

**THE BALANCED SCORECARD: THE EFFECT OF STRATEGY MAP
PROVISION ON PERFORMANCE EVALUATION JUDGMENTS AND
MENTAL MODEL DEVELOPMENT**

Kerry A. Humphreys
Australian School of Business
University of New South Wales

Ken T. Trotman
Australian School of Business
University of New South Wales

Preliminary Draft
Please do not quote without the permission of the authors
May 2008

Corresponding author:

Kerry A. Humphreys

Australian School of Business, University of New South Wales, Sydney NSW 2052

Tel: +61 2 9385 5833; Fax: +61 2 9385 5925; Email: k.humphreys@unsw.edu.au

THE BALANCED SCORECARD: THE EFFECT OF STRATEGY MAP PROVISION ON PERFORMANCE EVALUATION JUDGMENTS AND MENTAL MODEL DEVELOPMENT

ABSTRACT

Evaluating performance using the balanced scorecard is a complex task in which the decision maker is presented with multiple measures across four perspectives, and little guidance on how to weight and combine the multiple information cues in an overall judgment. Our study examines judgmental effects of providing superiors with causally linked strategic information (a strategy map), when making a balanced scorecard-based performance evaluation judgment. We find that participants provided with strategic objectives in narrative form, or both strategic objectives in narrative form and a strategy map do not demonstrate a common measures bias when all balanced scorecard performance measures are strategically linked. It is suggested that the inconsistency of this finding with the extant literature results from attributes of the task that were varied in the design of the current study, and indicates that task attributes may play an important role in determining when common measures bias is likely to occur. The current study also finds that participants provided with a strategy map containing causal linkages develop a more complete mental model of divisional strategy than those provided with strategic objectives only.

THE BALANCED SCORECARD: THE EFFECT OF STRATEGY MAP PROVISION ON PERFORMANCE EVALUATION JUDGMENTS AND MENTAL MODEL DEVELOPMENT

I. INTRODUCTION

Contemporary performance measurement systems aim to improve managerial performance evaluation by providing decision makers with a broader set of financial metrics than traditional accounting-based performance measures, as well as a set of future-oriented, non-financial measures (Sprinkle and Williamson, 2006). The balanced scorecard is the most well-known example of such a multiple performance measurement system. Evaluating performance using the balanced scorecard is, however, considered to be a complex task (Merchant, 2006). The decision maker is presented with multiple measures across four balanced scorecard perspectives, often without guidance on how to weight and combine the multiple information cues in an overall judgment and without normative criterion for evaluating the accuracy of the performance evaluations that result (Lipe and Salterio, 2000; Dilla and Steinbart, 2005; Hibbets, Roberts and Albright, 2006; Langfield-Smith, 2006).

Adaptive decision theory (Payne, Bettman and Johnson, 1993) suggests that in complex settings, a trade-off between the desire to maximize decision quality and the limited processing capacity of human decision makers will lead to decision makers seeking to minimize the cognitive effort required to make a decision. Consistent with this theory, Lipe and Salterio (2000, p. 293) found that superiors employed a “natural simplifying strategy” when making managerial performance evaluation judgments using the balanced scorecard. This simplifying strategy – known as ‘common measures bias’ – results in managerial performance evaluations based on measures common to both the divisions being evaluated, rather than balanced evaluations that reflect all divisional performance measures (both common and unique).

Over the past decade, a strong body of research into the judgmental effects of the balanced scorecard in performance evaluation has developed. A number of these studies (Lipe and Salterio, 2000; Ittner, Larcker and Randall, 2003; Banker, Chang and Pizzini, 2004; Libby, Salterio and Webb, 2004; Roberts, Albright and Hibbets, 2004; and Dilla and Steinbart, 2005) have found or assumed the existence of the common measures bias. As the balanced scorecard is designed to facilitate

performance evaluation that reflects all performance measures linked to a division's strategy, biased judgments have important implications for managers. The aim of this study is to examine the use of a strategy map when making these judgments and to measure the mental model completeness of participants, in order to better understand why the bias exists and how it can be reduced.

A strategy map conveys the causal linkages between the strategic objectives of an organization, which facilitates understanding of the strategy by managers and employees, and encourages consistency of strategic understanding across all staff – the first step towards organizational alignment (Kaplan and Norton, 2004). While Kaplan and Norton (2004, p. xiii) describe the strategy map as a “breakthrough” for the balanced scorecard, the cognitive impact of this tool on balanced scorecard-based judgments is largely untested in the accounting literature. As such, the current study contributes to our knowledge of “how and how well” individuals use multiple performance measures by investigating the impact the causal linkages communicated in a strategy map have on managerial decision making and cognition (Sprinkle and Williamson, 2006, p. 431).

The study makes two specific contributions to the accounting literature. First, we find that when all performance measures are linked to the strategic objectives of the organization and these strategic objectives are presented in narrative form in a decision making setting, managers both with and without a strategy map do not demonstrate common measures bias in performance evaluation judgments. Banker et al (2004) investigated whether participants provided with detailed strategy information made greater use of strategically-linked performance measures in a balanced scorecard than those provided with no strategy information, and found this to be the case. They did not find that the detailed strategy content prevented common measures bias from occurring, but suggested that “participants rely more on common measures in the absence of detailed strategy information” (p. 18). Given that the strategy map was designed to provide decision makers using the balanced scorecard with a map of the cause-and-effect linkages between strategic objectives (Kaplan and Norton, 2004), the present study sought to extend Banker et al by examining whether the cause-and-effect linkages communicated in a strategy map are able to mitigate common measures bias when all performance measures are strategically linked as

suggested by Kaplan and Norton (2004)¹. Hence, unlike Banker et al, all participants in the current study received information on the strategic objectives of each division and all the performance measures in the balanced scorecard related to one of these objectives. The treatment group in the current study also received a strategy map identifying the causal model underlying the balanced scorecard. It was found that with or without the causal linkages represented in a strategy map, when all performance measures in the balanced scorecard related to a strategic objective presented in the narrative (and the strategy map), common measures bias was not demonstrated in the managerial performance evaluations made by participants.

Second, the present study contributes to the accounting literature by extending research on managers' use of the balanced scorecard to investigate the impact of a strategy map on the cause-and-effect linkages contained in participants' mental models of divisional strategy. Mental models are simplified, cognitive representations of causal relationships that incorporate information currently active in working memory (Markman and Gentner, 2001; Hammersley, 2003; Hodgkinson, 2003). Those mental models that consist of high-quality information and include strong linkages help decision makers attend to important environmental cues (Eden and Spender, 1998; Hodgkinson, 2003). As such, causal linkages are a key feature of mental models (Brewer, 1993; Sterman, 1994; Markman and Gentner, 2001). Given the role of causal linkages in a strategy map and the importance of causal linkages in mental model development, we examine whether provision of a strategy map enhances mental model completeness. Developing a more complete mental model of organizational strategy across all employees is critical in a balanced scorecard decision setting where a shared understanding of organizational strategy is an expected outcome (Kaplan and Norton, 2004). The current study introduces a measure of mental model completeness – adapted from the psychology and operations research literature to examine this effect.

¹ The balanced scorecard “typically contains from 16 to 28 measures” and all measures are expected to be linked to the strategy of the organization (Dilla and Steinbart, 2005, p. 44; Kaplan and Norton, 1996a; Kaplan and Norton, 2004). However, the manipulation employed in Banker et al (2004) meant that task had only eight strategically-linked measures in each scorecard, with the other eight performance measures in each divisional balanced scorecard not linked to the divisional strategy.

II. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Performance evaluation judgments based on the balanced scorecard

Kaplan and Norton (1996a) suggested that the balanced scorecard be used in management review discussions, with the multiple performance measures encompassed in the scorecard not only providing guidance on “how past results have been achieved but also whether their expectations for the future remain on track” (p.15). Libby, Bloomfield and Nelson (2002), however, highlighted the negative effect the limited capacity of working memory has on decision makers’ ability to consider multiple factors when making a judgment. Bazerman (2006) also contended “that people have a *bounded awareness* that prevents them from noticing or focusing on observable and relevant data” (p. 171).

