

## An investigation of the effect of Balanced Scorecard implementation on financial performance

Stan Davis<sup>a,\*</sup>, Tom Albright<sup>b</sup>

<sup>a</sup> Babcock Graduate School of Management, Wake Forest University, P.O. Box 7659, Winston-Salem, NC 27109-7659, USA

<sup>b</sup> Culverhouse School of Accountancy, The University of Alabama, P.O. Box 870220, Tuscaloosa, AL 35487-0220, USA

Received 25 November 2002; accepted 30 November 2003

---

### Abstract

In this quasi-experimental study, we investigate whether bank branches implementing the Balanced Scorecard (BSC) outperform bank branches within the same banking organization on key financial measures. Although the BSC has gained popularity among managers as a performance measurement tool, little empirical evidence exists to substantiate claims that the BSC promotes superior financial performance when compared to a traditional performance measurement system. We find evidence of superior financial performance for branches implementing the BSC when compared to non-BSC implementing branches.

© 2003 Elsevier Ltd. All rights reserved.

**Keywords:** Balanced Scorecard; Performance measurement systems; Field study; Quasi-experimental design

---

### 1. Introduction

The purpose of this paper is to investigate the effectiveness of the Balanced Scorecard (BSC) in improving financial performance. The BSC has gained increasing popularity as an effective management tool that aligns employee actions and goals with corporate strategy since first being introduced in 1992. We present an empirical analysis that investigates the impact of BSC on a banking institution's financial performance.

Beginning in the early 1980s, management accounting researchers described the increasing irrelevance of traditional control and performance measurement practices. Weaknesses included failure to link performance measurement to strategic initiatives of organizations, an emphasis on accounting for external reporting rather than on accounting reports useful for internal decision making, and a failure to account

---

\* Corresponding author. Tel.: +1-336-758-4492; fax: +1-336-758-4514.

E-mail address: stan.davis@mba.wfu.edu (S. Davis).

for advances in technology that change how manufacturing firms operate (Palmer, 1992; Spicer, 1992). The growing importance of service industries and increased global competition has further intensified the need for alternative control and performance measures.

The BSC arose out of the need to improve the planning, control, and performance measurement functions of management accounting. Because of the rise in popularity of the BSC, and benefits attributed to its usage, Atkinson et al. (1997) state the BSC is a significant development in management accounting that deserves intense research attention. They suggest using multiple research methods, including case studies, behavioral experiments, and archival approaches. Although much has been written extolling the benefits of the BSC, few studies exist that directly assess financial performance benefits associated with the BSC or claims the BSC is superior to other performance measurement systems. This study seeks to determine whether an improvement in financial performance occurred after implementing a BSC and whether the change in financial performance is significantly greater than performance observed in a similar setting where a traditional performance measurement system using only financial measures is employed.

To achieve and sustain improved financial performance are among the proposed benefits identified by BSC advocates, yet no study has established a strong causal link between BSC usage and improved financial performance. Using quasi-experimental field-based research methods consistent with Yin (1994) and Cook and Campbell (1979), we gain insights into the effectiveness of the BSC by comparing the performance of BSC implementers to the performance of BSC non-implementers.<sup>1</sup> The data set is unique because we have a designed experiment setting (experimental and control groups with pre- and post-test data) in the context of a field study. Our study contributes to current literature by directly examining an actual BSC program and its ability to improve financial performance in an organization.

A primary tenet of the BSC is that success must be achieved on key non-financial measures (NFM) prior to realizing success on key financial measures. Employing the BSC method aids managers in identifying those key NFMs that are linked to success on selected financial measures. Previous studies seeking to establish linkages between specific NFMs and improved operational and financial performance have mixed results. This study differs from these prior studies in four fundamental ways. First, although earlier research examined relationships between NFMs and performance, few sought to establish an association between the implementation of a BSC performance measurement system that places more emphasis on a group of NFMs and improved financial performance. Second, many studies relied on survey and archival research methods to obtain information about performance measurement practices and organizational performance. This study utilizes a quasi-experimental approach to investigate the effects that a program focusing on NFMs has on organizational performance. Third, many studies relied on self-reported measures of organizational performance that asked respondents to rate a firm's performance as above or below industry averages. Other studies used company-wide financial performance measures rather than self-reported measurement assessments. This study benefits from the use of actual financial performance data for individual business units (BU) within the organization as a means to determine changes in financial performance—the objective identified by the scorecard designers as its dependent variable. Finally, most studies used cross-sectional analysis to compare performance and NFMs at a point in time. This study uses a longitudinal approach to determine if changes in financial performance are achieved with the implementation of a BSC program.

---

<sup>1</sup> Cook and Campbell (1979) define a quasi-experiments in field settings as “experiments that have treatments, outcome measures, and experimental units, but do not use random assignment to create the comparisons from which treatment-caused change is inferred”.

Financial performance is operationalized as a branch's rating on a composite measure of nine key financial performance measures. Because this composite measure is used for bonus calculations, both groups of branches at the field site (experimental and control) seek to improve their performance on this measure. We compare performance levels for a period before implementing the BSC with a period two years after implementing the program. The Wilcoxon rank test reports a significant increase in performance occurred during the observation period ( $P$ -value = 0.034) for the experimental branches. A similar performance comparison for control branches revealed an insignificant change in performance during the observation period.

We also compare the change in financial performance for the experimental division and the control division over the same time period. Results indicate the experimental division realized greater improvement in financial performance than the control division ( $P$ -value < 0.02). Thus, we provide evidence supporting the proposition that the implementation of the BSC program resulted in superior financial performance compared to what would have been achieved if the BSC program had not been implemented.

## 2. Literature review

The BSC translates the often-nebulous goals found in corporate mission statements into a strategic roadmap to be followed by employees. By detailing specific actions and outlining cause-and-effect relationships between those actions and key financial objectives, a BSC serves not only as a performance measurement system, but also as a means for communicating long-term strategic initiatives to business-units and achieving long-term financial success. It combines important practices and concepts from various disciplines and theories into a single performance measurement system for the purpose of improving financial performance.

Improved financial performance after the implementation of the BSC relies on the identification of key leading indicators of desired financial performance. These leading indicators, typically non-financial in nature, are logically derived from establishing causal links between improved performance on NFMs and improved performance on selected financial measures. Employing the BSC method aids managers in the selection of these key indicators through viewing the organization from four different perspectives: financial, customer, internal processes, and learning and growth (Kaplan and Norton, 1996). If the causal linkages between NFMs and financial measures are sound, focusing on improving leading indicator measures should lead to improved performance on the selected financial measures.<sup>2</sup>

A review of the current literature reveals several attempts to associate improved financial performance with improved performance on NFMs with mixed results.<sup>3</sup> The mixed results could stem from the lack of a coherent strategy and plan for placing emphasis on certain NFMs rather than others. To the extent that a BSC program provides an appropriate framework and process for integrating NFMs into a performance measurement system, organizations utilizing a properly designed and implemented BSC should experience better financial performance than organizations that do not utilize such a program.

