \text{sales previous year}}{\text{sales previous year}}$$

For your unit, the  
**revenue growth** is:

---

**Marketing expenses to revenues:**

$$\frac{\text{marketing expenses}}{\text{total revenue}}$$

For your unit, **marketing expenses to revenues** is:

---