One example cited by Bazerman (2006) is of bounded awareness of shared information found in psychology research by Slovic and MacPhillamy (1974). Slovic and MacPhillamy (1974) found that when subjects compared pairs of students with respect to their potential college performance based on a series of common (shared by both students) and unique (specific to each student) dimensions, the common dimensions were weighted more heavily in the comparison than the unique dimensions. It is suggested that unique variables are discounted because they require more cognitive effort to process – either requiring information to be “held in memory, inferred or transformed” in some way (p. 193).

In a balanced scorecard-based performance evaluation setting, Lipe and Salterio (2000) provided evidence of this bounded awareness of shared information – in this setting, termed ‘common measures bias’. Common measures bias is the use of a simplifying strategy where managers focus on measures that are common to two divisions when weighting and combining performance measures in performance evaluation judgments, rather than making a balanced evaluation of performance utilizing all available balanced scorecard cues.

This “natural simplifying strategy” (Lipe and Salterio, 2000, p. 293) has now been well-established in this decision setting (Banker et al, 2004; Libby et al, 2004; Dilla and Steinbart, 2005). Further, extant literature has looked at various strategies to alleviate common measures bias – in the case of Roberts, Albright and Hibbets (2004) asking participants to evaluate performance separately for 16 BSC measures, then mechanically aggregated the 16 ratings using pre-assigned weights; Libby, Salterio and Webb (2004) combined a requirement to justify the final performance

evaluation with the provision of a third-party assurance report on the BSC measures; and Dilla and Steinbart (2005) provided participants with balanced scorecard training.

In a case study of the unsuccessful adoption of a balanced scorecard by a global financial services organization (GFS), Ittner, Larcker and Meyer (2003) also found evidence of ‘common measures bias’ – specifically with respect to measures contained in the customer perspective of the scorecard. The use of the balanced scorecard in performance evaluation at GFS resulted in such a high level of perceived uncertainty amongst employees, regarding the criteria used to determine rewards, that the balance scorecard-based compensation plan was abandoned. This uncertainty highlights the “multifaceted and complex nature” of weighting performance measures when evaluating performance and calculating managerial bonuses (p. 754).

Ittner, Larcker and Randall (2003) question whether firms claiming to have balanced scorecards “are actually *using* the information, or have merely implemented measurement systems that capture information corresponding to the scorecard categories without making changes in the information used for decision-making and performance evaluation” (p. 733). Given the management time and cost involved in developing and implementing a balanced scorecard, these findings suggest that the balanced scorecard may not be fulfilling its ‘decision-facilitating’ role (Sprinkle, 2003; Sprinkle and Williamson, 2006).

Causes of deficiencies in balanced scorecard-based performance evaluation

Following Slovic and MacPhillamy (1974), potential causes for ‘common measures bias’ suggested in the psychology and business decision making literatures have focused on the complex nature of the decision making task. Payne et al (1993) suggested that ‘common measures bias’ occurs due to the “cognitive ease of comparison” of common dimensions (p. 51). Payne et al’s adaptive decision theory recognizes the trade-off between the desire to maximize decision quality and the limited processing capacity of human decision makers, which leads to them minimizing the cognitive effort required to make a decision.

Consistent with Payne et al (1993), Zhang and Markman (2001) suggested that unique variables require greater processing effort. This is due to difficulty in comparing unique variables and the need to evaluate on an absolute, rather than relative, scale. Zhang and Markman found that greater personal involvement by participants in comparative evaluations led to greater use of unique variables. It is

suggested that greater involvement led to increased motivation to process information in a more “effortful, careful, and detailed” manner (p. 15).

Causally-linked strategy maps and performance evaluation judgments

In a performance evaluation setting, Kaplan and Norton’s (2004) strategy map tool aims to facilitate communication of the relationship between information cues presented (being performance measures) and the judgment to be made (being an evaluation of performance in delivering divisional strategy). As Kaplan and Norton (2001) state “Balanced Scorecards should not just be collections of financial and nonfinancial measures, organized into three to five perspectives. The best Balanced Scorecards reflect the strategy of the organization.” (p.97). Thus, “A scorecard should contain outcome measures and the performance drivers of those outcomes, linked together in cause and effect relationships” (Kaplan and Norton, 1996b, abstract). This linking of performance measures with business unit strategy has been identified as a key strength of the balanced scorecard, and the cause-and-effect relationship chain that results is said to distinguish it from other performance measurement systems (Otley, 1999; Norreklit, 2003; Norreklit and Mitchell, 2007).

As such, managers who are presented with a strategy map depicting the cause-and-effect relationships between strategic objectives are expected to have a better understanding of the linkages between the “collection of financial and nonfinancial measures” in the balanced scorecard for the division (Kaplan and Norton, 2001, p.97). Further, this understanding of cause-and-effect linkages from the strategy map is expected to reduce the coordinative complexity² associated with the multiple relationships between performance measures, and between the measures and the performance evaluation judgment (Wood, 1986). Consequently, superiors are expected to be better equipped to make decisions that utilize all the measures of the balanced scorecard – both common and unique – without needing to employ a simplifying strategy. That is, to avoid demonstrating the common measures bias.

Banker et al (2004) found that providing managers with detailed information about business unit strategies resulted in the use of more strategically linked measures (either common or unique) than non-linked measures compared with those who

² In the management literature, Wood (1986) defines coordinative complexity as the nature (form, strength and sequencing) of relationships between information cues, acts and products (p. 68). In the accounting literature, Bonner (2008) locates this type of complexity within the ‘clarity of processing’ part of processing complexity.

received no information about the strategies. They also found that “the largest contrasts between ratings differences occur in the benchmark treatment, which suggests that participants rely more on common measures in the absence of detailed strategy information” (p.18). However, half the measures in Banker et al were not strategically linked. Further, and most importantly with respect to our research aims, by comparing participants receiving no strategy information with those receiving a full set of information (both a narrative description and a strategy map diagram), Banker et al did not investigate the role of causal linkages in this setting. A summary of key differences between prior research related to Hypothesis 1 (Lipe and Salterio, 2000; Banker et al, 2004) and the current study is provided in Figure 1.

[INSERT FIGURE 1 HERE]

Vera-Munoz, Shackell and Buehner (2007) suggest that “provision of a strategy map identifying the causal model underlying the BSC may improve users’ ability to understand and manage the relations embedded in the scorecard” (p.1018). As such, it is the causal links specifically in the strategy map that are expected to enhance the quality of balanced scorecard-based performance evaluation by preventing common measures bias. There is, however, a tension between the importance assigned to causally linked strategy maps and evidence of their use in practice. A survey of the financial services industry conducted by Ittner, Larcker and Randall (2003) found that many of the firms claiming to use a balanced scorecard place little or no reliance on a formal, causal business model (76.9%). While in German-speaking companies surveyed by Speckbacher, Bischof and Pfeiffer (2003), half of the companies that had implemented a balanced scorecard were unable to formulate cause-and-effect relationships among the scorecard objectives and measures. Given this tension and the limited literature with respect to the impact of causal linkages in this setting, this study seeks to investigate whether provision of a strategy map impacts performance evaluation judgments. Consequently:

***H1:** Superiors who receive a strategy map and narrative description of strategic objectives will produce a more balanced evaluation of managers’ performance (reflecting both unique and common measures) than superiors who only receive a narrative description of strategic objectives.*

Causally-linked strategy maps and mental model development

Mental models are simplified, cognitive representations of causal relationships that incorporate information currently active in working memory – that is, both external cues currently being attended to and knowledge stored in long-term memory structures (Markman and Gentner, 2001; Hammersley, 2003). Mental model development can be impaired as cognitive load increases (Sweller, 1988; 1994). Gary and Wood (2007) found in an organizational management setting that mental model accuracy was negatively related to the level of task complexity.