<sup>2</sup> An important caveat is the manner in which the BSC is implemented. According to Moon and Fitzgerald (1996), standards should be quantifiable and achievable. The rewards system should be clearly understood, rewards should be meaningful, and standards upon which employee performance is measured should be under the control of the employee. See Moon and Fitzgerald (1996) for a complete discussion of their framework.

<sup>3</sup> See, for example, Terziovski et al. (1997), Chenhall (1997), Lau (1997), Groves and Valsamakis (1998), Perera et al. (1997) and Ittner and Larcker (1998).

Recent literature includes three attempts to associate BSC usage and improved organizational performance. Hoque and James (2000) surveyed Australian manufacturing firms on their usage of nonfinancial measures typically found in discussions of BSC development. Organizational performance was a self-reported measure relative to peers within the same industry.<sup>4</sup> Their results indicate a significantly positive relationship between the usage of typical BSC measures and superior performance. The authors noted, however, that while their study relates the use of NFMs to performance, their survey fails to capture actual reliance on the BSC or the strength of the causal relationships that are so important to BSC implementation.

In an extensive field study of a large manufacturing organization, Malina and Selto (2001) investigate the effectiveness of the BSC in communicating strategic objectives and serving as a management control device. They find evidence of an indirect relationship between BSC's management control function and improved performance on BSC measures. Further, managers in their study perceived improved performance on the BSC would lead to improved efficiency and profitability.

Ittner et al. (2003) provide contradictory evidence to the two previously mentioned studies by finding a negative association between BSC usage and financial performance (ROA) in an expansive study of the financial services industry. They also find that while 20% of the respondents reporting using the BSC, over 75% of these firms reported not relying on business models that causally link performance drivers to performance outcomes.

Finally, one study has investigated directly the relationship between improved financial performance and the implementation of a performance measurement system that includes NFMs. Banker et al. (2000) examined the association between improved financial performance and NFMs in a hotel chain where a new incentive program included an emphasis on customer satisfaction performance measures. They found evidence of a relationship between customer satisfaction NFMs and future financial performance that suggests the implementation of the new incentive program positively impacted targeted NFMs and ultimately improved financial performance for the hotels on the new incentive program.

In a similar vein, the current study seeks to document whether improved financial performance is observed after the implementation of the BSC in a banking institution. The BSC is a management initiative that develops performance measures based on specific strategies for each business unit. The linkages between the NFMs chosen for the performance measurement system and improved financial performance are identified and communicated so that employees have an understanding of how their performance on various measures affects organizational financial performance. The field site chosen for this study implemented the BSC in only a portion of its branches. Branch managers of implementing branches systematically identified NFMs that were logically and causally linked to improve a specific financial measure; therefore we expect "BSC branches" to experience superior performance on the targeted financial measure as compared to "non-BSC branches".

### **3. Description of field site and its Balanced Scorecard**

The field research site is a banking organization located in the southeastern United States. The bank has a total of 30 geographical locations grouped into 14 reporting units, or branches. Branches are located in

---

<sup>4</sup> Hoque and James (2000) measured performance as a composite score on self-reported assessments of ROI, sales margin, capacity utilization, customer satisfaction, and quality relative to industry peers.

communities of varying sizes, ranging from small rural areas to cities. The bank employs approximately 375 persons, with approximately 250 working in branch operations and approximately 125 working at a central administrative location. Branches are divided into two geographical regions: the northern division (ND) and the southern division (SD). Each division consists of seven branches. The typical branch employs a branch president, a branch vice-president/chief loan officer, customer service representatives (CSRs), loan representatives, mortgage loan originators, head tellers, tellers, and administrative assistants.

Access was gained to the field site through an acquaintance of the researchers who is employed at the field site. The acquaintance was not involved in the development or implementation of the BSC program at the bank. The researchers were introduced to the individual responsible for implementing the BSC at the bank and directed all correspondences through this individual. The activities of one of the researchers can be categorized as a “participant-observation” (Yin, 1994) inasmuch as he served as an active participant in what Kaplan (1993, 1998) calls “action research”. This researcher’s role was limited to assisting in the development of individual scorecards for employees (NFM identification and causal link identification).<sup>5</sup>

The strategic direction of the bank is reviewed annually at a meeting of top bank officials and outside consultants. The purpose of the meeting is to outline the vision and mission of the bank and to ensure all top managers understand and agree on the direction of the organization. It is at this annual meeting that goals for individual branches are established. In setting individual branch goals, bank managers consider past performance, strategic emphasis, and local factors that may affect a branch’s ability to perform. Branch goals are then communicated with branch presidents and revisited if necessary.

The need for better communication of bank goals was evident after the SD president conducted a series of interviews with front-line employees and discovered a lack of knowledge of the bank’s mission and goals and a lack of understanding about how specific jobs contributed to the success of the bank. These interviews, coupled with the need to establish and sustain a positive trend in performance, prompted the SD president to seek and receive approval for implementing the BSC in the SD. The SD president was prompted to try the BSC after studying about the program in a graduate management course. To begin the educational process with his branch presidents, the SD president purchased copies of Kaplan and Norton’s (1996) book outlining the BSC and had each branch president read the book in early 1998. Bank management followed the model set forth by Kaplan and Norton (1996) as closely as possible in developing the BSC in the southern division.<sup>6</sup> An overview of the SDs BSC program is discussed in the following section.

### *3.1. Financial perspective objectives*

Financial perspective objectives for the SD branches were determined using the bank’s existing bonus payout program. In 1995, the bank identified nine key financial measures (KFMs) that were deemed to be important indicators of success for bank branches. Performance on these nine measures (shown in Fig. 1) determines a specific branch’s bonus level for the year. Bank management places various weights

<sup>5</sup> A natural concern when reporting results from ‘action research’ is the bias that may be introduced into the findings. The primary data analysis for this study involves comparing performance on financial measures between two groups of branches. The researchers were not involved in the selection of these financial measures or in the preparation of the statements that yielded these measures. To this extent, we contend our ability to bias these results is mitigated. Further, neither of the researchers was compensated for their work associated with any part of this project. Our involvement was limited to helping the bank develop effective balanced scorecards.

<sup>6</sup> The actual scorecard development began in mid-1998. The researchers’ involvement with the project began in late 1998.

### **Key Financial Measures and Bank Bonus System Description**

Loan Volume -- The outstanding loan balance for the bank (in dollars).