A key feature of mental models is the cause-and-effect relationships between variables (Brewer, 1993; Sterman, 1994; Markman and Gentner, 2001). Given that cause-and-effect linkages are a key feature of a strategy map (Kaplan and Norton, 2004), a strategy map is expected to influence development of cause-and-effect linkages in managers' mental models. As Birnberg, Luft and Shields (2006) state "Management accounting practices can influence how information is mentally represented and linked with other information in memory" (pp. 131). As such, in the complex task environment of a balanced scorecard-based performance evaluation, providing superiors with a map depicting causal linkages is expected to aid development of their mental model of the cause-and-effect relationships in the divisional strategies. Hence:

***H2:** Superiors who receive a strategy map and narrative description of strategic objectives will have more complete mental models of divisional strategies than superiors who only receive a narrative description of strategic objectives.*

III. METHOD

Overview of experiment

The current study follows the structure of Lipe and Salterio (2000), in order to maximize comparability of the performance evaluation task with extant research in this area. However, the experimental case has been adapted from both the case developed by Lipe and Salterio (also used in Libby et al, 2004; Roberts, Hibbets and Albright, 2005; and Dilla and Steinbart, 2005) and that employed by Banker et al (2004), as in both these cases not all performance measures were strategically linked –

that is, they included some non-strategically linked measures³. As the aim of the current study is to investigate whether the causal links in a strategy map will lead to a more balanced performance evaluation judgment (reflecting both unique and common measures) when all performance measures are strategically linked, the cases employed by Lipe and Salterio and by Banker et al were inappropriate for testing this aim. A summary of key differences between this prior research and the current study is provided in Table 1.

[INSERT TABLE 1 HERE]

The current case was developed to ensure that the full set of sixteen performance measures for each division are linked to the strategy of the division, so contain information that is relevant for evaluating the strategic performance of the division. With strategic linkages between the divisional strategies and performance measures, and a narrative description of the strategic objectives of each division provided to all participants, the strategy map then only provides those participants who receive it with information on the causal linkages between objectives.

To develop this case, the authors referred to the original Kaplan and Norton (1996a) case from which the Lipe and Salterio (2000) case was developed, *Kenyon Stores*. The extract from *Kenyon Stores*' internal scorecard presented by Kaplan and Norton (1996a, p. 111) shows two broad strategic themes – one focused on increasing revenue via greater market penetration, and the other focused on increasing productivity via greater asset utilization⁴. As such, a retail case study was developed based on these two broad themes and utilizing strategic objectives and performance measures from the Lipe and Salterio, wherever possible and appropriate. The resultant strategy maps developed for each division are presented in Figure 2.

³ Banker et al's (2004) experimental design required that each scorecard contain both strategically linked and non-strategically linked measures – with each perspective containing one common/unique measure linked to the division's strategy, while the other common/unique measure was not linked to their strategy. Lipe and Salterio, as highlighted by Banker et al (2004), also included some performance measures which were not linked to the articulated strategic objectives of the division. This was not a manipulation (as was the case for Banker et al), rather the strategic objectives to which some of the performance measures related were not presented in the case information.

⁴ The latter strategic theme was only partially reflected in the Lipe and Salterio (2000) case; and was largely absent from the Banker et al (2004) case. This strategic theme is, however, consistent with the documented strategies of several major retail organizations in practice. Retailers with asset productivity-focused strategies include Just Group Limited (2007 annual report), Gap Inc. (2006 annual report), David Jones Limited (2007 annual report), and Target Corporation (2006 annual report).

[INSERT FIGURE 2 HERE]

Participants

A total of 61 Executive MBA students completed the experiment. All were enrolled in an introductory accounting course. Participants had an average of 10.4 years of full-time work experience and 66 percent were male. The major areas in which participants had work experience were: General Management 19.7%; Technology 16.4%; Accounting and Finance 14.7%; and Project Management 8.2%⁵.

Research design

The study employs a 2 x 2 x (2) mixed experimental design. The two between-subjects factors are provision of a strategy map and basis for divisional out-performance, and the within-subjects factor is division. Provision of a strategy map was manipulated at two levels, being: i) the benchmark group ('strategy map absent') – a narrative description of strategic objectives was provided in Part One (see Figure 3); and ii) the treatment group ('strategy map present') – a narrative description of strategic objectives was provided in Part One, and a strategy map communicating causal linkages between objectives was provided with the balanced scorecard in Part Two (see Figure 4).

[INSERT FIGURE 3 HERE]

[INSERT FIGURE 4 HERE]

The second between-subjects manipulation was the basis for divisional out-performance⁶ (either on common or unique measures). Lipe and Salterio (2000) and much of the subsequent research in this area (Roberts et al, 2004; Dilla and Steinbart, 2005) manipulated all combinations of divisional out-performance on common

⁵ The remaining 41% of participants had work experience in the following areas: Consulting 6.6%; Human Resources 6.6%; Marketing 6.6%; Other 16.3%; and Undisclosed 4.9%.

⁶ 'Divisional out-performance' refers to the situation where the performance of one division consistently exceeds performance of another division across a particular set of performance measures (Lipe and Salterio, 2000).

measures and on unique measures, resulting in four combinations for out-performance⁷. However, as Libby et al (2004) highlighted, the only combination of results that “allows for the possibility of common measures bias” is where “performance on common measures favors one division and performance on unique measures favors the other division” and vice versa (p. 1082). Hence, the two patterns of divisional out-performance manipulated in the current study are: i) General Jeans out-performs Captain Kids on common measures (and Captain Kids out-performs General Jeans on unique measures); and ii) General Jeans out-performs Captain Kids on unique measures (and Captain Kids out-performs General Jeans on common measures).

The within-subjects factor is division (General Jeans and Captain Kids) with participants evaluating each manager’s performance for both divisions. Potential order effects related to division were controlled for, with half the participants completing the General Jeans performance evaluation and mental model completeness task first (and having this information presented first), and the remaining participants completing the Captain Kids performance evaluation and mental model completeness task first (and having this information presented first). Potential order effects related to the presentation of the performance measures were controlled for, with the ordering of common and unique measures varied for each scorecard category. Further, for half the participants in each cell, the ordering of common and unique measures was varied within each scorecard category.

Research instrument

In Part One, participants read through a business case which asks them to adopt the role of a senior executive of Rank Retail Incorporated (RRI), a firm operating in the fashion industry. Consistent with Lipe and Salterio (2000), the case describes the balanced scorecard concept (including listing the four categories of measures) and indicates that the firm’s Chief Financial Officer recently attended a

⁷ Lipe and Salterio (2000) manipulated all possible combinations of divisional out-performance, being: i) common and unique measures were consistently better for one division; ii) common and unique measures were consistently better for the other division; iii) common measures were better for one division and unique measures were better for the other division; and iv) unique measures were better for one division and common measures were better for the other division. Banker et al (2004), Dilla and Steinbart (2005) and Roberts et al (2004) then fully replicated this design in their studies investigating methods of overcoming ‘common measures bias’.

symposium on the balanced scorecard and has decided to implement the balanced scorecard in two of RRI's major divisions. These divisions are *General Jeans*, a retailer of youth fashion apparel, and *Captain Kids*, an established children's clothing retailer. The case provides a brief overview of the two divisions and outlines the strategic objectives of each (see Figure 3).

In the second part of the experiment, participants are presented with a balanced scorecard for each division, which provides annual actual versus target performance for sixteen performance measures⁸. The approach to assigning performance targets developed by Lipe and Salterio (2000) was adopted here. That is, performance data was selected to ensure common and unique items had the same level of performance above target in total. The sum of performance above target across all common measures was 85.18 for the better division (shown in the "percent better than target" column of Figure 5) and 51.99 for the worse division; while the sum of excess performance across all unique measures was between 84.96 and 85.08 for the better division and between 51.19 (shown in Figure 5) and 51.95 for the worse division. As was the case for Lipe and Salterio, the slight variation for the unique measures resulted from different units for the measures in each division.

[INSERT FIGURE 5 HERE]

All the performance measures for both divisions are presented in Figure 6. To assess the impact of causal linkages between the strategic objectives, those participants in the treatment group also receive a strategy map (see Figure 4).

[INSERT FIGURE 6 HERE]

Based on this information, participants are required to make the same judgment as in Lipe and Salterio (2000) of the manager's performance in delivering the strategy of the first division⁹. The balanced scorecard (and strategy map, where

⁸ The study design controlled for two order effects, being presentation order of the divisions (General Jeans vs. Captain Kids) and presentation order of the performance measures in the balanced scorecard (common/unique/common/unique vs. unique/common/unique/common ordering in each section).

⁹ Specifically, participants were asked to provide a rating on a scale from 0 ("reassign") to 100 ("excellent"), with anchor points in between including "very poor", "average" and "very good". As in Lipe and Salterio (2000), the qualitative descriptors for all the anchor points were presented below the rating scale.

applicable) is then presented for the second division, and participants are asked to evaluate the performance of the manager of this division.

The third part of the task asks participants to complete a knowledge test of the cause-and-effect relationships between strategic objectives for each division (see Figure 7 for participant instructions and an excerpt from the test). The final part of the task is a short debriefing questionnaire, where demographic information was collected, a manipulation check conducted and data obtained on task difficulty.

[INSERT FIGURE 7 HERE]

Pilot testing the instrument

Consistent with Lipe and Salterio (2000), a key objective of this experiment was to compare how subjects respond to common versus unique performance measures – in this case, when a strategy map is provided. Given the revised common strategic theme and, consequently, revised performance measures in our case, it was important to test whether the common and unique measures were considered to be equally relevant to the performance evaluation judgments made by participants. For this purpose, 12 experienced Accounting faculty were asked to read Part 1 of the study (containing background information on the company and an overview of the strategic objectives of each division) and then rate each performance measure “as to its relevance for evaluating the performance of the division manager”, using a 10 point Likert scale from 1 “low relevance” to 10 “high relevance” employed by Lipe and Salterio. There were no differences in the decision relevance of common and unique measures, either in total (average scores of 8.16 and 8.24 respectively, $t = -0.255$), or individually for General Jeans (8.17 and 8.40, $t = -1.063$) or Captain Kids (8.15 and 8.08, $t = 0.168$).

Dependent variables

Two dependent variables were measured, being difference in divisional performance evaluations and mental model completeness. All participants were first asked to evaluate the performance of the managers from both divisions on a scale from 0-100, with seven descriptive labels ranging from “Reassign” to “Excellent” performance (Lipe and Salterio, 2000; Lipe and Salterio, 2002; Libby et al, 2004). Consistent with Libby et al, we measure the difference in participant evaluations for

the two divisions as the dependent variable (where a significant positive difference indicates a common measures bias).

The second dependent variable, mental model completeness, was measured using an adapted cognitive mapping approach based on that originally developed by Axelrod (1976) and used more recently in organizational research (see Ford and Hegarty, 1984; Langfield-Smith and Wirth, 1992; Markoczy and Goldberg, 1995; Swan and Newell, 1998; Hodgkinson, Maule and Bown, 2004). A cognitive map is “a specific way of representing a person’s assertions about some limited domain” and is “designed to capture the structure of the person’s causal assertions” (Axelrod, 1976, p. 55). Cognitive mapping “looks at the pattern or structure of assertions of causality among variables and not at how or why variables or their causal relations come to be part of the map”. The respondent is given a set of variables and asked to identify the nature of the cause-and-effect relationships (positive, negative or no causality) between each combination of variables. Generally, establishing a cognitive map requires participants to indicate the nature of the cause-and-effect relations between every combination of variables considered – that is, “ $n(n-1)$ comparisons must be made, where n is the number of variables” (Ford and Hegarty, 1984, p.273-4).

Gary and Wood (2007), however, adapted this approach by asking participants to simultaneously indicate *both* the nature (positive, negative or no causality) and direction (A affects B, or B affects A) for each the relationship. Consequently, participants are only required to classify $[n(n-1)]/2$ variable combinations. This approach is utilized in the current study (please refer to Figure 7). Hence, with each division’s strategy map containing eight strategic objectives, this means that participants are asked to specify the nature and direction of the relationship for 28 pairwise combinations for each division. Mental model completeness is then evaluated based on the percentage of items that subjects answer correctly regarding the complete set of causal relationships contained in both divisional strategies, once all task instruments have been taken away.

IV. RESULTS

Participants were asked to rate their agreement with the statement that: “The two business units, General Jeans and Captain Kids, used some different performance measures” (based on a 9 point Likert scale, ranging from 1: strongly disagree to 9:

strongly agree). This manipulation check showed that participants recognized that the two divisions utilize different performance measures ($p < 0.001$).

We used a 2 x 2 ANOVA to test each of our hypotheses. With respect to Hypothesis 1, the difference in divisional scores was compared for the two treatment groups. A significant difference between the benchmark and treatment groups for provision of a strategy map was not found; hence, Hypothesis 1 was not supported (see Panel B of Table 2). However, this result was due to an unexpected finding, being the absence of a common measures bias for the benchmark group, which was inconsistent with the extant research on balanced scorecard-based performance evaluation. Both the benchmark and strategy map treatment groups demonstrated no significant difference between the scores assigned to the division with higher performance on common measures and the division with higher performance on unique measures. The mean difference in divisional performance evaluations for the benchmark and treatment groups were -0.8 and -0.3, respectively (see Panel A of Table 2). Consequently, because the evaluations of the benchmark group did not demonstrate any common measures bias, it was not possible for Hypothesis 1 to be supported – that is, for participants provided with a strategy map to produce a ‘more balanced’ evaluation of managers’ performance (reflecting both unique and common measures) as hypothesized.

[INSERT TABLE 2 HERE]

One possible reason for this unexpected finding is the change in the case study used in the current study. In the extant literature in this area, the WCS Incorporated case developed by Lipe and Salterio (2000) has either been employed in full or an adapted form. In contrast, the current study revised the retail-based case in two key ways, being strategically linking all performance measures and a change in focus for the common strategic theme to asset productivity (rather than increased price mark-ups). The latter change meant participants were presented with one profitability-focused theme (increasing sales) and one asset productivity-focused theme (increasing asset utilization) for each division, where previously they were asked to consider two profitability-focused themes for the divisions. This distinction may have more clearly differentiated the two strategic themes and consequently simplified the task, resulting

in consideration of both common and unique performance measures by participants in their performance evaluations. Both these changes are discussed further in Section V.

A statistically significant difference was found in the mental model completeness score for the treatment group provided with a strategy map, as shown in Panel A of Table 3 ($F = 5.699$, $df = 1$, $p = 0.021$). Panel B of Table 3 indicates that when a strategy map was provided to participants they were able to correctly identify an average of 51.0% of the strategic linkages for the two divisions (52.5% for General Jeans and 49.4% for Captain Kids). In comparison, when a strategy map was not provided to participants they were only able to correctly identify an average of 41.4% of the strategic linkages for the two divisions (44.5% for General Jeans and 40.2% for Captain Kids). This result supports Hypothesis 2.

[INSERT TABLE 3 HERE]

V. DISCUSSION, LIMITATIONS AND CONCLUSIONS

We conducted an experiment to determine the impact that the causal links contained in a strategy map have on common measures bias in divisional performance evaluations and on the completeness of mental models developed of divisional strategies. As predicted, providing a strategy map containing causal linkages does result in performance evaluations without common measures bias. However, we did not find a common measures bias in the control group, which is inconsistent with prior research based on the case developed Lipe and Salterio (2000). We also find that providing participants with the causal linkages contained in divisional strategy maps results in a more complete mental model of divisional strategy being developed. This finding is consistent with mental model theory, which highlights the role of cause-and-effect linkages in mental model development.