Non-Interest Deposit Volume -- The balance of all deposits (in dollars) by customers upon which the bank pays no interest to the depositor. This category primarily consists of non-interest bearing checking accounts.

Loan Yield (%) -- The portfolio interest rate earned by the branch for outstanding loans. This measure is closely (and inversely) related to the loan volume KFM since the cost of borrowing money affects the number of loans made by branches.

Non-Interest Income -- Income unrelated to interest revenue on outstanding loans. This category includes service charges on deposits, credit insurance income, and annuity sales commissions.

Net Charge-Offs -- The dollar amount of loans determined non-collectible and written off by the bank (net of collateral recoveries).

Cost of Funds (%) -- The average interest rate the bank pays on customer deposits.

Non-Interest Expense -- Expenses unrelated to interest paid on customer deposits in demand and savings accounts. These expenses include salaries and benefits, occupancy expense, equipment and data processing expense, and other miscellaneous operating expenses.

Document Exceptions -- The percentage of outstanding loans with covenant violations (dollar value of loans with covenant violations divided by total dollar value of loans outstanding).

Revenue / Salary Expense -- Total revenues of a branch divided by the total salary expenses, including benefits.

The bank pays yearly bonuses based on a particular branch's performance on these nine measures. To determine an overall performance level for a branch (and to compare performance across branches), each measure is placed on a scale that is adjusted for a given branch's size, market, and other factors that may affect its ability to perform.

Fig. 1. Key financial measures and bank bonus system description.

on the measures for each branch, given a particular branch's strategic focus and competitive setting. Branch performance on these nine KFMs is combined into a composite key financial measure (CKFM) to determine an overall branch's financial performance level.

CKFMs are comparable measures of performance across branches because each KFM performance level (for each branch) is placed on a scale with performance levels adjusted for a particular branch's size, market, potential, and other factors that affect its ability to perform. For example, a larger branch with an outstanding loan balance of \$57 million may have a loan balance KFM scale rating of 4, while a smaller branch with an outstanding loan balance of \$23 million also may have a loan balance KFM scale rating of 4. Since scale intervals and baselines are adjusted for each branch, both branches would be rated equally on the loan balance KFM.

Because improving performance on the CKFM is considered critical to the success of the bank, and because all branches are judged and rewarded based on their CKFM, the SD president made the CKFM measure the financial objective of the BSC. Therefore, each branch's BSCs were prepared with the goal of positively impacting their CKFM.

### 3.2. *Customer perspective objectives*

Providing a small-town community bank atmosphere with quality customer service is the hallmark of the bank's strategy. The bank focuses on building relationships with its customers that foster word-of-mouth advertising and increases its customer base.<sup>7</sup> To measure the quality of customer service of each branch, the SD president instituted periodic branch-by-branch customer service surveys and secret-shopper programs. These two measures serve as outcome (lag) indicators for the customer perspective. The bank eventually plans to track customer retention rates as a means of measuring the effectiveness of its customer satisfaction efforts. Prior to the BSC, customer satisfaction surveys were not conducted regularly on a branch-by-branch basis.

To achieve higher customer satisfaction at the branch level, branches focus on building customized customer service for each client. Measurable scorecard activities include greeting each customer by name (measured by secret shoppers), sending thank-you cards to customers, calling customers with new product information, and providing quality advice to new and existing customers.

### 3.3. *Internal business processes perspective objectives*

The primary process identified for improvement for SD branches was the cross-sell/referral process. As new and existing customers interact with branch personnel, opportunities exist to cross-sell bank products or to refer customers to other branch personnel for assistance in selling products.<sup>8</sup> Outcome measures for this perspective include the number of successful cross-sells and the number of successful referrals. Branch personnel track both the number of cross-sell/referral attempts as well as the number of successful cross-sells/referrals.

### 3.4. *Learning and growth perspective objectives*

The foundation of the bank's BSC is a training program designed to educate and ultimately empower employees to achieve the objectives of the other three perspectives. Employees receive training in customer service, product offerings, sales techniques and office technology efficiency. By enhancing an employee's ability to interact and recognize customer needs, cross-sells, referrals, and customer service should improve. Measures include the number of training hours received per month, and scores on in-house tests of product offerings.

An additional objective of the learning and growth perspective is to improve employee satisfaction and retention. Outcome measures related to this objective include employee satisfaction and employee turnover ratings. Management hopes that implementing the BSC will increase employee loyalty, satisfaction, retention, and productivity. Specifically, the bank hopes increased training, participation in the development of scorecard measures, upward evaluations (where employees evaluate their supervisors), and BSC-related incentives increase employee satisfaction ratings.

---

<sup>7</sup> A survey conducted in 1997 on behalf of the bank revealed that over 65% of customers opening new accounts heard of the bank from either a relative or friend.

<sup>8</sup> Certain bank personnel are not allowed to sell some products (e.g. loans cannot be sold by tellers). In these cases, tellers may refer customers to a loan representative. A cross-sell occurs when a customer purchases a bank product at the suggestion of a bank employee.



### 3.5. The reward system associated with the BSC

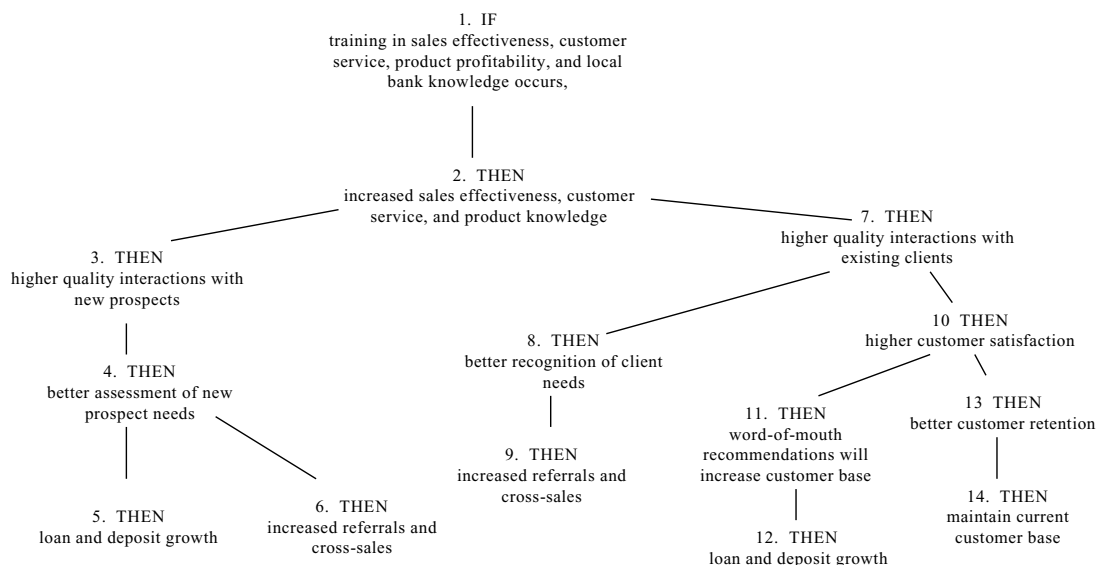
The SD president was limited in his ability to reward individual BSC performance because he was not allowed to change the existing bonus payout program for only his division. The BSC was introduced to employees as a means to improve their CKFM rating, and thereby improve their end-of-year bonus. It was also explained that an employee's performance on his or her individual BSC would be used in determining an annual raise and in considering promotion decisions. Branch presidents also developed small incentive schemes on a branch-by-branch basis that usually involved recognition at a weekly or monthly meeting and/or monetary rewards on a quarterly basis.