The absence of common measures bias in our study does not detract from the basic finding from extant studies (Lipe and Salterio, 2000; Banker et al, 2004; Libby et al, 2004; and Dilla and Steinbart, 2005) that common measures bias does exist under particular circumstances. However, it does suggest that there are circumstances where this bias does not exist and therefore future research that tests under what circumstances the common measures bias does and does not exist would be useful. The potential differences between our study and the extant literature relate to the task, the culture (Australian and North American), the participants and the provision of

strategy information. The latter appears unlikely to explain the difference, given the strategy information Banker et al provided to treatment participants (both a narrative description of strategy and a strategy map) exceeded that provided to our control group participants (a narrative description of strategic objectives). Cultural differences between Australia and North America are also unlikely to explain the difference. For example, in the last 25 years, judgment and decision making research in an audit setting has not identified any differences. It is possible that participant experience may be a factor in the current study, given that our average experience of participants was 10.4 years, compared with greater than 5 years in Lipe and Salterio (2000) and 6.4 years in Banker et al, and this issue could be further addressed in future research. However, task effects seem the most likely explanation for the differences, as discussed below.

The case employed in our study has two notable differences from the case developed by Lipe and Salterio (2000), upon which much of the prior research in this area (including Banker et al, 2004) has been based. These task differences are the inclusion of only strategically-linked performance measures in the scorecards of the two divisions and, more importantly, inclusion of a strategic theme for both divisions focused on improving asset utilization. Each will be discussed in turn.

As outlined in Section III, both Lipe and Salterio (2000) and Banker et al (2004) included some performance measures which were not linked to the strategic objectives of each division in their respective tasks. In contrast, in the case developed for this study, all performance measures presented to superiors were linked to the divisional strategies, which were communicated to all participants. Using a full set of sixteen, strategically-linked performance measures for each division was necessary in the current study to control for the type of additional strategy information provided to the treatment group – that is, to ensure that the only additional information provided to the treatment group by the strategy maps was the causal linkages between strategic objectives. It is possible, however, that the simple act of ensuring all performance measures were linked with divisional strategy, as Kaplan and Norton suggested (2004), may have facilitated a more balanced performance evaluation (that is, without common measures bias) by participants. While Banker et al investigated the impact of detailed strategy information on the use of strategically linked performance measures, they also included eight non-linked performance measures in each

scorecard; hence, the judgmental effects of a completely strategically linked balanced scorecard were not investigated.

The second difference identified above is the inclusion of an asset utilization-based strategic theme in the current study. As discussed in Section III, the current case included both a revenue growth-based and an asset utilization-based strategic theme, adapted from the *Kenyon Stores* case. Where in the previous cases developed by Lipe and Salterio (2000) and Banker et al (2004) participants were asked to consider two profitability-focused themes for the divisions, in the current study they were presented with one profitability-focused theme (growing revenue) and one asset productivity-focused theme (increasing asset utilization) for each division. This distinction may have more clearly differentiated the two strategic themes and consequently simplified the task, resulting in consideration of both common and unique performance measures by participants in their performance evaluations.

Our experimental design is subject to certain limitations. Most significantly, as shown above, we are unable to conclude on why a common measures bias is not found in the control group in our study, as we do not employ both the WCS Incorporated case developed by Lipe and Salterio (2000) and the RRI case developed here. As discussed though, given the need to use strategically-linked performance measures to investigate the impact of the causal linkages of the strategy map (for both H1 and H2) the WCS Incorporated case was not appropriate, when controlling for the type of strategy information provided to participants.

Given the potential problems caused by the existence of the common measures bias (Libby, Salterio and Webb, 2004), the fact that our results show there are circumstances where this bias does not exist suggests there are benefits in renewed efforts to establish the circumstances where the bias exists and where it does not exist. This will allow the development of balanced scorecards that may avoid such biases. It is suggested that an important element of future research is investigating task effects. This point has been made frequently in the audit judgment and decision making research, where Trotman (2005, p. 76) notes “there appears to have been a lack of attention to tasks in audit JDM experiments. However, this lack of attention is certainly not unique to the auditing literature. For example, a similar point was made concerning group decision making research over 25 years ago.” Earlier warnings on the importance of task are provided by Ashton (1990, p. 175) who states that “while subject characteristics such as knowledge are important...task characteristics are

critical. Indeed, the essence of my argument is that the nature of the task is the key to understanding the paradoxical effects of incentive, feedback, and justification (and perhaps other pressure inducers).”

In addition to considerations related to the common measures bias discussed above, our study contributes to the literature an adapted cognitive mapping approach for measuring the mental model developed by managers for judgment purposes. Using this measure we find that participants provided with a strategy map demonstrate a more complete understanding of the causal linkages that do (and do not) lead to strategic objectives being achieved. While the absence of common measures bias in our benchmark group meant this more complete model was not required to improve participants’ performance evaluation judgments, there are other significant potential implications of this finding. With respect to the literature, this approach may be employed in situations where common measures bias is observed to better understand the cause of this bias. With respect to practice, an often cited benefit of implementing the balanced scorecard and strategy map is the facilitation of a shared understanding of organizational strategy (Kaplan and Norton, 2004), yet recent evidence has indicated that only a portion of the organizations that have implemented the balanced scorecard have actually developed cause-and-effect models of the business and its strategy (Ittner, Larcker and Randall, 2003; Speckbacher, Bischof and Pfeiffer, 2003). Given the importance of developing consistent strategic understanding across all organizational staff, our finding highlights the importance of both developing strategy maps and providing them to all employees.

REFERENCES

- Ashton, R.H. (1990). "Pressure and Performance in Accounting Decision Settings, Paradoxical Effects of Incentives, Feedback, and Justification.", *Journal of Accounting Research*, 28 (Supplement), pp. 148-180.
- Axelrod, R. (1976). *Structure of decision*. Princeton, New Jersey: Princeton University Press.
- Banker, R.D., Chang, H., and Pizzini, M.J. (2004). "The Balanced Scorecard: Judgmental Effects of Performance Measures Linked to Strategy", *The Accounting Review*, 79 (1), pp. 1-12.
- Bazerman, M. (2006). *Judgment in Managerial Decision Making*. 6th Edition. Wiley.
- Birnberg, J.G., Luft, J.L and Shields, M.D. (2006). "Psychology Theory in Management Accounting Research." In Chapman, C.S., Hopwood, A.G., and Shields, M.D. (Eds), *Handbook of Management Accounting Research*, Oxford: Elsevier.
- Bonner, S.E. (2008). *Judgment and decision making in accounting*. 1st Edition. New Jersey: Pearson Prentice Hall.
- Brewer, W. F. (2003). Mental models. *Encyclopedia of cognitive science*. L. Nadel. (Ed) London, Macmillan, 3, pp.1-6.
- Dilla, W.N., and Steinbart, P.J. (2005). "Relative Weighting of Common and Unique Balanced Scorecard Measures by Knowledgeable Decision Makers", *Behavioral Research in Accounting*, 17, pp. 43-53.
- Eden, C., and Spender, J.C. (1998) *Managerial and Organizational Cognition: Theory, Methods, and Research*, London: Sage.
- Ford, J.D., and Hegarty, W. (1984). "Decision Makers' Beliefs about the Causes and Effects of Structure: An Exploratory Study", *The Academy of Management Journal*, 27 (2), pp. 271-291.
- Gary, M.S., and Wood, R.E. (2007). "Testing the Effects of a System Dynamics Decision Aid on Mental Model Accuracy and Performance on Dynamic Decision Making Tasks." *Paper presented at 2007 System Dynamics Society Conference*.
- Hammersley, J. S. (2003). "Pattern Identification and Industry Specialist Auditors". *PhD Dissertation submitted at the University of Illinois at Urbana-Champaign*.
- Hibbets, A.R., Roberts, M.L., and Albright, T.L. (2006). "Common-Measures Bias in the Balanced Scorecard: Cognitive Effort and General Problem-Solving Ability." *Working Paper presented at 2007 AAA Conference*.
- Hodgkinson, G.P. (2003). "The Interface of Cognitive and Industrial, Work and Organizational Psychology, *Journal of Occupational and Organizational Psychology*, 76 (1).
- Hodgkinson, G.P., Maule, A.J., and Bown, N.J. (2004). "Causal Cognitive Mapping in the Organizational Strategy Field: A Comparison of Alternative Elicitation Procedures", *Organizational Research Methods*, 7 (1), pp. 3-26.
- Hoque, Z., and James, W. (2000). "Linking the Balanced Scorecard Measures to Size and Market Factors: Impact on Organizational Performance", *Journal of Management Accounting Research*, 12, pp. 1-17.

- Ittner, C. D., Larcker, D.F., and Meyer, M.W., (2003). "Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard." *The Accounting Review*, 78(3), pp. 725-758.
- Ittner, C. D., Larcker, D.F., and Randall, T., (2003). "Performance Implications of Strategic Performance Measurement in Financial Services Firms." *Accounting, Organizations and Society*, 28, pp. 715-741.
- Kaplan, R.S., and Norton, D.P. (1992). "The Balanced Scorecard – Measures that Drive Performance." *Harvard Business Review*, January-February, p. 71-79.
- Kaplan, R.S., and Norton, D.P. (1996a). *The Balanced Scorecard: Translating Strategy into Action*, Boston: Harvard Business School Press.
- Kaplan, R.S., and Norton, D.P. (1996b). "Linking the balanced scorecard to strategy", *California Management Review*, 4, pp. 53-79.
- Kaplan, R.S., and Norton, D.P. (2001). "Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part I", *Accounting Horizons*, 15 (1), pp. 87-104.
- Kaplan, R.S., and Norton, D.P. (2004). "Strategy Maps: Converting Intangible Assets into Tangible Outcomes", Boston: Harvard Business School Press.
- Langfield-Smith, K. (2006). "A Review of Quantitative Research in Management Control Systems and Strategy." In Chapman, C.S., Hopwood, A.G., and Shields, M.D. (Eds), *Handbook of Management Accounting Research*, Oxford: Elsevier.
- Langfield-Smith, K., and Wirth, A. (1992). "Measuring Differences between Cognitive Maps", *The Journal of the Operational Research Society*, 43 (12), pp. 1135-1150.
- Libby, R., Bloomfield, R., and Nelson, M.W. (2002). "Experimental research in financial accounting". *Accounting, Organizations and Society*, 27, pp. 775-810.
- Libby, T., Salterio, S.E., and Webb, A. (2004). "The Balanced Scorecard: The Effects of Assurance and Process Accountability on Managerial Judgment", *The Accounting Review*, 79 (4), pp. 1075-94.
- Lipe, M.G., and Salterio, S.E. (2000), "The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures", *The Accounting Review*, 75 (3), pp. 283-298.
- Lipe, M.G., and Salterio, S.E. (2002), "A note on the judgmental effects of the balanced scorecard's information organization", *Accounting, Organizations and Society*, 27 (6), pp. 531-540.
- Markman, A.B., and Gentner, D. (2001). "Thinking", *Annual Review of Psychology*, 52, pp. 223-247.
- Markoczy, L., and Goldberg, J. (1995). "Method for Eliciting Comparing Causal Maps", *Journal of Management*, 21 (2), pp. 305-333.
- Merchant, K.A. (2006). "Measuring general managers' performance: Market, accounting and combination-of-measures systems", *Accounting, Auditing and Accountability Journal*, 19 (6), pp. 893-917.

- Norreklit, H. (2003). "The Balanced Scorecard: what is the score? A rhetorical analysis of the Balanced Scorecard", *Accounting, Organizations and Society*, 28, pp. 591-619.
- Norreklit, H., and Mitchell, F. (2007). "The balanced scorecard". In Hopper, T., Northcott, D., and Scapens, R., *Issues in Management Accounting, Third Edition*, Essex: Pearson Education.
- Otley, D. (1999). "Performance management: a framework for management control systems research", *Management Accounting Research*, 10, pp. 363-382.
- Payne, J., Bettman, J., and Johnson, E. (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.
- Roberts, M.L., Albright, T.L., and Hibbets, A.R. (2004). "Debiasing Balanced Scorecard Evaluations", *Behavioral Research in Accounting*, 16, pp. 75-89.
- Slovic, P., and MacPhillamy, D. (1974). "Dimensional Commensurability and Cue Utilization in Comparative Judgment", *Organizational Behavior and Human Performance*, 11, pp. 172-194.
- Speckbacher, G., Bischof, J., and Pfeiffer, T. (2003). "A descriptive analysis on the implementation of Balanced Scorecards in German-speaking countries", *Management Accounting Research*, 14, pp. 361-387.
- Sprinkle, G.B. (2003). "Perspectives on Experimental Research in Managerial Accounting", *Accounting, Organizations and Society*, 28, pp. 287-318.
- Sprinkle, G.B., and Williamson, M.G. (2006). "Experimental Research in Management Accounting." In Chapman, C.S., Hopwood, A.G., and Shields, M.D. (Eds), *Handbook of Management Accounting Research*, Oxford: Elsevier.
- Sterman, J. D. (1994). "Learning in and about Complex Systems", *System Dynamics Review*, 10 (2/3), pp. 291-330.
- Swan, J., and Newell, S. (1998). "Making sense of technological innovation: The political and social dynamics of cognition". In Eden, C., and Spender, J.C. (Eds), *Managerial and organizational cognition: Theory, methods and research*. London: Sage, pp. 108-129.
- Trotman, K.T. (2005). "Discussion of Judgment and Decision Making Research in Auditing: A Task, Person, and Interpersonal Interaction Perspective", *Auditing: A Journal of Practice & Theory*, 24 (Supplement), pp. 73-87.
- Vera-Munoz, S.C., Shackell, M.B., and Buehner, M. (2007). "Accountants' usage of causal business models in the presence of benchmark data: A note". *Contemporary Accounting Research*, 24, September, pp. 1015-1038.
- Wood, R.E. (1986). "Task Complexity: Definition of the Construct", *Organizational Behavior and Human Decision Processes*, 37, pp. 60-82.
- Zhang, S., and Markman, A.B. (2001). "Processing Product Unique Features: Alignability and Involvement in Preference Construction", *Journal of Consumer Psychology*, 11 (1), p. 13-27.

TABLES AND FIGURES

FIGURE 1
Strategy information provided to participants – a comparison
of the current study with related extant literature

No strategy information	Narrative strategy information	Narrative strategy information and strategy map
Banker et al (2004): Benchmark group	Lipe and Salterio (2000): All participants – <i>Strategic objectives relate to some performance measures</i> Humphreys and Trotman (2008): Benchmark group – <i>Strategic objectives for all performance measures</i>	Banker et al (2004): Treatment group – <i>Some variation in content of narrative and map</i> Humphreys and Trotman (2008): Treatment group – <i>Strategic objectives identical in narrative and map (map adds causal linkages and balanced scorecard categories)</i>

TABLE 1
Key differences between current study and related extant literature

RESEARCH STUDY	Strategy information provided*	Strategic linkage of performance measures	Focus of divisional strategic themes
Lipe and Salterio (2000)	Some narrative strategy information	Some strategically linked	Revenue growth; and Productivity
Banker et al (2004)	No strategy information VS Narrative strategy information and strategy map	Half strategically linked; Half non-strategically linked	Revenue growth; and Productivity
Humphreys and Trotman (2008)	Narrative strategy information VS Narrative strategy information and strategy map	All strategically linked	Revenue growth; and Asset utilization

* Please refer to Figure 1 for more detail regarding this dimension.