### 3.6. Causal links among scorecard measures

The SD and branch presidents met with branch personnel to explain the BSC program and elicit help in scorecard development. The process involved communicating branch goals and educating employees about their contribution to achieving the bank's strategic goals. Fig. 2 presents an example of causal chain logic employed by one of the branches.

Employee training plays an important role in the SD's BSC. Employees equipped with more skills and knowledge are better able to assess a customer's needs and make recommendations. As interactions with customers improve, more successful referrals and cross-sales, more satisfied customers, and more

**Sample Cause & Effect Relationship Diagram**



This sample illustration represents five different causal chains that are directed at improving branch performance on important KFM variables. Causal chains can be read by following the linkages for these sets of numbers: (1,2,3,4,5), (1,2,3,4,6), (1,2,7,8,9), (1,2,7,10,11,12), and (1,2,7,10,13,14).

Fig. 2. Sample cause-and-effect relationship diagram.



new customers (resulting from word-of-mouth recommendations and higher quality interactions with potential customers) will take place.

Linkages illustrated in Fig. 2 offered opportunities to establish quantifiable measures that appear on balanced scorecards. The first link is measured by tracking the number of hours each employee spends in training classes. Training session effectiveness (link 2) is measured through scores on in-house tests. Successful referrals and cross-sales (links 6 and 9) are tracked and compared to the overall number of referrals or cross-sale attempts. Customer satisfaction (link 10) is measured on a branch-by-branch basis through periodic surveys. The number of new customers resulting from current customer referral (link 11) is tracked by soliciting this information from new account holders. Finally, a customer turnover measure is used to determine how well the bank is retaining current customers (links 13 and 14).

## 4. Research method

### 4.1. Research design

We believe a field study method<sup>9</sup> is appropriate for studying management initiatives such as the BSC where success (i.e. the dependent variable) is idiosyncratic to the implementing organization. Kaplan and Norton (1996) state that BSCs vary between organizations and between business-units within organizations because of varying objectives. As such, studies on the effectiveness of the BSC should focus on the targeted lagging measures at the appropriate level within the organization. We believe the field site in this study allows for meeting these requirements.

Given data availability constraints and the research question posed, we chose to follow closely a quasi-experimental research design discussed by Cook and Campbell (1979) to determine if a change in financial performance took place after the implementation of the BSC. Fig. 3 presents five common categorizations of quasi-experimental research designs as presented by Cook and Campbell (1979). Different data collection scenarios include situations where no pre-test observations are available, no control group exists, or no pre-test observation exists for either a control group or experimental group. Field settings where control groups and pre-/post-treatment observations are available allow for stronger causal statements. This study resembles the scenario where both a control group and pre-/post-treatment observations are present (Design 4 of Fig. 3). The treatment for this study is the implementation of the BSC, which officially began in July 1999. The Northern Division did not participate in any BSC activities during the time periods covered by this study.

The research design of this study marks a major difference between the current study and previous related studies. The previous studies attained data pertaining to NFM performance (Banker et al., 2000) and NFM usage (Hoque and James, 2000) and associated NFM data with financial performance. In seeking to determine whether the implementation of an incentive program relying on NFMs improved financial performance, Banker et al. (2000) associate changes in NFM measures and financial measures for a hotel chain in a quasi-experimental setting similar to the third design shown in Fig. 3, where no control group existed.<sup>10</sup> Hoque and James (2000) associate usage of common scorecard NFMs and performance (a self-reported measure) from data gathered through surveys of Australian manufacturing firms. The

<sup>9</sup> Our study is most accurately categorized as the “Chicago School” of direct observation as discussed by Young (1999). As noted earlier, one researcher was a participant-observer in the process of BSC development.

<sup>10</sup> Banker et al. (2000) use data gathered over 72 months, and not single observations, as setting three of Figure 3 indicates.

**Quasi-Experimental Designs**  
(from Cook and Campbell, 1979)

Design Description	Notation*				Comments
(1) One-Group Posttest-Only Design	(T)	X	O		Major weaknesses include the lack of pretest comparison data and the lack of a control group.
(2) Posttest-Only Design with Nonequivalent Groups	(T) (C)	X	O O		Addition of a comparison group strengthens the design, yet the absence of pretest data causes significant interpretation problems.
(3) One-Group Pretest-Posttest Design	(T)	O <sub>1</sub>	X	O <sub>2</sub>	Absence of a control group lessens interpretability. Observed differences may be attributable to many causes. Strong causal inferences are difficult to make.
(4) Untreated Control Group Design with Pretest and Posttest	(T) (C)	O <sub>1</sub> O <sub>1</sub>	X	O <sub>2</sub> O <sub>2</sub>	A stronger design offering pretest and posttest observations. Results are usually interpretable, given a thorough investigation for confounding events and alternative explanations.
(5) Untreated Control Group Designs with Pretest Measures at More Than One Time Interval.	(T) (C)	O <sub>1</sub> O <sub>1</sub>	O <sub>2</sub> O <sub>2</sub>	X O <sub>3</sub>	A stronger design than others listed because the additional pretest data offers insights into data behavior prior to the treatment.

\* T and C stand for Treatment and Control groups, respectively. O<sub>i</sub> stands for an observation in time period *i*. X designates an intervention (or treatment) taking place during the time interval between two observations.

Fig. 3. Quasi-experimental designs (from Cook and Campbell, 1979).

current study, however, benefits from the presence of a control group of branches from the same banking organization that did not implement the BSC. Having a similar group of branches from the same general geographic region greatly enhances our ability to draw conclusions from the observations of this study. Further, the primary dependent measure in this study is the actual financial measure targeted by developers of the BSC at the bank. Previous studies on NFM and financial measure research use organization-wide financial measures or self-reported measures that may not be the stated target of improvement.

Instead of studying the relationship between a given NFM and improved financial performance, as prior studies have done with mixed results, we study whether the implementation of a performance measurement system that incorporates several NFMs in a systematic manner (the BSC) leads to improved financial performance. Ideally, we would incorporate tests of relationships between specific NFMs and various outcome measures as suggested by the linkages identified by going through the BSC process. Unfortunately, we were not able to obtain the detailed data for any of the NFMs appearing on this field site's BSC, thereby making similar causal inferences between financial and specific non-financial measures at this site impossible. We recognize that the inability to attain specific NFM data from the site is a limitation of the study.<sup>11</sup>

<sup>11</sup> However, as noted in Malina and Selto (2001) and Norreklit (2000), a model that captures all the relationships expressed in a BSC would be sizeable task and may not establish explicit causality between specific elements due to interdependencies and correlations among many variables.

#### 4.2. Branch selection

Unlike laboratory experiments where random assignment of subjects takes place, the use of non-conforming groups in a quasi-experimental design means participating subjects (branches) may differ between groups in ways other than the presence (or absence) of a treatment. To the extent possible, these differences should be identified and controlled for to improve the study's internal validity (Cook and Campbell, 1979). The bank has a total of 14 branches divided into two geographical divisions of seven branches. After studying individual branch profiles and recent histories, five branches were removed from the analysis because of extraneous events unrelated to the implementation of the BSC.<sup>12</sup> Three of these five branches were removed because they switched divisions during December 1999. Two branches moving from the ND to the SD were excluded from the analysis because they had not been exposed to the BSC during the observation period. The branch moving from the SD to the ND was excluded because it was exposed to the BSC and had completed most of the implementation procedures during the last six months of 1999.

Another branch on the SD was excluded from the analysis because it did not fully implement the BSC during the observation period. This branch began the process along with the other SD branches, but stopped the implementation process when the branch manager left the bank for employment with another organization. The final branch was excluded from the analysis because a loan officer was hired in 1999 that brought several customers (borrowers) with him to the branch, thereby drastically improving branch performance during the observation period. Discussions with bank personnel revealed no other relevant events during the observation period that merited removing other branches from the analysis. Removing unusual observations from the sample resulted in four branches in the experimental (BSC) group and five branches in the control (non-BSC) group. Table 1 provides information pertaining to branches excluded from the analysis.

#### 4.3. Dependent measure

As discussed earlier, we use a branch CKFM measure as the dependent variable in our statistical tests because this is the stated objective of the SD's BSC program. The CKFM represents an equitable measure of performance across all branches in both divisions and is a measure all branches seek to maximize. Bank management accounts for differences in size, market, economic, and strategic differences by adjusting scale intervals and KFM weights for each individual branch. Descriptive information for each branch is shown in Table 2.

#### 4.4. Statistical analyses

The goal of the statistical analysis is to determine if CKFMs for experimental division branches increase significantly after implementing the BSC and whether the change in performance is significantly greater than the change experienced by the control group of branches. CKFM levels from June 1999 and June

<sup>12</sup> The goal is to have a homogeneous group of branches with the exception of the implementation of the BSC. While lack of random assignment of subjects to groups is a limitation of any quasi-experimental study, we seek to mitigate this limitation through two methods suggested by Cook and Campbell (1979): elimination of subjects with known alternative explanations for changed performance and by using "gain score analysis".

Table 1  
Information on branches excluded from the analysis

Branch	Division	Change in CKFM <sup>a</sup>	Reason for being excluded from the analysis
1	Southern	(0.77)	This branch began the implementation process with the other southern division branches, but the process was stopped when the branch manager left for employment with another organization.
2	Northern	10.13	This branch switched from the southern division to the northern division on January 1, 2000. This branch had completed a large percentage of their BSC implementation at the time of the change in their division.
3	Northern	11.84	This branch's performance was improved during the observation period due to the hiring of a competing bank's top managers. The new bank managers brought several clients with them to this branch, thereby improving this branch's performance.
4	Southern	(2.65)	These branches were switched from the northern division to the southern division on January 1, 2000. These branches were not part of the BSC project.
5	Southern	(0.12)	

<sup>a</sup> The change in CKFM (composite key financial measure) is from June 1999 to June 2001.

2001 are used for statistical comparisons. Data from these two time periods are used to obtain performance levels prior to observing the effects of the BSC and a corresponding period two years later when the effects of the new program should be observed.

Due to the small number of observations in our sample, we use the nonparametric Wilcoxon signed rank test to determine whether a difference exists between CKFM observations from June 1999 to June 2001. We first perform this test on both groups of branches separately. The purpose of these tests is to determine whether a significant change in performance took place within either group. We expect performance increased significantly from June 1999 to June 2001 for experimental division branches; however, we make no prediction concerning the performance of control division branches during this period.

The second comparison test is used to determine if the change in performance level from June 1999 to June 2000 is greater for experimental division branches than for control division branches. We believe this test is necessary to determine if performance improvements reported by experimental branches are significantly greater than performance improvements that would normally have resulted if the BSC had not been implemented. Further, we recognize the effects of the BSC may not result in improved financial performance during times of economic decline. The positive effects of the BSC may lessen the negative impact of a downward trend. This possibility underscores the importance of a control group in a study seeking to determine the effects of a new program.

## 5. Results

### 5.1. Statistical tests of differences within each group of branches

Fig. 4 presents a line graph of group median CKFMs from June 1999 through June 2001 for BSC and control division branches. Experimental and control branch performance was roughly equal at the