FIGURE 2
Strategic Objectives (with causal links) for General Jeans

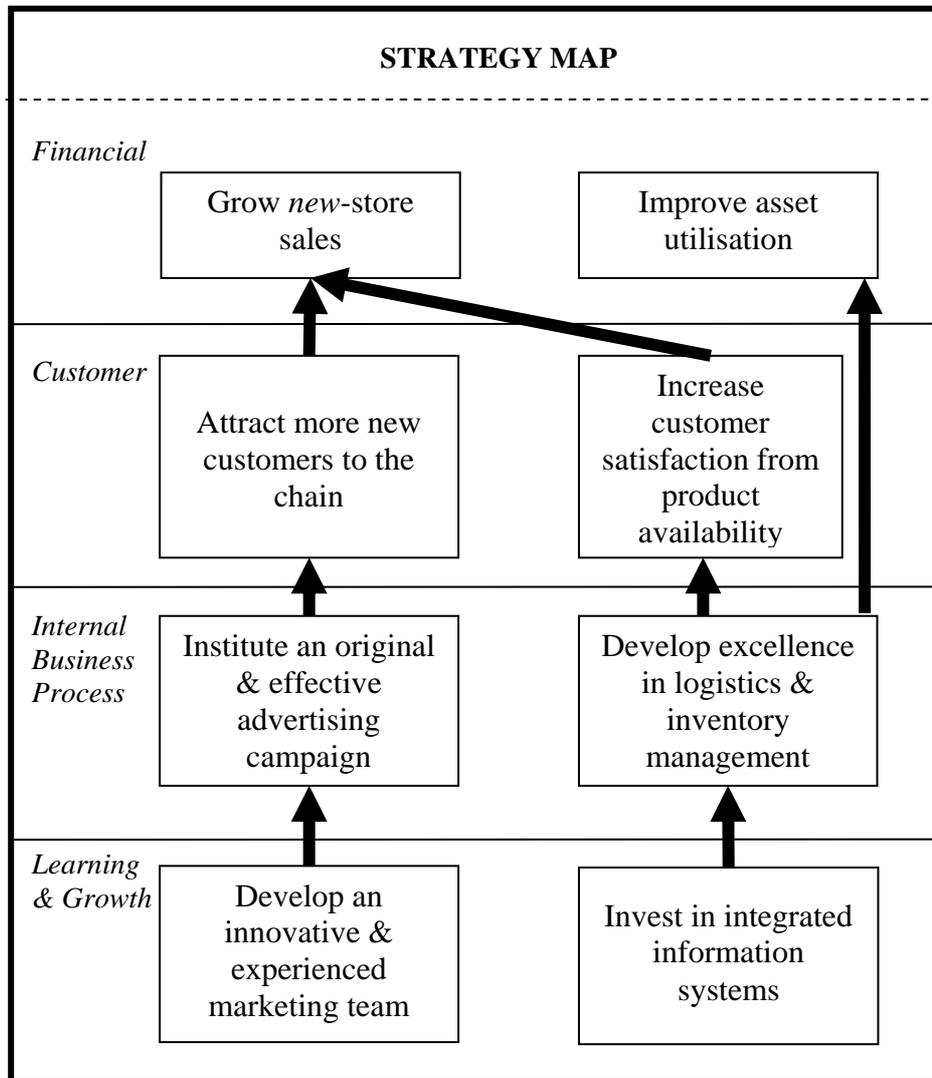


FIGURE 2 (continued)
Strategic Objectives (with causal links) for Captain Kids

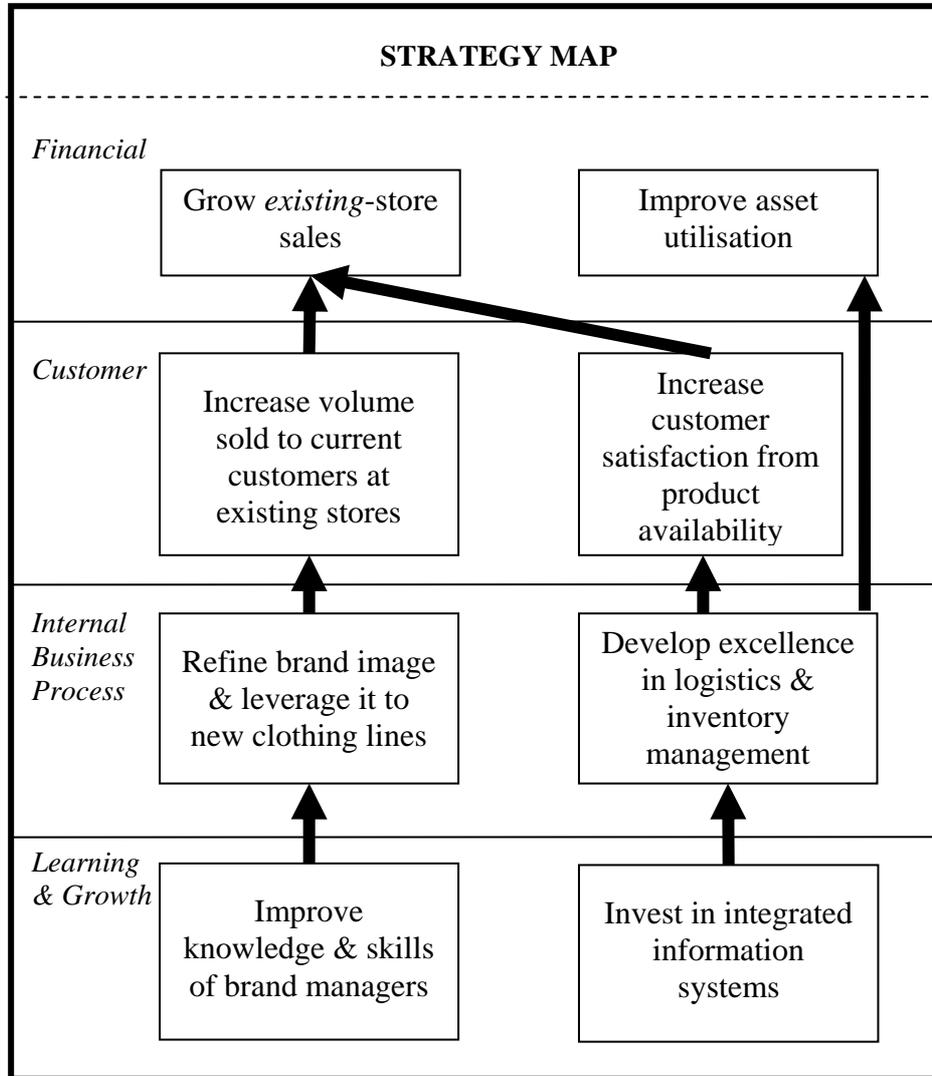


FIGURE 3
Background and Overview of Strategic Objectives for
General Jeans and Captain Kids

General Jeans

General Jeans is a retailer of youth fashion apparel – it specialises in providing basic, reasonably priced, casual clothing (including jeans and t-shirts) for teenagers and young adults.

The division is investing in information systems integration and the development of an innovative and experienced marketing team, and is ensuring its assets are effectively utilised. It is also opening new stores, and is focused on growing sales from these stores.

General Jeans' management has determined that to be successful it must develop excellence in logistics and inventory management, as well as original and effective advertising campaigns. The division aims to attract more new customers to the chain, and satisfy their clothing needs with greater product availability.

Captain Kids

Captain Kids is an established children's clothing retailer, specialising in fashionable apparel for infants and children.

The division intends to introduce new clothing lines, by leveraging its distinctive brand image. It also aims to satisfy its customers through greater product availability, making the existing stores 'one-stop shops' for the current Captain Kids customer base of style-conscious, time-constrained parents to purchase in greater volumes.

Captain Kids' management has determined that to be successful it must develop excellence in logistics and inventory management. It is concerned with effectively utilising its assets and growing sales from its existing stores. At the same time, the division is investing in building the knowledge and skills of its brand managers, and in information systems integration.