Table 2  
Descriptive statistics, KFM weightings, and performance data by branch

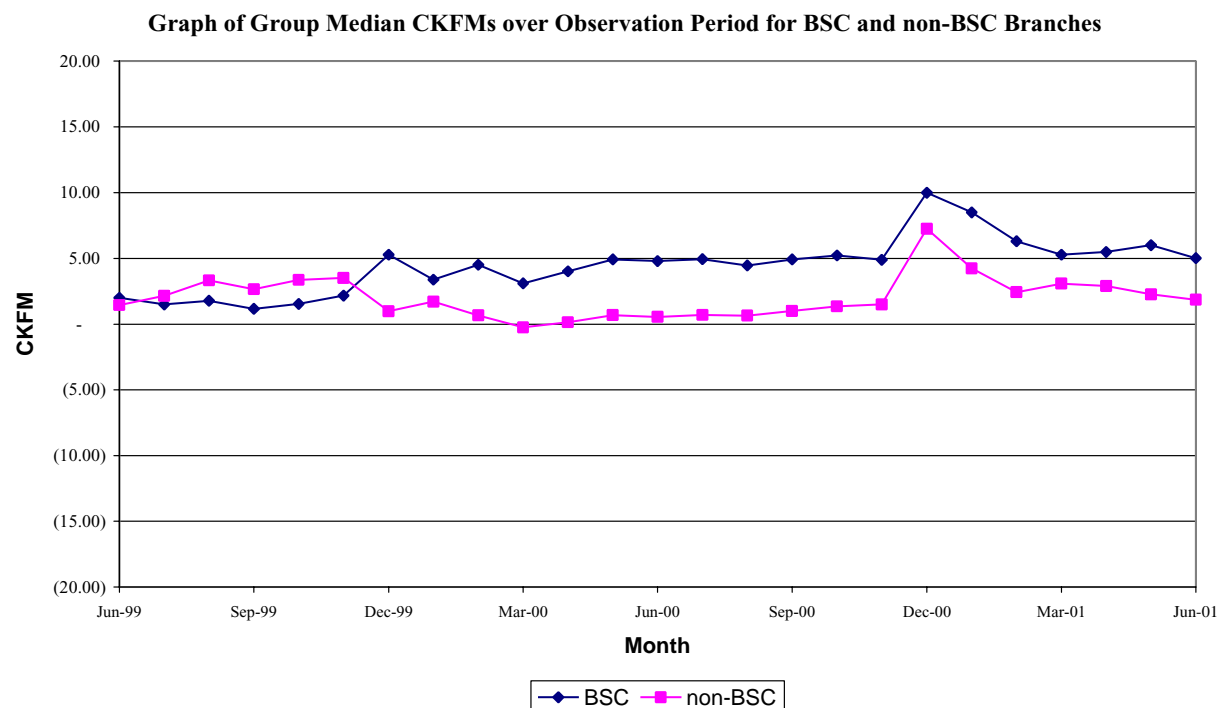
	BSC branches				Non-BSC branches (used as control group)				
	A	B	C	D	E	F	G	H	I
No. of employees	14	8	26	27	10	9	13	19	20
Assets <sup>a</sup>	35089.5	13341.6	58827.9	80039.4	6958.0	8934.1	38642.3	44664.3	40682.7
Total revenues <sup>a</sup>	3405.7	882.1	5479.0	6541.2	746.7	974.4	2461.7	4112.5	3742.3
KFM weighting <sup>b</sup> (percentages)									
Loan volume	13	43	27	28	8	12	32	27	23
Non-interest deposit volume	2	22	3	3	3	5	4	3	4
Loan yield (%)	10	2	9	9	16	6	5	7	8
Non-interest income	17	1	10	9	20	21	12	12	12
Net charge-offs	20	4	16	15	21	22	14	18	19
Cost of funds (%)	5	4	4	5	5	6	5	4	4
Non-interest expense	8	8	8	8	8	8	8	8	8
Document exceptions	10	2	8	8	5	5	5	6	7
Revenue/salary expense	15	14	15	15	14	15	15	15	15
CKFM levels									
As of June 1999	(2.74)	0.95	3.15	3.09	(0.31)	6.67	10.82	0.00	(2.13)
As of December 1999	(8.20)	2.46	1.88	3.58	(0.38)	7.18	12.62	3.51	(4.67)
As of December 2000	0.16	8.60	2.69	7.06	(2.11)	1.50	10.50	3.12	(5.20)
As of June 2001	0.29	10.65	5.30	6.70	0.73	2.55	11.59	2.26	(5.38)
Change in CKFM levels									
From June 1999 to December 1999	(5.46)	1.51	(1.27)	0.49	(0.07)	0.51	1.80	3.51	(2.54)
From December 1999 to December 2000	8.36	6.14	0.81	3.48	(1.73)	(5.68)	(2.12)	(0.39)	(0.53)
From December 2000 to June 2001	0.13	2.05	2.61	(0.36)	2.84	1.05	1.09	(0.86)	(0.18)
From June 1999 to June 2001	3.03	9.70	2.15	3.61	1.04	(4.12)	0.77	2.26	(3.25)

<sup>a</sup> As of 12/31/1999; in thousands.

<sup>b</sup> Descriptions for these measures are given in Fig. 2.

beginning of the observation period. Control branches performance began the 24-month period with a median of 1.44 on the CKFM scale compared to a median of 2.01 for experimental branches. At the end of 1999, control and experimental groups had median CKFMs of 3.51 and 2.17, respectively. Over the course of the 24-month period, control branch performance declined gradually, ending with a median of 2.26 on the CKFM scale as of June 2001. Although experimental branch CKFM performance improved to median of 6.00 in June 2001, the change in performance for experimental branches was not a gradual increase (as compared to the gradual changes experienced by control branches).

The reason for observing such a stark difference in CKFM levels for experimental division branches between December 1999 and January 2000 relates to the manner in which the CKFM is calculated. Balance sheet measures, such as the loan volume KFM, represent an average balance for the calendar year-to-date. Therefore, the December figure used for the loan volume KFM represents an average balance for the year, and not the ending loan volume balance. The January figure used for the loan volume KFM



Monthly median performance in BSC and Non-BSC Branches. The BSC program began in July 1999 in the BSC branches.

Fig. 4. Graph of group median CKFMs over observation period for BSC and non-BSC branches.

represents only January's year-to-date balance, thereby allowing for drastic differences if a branch's loan volume balance has grown gradually over the past few months of the calendar year. Further, income statement-related KFMs are annualized on a monthly basis, allowing for large fluctuations in income statement-related KFMs early in the year.

Fig. 4 indicates an increase in performance for BSC branches. The purpose of the statistical tests is to determine whether a significant increase in performance occurred from June 1999 to June 2001 for BSC and control division branches. In addition, the tests examine whether BSC branch performance increased significantly more than control division performance. Table 3 presents results from a Wilcoxon rank test used to determine whether a difference in performance is observed for either experimental or control division branches. Panel A presents CKFM levels for experimental division branches. The Wilcoxon  $z$ -statistic equals 1.826 (one-tailed  $P$ -value = 0.034), indicating a significant difference exists (in the expected direction) between the average performance in June 1999 and the average performance in June 2001 for experimental division branches. We use a one-tailed  $P$ -value for this test because we predict an improvement in performance following BSC implementation. The median CKFM performance level increase of experimental division branches was 3.32 scale points over the observation period.

Table 3, Panel B presents CKFM levels for June 1999 and June 2001 for control division branches. The  $z$ -statistic from the Wilcoxon test is 0.674 (two-tailed  $P$ -value = 0.500), indicating a non-significant difference exists between the performance for these two time periods for control division branches. A two-tailed  $P$ -value is used for this test because no prediction was made concerning the change in perfor-

Table 3  
Results of statistical *t*-tests comparing CKFMs of BSC and non-BSC branches

Panel A BSC branches			Panel B non-BSC branches			Panel C change in CKFM value from June 1999 to June 2001	
	CKFM level for June 1999	CKFM level for June 2001		CKFM level for June 1999	CKFM level for June 2001	BSC branches	
Branch A	(2.74)	0.29	Branch E	(0.31)	0.73	Branch A	3.03
Branch B	0.95	10.65	Branch F	6.67	2.55	Branch B	9.70
Branch C	3.15	5.30	Branch G	10.82	11.59	Branch C	2.15
Branch D	3.09	6.70	Branch H	0.00	2.26	Branch D	3.61
			Branch I	(2.13)	(5.38)	non-BSC Branches	
Median	2.02	6.00	Median	0.00	2.26	Branch E	1.04
Median change		3.32	Median change		0.77	Branch F	(4.12)
Wilcoxon <i>z</i> -statistic		1.826	Wilcoxon <i>z</i> -statistic		0.674	Branch G	0.77
<i>P</i> -value <sup>a,b</sup>		0.034	<i>P</i> -value <sup>a,c</sup>		0.500	Branch H	2.26
						Branch I	(3.25)
						Wilcoxon <i>z</i> -statistic	2.205
						<i>P</i> -value <sup>a,b</sup>	0.014

<sup>a</sup> Because of the small sample sizes, nonparametric Wilcoxon tests were performed.

For Panel A,  $H_1: P[X > Y] > 0.5$ , where  $X$  represents June 2001 CKFM values and  $Y$  represents June 1999 CKFM values.

For Panel B,  $H_1: P[X > Y] > 0.5$ , where  $X$  represents June 2001 CKFM values and  $Y$  represents June 1999 CKFM values.

For Panel C,  $H_1: P[X > Y] > 0.5$ , where  $X$  and  $Y$  represent changes in CKFM values for BSC and non-BSC branches, respectively.

<sup>b</sup> A one-tailed *P*-value because an increase in performance is expected due to the implementation of the BSC.

<sup>c</sup> A two-tailed *P*-value because no prediction is made about the direction of change in performance for the non-BSC branches.

mance for control division branches. The median CKFM performance level increase of control division branches was 0.77 scale points over the observation period.

## 5.2. Statistical tests of the changes in performance between BSC and non-BSC branches

Table 3, Panel C presents the results of a Wilcoxon sign rank test comparing the changes in performance for the two groups of branches. The Wilcoxon *z*-statistic is 2.205 (one-tailed *P*-value = 0.014), indicating a significant difference exists (in the expected direction) in the change in performance for these two groups of branches. The results of our tests indicate experimental branch performance improved subsequent to implementing the BSC, while control branch performance did not improve. The final test indicates experimental branch performance improved significantly more than control branch performance. Our results indicate the treatment effect of the BSC produced a significant and positive effect on the dependent measure, CKFM.<sup>13</sup>

<sup>13</sup> The statistical analyses presented above rely on comparison of data from two points in time in order to observe whether a difference in performance exists. We conducted similar statistical tests using data from all 24 months to ensure our results are not a function of being drawn from only one month's observation. These tests included repeated-measures ANOVA and a nonparametric (Friedmans) repeated measures test for differences. These additional tests support the significant findings reported above.



## 6. Discussion and conclusion

The purpose of this study was to determine whether an improvement in financial performance occurred after implementing a BSC and whether the change in financial performance is significantly greater than performance observed in a similar setting where a traditional performance measurement system using only financial measures is employed. We provide evidence supporting the proposition that the BSC can be used to improve financial performance; the findings indicate branches in the BSC group outperformed non-BSC branches on a common composite financial measure. The research method and design of this study allow for a causal statement concerning the association between BSC implementation and financial performance improvement. The ability to use the actual targeted financial measure of the BSC at the business-unit level as the dependent variable in this study provides a more direct test of whether the BSC has achieved its intended purpose than using corporate-wide measures. Further, the presence of a control group of subjects strengthens the study because of the ability to compare performance of BSC-adopters to performance of a similar group of branches within the same organization that do not adopt the BSC.

Our findings are consistent with similar studies in the manufacturing industry (Hoque and James, 2000) and the hotel industry (Banker et al., 2000) that show the inclusion of NFMs in a performance measurement system is associated with improved financial performance. Our findings differ, however, from Ittner et al. (2003). Whereas they find a negative association between BSC usage and ROA, we report a positive improvement in financial performance on a targeted financial measure for a group of bank branches implementing the BSC. Ittner et al. (2003) report two findings that may explain these divergent findings. First, firms that reported using the BSC overwhelmingly reported they did not rely on causal business models. Further, Ittner et al. (2003) report a positive relationship between ROA and firms that rely on business modeling. Given that developing and understanding causal assumptions between selected measures is an integral component of a properly designed BSC, the lack of an association between BSC usage and financial performance is not necessarily unexpected.

This study contributes to the existing performance measurement and BSC literature by providing evidence of the ability of the BSC to improve financial performance. A criticism of management accounting research has been its inability to assess whether new management initiatives, such as the BSC, are better or just different (see Foster and Young, 1997). Our findings provide support for the proposition that the BSC method can promote improved financial performance when compared to a traditional performance measurement system focusing solely on financial measures. By incorporating a group of NFMs into the performance measurement system in a logical and systematic manner, a group of branches outperformed another “control” group on a financial metric important to both sets of branches. To date, studies seeking to associate a NFM and financial performance measures have yielded mixed results. Our findings suggest that one possible explanation may be the lack of a coherent linkage between the measures chosen for the performance system and the targeted financial measure of interest.

### 6.1. Limitations and areas for future research

The presence of a comparison group in a field setting such as this is unique. In this case, the control group helps mitigate one of the major limitations of the study—the inability to directly test causal chains of non-financial measures within a given BSC to improved financial performance. While we were not able to attain the detailed performance data for non-financial BSC measures, we compared CKFM measures for both BSC and non-BSC branches before and after program implementation to determine if a change had

taken place. To the extent that we only include branches with similar histories over the observation period in our analysis, we believe the observed difference in performance is attributable to implementing the BSC.

Aside from the limitation in attaining data pertaining to NFM performance previously discussed, this study has the following limitations. Observing performance over a 24-month timeframe is a limitation of our research. This study addresses whether a difference in performance occurred after the implementation of the BSC. Using a longer timeframe would allow for determining whether this difference is sustainable for longer periods of time, continues to improve over time, or reverts to pre-BSC implementation levels. The possibility exists that the findings of this study are a result of a Hawthorne Effect, a phenomena where subjects improve their performance on a given task simply because of the special attention they have received during a study. While we cannot entirely dismiss this possibility as an explanation for the findings in this study, [Ittner and Larcker \(1998\)](#) note that a one-year period is not an unreasonable period for observing effects of emphasizing non-financial metrics on financial measures in banking institutions and [Banker et al. \(2000\)](#) use a six-month lag in their study of hotels. The improvement observed in this setting is maintained two years after implementation of the BSC.