FIGURE 4
The Strategy Map and Balanced Scorecard for General Jeans Employed in the Experiment

BALANCED SCORECARD

STRATEGY MAP		MEASURE	TARGET	ACTUAL	PERCENT BETTER THAN TARGET
<p><i>Financial</i></p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;">Grow <i>new-store</i> sales</div> <div style="border: 1px solid black; padding: 5px; width: 45%;">Improve asset utilisation</div> </div>		<p><i>Financial</i></p> <ol style="list-style-type: none"> Asset turnover (times) 12.0 13.0 8.33 Percentage of sales from new stores 30.00% 31.32% 4.40 Inventory turnover (times) 7.0 7.6 8.57 Sales growth per new store 16.00% 16.96% 6.00 			
<p><i>Customer</i></p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;">Attract more new customers to the chain</div> <div style="border: 1px solid black; padding: 5px; width: 45%;">Increase customer satisfaction from product availability</div> </div>		<p><i>Customer</i></p> <ol style="list-style-type: none"> Sales to new customers 34.0% 36.7% 7.94 'Mystery shopper' product availability rating (/10) 7.5 8.5 13.33 New relationships with target customers 120 123 2.50 Customer satisfaction rating 92% 95% 3.26 			
<p><i>Internal Business Process</i></p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;">Institute an original & effective advertising campaign</div> <div style="border: 1px solid black; padding: 5px; width: 45%;">Develop excellence in logistics & inventory management</div> </div>		<p><i>Internal Business Process</i></p> <ol style="list-style-type: none"> Out of stock ratio 30% 25% 16.67 Advertising campaign awareness rating (/10) 8.0 8.9 11.25 Correct deliveries from suppliers 80.0% 92.5% 15.63 New advertising campaign awards 25 27 8.00 			
<p><i>Learning & Growth</i></p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;">Develop an innovative & experienced marketing team</div> <div style="border: 1px solid black; padding: 5px; width: 45%;">Invest in integrated information systems</div> </div>		<p><i>Learning & Growth</i></p> <ol style="list-style-type: none"> Awards won by marketing team recruits 49 53 8.16 Information systems projects delivered on time 72.0% 81.6% 13.33 Retail experience of new marketing managers (years) 17.0 17.5 2.94 New IT investment (\$m) 66 70 6.06 			

FIGURE 5
The Balanced Scorecard for General Jeans Employed in the Experiment ^a

MEASURE	TARGET	ACTUAL	PERCENT BETTER THAN TARGET
<i>Financial</i>			
1. Asset turnover (times)	12.0	13.0	8.33
2. Percentage of sales from new stores	30.00%	31.32%	4.40
3. Inventory turnover (times)	7.0	7.6	8.57
4. Sales growth per new store	16.00%	16.96%	6.00
<i>Customer</i>			
1. Sales to new customers	34.0%	36.7%	7.94
2. 'Mystery shopper' product availability rating (/10)	7.5	8.5	13.33
3. New relationships with target customers	120	123	2.50
4. Customer satisfaction rating	92%	95%	3.26
<i>Internal Business Process</i>			
1. Out of stock ratio	30%	25%	16.67
2. Advertising campaign awareness rating (/10)	8.0	8.9	11.25
3. Correct deliveries from suppliers	80.0%	92.5%	15.63
4. New advertising campaign awards	25	27	8.00
<i>Learning & Growth</i>			
1. Awards won by marketing team recruits	49	53	8.16
2. Information systems projects delivered on time	72.0%	81.6%	13.33
3. Retail experience of new marketing managers (years)	17.0	17.5	2.94
4. New IT investment (\$m)	66	70	6.06

^a A similar scorecard was prepared for Captain Kids, the other division in the experiment.

FIGURE 6
Common and Unique Performance Measures for
General Jeans' and Captain Kids' Balanced Scorecards

Type	Measure
<i>Financial Perspective</i>	
Common	Asset turnover (times)
Common	Inventory turnover (times)
Unique – General Jeans	Percentage of sales from new stores
Unique – General Jeans	Sales growth per new store
Unique – Captain Kids	Average sales growth per store
Unique – Captain Kids	Percentage of existing store sales from new clothing lines
<i>Customer Perspective</i>	
Common	'Mystery shopper' product availability rating (/10)
Common	Customer satisfaction rating
Unique – General Jeans	Sales to new customers
Unique – General Jeans	New relationships with target customers
Unique – Captain Kids	Average sale per customer
Unique – Captain Kids	Repeat sales
<i>Internal Process Perspective</i>	
Common	Out of stock ratio
Common	Correct deliveries from suppliers
Unique – General Jeans	Advertising campaign awareness rating (/10)
Unique – General Jeans	New advertising campaign awards
Unique – Captain Kids	Brand recognition rating
Unique – Captain Kids	New clothing lines introduced
<i>Learning and Growth Perspective</i>	
Common	Information systems projects delivered on time
Common	New IT investment (\$m)
Unique – General Jeans	Awards won by marketing team recruits
Unique – General Jeans	Retail experience of new marketing managers (years)
Unique – Captain Kids	Average hours of brand manager training
Unique – Captain Kids	Brand manager skill rating (/100)

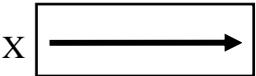
FIGURE 7

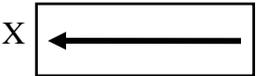
Extracts from the Mental Model Completeness test Employed in the Experiment

Panel A: Participant instructions

Please think about the relationships between the variables that form part of the respective strategies of General Jeans and Captain Kids. **Relying only on the Retail Rank Incorporated case study you have read today**, define the relationship between each pair of variables.

You can define each relationship in one of the following three ways:

- 

1. This arrow indicates that *a change in X* leads directly to *a change in Y*.
- 

2. In contrast, this arrow indicates that *a change in Y* leads directly to *a change in X*.
- 

3. If there is no direct relationship between the pair of variables, write 'NONE' between the two variables.

Please ensure to focus only on **DIRECT** relationships – if a variable only has an indirect effect on another variable, you should write 'NONE'.

Panel B: Extract from General Jeans' test

1	Marketing team innovation & experience		Advertising campaign originality & effectiveness
2	<i>New-store sales</i>		Marketing team innovation & experience
3	Asset utilization		New customers

TABLE 2
Experimental Results for Managers' Performance Evaluations

Panel A: Mean Evaluations (Standard Deviations) of the Performance of General Jeans and Captain Kids Division Managers by Experimental Condition

<u>Divisional Out-performance</u>		<u>Strategy map absent</u>	<u>Strategy map present</u>
Common Measures	Favor General Jeans	73.1 (10.8) ¹⁰	74.5 (11.4) ¹¹
	Favor Captain Kids	74.0 (11.2) ¹²	72.4 (7.3) ¹³
	<i>Average:</i> General Jeans and Captain Kids	73.5 (10.8)	73.5 (9.4)
Unique Measures	Favor General Jeans	74.8 (10.7)	74.5 (8.3)
	Favor Captain Kids	73.9 (12.6)	72.9 (7.9)
	<i>Average:</i> General Jeans and Captain Kids	74.3 (11.5)	73.7 (8.0)
Difference^a	Average Common <i>less</i> Unique	-0.8 (9.5)	-0.3 (9.3)

^a A significant positive difference indicates a common measures bias

Panel B: 2 x 2 ANOVA of the Differences in Managers' Performance Evaluations by Condition

Factor	<u><i>df</i></u>	<u><i>SS</i></u>	<u><i>MS</i></u>	<u><i>F</i></u>	<u><i>p-value</i></u>
StrategyMap	1	52.734	52.734	0.580	0.450
OutPerformance	1	16.276	19.276	0.179	0.674
StrategyMap x OutPerformance	1	4.401	4.401	0.048	0.827
Error	56	5092.500	90.938		

¹⁰ n = 16 – however, one participant has been excluded, as they did not respond to this Part of the study.

¹¹ n = 15

¹² n = 15

¹³ n = 15

TABLE 3
Experimental Results for Mental Model Completeness

Panel A: Mean Evaluations (Standard Deviations) of Mental Model Completeness by Experimental Condition

	Strategy map absent ¹⁴	Strategy map present ¹⁵
General Jeans	44.5% (4.4)	52.5% (4.6)
Captain Kids	40.2% (3.6)	49.4% (5.2)
TOTAL Mental Model Completeness	41.4% (7.2)	51.0% (9.1)

Panel B: 2 x 2 ANOVA of Mental Model Completeness by Condition

Factor	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p-value</i>
StrategyMap	1	385.421	385.421	5.699	0.021
OutPerformance	1	8.119	8.119	0.120	0.730
StrategyMap x OutPerformance	1	124.950	124.950	1.848	0.180
Error	53	3584.250	67.627		

¹⁴ n = 31 – however, four participants have been excluded, due to incomplete responses provided.

¹⁵ n = 30