Without access to more historical data, it is difficult to show that our results are not a reversion to mean performance. The possibility exists that the observations shown here are either a reversion to the mean or normal changes that are part of a longer timeframe trend. The comments of the SD president, bank VP of Finance, and ND president all indicate that the change in performance for the BSC branches during 2000 was an unusual event, indicating that it was likely not a reversion to the mean performance levels. Also, the ND president requested assistance from the SD president in setting up a BSC program for his division in late 2000, further supporting the contention that the changes observed by the SD during the observation period were due to the BSC.<sup>14</sup>

Another limitation is the generalizability of the study. We believe field study method is an appropriate tool for studying the effectiveness of the BSC on financial measures, but because each organization's BSC is unique, care must be taken in generalizing our findings. We do not suggest that every BSC implementation will result in improved financial performance, or that every BSC must be identical to the one implemented by the organization in this study. Rather, we believe our findings indicate that the BSC method can be effectively used to improve financial performance (a previously untested proposition) compared to a performance measurement system focusing solely on financial measures, and that the findings generalize to the phenomenon being studied (the BSC).<sup>15</sup>

The organization in this study systematically chose NFMs for their performance measurement systems that were causally linked to improved financial performance and experience the desired results. A goal of future research should be to continue investigating BSC implementations with the purpose of developing theoretical expectations for various aspects and differences between BSCs. Including studies where BSC implementations are considered a failure will further broaden our understanding of the phenomenon. We suggest that future research focus on studying the effectiveness of the BSC in similar and dissimilar field settings as well as focus on the expected time period lag prior to observing results and the circumstances affecting the length of time before results are observed. Research should determine whether performance improvements are sustained or are short-term anomalies. Further, research should

<sup>14</sup> The primary nonequivalent factors between the groups are the divisional designations and the implementation of the BSC. The presence of pre-treatment data and change tests aid in the control for divisional differences. As in any study, we note the possibility of an unidentified nonequivalent factor as explanation for the observed difference between the groups.

<sup>15</sup> For a more detailed discussion on drawing generalizations from field studies, see [Yin \(1994\)](#).

explore how the benefits of the BSC are affected by different industry characteristics, including type of industry, level of competition, and type of strategy. Finally, several behavioral issues are raised at the individual and organizational level with the implementation of a new performance measurement system that shifts importance from financial measures to NFMs (see [Lipe and Salterio, 2000, 2002](#); [Roberts et al., forthcoming](#)). These BSC-related issues merit investigation as well in order to increase our understanding of the venues where the BSC is most effective.

## Acknowledgements

The authors thank the participating company and its employees for their cooperation during this project. The authors also thank Gary Taylor, Brian Gray, Diane Johnson, Sue Haka, Doug Cagwin and participants at the S. Paul Garner Center for Current Accounting Issues Research Workshop at the University of Alabama for their suggestions and insights on the paper. This research was supported by the Babcock Graduate School of Management, Wake Forest University, Research Fellowship Program.

## References

- Atkinson, A.A., Balakrishnan, R., Booth, P., Cote, J.M., Groot, T., Malmi, T., Roberts, H., Uliana, E., Wu, A., 1997. New directions in management accounting research. *J. Manage. Acc. Res.* 9, 79–108.
- Banker, R.D., Potter, G., Srinivasan, D., 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. *Acc. Rev.* 75 (1), 65–92.
- Chenhall, R.H., 1997. Reliance on manufacturing performance measures, total quality management and organizational performance. *Manage. Acc. Res.* 8 (2), 187–206.
- Cook, T.D., Campbell, D.T., 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Houghton Mifflin Company, Boston, MA.
- Foster, G., Young, S.M., 1997. Frontiers of management accounting research. *J. Manage. Acc. Res.* 9, 63–77.
- Groves, G., Valsamakis, V., 1998. Supplier–customer relationships and company performance. *Int. J. Logist. Manage.* 9 (2), 51–64.
- Hoque, Z., James, W., 2000. Linking Balanced Scorecard measures to size and market factors: impact on organizational performance. *J. Manage. Acc. Res.* 12, 1–17.
- Ittner, C., Larcker, D.F., 1998. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *J. Acc. Res.* 36 (Supplement), 1–35.
- Ittner, C., Larcker, D.F., Randall, T., 2003. Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society* 28, 715–741.
- Kaplan, R.S., 1993. Research opportunities in management accounting. *J. Manage. Acc. Res.* 5, 1–14.
- Kaplan, R.S., Norton, D.P., 1996. *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business School Press, Boston, MA.
- Kaplan, R.S., 1998. Innovation action research: creating new management theory and practice. *J. Manage. Acc. Res.* 10, 89–118.
- Lau, R.S.M., 1997. Operational characteristics of highly competitive manufacturing firms. *Product. Inventory Manage. J.* 38 (4), 17–21.
- Lipe, M.G., Salterio, S.E., 2000. The Balanced Scorecard: judgmental effects of common and unique performance. *Acc. Rev.* 75 (3), 283–298.
- Lipe, M.G., Salterio, S.E., 2002. A note on the judgmental effects of the Balanced Scorecard's information organization. *Acc. Organ. Society* 27 (6), 531–540.
- Malina, M.A., Selto, F.H., 2001. Communicating and controlling strategy: an empirical study of the effectiveness of the Balanced Scorecard. *J. Manage. Acc. Res.* 13, 47–90.

- Moon, P., Fitzgerald, L., 1996. Delivering the goods at TNT: the role of the performance measurement system. *Manage. Acc. Res.* 7 (4), 431–457.
- Norreklit, H., 2000. The balance on the Balanced Scorecard—a critical analysis of some of its assumptions. *Manage. Acc. Res.* 11, 65–88.
- Palmer, R.J., 1992. Strategic goals and objectives and the design of strategic management accounting systems. *Advances Manage. Acc.* 1, 179–204.
- Perera, S., Harrison, G., Poole, M., 1997. Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: a research note. *Acc. Organ. Society* 22 (6), 557–572.
- Roberts, M., Albright, T., Hibbets, A., forthcoming. Debiasing Balanced Scorecard evaluations. *Behav. Res. Acc.*
- Spicer, B.H., 1992. The resurgence of cost and management accounting: a review of some recent developments in practice, theories, and case research methods. *Manage. Acc. Res.* 3, 1–37.
- Terziovski, M., Samson, D., Dow, D., 1997. The business value of quality management systems certification—evidence from Australia and New Zealand. *J. Operations Manage.* 15 (1), 1–18.
- Yin, R.K., 1994. *Case Study Research: Design and Methods*, 2nd ed. Sage Publications, Thousand Oaks, CA.
- Young, S.M., 1999. Field research methods in management accounting. *Acc. Horizons* 13 (1), 76–84.